AI IN THE FINANCIAL SECTOR

Israel's Interagency Taskforce on AI in Finance

Final Report

December 2025

Table of Contents

Preface 3
Executive Summary 6

Part A – Background, Developments, and Regulatory Trends 30

A. The Entry of Artificial Intelligence into Our Lives 31
B. Regulation of Artificial Intelligence in International Organizations and Developed Countries 38
C. Government Engagement on Artificial Intelligence in Israel 42
D. AI applications in the Financial Sector 46
E. Benefits and Challenges Arising from the Use of Artificial Intelligence 49
F. Artificial Intelligence in Financial Regulation Worldwide 55
G. AI Activity in the Financial Sector in Israel 60
H. The Inter-Ministerial Taskforce for Artificial Intelligence in the Financial Sector 62
1. Background to the Taskforce's Establishment 62
2. The Taskforce's Work 63
3. The Interim Report 65

Part B – General Approach, Guiding Principles, and Definition of Artificial Intelligence in the Financial Sector 66

A. The General Approach – Encouraging Innovation and Mitigating Risks 67
B. Guiding Principles for AI Use in the Financial Sector 69
C. Defining Artificial Intelligence 75

Part C – Cross-Cutting Issues 83

"Black Box" and Explainability 85
Human in the Loop 93
Notification and Disclosure 102
Privacy and Protection of Personal Information 113
Bias and Discrimination 157
Liability and Accountability 171
AI Governance 185
"Connecting the Dots" – A Comprehensive & Risk-Based View 198
Financial Stability 205
Competition 211
Operational Risks Including Cybersecurity Risks, Third-Party Risks, Fraud and Disinformation.....223

Part D – Specific Areas Examined by the Team 237

Investment Advice and Portfolio Management 238
Credit in the Banking System 269
Insurance Underwriting 282

Part V – Additional Actions to Promote Financial Regulation in the Field of Artificial Intelligence 291

Appendices 308

**** This translation is provided as a non-binding and for convenience only. In the event of any discrepancy or inconsistency, the Hebrew version shall prevail ****

Preface

Recent developments in artificial intelligence (AI) have attracted unprecedented attention, given their potential to transform economic and social activity. In particular, advances in generative AI and large language models (LLMs), widely accessible since late 2022, alongside expectations regarding artificial general intelligence (AGI), have underscored the potentially significant near-term impact of these technologies.

The use of AI applications has expanded rapidly. Investment and research activity have increased materially, alongside growing efforts to integrate AI into a broad range of business processes. As a result, the opportunities and risks associated with AI, and the appropriate regulatory response, are now central to policy discussions globally.

AI is not new to the financial sector. However, recent technological advances are expected to accelerate its adoption across a wider range of use cases. These include core decision-making functions in credit, insurance, and investment; customer-facing activities such as needs assessment, product matching, and service delivery; and internal processes including risk management and fraud detection.

If its potential is realized, AI may deliver significant benefits to the public, including improved quality of financial products and services, reduced costs, increased competition, and enhanced accessibility. At the same time, its deployment raises a range of risks and challenges. These include novel issues, such as limited transparency in how AI systems operate and generate outputs, as well as the amplification of existing risks, including discrimination and manipulation.

An inter-agency taskforce was established to examine the use of AI in the financial sector and to address these opportunities and challenges at an early stage. The taskforce comprises representatives from the Ministry of Justice, the Ministry of Finance, the Banking Supervision, the Capital Market, Insurance and Savings Authority, the Israel Securities Authority, and the Competition Authority. Its work reflects the view that a coordinated, cross-authority approach is required to assess the implications of AI and support appropriate policy development.

The taskforces's work builds on prior government activity in this area. This includes Government Resolution No. 212 (2021) and Government Resolution No. 173 (2023), which addressed the development of a national AI strategy; a 2023 policy paper issued by the Ministry of Innovation, Science and Technology and the Ministry of Justice, setting out a regulatory approach to AI; and a joint report by the Ministry of Justice and researchers from Tel Aviv University on AI applications in the financial sector. Consistent with this policy framework, the team adopts a sector-based approach to regulation (rather than a cross-cutting model), aligned with international developments and based on flexible, principles-based tools intended to support innovation. The present work applies this approach to the financial sector.

The taskforce's activities included analysis of underlying AI technologies, assessment of their applications in financial services, examination of legal and regulatory implications, and engagement with stakeholders. This engagement included a public consultation and discussions with representatives from industry, academia, government, and international regulatory bodies.

The taskforce's overall approach is to support the responsible adoption of AI in the financial sector, while ensuring that regulatory frameworks remain effective. Although the deployment of AI raises complex challenges, and much of this report addresses these challenges, the underlying premise is that they can be managed appropriately. In this context, existing regulatory frameworks, together with targeted adaptations over time, are considered capable of supporting innovation while maintaining key supervisory objectives, including consumer protection and financial stability.

The report also reflects the view that, at this stage, regulatory intervention should remain flexible and proportionate. Policymaking in this area involves balancing competing considerations: promoting innovation and capturing the benefits of AI, while addressing associated risks; providing regulatory clarity, while accommodating rapid technological change; and responding to growing interest in AI, despite its still-evolving adoption across the financial sector in Israel and globally.

A comparative review indicates that most jurisdictions have not yet established dedicated regulatory regimes for AI in the financial sector. Instead, authorities are continuing to assess the implications of AI while applying and clarifying existing regulatory frameworks. At the same time, in certain jurisdictions, most notably the European Union, horizontal AI regulation has been introduced and will also apply to financial services. In parallel, many financial regulators have issued statements clarifying how existing rules apply to AI.

Against this backdrop, the taskforce's recommendations focus on identifying key issues, analyzing specific use cases, and setting out policy directions and next steps, rather than proposing detailed prescriptive regulation at this stage.

Given the pace of technological change, regulatory approaches will need to evolve over time. The taskforce observed that recent policy papers and research have quickly become outdated in light of rapid developments in LLMs and general-purpose AI. Regulators are also expected to deepen their understanding as practical applications expand. This report should therefore be regarded as an initial framework for ongoing work in this area.

Structure of the Report

Part A outlines the background to the establishment of the team and reviews key developments and regulatory trends relating to AI generally, and in the financial sector in particular. It also summarises relevant government initiatives and international regulatory approaches.

Part B sets out the overarching approach and guiding principles for addressing AI-related challenges in the financial sector. These include a risk-based approach, technological neutrality, and regulatory

consistency across similar activities and risks. This section also discusses the definition of AI and the challenges associated with it.

Part C examines cross-cutting issues arising from the use of AI applications and presents the team's recommendations. These include issues linked to the distinctive characteristics of AI, such as limited explainability ("black box") and the role of human oversight, as well as broader legal considerations, including discrimination and privacy. The section also addresses AI governance and risk management, and considers system-wide implications, including effects on competition, financial stability, and risks related to fraud, cyber threats, and disinformation. For each issue, the report outlines key questions, international approaches, potential policy responses, and recommendations.

Part D focuses on three specific areas: investment advice and portfolio management, credit, and insurance underwriting. In each case, the relevant financial regulators set out potential use cases, the applicable regulatory frameworks, key issues, and preliminary policy considerations.

Part E sets out additional measures to support the integration of AI within financial regulation. These include tools to promote innovation and regulatory certainty, as well as the potential use of AI by regulators.

The taskforce thanks the members of the steering committee for their guidance and support: Adv. Meir Levin, former Deputy Attorney General (Economic Law – Counseling and Legislation); Mr. Amit Gal, Head of the Capital Market, Insurance and Savings Authority; Adv. Sefi Zinger, Chair of the Israel Securities Authority (and Ms. Anat Guetta, former Chair); Mr. Danny Hahiashvili, Supervisor of Banks (and Mr. Yair Avidan, former Supervisor); Adv. Michal Cohen, Commissioner of Competition; and Adv. Assi Messing, former Legal Advisor to the Ministry of Finance.

The taskforce also acknowledges the contributions of those involved in drafting the report and supporting its preparation, as well as participants in the team's discussions, reviewers, and stakeholders from industry, academia, civil society, and regulatory bodies in Israel and abroad, whose input informed this work.

The taskforce further thanks Prof. Assaf Hamdani, Prof. Kobi Kastiel, Prof. Dan Amiram, and Dr. Ari Achiaz (Tel Aviv University) for their foundational academic work and ongoing support.

Team Members

Adv. Amir Wasserman • Israel Securities Authority • Co-Head of the taskforce

Adv. Sarit Felber • Ministry of Justice (Economic Counseling and Legislation) • Co-Head of the taskforce

Adv. Tom Goldreich • Israel Securities Authority • Taskforce Coordinator

Ms. Ilanit Madmoni • Bank of Israel (Banking Supervision)

Adv. David Koppel • Ministry of Finance

Ms. Aviva Weiss • Capital Market, Insurance and Savings Authority

Adv. Eyal Shapira • Competition Authority

Executive Summary

The Benefits, Risks, and Current Status of AI in the Financial Sector

The rapid development of Artificial Intelligence (AI) technology has expanded and accelerated the discussion surrounding its possible regulation. Alongside assessments that the deployment of AI will be no less than revolutionary for many aspects of life, including in finance, for which significant advantages are anticipated, serious concerns have also taken root regarding the entry of this new technology.

The primary potential benefits from the integration of AI into the financial sector lie in the its ability to increase operational efficiency and lower costs, improve financial products and services, advance inclusion and financial accessibility, and enhance compliance.

Key criticisms of the introduction of AI to the financial sector draw attention to sweeping challenges, which are not unique to the financial sector, such as the "black box" and explainability problem, the need for human involvement, effective disclosure of activity, information and privacy protection, and the prevention of discrimination and bias in the system's operations. In addition to the above, several other challenges are specific to the financial sector, such as systemic risks to financial stability, cyber risks, fraud and disinformation, and the risk of impairing competition in the field.

Engagement in AI regulation poses several key questions to policy-makers pertaining to, among other things, the appropriate timing to impose regulation applying to AI technology, the character and stringency of that regulation and whether the regulation should be integrated into existing regulation in various areas of life or if dedicated AI regulation is required instead.

Various approaches with respect to these issues are developing throughout the world – starting with the European Union, with comprehensive legislation entered into force regulating AI development, operation and use (the EU Artificial Intelligence Act), and extending to countries like the United States and the United Kingdom, where to date, no mandatory federal legislation has been established, and instead a softer approach has been adopted, which favors encouraging AI activity while allowing regulators to act in acting on various issues within the purview of their existing authorities.

The approach in Israel as presented in a government policy statement, is that regulation of AI shall be on an industry basis (as opposed to sweeping regulation encompassing AI technology in all its manifestations), treading softly by incrementally adopting regulatory measures, where and when necessary.

The general state of affairs with respect to the adoption of AI applications in the financial sector in Israel, is in line with findings in other countries, according to which, the technology raises great interest, but the pace of adoption is slow and cautious. The assessment of participants in this field, however, is that AI will play an ever-growing role in financial activity, an assertion that finds expression in the

investment in increasingly greater inputs to examine potential AI applications and integrating them in operations.

In our paper and its recommendations, the task force took into consideration: the government's general policy regarding AI regulation, the pace of AI technology adoption in the financial sector at this time, the need for caution in policy-making so as to not unintentionally create barriers to entry, and the fact that in most of the world, rigid and specific regulation has not been set yet with respect to AI in the financial services industry. In light of the above, the task force focused on analyzing key issues encountered in determining the contours of regulatory policy and in presenting essential measures that must be addressed in the future, as opposed to drafting recommendations for stringent regulation at this time. This report is only a starting point for dealing with the subject. There will be a need to regularly update the above regulatory concepts and tools as future developments unfold.

Guiding Principles for the Treatment of AI in the Financial Sector

Generally speaking, the approach proposed by the task force is to encourage innovation in order to realize the benefits engendered in AI, while addressing the risks. Without derogating from the many concerns voiced worldwide regarding the implications of AI technology, the task force is of the opinion that in the financial sector, an appropriate response to these risks is possible, given that the industry is already closely regulated. Additionally, the task force believes that the integration of AI in the financial sector is almost unavoidable, and that a significant potential for improving financial services, lowering their costs and enhancing competition in the industry, can be realized for the benefit of consumers.

The report discusses in detail nine guiding principles pertaining to the character and manner of implementing the desired AI regulation in the financial services industry:

- **Flexible and adaptive regulation**;
- **Alignment with International regulation,** ensuring coherence with globally accepted regulatory frameworks;
- **Promotion of innovation and technological integration**, *inter alia* through the removal of barriers obstructing market development;
- Promoting regulatory tools enabling **experimentation and learning**;
- Imposing new **regulatory measures only in places where the circumstances so warrant** (as opposed to a default situation in which AI use necessarily triggers regulatory change);
- **Risk-based regulatory approach**;
- Incorporating **consumer protection, social and human rights considerations** into the regulation of AI in finance;
- **Regulatory uniformity** with respect to its application to similar services and risks;

- **Technological neutrality** - regulation derived from the nature of the activity regardless of the technology involved, unless the activity otherwise warrants special treatment.

Defining Artificial Intelligence

The task of defining AI is in itself a complex issue, and one can see a broad array of attempts to do so throughout the world. In our case, the purpose of the definition is regulatory, designed to address the needs of regulation and supervision, and as such, it carries important practical implications.

In Israel, as of now, no legally binding definition for the term Artificial Intelligence has been set. The report includes an analysis of the key difficulties in determining a definition of AI, including an examination of definitions given abroad and a discussion of the various issues pertaining to a possible definition, such as whether there is a need to refer to a specific technology in the definition, or whether and how to distinguish between AI systems and more "traditional" computerized systems?

The definition of AI in the financial sector proposed by the task force is based, *inter alia*, on the definitions adopted by the OECD and the European Union, with the necessary adjustments:

An AI system is a machine-based system that operates with varying levels of autonomy, and designed to generate output such as content, predictions, recommendations, or decisions, that may impact investors, customers, or the activity of the regulated firm.

Alongside this basic definition, additional definitions can be modularly determined, in order to address the difficulties of overreach underreach of the definition. In this manner, a definition can be determined that includes the element of materiality, a definition of AI that features machine learning, or a definition of generative AI systems or general-purpose AI.

In this context, it is proposed that, in accordance with regulatory guidance, a financial entity may determine that a computer-based system does not, by its nature, constitute an artificial intelligence system. This discretion may enable addressing situations in which it is clear that the system in question is not an artificial intelligence system, even though the definition may formally apply to it.

Key Challenges and Considerations for AI Integration in Financial Institutions

The entry of AI presents challenges stemming from, among other things, the technology's unique characteristics, such as the "black box" issue, the fact that the technology's applications are still in their initial stages and that the technology's maturity is limited. In addition, the current regulatory framework focuses on human discretion, and has not been adjusted to algorithmic decision-making. Within the framework of examining these issues, the report includes proposed boundaries with respect to AI activities and tools, which are designed to ensure safe and responsible usage of the AI systems in finance. The guidance offered in this part of the report is directed at the regulated entities seeking to operate an AI system and for the regulators charged with supervising this activity.

"Black Box" and Explainability

The term **"explainability"** refers to the ability to provide explanation, rationale, or justification for the output generated by the AI system. The focus on the need "to explain" AI activities and output stems from the fact that these systems, in their advanced form, are characterized as a **'black box'**. Given its complexity, the scale of the data on which they rely, and the manner in which relationships between the data are formed, it is not possible to reliably track the manner in which the system arrives at its outputs. The report distinguishes between **general explainability**, which refers to the ability to understand the system's features and generally how it works (this type of explainability is sometimes referred to in international reports as transparency), and **specific explainability,** which refers to the ability to provide an explanation for the way in which a specific decision has been made. There are various justifications for regulatory demands for explainability – general and specific alike - with respect to the parties involved: the financial institution, the regulator, the consumer and the public. Explainability serves regulatory, consumer protection and social interests. For example, explainability regarding the manner in which the AI system made a decision pertaining to extending a person credit, shall facilitate the financial institution's use of the system as part of its internal control system, enable regulators to verify that the regulated entity complies with legal requirements, and will allow consumers to understand the decision made in their particular case and to challenge it, if necessary. It will also facilitate public scrutiny of the overall system. The demand for explainability, however, could face hurdles stemming from, among other things, the technology's level of maturity or from the imperative to protect trade secrets. A sweeping explainability requirement, i.e., a specific explainability requirement with respect to each and every decision (or even a requirement to have specific explainability capabilities) could preclude the very introduction of advanced AI technology. Recently, particularly large models have begun to provide a form of explanation or a list of key sources underlying their decision-making process, with the aim of enabling users to understand (even if not fully) the process that took place within the system when generating the output they received. This may assist in understanding the model’s decision-making process

Explainability – key recommendations:

- **Specific explainability should not be mandated in each and every case involving an AI system.** The considerations in determining whether such an obligation should apply include: whether there is a reasoning obligation under law; the degree to which the decision is important for the individual or the financial institution (high-risk decisions); and the degree to which AI is involved in the decision-making.
- **An explainability requirement (general and specific) should be considered with respect to high-risk systems.** In this context, the nature of the decision should be vetted, focusing on decisions adversely affecting the customer, including cases in which a customer requests to study

the decision or challenge it (for example, a rejection versus acceptance of a loan request). To the extent that the AI system does not provide a material component in the process of generating the recommendation or decision, explainability will not be required with respect to its manner of operation**.**

- **Even when a specific explainability requirement should presumably be imposed, "compensatory" measures should be recognized –** such as tight controls on the system's output or service alternatives offered to the client in lieu of the AI system. **Given proper compensatory measures, specific explainability shall not be required.**
- **In certain cases, partial individual explainability may suffice**. Technological developments in this field seek to achieve individual explainability (even if not in full). Such a partial explanation may, in many cases, be valuable in understanding the rationale or the components that led to the output. Accordingly, the financial regulators may adjust the required level of explainability according to the circumstances and, in some cases, suffice with partial individual explainability.
- **Specific explainability shall not be required in cases in which the technology serves solely to assist decision-making.**
- **It is important that the financial institution operating an AI system has an understanding as to how it operates, its limitations, etc.**

Human In the Loop

Human intervention, involvement and supervision over algorithmic decision-making is one of the solutions most commonly proposed to address the inherent challenges of AI activity. Human involvement is perceived as a possible means to address potential malfunctions and failures of an algorithmic system, particularly during the current stage in which the ability of AI systems to make precise, correct and safe decisions is questionable. Another key justification for human intervention requirements focuses on concerns regarding the loss of individual autonomy, in which an individual "loses" control over decisions made with respect to them, *inter alia*, given the inability to communicate with or understand the party that made the decision. This is in addition to the system's inability to consider always characteristics within a comprehensive context. Consequently, it is argued, there is a need to set an obligation to provide alternatives which include human involvement in certain cases.

Lastly, it could be asserted that an AI system activity without human involvement in areas in which professional licensing is mandated, such as investment or retirement savings advice, may violate the law and licensing requirements, and that without human involvement, it is impossible to achieve the objectives of the duty of loyalty towards the individual or to ensure threshold qualification standards under the law are upheld.

Alongside these justifications, human involvement requirements have also attracted criticism. One key criticism is that there is no certainty that human involvement will actually improve the system's output. Another criticism pertains to the uncertainty surrounding the division of liability between the human and the decision-making system. Finally, the imposition of mandatory human involvement raises concerns that it will eliminate or reduce the advantages engendered in the integration of automated systems.

Human in the loop – key recommendations:

- Human involvement can be applied in real time (for example, through examining an appeal by a client) or in retrospect. **"Real- time" human involvement will be required only for decisions material to an individual and entail a high-risk, provided that the technology plays a substantial role in the decision-making process and in the absence of adequate compensatory alternatives.** This is in order to prevent blanket requirements that can generate costs and raise barriers to entry. The considerations presented above with respect to explainability – positive versus negative decisions, decision-supporting systems as opposed to decision making systems – are also appropriate within the context of the human involvement issue. As for other decisions, human involvement can be considered when challenging or examining decisions made by an AI system. This recommendation is also relevant to generative AI, for example, when there is significant interaction with customer service.
- **In AI systems that support human decision-making (as opposed to systems that make decisions independently and automatically), it is possible to reduce requirements for human involvement in specific decisions and instead focus on oversight of the system's operation and its outcomes,** given that a human actor ultimately makes the decision.
- **A blanket** requirement to provide human involvement alongside activities carried out by artificial intelligence systems should not be imposed. At the same time, where there is an option to receive a human decision or service, this should be brought to the customer's attention so that they may choose whether they prefer a human or algorithmic alternative.
- There is a basis for requiring human oversight over the system's operation as a whole. Such a requirement is intended to ensure the proper functioning of the system in aspects such as compliance of decisions with applicable law and regulation, the absence of errors and biases, etc. This does not mean that continuous oversight must be human at every stage, but rather that human involvement is also required.
- **The obligation to deploy systems that allow control and human involvement is imposed on the regulated financial entity.**

Notification and Disclosure

One of the regulatory obligations discussed with respect to the integration of AI systems in the provision of various services and products is the requirement to notify the very use of an AI system in the process, i.e., notification that an AI system is being used. The notification requirement could, among other things, provide the consumer the option to choose the manner in which he receives the service, and can inform the regulator as to the type and scope of usage.

Other than notification about the use of an AI system, specific regulatory disclosure requirements may also be imposed with respect to financial products and services. In this context, the regulatory requirements could incorporate a wide range of possibilities, both with respect to the contents of the disclosure (for example, general or specific content, content designed to primarily reflect risks and warnings or content designed to reflect advantages, etc.) and with respect to the format of the disclosure (for example, disclosure included in the client's agreement with a firm, within a periodic disclosure, standard disclosure or prominent disclosure, etc.).

A third aspect of disclosure requirement is the reduction of the misleading information surrounding AI deployment. Complex new technologies could constitute fertile ground for publications glorifying the parties using them, as a marketing tool. Mandatory disclosure requirements could reduce this risk by mandating truth in such publications.

In the financial world, disclosure requirements are one of the most commonly implemented regulatory techniques, which can be found with respect to virtually all services and products offered to customers. It is not surprising therefore that in the vast majority of leading documents dealing with regulation and ethics pertaining to AI worldwide, the establishment of disclosure requirements, transparency or explainability is recommended. In this context, the task force distinguished three dimensions of disclosure: basic notification about the use of an AI system in products and services (for example, notification in a bank's customer service that a chat-bot interaction is not conducted with a human being, but rather with an AI system); disclosure to an investor or customer pertaining to the AI system's operating features (for example, the scope of AI involvement in a credit decision, the system's limitations, the existence of service alternatives, etc.); and disclosures dealing with misleading marketing information regarding artificial intelligence.

Notification and disclosure – Key recommendations:

- **Mandatory notification regarding the very use of an AI system, particularly in the technology's initial entry phase, is warranted.** The notification requirement at this stage is necessary not only because the technology's efficiency has yet to be determined, but also because services are still being rendered without using AI systems and customers have the option to choose. **The notification must be worded in clear and neutral language**, which does not create an ungrounded preference for one service over another. **The notification requirement takes on**

added importance where a customer may mistakenly believe that they are interacting with a human, while in practice the interaction is conducted with an artificial intelligence system. There is no need, however, for a notification requirement in cases in which the use of an AI system is obvious. There will be a need to re-examine the requirement once the technology's initial entry stage has ended.

- **Disclosure requirements regarding the AI system's characteristics and implications should be integrated into existing disclosure requirements regarding the service or product offered to customers.** In this context, there is a need to consider disclosure requirements when the AI system materially influences the service or product, and respectively, the disclosure requirements should deal with the characteristics unique to AI technology. An example of such disclosure requirements is presented in the report regarding algorithmic portfolio management, which deals with the essence of the service and its unique characteristics, the potential risks stemming from it and the system's methodology.
- **Restrictions on misleading marketing publications should be imposed –** should a phenomenon of misleading marketing information ("AI washing") develop**,** it can be dealt with using the currently available regulatory tool box, including staff bulletins and directives addressing the disclosure required of regulated entities, enforcement actions, and financial education actions undertaken towards customers and investors.

The Right to Privacy and Personal Data Protection

AI technology is based on extensive information collection and processing, including personal data, i.e., information pertaining to an identified or identifiable person. This data provides the "fuel" or "oxygen" of AI systems. It is the primary essential resource propelling their development and operation. As such, the development and use of AI systems entails the endeavor to build a base of information, which is as comprehensive, varied and far-reaching as possible, whether this stems from an aspiration to generate precise output or from the desire to develop innovative products which do not necessarily rely on "standard" information. Additionally, the output of AI systems – the predictions, recommendations and even decisions – are directed in many cases at humans, and serve as means to predict behavior, identify patterns of human behavior, personality traits, mental and physical condition, etc. Consequently, the output contains specific personal insights, sometimes highly sensitive, into various aspects of an individual's private life. These features of AI systems pose significant challenges and tension pertaining to the protection of the right to privacy and private information, throughout the life cycle of these systems, including the implementation of the key tenets of privacy protection law, such as the principle of adhering to the intended purpose of the data collection and minimizing the amount of information collected and processed. The report includes the following: a review of the relevant arrangements in

Israeli privacy protection laws; a discussion focusing on key challenges, while applying examples from the financial world; a comparative analysis of issues arising worldwide, particularly as reflected in EU regulation dealing with the processing of personal data (the General Data Protection Regulation (GDPR)); and recommendations with regards to each challenge.

Privacy and personal data protection – Key recommendations:

- **The distinction between personal and anonymous data** – privacy protection laws apply to personal data, i.e., information pertaining to an identified or easily identifiable person. The question of whether it is possible to identify a person based on certain data, including data from which obvious identifying components have been removed, becomes even more complicated with the emergence and development of AI systems, which upgrades the capacity for the re-identification of anonymous data. This engenders significant risks to privacy and personal data protection.

  It is proposed, *inter alia*, to consider directing regulated entities to perform a risk assessment regarding the re-identification of personal data and to undertake data protection measures with respect to "anonymized" data. It is also proposed that financial regulators, after consulting with the Privacy Protection Authority, conduct a comparative analysis of common and advanced private data anonymization methods as a tool to mitigate re-identification.

- **Informed consent for data processing in AI systems –** The Privacy Protection Law prohibits the violation of a person's privacy unless with his consent or by virtue of the law, and stipulates that consent must be an "informed consent". The digital environment already makes the possibility to confer genuine "informed" consent for the processing of personal data and for the purpose of that processing difficult. The implementation of an informed consent requirement as a precondition for processing personal data involving a violation of privacy for the sake of the development, training and use of data in AI systems presents an additional significant challenge. This derives from the AI systems' possible mode of operation as a "black box", and from the difficulty in presenting to the individual the purpose for which his personal data will be used, which actions will be taken with the data and precisely which data about him will be used for these purposes. Furthermore, the consent "binds" the agent processing the data to the permitted usage, which is not always known at the time the consent is given. Another issue that arises pertains to the use of personal data collected within the framework of previous service contracts, that have lapsed, and did not include consent to process data for the purpose of developing or training AI systems ("the inventory problem").

  Most of the task force's recommendations in this regard deal with measures to strengthen informed consent in cases in which personal data is used in an AI system, to the extent that

the activity is high-risk, such that its purposes of use are more complex or deviate from the reasonable expectations of the data subject. **Alongside these recommendations, it is proposed to require additional legal foundations for processing personal data for the purpose of AI system development and training**, in a manner facilitating technological development on the one hand and providing protection for the data subjects on the other. With respect to legacy data ("the inventory problem"), it is proposed to clarify that the need to obtain renewed consent will depend on the nature of the consent originally given, its purpose, and the activity or service proposed to be carried out, and it should not be inferred that the mere use of AI systems necessarily requires obtaining new consent from every customer.

- **Inferred personal data from AI system outputs –** Artificial intelligence systems significantly expand the ability to infer new data about subjects, and aim, among other things, to identify unexpected patterns of behavior or correlations that were previously not evident to the human eye. Improved computing and processing capabilities, in tandem with the availability of big data and the development of AI technologies, dramatically improve the ability to glean rich and nuanced insights about people, including with respect to personality and behavioral traits, ethnic origins, sexual preferences and opinions, as well as predictions regarding future behavior. These capabilities pose an additional challenge to privacy protection, to the question of whether to classify AI outputs as "personal data" and whether it is necessary to designate additional obligations and rights with respect to this type of information or if proper adjustments can be made in the implementation of the law on this matter.

  The assumption adopted in the report is that AI system outputs pertaining to or directed at a specific individual constitute private data under existing law, which is entitled to the protection of privacy protection law. In the Report, it is recommended to adopt a rule with respect to reasonable inferred personal data, which would apply with regards to high-risk decisions and would meet standards of acceptability, relevance and reliability.

- **Implementation of the adherence-to-purpose principle** – this basic principle of privacy protection law stipulates that private data can be collected and used only for the purpose for which it was collected or provided. The questions regarding the implementation of this principle for AI systems pertain – much like the problems described above regarding consent requirements – *inter alia,* to the stipulation that the purpose of processing the data must be clear, detailed and explicit, and the issue regarding the use of the existing "inventory" of data collected for other purposes. An additional issue pertains to the scraping of information appearing online for the purpose of developing and training AI systems, in a manner that could deviate from agreed boundaries or reasonable user expectations, and be used contrary to the adherence-to-purpose principle.

It is proposed that during the systems' development stage, the level of detail of the purpose for the processing of private data and its use shall be in accordance with the assessment of the system's perceived risks. The said explanation must be more detailed and clearer as the purpose of the private data's use becomes increasingly complex or as it deviates from the client's reasonable expectations. During the systems' use, the assumption is that the difficulty to clearly specify the purpose of the data's use diminishes as a matter of course. **It is also proposed to consider relief in the implementation of the adherence-to-purpose principle in order to develop AI systems subject to restrictions and control mechanisms, pertaining, *inter alia*, to personal data protection, and to examine the possibility of regulating through a directive the conditions and circumstances in which data can be scraped from the Internet in a manner that does not violate the provisions of the Privacy Protection Law or any other law.**

- **Data minimization principle –** A core principle of privacy protection, according to which only the essential relevant information should be collected for processing, while the collection and processing of surplus information (i.e., information deviating what is necessary to serve the original purpose) is prohibited. An inherent conflict exists between the data minimization principle and the development of AI systems that are based on as broad an array of information as possible.

  The recommendations seek to examine how to implement the data minimization principle with respect to AI systems; for example, by emphasizing the minimization of the personalization of information, while other information will not require the same manner degree of minimization, or by undertaking organizational and technological measures that enable the cleaning and deletion of data which is no longer necessary. It is also proposed to check ahead of time the limitations placed on the use of certain types of information for the development of AI systems in certain fields (for example, restrictions on the use of facial images) and it is also proposed to create a separation between the data that is collected and used, to store it separately, and to define retention periods, access permissions, and the like, in accordance with the evolving needs of each stage.

In addition to the recommendations regarding the above challenges, **it is proposed to weigh the intention of regulated entities to adopt organizational tools to manage and regulate the privacy risks inherent in the development of AI systems, such as conducting a privacy impact assessment, implementing and integrating privacy-by-design concepts, and this in addition to appointing someone specifically charged with privacy protection.**

Bias and Discrimination

A key challenge with respect to AI-based systems is the discrimination and bias risk. While decision-making by algorithms as opposed to human beings may mitigate the human predisposition towards unconscious or unintentional bias, AI systems are still exposed to bias or discrimination risk for various reasons: the use of databases which are not adequately representative, or in which the data reflects existing social biases and discrimination, as well as algorithms that may make use of discriminatory affiliation data (such as ethnicity, gender, race, etc.) or of "proxy" data correlated with discriminatory data. The broad scope of AI system operations greatly increases the risk of these phenomena, relative to a specific human decision.

The report reviews the law applicable to the prohibition of discrimination in Israel as well as the treatment of potential discrimination of AI-system involvement in the regulation of financial operations abroad. The concern for bias and discrimination can arise, for example, with respect to the extension of credit or the pricing of insurance policies. Regulators abroad have published positions making clear the need to ensure fairness and non-discrimination when using AI systems.

Bias and discrimination – Key recommendations:

- **Existing non-discrimination laws shall continue to apply even when using AI systems. The actions and decisions of parties are subject to the prohibition of discrimination without regard to the technology they use, and that is the case with AI systems as well.**
- The use of AI systems may generate inequality risks, *inter alia* given concerns that the infringement of equality may be latent, unconscious or difficult to pinpoint. **It is proposed that financial regulators consider actions to direct the regulated entities subject to the authorities given to them.** Key issues in this context include placing controls and taking measures to identify and prevent biases throughout the system's entire lifespan, particularly regarding the results reached by the system.
- **Placing controls and taking measures throughout the entire lifespan of the system, include the following key aspects –** the databases used for training, the validation, and testing of the systems must be representative and diverse**;** actual activity must be examined in order to detect discriminatory treatment towards different populations with similar data**.**
- **Outcome-based oversight –** it is proposed to employ outcome-based oversight tools, inter alia, for the purpose of examining claims of discrimination (or the absence thereof). Such tools may assist supervised entities and regulators in identifying discrimination in systems whose operation may, on its face, appear neutral and without the use of discriminatory data. This is not intended to replace the use of other oversight tools, but rather to supplement them**.**

Liability and Accountability

Legal regulation is generally based on the responsibility of a human party to fulfill his obligation, and examines his behavior and discretion in the case of a violation claim. In the age of AI, the role of the human agent is reduced and at times, even ill-defined, disrupting the familiar legal structure for assigning liability, which is historically based on finding the person or entity central to an event and holding them responsible for their actions. The issue of liability becomes more acute given the automated and autonomous features of the AI systems, as well as the fact that at times, numerous systems and factors are involved in generating the output used by the public. Accordingly, for example, questions arise as to the behavioral standard expected of the AI system, such as to whether awareness or belief can be attributed to a computer system.

Various approaches can be taken when placing liability between the various parties involved in AI system operations. The system provider, deployer and end-user are all at issue in this context. The report includes an analysis of the considerations for assigning liability to the parties involved in the development and deployment of AI, as well as a comparative review of approaches used abroad with respect to assigning responsibility for AI systems.

Liability and accountability – key recommendations:

- **The principle assigning liability to the regulated entity should be maintained.** There is no place to deviate from the accepted principle of financial regulation, under which legal liability is attributed to the regulated entity. Within the contractual framework in which the regulated entity arranges its activity *vìs a vìs* third parties, the regulated entity can determine that it is entitled to compensation from the third party in cases in which damage was sustained (for example, as a result of the failure of a computer system the regulated entity acquired from the third party), however the legal liability towards the customer remains with the regulated entity.
- **It is proposed to consider the engagement of a financial institution with a third party with respect to AI systems to be outsourcing and to implement the relevant regulation accordingly.** Outsourcing is a well-established, regulated practice, for financial institutions, that does not detract from the liability placed on them to fulfill their regulatory obligations under all the laws and provisions applicable.
- **Under certain circumstances, financial regulation imposes specific liability on certain organs or executives of the financial institution,** while generally the overall responsibility for the actions of a regulated entity rests with its board of directors. Under the above approach, and in keeping with regulation overseas, **it is proposed not to deviate from these liability rules with respect to the deployment of AI systems –** as they pertain to these systems' integration in the financial institution, to how these systems are used and to the relevant risk management

assessments. None of this detracts from the possibility that additional specific liability will be applied under AI governance provisions that will be determined in the future (see below).

- In light of the developments occurring worldwide with respect to the obligations imposed on AI system providers, **in the future there will be a place to consider whether regulation applying to AI providers shall be taken into account either as a consideration or a criterion for operating AI systems by the financial firms deploying such systems.**
- **Prohibition of limiting liability towards customers –** it is proposed, as a rule, to prohibit a financial institution from removing or limiting its liability towards customers regarding the services it renders through an AI system. The above does not remove any customer liability in cases in which the circumstances so warrant, for example, in cases in which the customer knowingly took an investment risk which materialized during the course of using an AI-based service, or where the system was knowingly misused or used by the customer in a manner that falls outside the bounds of reasonable use. In addition, nothing in the foregoing is intended to prevent the supervised entity from providing clarifications regarding reasonable limitations of the AI system.

AI Governance

The term AI governance describes the framework for mitigating the challenges and reducing the risks inherent in AI use. The role of AI governance rules is to create the appropriate mechanisms for the management and control over AI in an entity, while regulating decision-making processes, risk management, and the control and supervision of operating AI systems in organizations. The principles-based AI governance rules constitute a means of corporate risk management and contribute to an entity's resilience and stability. They are designed to "translate" the risks and challenges involved in using AI systems to methods, processes and practical measures that will achieve responsible and fair use of AI systems.

The report lays out a line of subjects and measures for an AI governance framework of financial institutions, based on best practices being developed abroad. The measures are presented as a tool box for regulators and regulated entities, which can be used under various considerations and circumstances; for example, the nature of the service rendered through AI systems, its materiality, potential harm to individuals, etc. As stipulated above, the measures should be applied only after examining existing governance arrangements and assessing the need for establishing these measures or integrating them in existing governance frameworks (in full, in part, or the combination thereof). The measures outlined in this chapter do not constitute an exhaustive list, and, unless and until otherwise determined by the financial regulators, they are not binding

The AI governance tool box in the report is presented according for the following categories:

Procedures, policy statements, processes and practices – dealing with the entire life cycle of an AI system;

Decision-making and liability procedures – including the board of directors' and executive management's liability related to aspects of the adoption, supervision, monitoring and use of AI systems;

Monitoring, supervision, validation – prior to launching the use of an AI system, and regularly thereafter;

Emergencies and discontinuation of activity – measures to discontinue the system's operations and to back up the system;

Outsourcing and third-party suppliers

Data governance is tangential to AI governance. This area takes on added significance with respect to AI systems, for which operations and outputs largely depend on information and data fed into them. The report refers to possible tools, similar to those recommended for AI governance, pertaining to **the sources of information and use of information** fed into AI systems.

Risk-based approach

The treatment of risks arising from AI systems should be undertaken with a comprehensive outlook. This stems from the fact that the various risks could be intertwined, and the means to address these risks could be similarly intertwined as well. The principle of risk-based regulation pertaining to AI appears as a common thread in policy documents worldwide and in Israel and its advantages are clear: enhancement of regulatory effectiveness and the efficiency of resource allocation by both regulators and the regulated entities, given the focus on material risks; flexibility in adapting the regulation to varying risks arising over time and in specific cases; encouragement of business activity and innovation assuming that a risk-based regulation does not over-burden the regulated entities.

A salient example of AI regulation using a risk-based approach is the EU's Artificial Intelligence Act. This legislation characterizes four categories of AI systems risks, and the restrictions imposed are a result of this ranking – ranging from no regulation for "minimal-risk" applications to prohibition for unacceptable-risk applications.

The task force's key recommendations with respect to the risk-based approach:

- As stated above, the general concept is that regulatory treatment of AI systems should take a risk-based approach.
- Multi-tiered risk grading is a useful tool for classifying the level of risk assigned to financial sector AI systems and for determining the appropriate measures to address the risk. Classification of AI systems can be based on a **two-tiered risk scale – low-medium and high.**

- **Risk classification in financial institutions should be conducted on a dual basis:** first, the risk level for **customers, investors or the public** and second, the risk level for the **financial institution**. Risk classification should, first and foremost, be carried out by the supervised entities. This does not preclude a financial regulator from requiring that certain systems be classified as high-risk, or from requiring a supervised entity to change the classification level and, accordingly, the measures required in respect of a particular system.
- **The measures taken with respect to an AI system shall be adapted to its risk level.** For example, with respect to high-risk systems, more significant governance requirements can be imposed in order to address inherent risks, while for lower risk systems, more limited requirements may suffice.
- **The report includes an illustration of how to treat the various issues arising from the operation of AI systems in the financial sector based on risk classification**.

Issues Pertaining to the Overall Impact of AI Use on the Financial Market

The introduction of AI to the financial sector could have implications for the entire market. The impact of AI in this context does not relate solely to the institutions seeking to operate AI systems, but to the entire financial system, as it pertains from a prudential perspective to financial stability, the market's competitive structure, and systemic operational risks, including disinformation risks. This part of the report is designed primarily for regulators and the rest of the policy-setters engaged in financial sector regulation.

Financial stability

Addressing financial stability risks is high on the list of regulatory priorities, in order to safeguard the public's savings, maintain the financial system's proper functioning and to prevent financial crises that could harm the entire economy. The recognition of the importance of prudential risk has become more acute, following the global financial crisis of 2008, which illustrates the need to take a broad perspective in the assessment of financial systems (as opposed to examining specific components or institutions within the system), while identifying trends and risks that may materially affect the system's functioning. Following the crisis, attention to and controls over systemic risk has taken root worldwide, as is the case in Israel, in which the Financial Stability Committee was established to support the stability of the financial system and its proper functioning.

Prudential risks resulting from the entry of AI are usually attributed to the following aspects as described in the report:

Herding and volatility – the use of AI models by a large number of players in the financial system could lead to herd behavior, i.e., massive similar activity by myriad players.

Concentration and third-party dependency – a concentration of financial activity by a small number of parties possessing the information and technological capabilities for using AI, gives rise to concerns over an increase in the risk of operational or other failings and to dependency that is likely to develop on a small number of AI companies.

Interconnectedness and contagion – these risks pertain to the concern that shocks incurred by an institution or a certain portion of the financial system will quickly spread to other institutions and parts in the system.

The above risks are joined by the traits of AI systems, such as the "black box" issue which complicates the analysis of the system's activity and, as a consequence, serves to increase its inherent risk, or the risk of over-reliance on AI systems and the potential widespread use of them relative to regular systems. The fact that these systems are able to make a large number of decisions in a short time – which may accelerate the pace and spread of crises - must also be taken into consideration. A number of examples of extreme situations and the potential risks resulting from growing AI use in the financial sector are presented in the report.

Although the issue of prudential risk was raised by many countries and organizations abroad, other countries, in the meantime, appear to be content with monitoring and examining the potential implications of artificial intelligence on financial stability.

Financial stability – key recommendations:

Given the limited adoption of AI applications and the absence of concrete measures in other countries, the task force is of the opinion that at this stage, **there is a need to continually monitor this issue, while the need to take action should be revisited in the future.** Actions can include: a mapping of activities potentially entailing significant risk; an examination of the existing framework to address financial institutions with operational and cyber risks; and the formulation of a strategy to reduce systemic risks derived from reliance on third parties critical to the functioning of the financial system.

In Israel, the operation of the Financial Stability Committee can also serve to continue the inquiry into the implications of AI use on the financial system's robustness, as well as its ability to contend with systemic risks to financial stability.

Competition

Artificial Intelligence has the potential to enhance productivity in the financial services industry, to drive innovation among existing and new players and to facilitate the creation of new products and services beneficial to the public.

Alongside the positive impact that AI may have, the report analyzes the competition problems which may arise along the AI supply chain, including the suppliers of know-how and the inputs for developing

AI models, the players active in developing and designing these models, the business users which integrate the systems in work processes, products and services.

The key competitive concerns pertain to both AI model development and to the players that will be able to make use of these models. In the field of AI model development, concerns arise, *inter alia*, with respect to barriers to access to needed inputs, a concern as to the reinforcement of the dominant position held by large technology companies through AI models and a concern that "winner-takes-all" or "winner-takes-most" markets could develop as a result. With respect to the financial players that will use the models, the fact that the financial sector is already highly concentrated and laden with switching barriers also sparks concern that existing economies of scale and scope could become even more consequential.

Competition – key recommendations

- **The application of competition law even where anti-competitive practices are executed through AI models or resulting from their operation should continue.** To illustrate, preventing competitor access to key inputs essential to developing AI models by a monopoly, could constitute a violation of the Economic Competition Law, 5748- 1988.
- **The Competition Authority should, within its vetting of mergers brought to it, consider concerns of the consolidation of market power or the erection of entry barriers through AI-** for example, mergers which create market power on the basis of informational advantages on customers held by the merging parties.
- **Access given to financial institutions to aggregate anonymous information should be expanded.** In order to develop AI models, access to a great amount of information over time is necessary, and for this reason, it is proposed to facilitate this access, for example, with respect to the database of credit information or with respect to data that can be made accessible by virtue of the Open Banking Reform. The Bank of Israel has reservations regarding this recommendation.
- **It is proposed to examine ways to prevent the consolidation of market power in the AI service provision market.** Some possibilities that may be considered in this context include setting regulatory provisions according to which a single AI service provider will be prohibited from providing service to players cumulatively holding more than a certain market share of a specific market, or an obligation that players in a certain market or in a market with certain traits, be required to work with more than one provider of AI services.

Operational risks including cyber, third-party, fraud and disinformation

The development and availability of AI tools contribute to the increase of operational risks threatening financial institutions, including cyber risks, third-party risks, fraud and disinformation.

AI tools may be less expensive, more accessible and more effective against attackers seeking to harm a financial system or the public.

Among the operational risks that may materialize as a result of the introduction of AI to the financial sector, the task force thought it appropriate to dedicate a special section to discuss **disinformation risks** that currently raise considerable concern. It has therefore conducted a more detailed analysis of the implications of such risks for the financial sector.

Manipulation and the systematic undermining of information, often referred to as "disinformation" or "false information", are fairly common phenomena in the digital age; the objective of which is to influence the content disseminated online or through other digital means, and as a result, influence the decisions made based on that information.

The risk of disinformation is perceived as one of the significant risks inherent in AI development. It could be particularly relevant for the financial services industry, since the financial decisions made by the general public, as well as market trading and the activity of large financial institutions, are dependent on large amounts of frequently received reliable information and on the public's trust of financial institutions and the financial system overall. Consequently, these areas of activity (like other areas of activity based on receiving and processing of vast amounts of information) are particularly vulnerable to the phenomenon of disinformation, and the materialization of large-scale disinformation risk could not only harm the investors' and clients' financial resources, but could also impinge on systemic stability.

The report presents areas in which special regulatory attention should be paid: financial activities in which quick decisions may be made by the general public; financial activities characterized by information-based continuous pricing; financial institutions exposed to contagion and interconnectedness risks; financial institutions that could make use of AI systems for rapid financial transactions; and the use of AI to bypass security mechanisms.

Operational risks – key recommendations:

It is proposed to map out the activities that generate particular exposure to **disinformation risk** and to check if existing tools suffice to address this risk, or if there is room for improvement. Among the regulatory tools proposed for consideration in this regard, for the purpose mapping risks and addressing them:

- **directing regulated entities to assess disinformation risk**, as part of their cyber risk assessment, **and the adequacy of existing tools and controls at their disposal to address this risk;**
- **determining mechanisms or testing existing mechanisms to address disinformation risks in the relevant institutions,** such as directives and procedures dealing with cyber risks among

regulated entities, or the circuit breaker mechanism with respect to trading activity on the stock exchange;

- **increasing public awareness of disinformation risks**, through public explanations, warnings and other financial education activities;
- **instituting proactive preparedness procedures for financial regulators in the handling of disinformation events** (like the proactive preparedness in emergency cases).
- **adding disinformation risk materialization scenarios to the procedures for handling** cyber risks and emergency preparedness as well as **to the drills carried out by the financial regulators and the financial system with respect to emergency situation protocols**.
- **improving investigation powers with respect to disinformation events resulting from AI activities.** Particular importance is placed on this with respect to securities trading that could be particularly vulnerable to these events. In this context, there is a need to continue advancing proposed legislation currently under consideration aimed at assisting the enforcement of securities offences under the Penal Law and the Computers Law.
- **examining the determination of powers in general law to remove false content and referral to the competent authority within this context.**

Selected Areas of Financial Activity

The task force examined three specific areas of activity – investment advice and portfolio management, credit in the banking system and insurance underwriting. Each one of the financial regulators presents the manner in which it views potential applications of AI in the area within its purview, what it considers to be the relevant regulatory framework, major issues and preliminary recommendations.

Investment advice and portfolio management

Artificial Intelligence may have many and varied applications in investment advice and portfolio management activities. These applications, which are reviewed in the report, can encompass virtually all professional activity related to providing investment advice or managing customer finances, as well as the additional business activities of licensees, such as marketing, customer relations or setting up compliance mechanisms.

The Israel Securities Authority (ISA) has been called upon in past years to deal with algorithm-assisted investment advice and portfolio management. Within the framework of the Directive to Licensees on the Provision of Services Using Digital Tools, the ISA determined how AI can be incorporated in the three stages of the service – assessing client needs and entering into a service agreement; determining investment policy and the client's risk profile; and providing regular service. The directive addresses, *inter alia*, issues of proficiency, controls, liability, explanations to clients and human involvement.

The report includes a review of this directive and the relevant comparative law, as well as an analysis of the advantages and risks involved in integrating AI in investment advice and portfolio management activities. The key advantages include the extending and increasing the accessibility of these services to greater numbers of clients; the expected cost reduction for service provision; and the potential for service providers to enhance the services provided and improve the processes in their provision. The key risks include the failure to meet the duty of trust and care and the materialization of potential conflicts of interest; the emergence of gamification risk; concerns regarding the impingement of service quality; and concerns of industry concentration and reliance on a small number of systems.

Investment advice and portfolio management – key recommendations:

- **Active encouragement of the integration of AI technology among investment advice and portfolio manager licensees in order to increase the number of customers receiving these services –** *inter alia*, through individual action towards existing and potential players, the publication of a manual encouraging and explaining how to implement the service, and by taking additional measures to promote regulatory certainty and innovation.
- **Updating the Directive on the Provision of Services Using Digital Tools.** It is proposed to consider adding a chapter to the directive specifically dealing with current AI technology clarifying the role of the appointed licensee (charged with providing the technological services), considering instructions dealing with AI systems that are not operated in client interfaces, as well as instructions on the use of chatbots (see below).
- **Examination of chatbot activity in the provision of investment advice and portfolio management as well as the regulation applied.** The report explains the regulatory challenges stemming from this activity, in light of the development of Large Language Models.
- **Clarification of the law with respect to general-purpose AI systems,** that is to say, with respect to systems designed to serve various purposes, including investment advice and portfolio management.
- **Promotion of research examining customer interaction with computerized systems.** Research pertaining to the interaction of customers using AI systems and automated decision-making is necessary in order to develop the market, understand the effects of technology on investor behavior and assess existing and the required regulation regarding this activity.
- **Regulation of brokerage services.** Given the close connection between brokerage and investment advice and portfolio management activities, as well as the potential extensive use of AI in brokerage activity, there is a need to regulate directly this segment of the securities industry to facilitate its development and ensure comprehensive regulation of it.

Consumer credit

Artificial Intelligence applications can serve a large array of activities in the area of bank credit, such as client needs assessment, credit underwriting, credit rating determination and regular credit management. The assumption is that the use of AI constitutes the natural evolution of existing statistical models in the field of credit.

The report describes the field of credit and the regulation currently governing it, potential AI applications in this field in the Israeli market, the implementation of existing regulation of these, and the issues emerging in them.

Consumer credit – key recommendations:

- **Use of existing risk management and consumer protection regulation –** it is proposed to adhere to existing regulation where it is suitable for adequately dealing with the challenges of AI, along with clarifications that will lend certainty to the market from the standpoint of explainability, discrimination, and similar issues. It is proposed to adhere to the accepted risk management approach for supervision of this field. Existing provisions with respect to consumer protection, risk management and model management provide an adequate regulatory framework, even though clarifications from the regulator may be required with respect to factors increasing and lowering risks and the tools to handle them, as detailed in chapter on risk management in the report.
- **Discrimination and exclusion risks** – Pursuant to the challenges pertaining to discrimination discussed above, there are specific provisions addressing this in the credit field, for example with respect to data that may or may not be used in determining credit ratings. It is proposed to consider releasing a clarification that the use of AI does not detract from existing law, and that financial institutions must promise, prior to using any AI instrument, that they have the means to ensure compliance to the regulation of this field – such as the non-use of proxy variables for data, which under credit data law, are prohibited from consideration.
- **Excessive lending concerns** – one of the concerns described in the report is the use of technology for the aggressive marketing of credit to customers. In this context, it is proposed to consider mandatory disclosure regarding the use of AI for this purpose, which would serve customers and the regulator alike.

Insurance underwriting

Underwriting is the process in which insurance companies assess the risk of certain insurance coverage. Key factors in this process include vetting the client's characteristics, the characteristics of the insurance coverage acquired by it, pricing the policy premium and the characteristics of the reinsurance. The underwriting process is designed to verify that the premium fits the insurance risk, and that it is

conducted using actuarial models based on statistical tools and algorithms, such that AI technology will likely be very useful to the process.

The report describes the insurance underwriting process, which is divided into the construction of the overall underwriting model and the underwriting of an individual client or transaction, along with the potential use of artificial intelligence within it. The issues that arise regarding this process are also widely discussed in the report, including black box risks and lack of explainability, privacy, discrimination, model risks, and information system and cyber risks. The tools to address these challenges include risk management processes, which are part and parcel of the work of insurance companies, as well as additional tools discussed in the report, such as AI governance, and mandatory notification, and disclosure.

Insurance underwriting – key recommendations:

- **Use of existing risk management and consumer protection regulation –** as a rule, existing regulation is technologically neutral, and consequently, an insurance company electing to adopt AI should develop and implement the technology according to existing regulation.
- **Updating regulation where the technology requires special attention –** key areas for which the need to update regulation should be examined include model risk management, explainability, notification and disclosure (for example notification about the use of AI as a means of contact with customers using a chatbot). Additionally, the supervision and controls placed on insurance companies and the tools at the regulator's disposal - for example, whether the supervisory tools used in vetting the entry of a new insurance product suffice to examine the decisions made in the AI system.

Additional Actions to Promote Financial Regulation with Respect to AI

The final chapter of the report does not deal with a certain issue related to AI activity (explainability, discrimination, and others) or with a particular type of activity (investment advice, credit and others), but rather with structural actions required to make financial regulation with respect to AI use advanced and effective. The general recommendations in this chapter focus on three goals: encouraging innovation; promoting regulatory certainty and legal adjustments; and enhancing supervision.

The task force's key recommendations:

Encouraging innovation

- Updating the legislation proposed by the government on "regulatory sand boxes" for the financial sector, which establishes dedicated uniform regulation for all financial regulators to create expand a testing environment, and facilitate its future expansion to enhance its potential effectiveness.

- Alongside advancing the said law, it is proposed to establish independent "sand boxes" by each financial regulator in its area of responsibility, and to initiate legislative amendments to the relevant financial laws to enable this in cases in which existing law does not allow for it.
- Establishing innovation hubs by the financial regulators, while dedicating resources and focusing efforts in the near term on AI applications.
- Examine possible solutions for problems of information accessibility, particularly those encountered by small or new entities, such as the establishment and opening of data bases as a basis for AI activity.

Promoting regulatory certainty and legal adaptation

- Regulatory certainty with regard to AI should be augmented through tools such as policy papers, pre-ruling responses, interpretative position statements, Q&A, publications of manuals and warnings.
- Regulation in primary and secondary legislation should reflect the principle of technological neutrality and permit flexibility in operating through various technological means. In addition, and without derogating from the existing authority held by financial regulators, the report proposes considering the grant of regulatory powers to the competent authorities to issue directives as necessary to address technological changes, to grant exemptions and to determine, for the purpose of adapting the law to activity and enhancing legal certainty, determine temporary and sunset provisions where necessary.

Promoting AI-based supervisory activity:

- Promote supervisory activity based on AI technology (Suptech), as part of the regulator's comprehensive strategy pertaining to IT systems and their deployment for supervisory and enforcement purposes.
- Allocate government resources, outside the regular fiscal framework, to promote AI-based supervisory activity.
- Consider updating powers to receive information for analytic, supervisory and enforcement purposes.
- Provide systematic and centralized treatment of sweeping issues related to the use of artificial intelligence by the state and public authorities and design a format for uniform treatment.
- Harness resources assigned to financial regulators, encouraging coordination between them and continuing collaborative work on the subject.

Part A

Background, Development,

and Regulatory Trends

A. The Entry of Artificial Intelligence into Our Lives

While it seems that artificial intelligence has burst into our lives only recently, the engagement in attributing human capabilities to machines, including human intelligence, is rooted in the 1950s.[1] The term "Artificial Intelligence" (AI for short) was apparently coined for the first time by a group of American researchers in the framework of proposed research at Dartmouth College in 1955.[2] Even then, the researchers assumed that every form of learning or expression of human intelligence can be described in a manner that enables a machine to imitate it.[3]

Over time, with the development of computing power and the ability to access and process increasing amounts of data at lower costs,[4] significant progress has been made in the field of artificial intelligence, particularly with respect to academic research. In the 1990s, machine learning—an important subfield of artificial intelligence—began to grow substantially.[5] In the first decade of the 2000s, deep learning, a specialized subset of machine learning, started to gain wider recognition.[6]

We have recently witnessed another significant development in the world of artificial intelligence, in the form of generative AI, a technology that enables the creation of new content, including text, images, predictions and more. The main difference between generative AI and traditional (or "regular", that is, non-generative) artificial intelligence lies in the fact that the former, as its name suggests, is capable of generating original content, and in many cases, it while traditional artificial intelligence is primarily intended for identifying patterns, classification and predicting forecasts based on existing information. It is often difficult to discern whether content was created by a human being or is AI-generated by a computer. Generative AI is based on a Large Language Models (LLMs), which are themselves based on Natural Language Processing (NLP).

---

[1] Artificial Intelligence in The Financial Sector – Open Innovation and Ethical Commitment, Everis (2021). https://www.nttdata.com/global/en/-/media/nttdataglobal/1_files/services/data-and-intelligence/artificial_intelligence_in_the_financial_sector.pdf?rev=704e0e643b6b4f3bbd1ef22b04cdcd3e

[2] A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, John McCarthy, Marvin L. Minsky, 2 Nathaniel Rochester, and Claude E. Shannon (August 31, 1955), http://jmc.stanford.edu/articles/dartmouth.html?force_isolation=true; Some challenge this assertion, noting that the subject had been studied earlier, even if not under the term "artificial intelligence"; see, for example, https://sites.harvard.edu/sitn/2017/08/28/history-artificial-intelligence/

[3] Ibid., p. 1. "The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves".

[4] OECD (2021), Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers, https://doi.org/10.1787/98e761e7-en

[5] The field of machine learning focuses on the development of algorithms and models that enable computer-based systems to "learn" from data provided to them as input, and based on that input, to generate content, assessments, recommendations, or decisions independently.

[6] Deep learning is concerned with the development of algorithms and models aimed at stimulating the functions of the human brain by means of artificial neural networks that process information through multiple layers.

The watershed that symbolized the breakthrough of artificial intelligence in the eyes of the general public was the launch of ChatGPT by OpenAI. Following its debut, this application recruited no less than one hundred million users during November-December 2022, and for a short period of time, it was the highest-ranking application in terms of the number of users enlisted in the shortest amount of time.[7] Several technology giants have since presented innovations and developments in the field of artificial intelligence, and particularly generative AI, signaling the dawn of a real revolution in information technology. Entities engaged in the subject assess that, in the coming decade, the use of AI tools in organizations will improve employee productivity by approximately 40%.[8] According to the strategic consulting firm, McKinsey & Company, in the case studies examined, generative AI alone could add approximately $3.5 trillion on average to the global economy annually. By way of comparison, this amount is roughly equivalent to GDP of the United Kingdom in 2023-2024.[9]

As artificial intelligence technology develops, one can see its gradual incursion into many and varied areas of life. For example, AI systems advise on choosing films or series according to the viewer's historical preferences, assist in purchasing products on the internet, determine to which content and advertisements users will be exposed on social networks, analyze medical imaging, guide autonomous vehicle operations and more. The entry of AI systems can potentially advance humanity toward achieving unprecedented societal goals: autonomous cars will be able to reduce the number of road accidents and traffic congestion; medical equipment will be able to assist in diagnosing diseases at earlier stages and may even assist in developing medicines for their cure; commercial systems will be able to improve business processes, optimizing supply chains, increasing employee effectiveness and assisting in tasks that until now have been performed exclusively by human beings.

Alongside the many advantages, the deployment of AI systems, and particularly generative AI, raises many concerns. A significant loss of jobs - including in manufacturing, operations, consulting, transportation, customer service, medicine and more appear imminent. The infringement of privacy is another concern. Since AI systems are based on vast amounts of information collected from various databases, their cross-referencing could lead to the exposure of otherwise private information about users. Other potential misuses of AI technology include the government surveillance of its citizens, the creation of autonomous weapons, crime, discrimination and more.

In order to be able to enjoy the advantages engendered by artificial intelligence, it must therefore be ensured that it is used in a responsible, fair and proper manner. This, *inter alia*, through enabling

[7] For comparison, Facebook (now Meta) reached 100 million registered users within four and a half years of its launch, YouTube reached 100 million users within four years, and Instagram achieved the same milestone within two and a half years.
[8] https://mitsloan.mit.edu/ideas-made-to-matter/how-generative-ai-can-boost-highly-skilled-workers-productivity
[9] https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#key-insights

regulation that supports innovation on the one hand, while reducing the risks associated with AI use on the other.

In the period that has elapsed since the publication of the Interim Report, there has been a significant continuation of substantial capital investment in the field of artificial intelligence, the development of more advanced capabilities, and the integration of AI systems in diverse business activities. Among the developments in the field, one can count the improvement of the reasoning and analysis capabilities of models and the development of AI agents, that is, autonomous systems or entities that operate in a designated environment, process data, make decisions independently and perform actions according to goals defined for them;[10] The development of artificial intelligence models outside the United States is also noteworthy, particularly the DeepSeek model that was launched in China and is operated using significantly less computing power than that customary in the U.S.

Box 1: Artificial Intelligence – Glossary of Terms

Artificial Intelligence (AI): a field in computer science that deals with the development of computerized systems capable of performing tasks that simulate human cognitive abilities such as learning, pattern recognition, generating predictions or problem-solving and decision-making.

Machine Learning: a central subfield of AI that focuses on the ability of computing systems to "learn" without explicit programming. This is accomplished through developing statistical models that enable computerized systems to perform complex tasks without explicit programming of every possible scenario, through exposure to large quantities of data (pattern recognition, prediction generation and more). Machine learning applications are diverse and include, *inter alia*:

- Assistance in making or indpendently making a specific decision – for example, whether to approve or reject a customer's credit application in light of the probability of them repaying the loan or whether to authorize a transaction or financial operation after concerns of fraud have been examined.
- Assistance in performing ongoing and extensive tasks – for example, for autonomous driving systems or chatbots.

Deep Learning: an advanced subfield of machine learning. It is based on multi-layered artificial neural networks, which imitate the structure of the human brain. This method enables systems to learn abstract representations of data, identify complex patterns, and perform tasks such as image recognition, natural language processing, and automatic translation at a particularly high level. Deep learning has led to

[10] Naveen Krishnan, AI Agents: Evolution, Architecture, and Real-World Applications (March 2025), https://arxiv.org/pdf/2503.12687

significant breakthroughs in many fields of artificial intelligence, and currently serves as the foundation for the most advanced AI systems.

Algorithms: the core component of machine learning. Algorithms constitute a series of instructions that the computer executes in order to solve a problem or complete a task defined for it. Unlike the traditional approach where the operation of algorithms was designed and defined in advance, in machine learning, algorithms are designed to learn from examples, analyze data, identify patterns and use the learned information to generate insights or predictions.

Big Data: diverse databases, structured and unstructured, that accumulate quickly in large volumes that cannot be processed using traditional tools such as a personal computers and Excel spreadsheets. Collection of big data is performed quickly, drawing on diverse information sources such as existing data in the organization, the internet, digital "footprints" and the like. Organizations and companies analyze big data to extract value from them – to learn how to optimize their work, improve offers to customers and more.

What is between AI, ML and other predictive models: The use of models for prediction is not new, but rather a long-standing practice in the world of research and data analysis. Statistical methods and mathematical tools, such as linear regression, have been and continue to be used for identifying relationships between variables and forecasting outcomes. Linear regression, for example, enables the examination of correlative or causal relationships between variables and assists in building models for generating predictions. However, significant technological developments, primarily a massive increase in the availability of digital data (big data) and a dramatic rise in computing power, have led to a breakthrough in the field of machine learning, enabling the analysis of databases on an unprecedented scale and the establishment of more complex and rapid relationships between variables. This has led to the development of more sophisticated models and their integration in a wide range of fields.

The role of the human factor: In traditional models, a human expert defines the relationship between input variables and the target variable (the variable that one seeks to predict) in advance – that is, to what extent certain data are related to the target variable. In machine learning models, in contrast, the system is fed input variables and the target variable and then through advanced statistical algorithms the system "independently" learns the relationships between the input target variables and creates a prediction model accordingly.

Flexibility and adaptation: Machine learning models are characterized by much greater flexibility in identifying complex relationships that a human would have difficulty identifying manually. In addition, as their name suggests, machine learning models, learn. Machine learning models are capable of improving dynamically as they are exposed to new information. They are capable of updating the prediction generation process and the model structure itself, including adding new variables, rebalancing the weights of existing variables, identifying and removing variables that are no longer relevant and the

like. However, machine learning models are typically characterized as "black boxes". Their decision-making mechanism is complex and sometimes difficult for humans to unravel and understand. This limitation could make it difficult to use them in certain contexts, where transparency is essential for the decision-making process.

Supervised, unsupervised, and reinforcement learning:

- **Supervised learning** – A machine learning method in which the relationship between input data and the target variable is empirically modeled. In the supervised learning process, the model is fed "labeled" data – that is, the input data are already labeled according to their relationship to the target variable. For example – if one wishes to train a model to predict fraud, the database will include input data, such as transaction characteristics (location, time, amount, etc.), which in retrospect, turned out to be were found to be either legitimate or fraudulent. The model will derive the weight to be given to each variable from this data in an attempt to optimally classify a certain transaction as fraudulent. Afterward, it will be possible to feed new input data to the model and the model will label them or predict through them the probability of fraud based on what it has learned. Supervised learning is common in fraud detection, spam mail detection, risk management and credit- rating determination.
- **Unsupervised learning** – A model of this type is fed unlabeled data and the model is programmed through statistical methods to find patterns in them without human supervision, grouping together data points that share similar characteristics and separating them from those that differ. In the fraud example, for instance, transactions or operations that ultimately turned out to be fraudulent would not be labeled as such; only their characteristics would be provided as input. This technique is used, *inter alia*, for identifying anomalies within databases – which at a later stage can help in detecting fraud or other irregular activities.
- **Reinforcement learning** – A method in artificial intelligence in which an agent (such as software or a robot) learns to operate in a certain environment in order to maximize reward. The agent performs many actions, receives feedback in the form of a reward (positive or negative), and learns over time on its own which actions render the best results. For example, an agent playing chess learns which strategies and moves lead to more victories, just as human chess players improve their game with experience. This method enables the agent to learn and improve without explicit programming for every possible strategy. Reinforcement learning is employed in a variety of fields such as games, robotics and autonomous navigation.

Stages in developing a supervised machine learning model:

Algorithm selection

Selection of the type of algorithm according to the intended use for which the model is developed. Relevant considerations include the volume of data, the nature of the relationships, the need for explainability, and the like.

Training

Data are input into the algorithm, on the basis of which the algorithm develops a predictive model.

Model validation

Data are then input into the developed model — different from those used for training — in order to validate the model. This is done by examining the results obtained from the new data inputs. The validation stage assists in assessing the reliability of the model.

Model tuning and testing

In the next stage, the models are tuned and tested. Among other things, it is examined whether the model suffers from "overfitting," that is, whether the model is overly adapted to the data on which it was trained and is therefore not effective in producing accurate predictions with respect to other data. Model validation is carried out by inputting entirely new data into the model. At this stage, the model may be operated as a "shadow model" — running alongside existing systems, comparing outputs, making adjustments to the model, and planning risk management policies, among other things.

Artificial General Intelligence (AGI): As a rule, AI systems are developed for the purpose of performing certain distinct tasks. Artificial general intelligence refers to AI systems capable of performing a wide range of tasks and operations, without being explicitly programmed to perform certain tasks or operations. That is, a system that mimics human behavior, which is autonomous and "learns" the performance of the task from the environment without requiring human supervision. It is possible that in the future, there will be no need for designing a unique model for investment advice but rather a general system will be able to learn the task and perform it.

Generative AI: A type of machine learning that is capable of generating content in various media: writing, images, music and more. Generative AI is the product of an algorithm that trains on enormous quantities of unprocessed data – web pages, books, images, films, audio clips and the like, and learns patterns that enable the model to generate appropriate content according to given questions or prompts. Generative AI is one of the central technologies incorporated in disinformation operations. Through this technology, it is possible to generate, for example, deepfake images and videos of a quality that makes it difficult to distinguish them from real images. Certain machine learning techniques are capable of detecting whether a certain product is real or fake, also based on training. However, as generative AI advances and becomes more sophisticated, the difficulty of identifying fake content increases.

Natural Language Processing: A branch in the field of artificial intelligence dealing with processing written and spoken text while understanding the meaning emanating from the words. The algorithm searches for patterns in the way text is constructed and how the use of words and text structure create certain meanings. This process assists in the creation of independent content that serves in chatbots and similar applications. The technological frontier in this field today is Large Language Models (LLMs).

Alignment: One of the challenges in developing AI systems is the alignment challenge, that is, the need to ensure that the autonomous system will operate according to goals defined for it and to a predetermined set of human values governing its operation. A common example is, for instance, ensuring that generative AI refrains from providing unethical independent information, such as instructions for harming others or content using inappropriate language.

B. Regulation of Artificial Intelligence in International Organizations and Developed Countries

Driven by rapid technological development and the growing recognition of the extensive influence of artificial intelligence on many spheres of life, regulators, governments and international organizations have begun creating regulatory frameworks to oversee the AI domain.

The European Union began dealing with the regulatory aspects of artificial intelligence as early as 2018,[11] and recently enacted the EU Artificial Intelligence Act (hereinafter: "EU AI Act"),[12] the most comprehensive regulatory framework to date. Starting in the second half of 2024, this legislation is incrementally entering into force, scheduled to be fully adopted in all member states of the European Union by 2027. This regulatory framework seeks to establish harmonious legislation regarding the development, operation and use of artificial intelligence in the European Union member states. The new legislation imposes obligations on system developers (code writers) and on system operators,[13] *inter alia* according to the level of risk associated with the AI applications being employed. A preliminary version of the legislation was presented in 2021 and the law was approved and published in its final and binding form in July 2024. During the legislative process, quite a few changes were made to the proposal, some of which stemmed from the considerable technological progress that had already been made in the field, since the bill was first presented.

The Organization for Economic Cooperation and Development (OECD) began dealing with certain aspects of artificial intelligence in 2019, when it published a policy paper dealing with the use of artificial intelligence in the public sector.[14] This was the first document of general principles formulated by a leading international organization, and within its framework, principles were established in the field of AI use based on a human-centric approach. Pursuant to these principles, the paper articulated several values to guide operations, including: the contribution of artificial intelligence to sustainable development and social welfare; respect for values that place the human at the center and the value of fairness. The use of AI systems and their manner of operation must be transparent, AI systems must be resilient and secure, while accountability for the results of artificial intelligence is imperative. In addition, the OECD document proposed policy tools designed to promote the deployment of artificial intelligence in a trustworthy and ethical manner. Subsequently, the organization published a series of

---

[11] European Commission, High-Level Expert Group on Artificial Intelligence - Draft Ethics Guidelines for Trustworthy AI. (December 2019) - https://digital-strategy.ec.europa.eu/en/library/draft-ethics-guidelines-trustworthy-ai.

[12] Regulation (Eu) 2024/168 of the European Parliament and of the Council of 13 June 2024, laying down harmonised rules 12 on artificial intelligence and amending Regulations (EC) No 300/2008), (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828) Artificial Intelligence Act, https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

[13] Deployers – see further discussion of the issue of liability (in Part C – Cross-Cutting Issues below).

[14] Berryhill, J., et al. (2019), "Hello, World: Artificial intelligence and its use in the public sector", OECD Working Papers on Public Governance, No. 36, OECD Publishing, Paris, https://doi.org/10.1787/726fd39d-en.

policy papers dealing with artificial intelligence in society,[15] in the financial sector (hereinafter: the "2021 OECD Report" and the "2023 OECD Report", respectively),[16] competition,[17] environment and sustainability,[18] employment,[19] risk management,[20] technology[21] and more. Safeguarding human rights under the human-centric approach, alongside promoting innovation in an ethical and controlled manner is the common thread running through all these publications. The OECD recommendations carry special weight, since they are regarded to be an international standard reflecting agreed policy even among countries adopting diverse approaches to artificial intelligence.[22] Additional international organizations such as the Group of Seven (G7) and the Group of Twenty (G20) forums have also published statements and policy documents [23]on the processes governing the use of artificial intelligence tools.

In the period that has elapsed since the publication of the Interim Report, some of the provisions of the EU AI Act have entered into force, including regulation applying to high-risk systems, including systems dealing with credit scoring, credit worthiness and risk management, as well as the pricing of life and health insurance policies. The provisions established regarding General Purpose AI (GPAI) systems entered into force in August 2025. In July 2025, the European Commission published three key documents on GPAI: (1) a guide for developers of general-purpose artificial intelligence-based models regarding the scope of their obligations under the EU AI Act;[24] (2) a voluntary code for developers of GPAI-based models developed by industry experts providing practical information regarding issues of transparency, copyright infringement, security, etc.;[25] and (3) a form requiring developers of GPAI-based

---

[15] OECD (2019), Artificial Intelligence in Society, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en

[16] See *supra* note 4, and OECD (2023), "Generative artificial intelligence in finance", OECD Artificial Intelligence Papers, No 9. OECD Publishing, Paris, https://doi.org/10.1787/ac7149cc-en.

[17] OECD (2017), Algorithms and Collusion: Competition Policy in the Digital Age, https://www.oecd.org/content/dam/oecd/en/publications/reports/2017/05/algorithms-and-collusion-competition-policy-in-the-digital-age_02371a73/258dcb14-en.pdf

[18] OECD (2022), Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint 18 | en | OECD, https://www.oecd.org/publications/measuring-the-environmental-impacts-of-artificial-intelligence- compute-and-applications-7babf571-en.htm

[19] OECD (2023), OECD Employment Outlook 2023 : Artificial Intelligence and the Labour Market | OECD iLibrary, https://www.oecd-ilibrary.org/sites/08785bba-en/index.html?itemId=/content/publication/08785bba-en

[20] OECD (2023), Common guideposts to promote interoperability in AI risk management | OECD Artificial Intelligence Papers | OECD iLibrary, https://www.oecd.org/en/publications/common-guideposts-to-promote-interoperability-in-ai-risk-management_ba602d18-en.html ; OECD (2023), Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI | en | OECD, https://www.oecd.org/sti/advancing-accountability-in-ai-.2448f04b-en.htm.

[21] OECD (2023), AI language models: Technological, socio-economic and policy considerations | en | OECD, https://www.oecd.org/en/publications/ai-language-models_13d38f92-en.html

[22] See *infra* note 32, at page 36.

[23] See for example: G20, AI Principles (June 2019), https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf?force_isolation=true ; G7, "Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems On the basis of the Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI systems, the Hiroshima Process International Code of Conduct for Organizations (October 2023)", https://ec.europa. eu/newsroom/dae/redirection/document/99644?force_isolation=true; G7, Leaders' Statement on the Hiroshima AI .Process, https://www.mofa.go.jp/files/100573466.pdf.

[24] https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act?force_isolation=true.

[25] https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai?force_isolation=true.

models to provide information regarding the databases and information on which the AI systems are based and from which they draw information for their operation.[26]

In the United States, although various AI-related bills have been introduced in the Congress, at the time of this writing, no significant federal regulation resembling that of the EU has been enacted. In 2020, the federal government published a guide for regulation of AI applications for all government ministries and departments.[27] In 2022, it published guidelines for a "Bill of Rights" regarding the use of artificial intelligence, describing the safety measures that should be taken to protect the rights of the American public.[28] At the end of 2023, a presidential order was issued regarding the safe and reliable use and development of AI technology (hereinafter: "the Presidential Order").[29] In parallel, states in the United States began to legislate local laws on issues involving artificial intelligence and its application in various areas of life. Unlike the European Union, the American administration, has not adopted comprehensive, cross-sectoral legislation in the field of artificial intelligence, but rather seeks to grant regulators and the competent authorities the authority to act in their respective fields, under the principles published in the Bill of Rights and the Presidential Order, alongside use of soft regulation (such as standardization, involvement in self-regulation of the main players in the field, etc.).

In the period that has elapsed since the publication of the Interim Report, and following the ascension of the Trump administration, the Presidential Order published during the Biden administration was canceled, and in its place a new presidential order, Executive Order 14179 (*Removing Barriers to American Leadership in Artificial Intelligence*) was published, favoring a regulatory regime that encourages innovation while removing existing barriers and restrictions.[30] Following this order, the "American AI Action Plan" was published in July 2025. The Action Plan seeks to promote the AI industry and American leadership of it under the pillars of innovation, infrastructure, international diplomacy and security.[31] *Inter alia*, in its emphasis on the commercial development of AI, the Action Plan focuses on deregulation rather than regulation designed to prevent or prohibit AI abuses. As mentioned above, no comprehensive artificial intelligence law has been legislated in the US at the federal level, however there are several states, such as Colorado, Texas and Utah, which published AI laws for

---

[26] https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act.

[27] The White House, Guidance for Regulation of Artificial Intelligence Applications (November 2020), https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf?force_isolation=true.

[28] The White House, Blueprint for an Ai Bill Of Rights: Making Automated Systems Work for the American People (October 28 2022), https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/.

[29] The White House, Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (29 October 2023), https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

[30] Removing Barriers to American Leadership in Artificial Intelligence – The White House, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/.

[31] The American Action Plan, p. 1. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf.

their jurisdictions. These laws deal with aspects such as transparency, prevention of discrimination, privacy and use of information.

In the United Kingdom,[32] the government published a national strategy on artificial intelligence, and has also established a unit under the Ministry of Science, Innovation and Technology to supervise the adoption of artificial intelligence under the established strategy.[33] The British approach values innovation and global leadership in the field of artificial intelligence, through promoting long-term investment in technology and the integration of technology at the state level, while safeguarding human rights and the proper use of AI technology. The U.K. approach is similar to that of the United States. Currently, the British government has no intention of enacting cross-sectoral legislation, opting instead for sectoral regulation that allows flexibility for regulators and competent authorities to act in their areas under their purview according to established principles.

There is indeed similarity between the approaches described above regarding the need for managing artificial intelligence risks, avoiding the infringement of human rights and ensuring the responsible and proper use of AI technology. However, the ways to achieve these goals differ between the different approaches, and it appears that at this point in time, there is a conceptual gap regarding the role of rigid mandatory regulation to deal with these risks.

The engagement in artificial intelligence poses several key questions to policy-makers worldwide. Fundamental questions have been raised related to the appropriate timing to impose regulation that will apply to artificial intelligence technology (whether there is a need to establish rules now or should this be delayed until the technology is further developed, the nature and strength of such regulation and the determination of whether AI regulation should be integrated into existing regulation in various areas or should dedicated regulation be established that will apply on a cross-sectoral basis. This last question also has implications for regulatory structure. In Spain, for example, the federal government chose to establish an independent artificial intelligence authority, AESIA, whose role is to concentrate supervision, professional consultation, and public awareness in artificial intelligence, even before the E.U. legislation entered into force.[34]

---

[32] HM Government, National AI Strategy (September 2021), https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf?force_isolation=true.

[33] Office of Artificial Intelligence, https://www.gov.uk/government/organisations/office-for-artificial-intelligence?force_isolation=true.

[34] https://www.boe.es/diario_boe/txt.php?id=BOE-A-2022-20639.

C. Government Engagement in Artificial Intelligence in Israel

In Israel, government engagement in the field of artificial intelligence was initiated several years ago. In August 2021, the Government of Israel adopted Resolution No. 212 entitled the *Plan for Promoting Innovation, Encouraging Growth of the High-Tech Sector and Strengthening Technological and Scientific Leadership* (hereinafter: "the Government Resolution"), within which the Ministry of Innovation, Science and Technology was instructed to formulate a national plan to deal with the various aspects requiring attention, including regulation, information policy and data collection, ethics, international cooperation, and the integration of AI systems in the public sector.[35] The Government Resolution was based, *inter alia*, on other government or public initiatives on this subject.[36] In February 2023, the government approved Resolution No. 173 titled *Strengthening Israel's Technological Leadership*, which includes a national plan for artificial intelligence including accelerating basic and applied research in the field, developing infrastructure, integrating AI applications in the public sector, and budgeting an investment of up to 500 million shekels for AI.[37]

In December 2023, the Ministry of Innovation, Science and Technology, in cooperation with the Office of Legal Counsel and Legislative Affairs (Economic Law Department) at the Ministry of Justice, published a comprehensive policy document concerning *Israel's Policy on Artificial Intelligence, Regulation and Ethics* (hereinafter: "the Policy Document").[38]

The Policy Document proposes several principles to shape regulation policy. The basic assumption underlying these principles is that the adoption of uniform and coherent government policy for AI regulation is crucially important, and that such policy should be based, *inter alia*, on the following components:

- **Sectoral regulation** –AI regulation will be formulated in the sector where it is required under the leadership of the regulators responsible for sectoral oversight. This is contrast to sweeping cross-sectoral regulation covering the entire field of artificial intelligence in a wide range of sectors. The concept underlying this principle is that given disparities in potential risks, threats to individual rights and to public interest, etc., different sectors require different regulation in this field.
- **Non-uniform risk-based regulation** – The regulation will not apply uniformly to applications whose associated risks differ in kind and degree. For this purpose, the regulation must be adapted

[35] https://www.gov.il/he/pages/dec212_2021.
[36] Among others, The National Initiative for Secured Intelligent Systems to Empower the National Security and Techno-Scientific Resilience: A National Strategy for Israel. Special Report to the Prime Minister (2020); Telem Forum committee of artificial intelligence and data science (2023).
[37] https://www.gov.il/he/pages/dec173-2023?fireglass_rsn=true#fireglass_params&tabid=a1d3443a368e2634&s
[38] Israel's Policy on Artificial Intelligence Regulation and Ethics (December 2023), https://www.gov.il/BlobFolder/policy/ai_2023/en/Israels%20AI%20Policy%202023.pdf.

to the risks arising from the use of a specific application and must be the product of risk management conducted by the relevant regulator. To illustrate, the risk arising from the use of an AI-based navigation application is not similar in the risk associated with it to the use of an application that processes imaging and automatically generates medical diagnoses.

- **Alignment with international standards** – The proposed regulation will be consistent with regulatory practices customary in leading countries in the field of artificial intelligence and with principles adopted by leading international organizations. This approach aims to prevent the creation of regulatory barriers for Israeli companies seeking to operate AI applications abroad, as well as for foreign companies wishing to operate AI applications in Israel. Furthermore, it is proposed that government entities and the relevant regulators be actively involved in international forums that formulate and shape AI regulation.
- **"Soft" regulation** – It was proposed that the regulation be flexible, facilitating technological progress without requiring frequent legislative updates chasing technological developments. To this end, it is recommended to rely on the establishment of principles, standardization, regulatory guidance for voluntary adoption, and self-regulation.
- **Development of regulation in a modular manner using regulatory experimentation, while maintaining technological neutrality** – To ensure the safe deployment of AI applications for public use, and facilitate regulators in obtaining in-depth knowledge of the technology and its development trajectories, regulatory sandboxes and pilot programs designed in cooperation with the relevant regulators is proposed should be initiated.[39]
- **Involvement of stakeholders in drafting regulation** – Regulation in the field of artificial intelligence will be developed and promoted in partnership with stakeholders and persons knowledgeable in the field, including individuals from industry, academia, and civil society.
- **Establishment of a regulators' forum and a forum for public participation** It was proposed to establish a professional forum, which will include representatives from the relevant regulatory agencies, alongside experts in technology, policy and law. This forum will discuss common issues, formulate uniform and coherent regulation, and promote government coordination in dealing with AI-related matters. It was also proposed to establish a forum that will comprise representatives of regulatory and policy-making bodies, alongside representatives of industry,

[39] A regulatory sandbox is a policy tool that enables regulators to create a regulatory "testing environment." Within this framework, the regulator makes adjustments to existing regulation in order to allow regulated entities to introduce new technologies in a controlled, limited manner and for a restricted period of time. In addition, the sandbox assists regulators in addressing information and expertise gaps between themselves and the market, as well as coping with the rapid pace of technological development. Temporary, measured, and flexible regulation enables a process of learning "on the go" and provides tools for the cautious and controlled removal of targeted barriers. For their part, regulated entities benefit from the opportunity to operate in a "friendly" environment, without bearing regulatory burdens that are not essential to their activities (Policy Document, at p. 100).

academia, civil society and the public, to identify key issues which regulators and policy-makers must address, *inter alia* for making necessary adjustments, removing barriers, and formulating responses to these issues.

- **Usage survey** – To ensure that regulation is established on a factual basis and adequate understanding by regulators of current and future usage, regulators and relevant government entities will map and survey the uses of AI applications in the sectors under their purview.
- **Making information and tools accessible for responsible use of artificial intelligence, including tools for risk management** – It was recommended to consider developing and adopting a uniform risk management framework for the use of artificial intelligence. Such a framework will establish a common language among government bodies and regulators, and between these entities and the public at large. Use of a uniform language will increase legal certainty, will help the public, particularly the private sector, assess the risks associated with AI applications, and will assist regulators examine both the applications and the risks ensuing from their use.

Following the drafting of the Government Resolution and prior to the publication of the Policy Document, the National Plan for Artificial Intelligence was launched in July 2022. This is a long-term plan whose objective is to facilitate the development and integration of artificial intelligence in the public sector and the Israeli economy through promoting research and investing in human capital and infrastructure necessary to ensure the technological leadership of the State of Israel.[40]

As part of the government effort to regulate the use of artificial intelligence in specific sectors, treatment of certain regulatory aspects in the fields of autonomous vehicles, privacy (which we will discuss in Part C below), health,[41] and more has been promoted in recent years.

In the period that has elapsed since the publication of the Interim Report, the Government of Israel adopted a resolution on establishing a National Artificial Intelligence Headquarters under aegis of the Prime Minister's Office (Government Resolution No. 3375 of September 21, 2025). The roles of the National Artificial Intelligence Headquarters include, *inter alia*: spearheading the formulation and implementation of a national strategy in the field of artificial intelligence, including policy regarding research infrastructure; formulating a national plan to accelerate artificial intelligence development with the purpose of strengthening Israel's global standing in the field; formulating policy; and providing professional guidance for government ministries concerning the regulatory aspects of artificial

[40] https://aiisrael.org.il/.

[41] Ministry of Health, Guiding Principles for the Development of Machine Learning-Based Technologies (April 2023), https://www.gov.il/en/pages/digital-medical-technology-gmlp-1.

intelligence, with an eye to removing barriers and enhancing AI development, deployment and use in the economy.[42]

[42] Government of Israel, Decision No. 3375 (Sept. 21, 2025), https://www.gov.il/he/pages/dec3375-2025.

D. AI applications in the Financial Sector

Artificial intelligence initially entered the financial sector in the early 1980s, when a hedge fund specializing in quantitative investing was founded. Its approach was based on advanced mathematical models designed to determine the optimal point for executing profitable transactions.[43] An AI-based financial planning system was also created for high-income earners around the same time.[44] At that time, developments based on artificial intelligence could also be found in the worlds of taxation, stock market trading, financial planning and rating creditworthiness.[45] In the 1990s, progress was made in applications for the detection of fraud and money laundering operations.[46] In 1998, the U.S. Securities and Exchange Commission (SEC) created a new regulatory framework designed to reduce barriers of entry for new trading systems and encourage competition and innovation.[47] In practice, the new regulation enabled the entry of high-frequency (HFT) and algorithmic trading systems, which commanded an increasingly central place in stock market trading in the United States. Today, estimates are that about half of trading volume in the United States and about a third of trading volume in Europe is performed by HFT systems or algorithmic trading.[48]

Financial entities worldwide have been using AI-based applications in their diverse activities for some time; such as in insurance (claims settlement, insurance underwriting); in banking and credit (credit underwriting, determining credit scoring; in financial advice (investment advice, retirement-savings advice); and in portfolio management and securities trading (algorithmic trading, HFT, transaction viability analysis, liquidity management). In addition, financial entities use AI-based tools for activities supporting or ancillary to their core activities, such as in the detection of financial fraud, money laundering and terrorism financing activity, marketing and sales, customer service, risk management and compliance, cybersecurity, and more. These things are true, although apparently to a lesser extent, with respect to the financial system in Israel as well, where one can detect an increasing use of AI applications among financial sector companies. For a more detailed review of activity in the financial sector in Israel (see the chapter, *Artificial Intelligence Activity in the Financial Sector in Israel* below).

Despite the enthusiasm and buzz surrounding artificial intelligence in general and AI applications among financial entities in particular, cases in which advanced AI applications perform the core activity of the financial entity end-to-end, autonomously and without any human involvement are few and far between.

---

[43] Artificial intelligence in finance - Bonnie G. Buchanan, PhD, FRSA, The Alan Turing Institute (April 2019), p.4, doi.org/10.5281/zenodo.2612537.
[44] Ibid.
[45] Ibid.; see also Everis (*supra* note 1), p. 15.
[46] Ibid., p. 5.
[47] Regulation of Exchanges and Alternative Trading Systems (1998), https://www.sec.gov/rules-regulations/1998/12/regulation-exchanges-alternative-trading-systems.
[48] https://www.nasdaq.com/glossary/h/high-frequency-trading;%20Zaharudin%20KZ,%20Young%20MR,%20Hsu%20WH%20(2022)%20High.

While there are numerous AI applications in the financial sector, these serve primarily as decision-support tools (assisting rather than replacing human decision-making) while others perform operations autonomously, but only with respect to specific segments of the financial product or service value chain, and usually without directly interacting with the financial entity's customer.[49] Financial entities that extend credit to their customers, for example, use AI-based applications for underwriting the loan and its terms. The final decision, however, of granting or denying credit and the precise terms set is determined by human agents of the financial entity, who are also the party in contact with the customer. Another example: investment advisors who use AI-based applications to structure portfolios for their clients, according to the latter's risk appetite and personal preferences, the final decision regarding asset allocation remains in the hands of the human advisor. It can be expected that in the future, as technology advances and current limitations and challenges diminish, human involvement in the provision of financial services will decline. This will likely generate even greater automation and extend the use of AI applications to additional segments of the financial services industry. Recent research indicates that artificial intelligence is used today primarily to streamline and enhance output in back-office operational activities as well as to enhance compliance and risk management functions.[50] A survey conducted by the McKinsey & Company among international financial institutions, has found the optimization of operational services (such as customer service, back office, etc.) currently accounts for the primary use of AI applications. This finding has held repeatedly for four consecutive years.[51]

Customer service is one area in which the deployment of AI applications in financial entities is widespread. Many financial entities have integrated (or are in the process of integrating) chatbots, an AI-based application that provides automatic responses, at varying levels of sophistication, to service customer inquiries. It is capable of providing a response simulating human responses. Advanced chatbots use Large Language Models (LLM) based on Natural Language Processing (NLP). At times, it is difficult to distinguish between an AI-generated automated response and a human response. In some cases (depending on the level of chatbot sophistication), the chatbot is capable of providing a complete response to service inquiries from customers without the customer being aware that they had been assisted by an artificial intelligence tool rather than a human being. According to assessments made in 2022 by the Consumer Financial Protection Bureau (CFPB) in the United States, an estimated 37% of Americans have engaged with a chatbot when requesting banking services from the bank, with this proportion expected to grow significantly.[52] Cost savings is one of the primary motivations for adopting

[49] 2023 OECD Report, p. 10.
[50] Ibid.
[51] McKinsey & Company, The state of AI in 2022—and a half decade in review (December 2022), https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review.
[52] CFPB, Chatbots in consumer finance (June 2023) , https://www.consumerfinance.gov/data-research/research-reports/chatbots-in-consumer-finance/chatbots-in-consumer-finance/#note.

chatbots in the financial system. According to industry estimates, the adoption of AI applications for customer service systems may save the U.S. banking system alone approximately $8 billion.[53] As mentioned above, it appears that the adoption of advanced AI applications among financial entities (especially "traditional" financial entities) is undertaken slowly and is highly controlled. This is particularly true regarding applications enabling full automation, devoid of any human involvement whatsoever, of the financial entity's core activity. A survey conducted by the European Securities Authority (hereinafter: "the 2023 ESMA Report")[54] found that currently the use of AI tools in companies providing robo-advisory services and credit scoring services is limited. To the extent that they are used, such use is confined to particular segments of the range of services provided by the financial entity as opposed to full automation of the entire service (Not for nothing, investment advisors in the European Union stated that to date, they have not been required to substantially modify their investment practices as a result of integrating artificial intelligence).[55]

There are several reasons for the relatively sluggish and controlled adoption of AI applications in the financial sector, including: the strict regulatory environment in which the sector operates in most jurisdictions; the emergence of previously undetected risks; as well as the intensification of existing risks (for example, risks regarding privacy protection and information confidentiality, cyber risks, information quality and concerns related to biases and discrimination, compliance, lack of explainability of how the technology and its products work, and the like). In addition, the shortage of human capital with adequate competence and expertise, the use of legacy systems and the difficulty of adapting these systems to this type of technology also contributes to the slow adoption of AI tools.

In summary, the potential for the deployment of artificial intelligence in the financial sector is significant, and technological developments have indeed stoked interest among financial entities. At the same time, however, the AI applications to date are based primarily on technologies from previous generations. The deployment of more advanced AI applications is still limited at this stage, and relates less to core activities vis-à-vis customers. This situation appears to prevail in financial services in numerous Western countries as well as in Israel.

---

[53] Ibid.

[54] ESMA, Artificial Intelligence in EU Securities Market (February 2023), p. 9, https://www.esma.europa.eu/sites/default/files/library/ESMA50-164-6247-AI_in_securities_markets.pdf; See also: 2023 OECD Report, p. 4.

[55] The 2023 ESMA Report, p. 6.

E. Benefits and Challenges Arising from the Use of Artificial Intelligence

The technological progress in the field of artificial intelligence in general and machine learning in particular engenders significant advantages in a wide range of fields. In the not-so-distant future, it is possible that a significant part of daily operations will be performed or supported by AI technology. However, alongside the many advantages, significant challenges also arise from the technological innovation that artificial intelligence has ushered in.

We will first discuss the main potential benefits from the entry of AI-driven technology to the financial sector:

(1) **Operational streamlining and cost reduction** – The integration of artificial intelligence in the operational systems of financial entities and the automation of complex operations requiring numerous resources, including human capital, is expected to lead to greater efficiency and lower costs. For example, customer service performed by a chatbot is expected to lead to a reduction of human customer service systems and the channeling of resources to other operations. Estimates are that use of AI operational applications will increase significantly in the coming years. The reduction of expenses incurred by financial entities may carry over to customers and investors who will enjoy lower-cost services and products.

(2) **Enhancing product and service quality** - Artificial intelligence may enhance the quality of products and services offered in the financial sector, through the improvement of decision-making processes, better adaptation of products and services to customer and investor needs, and the expansion of financial activity. For example, credit extension activity could improve with the aid of artificial intelligence which will take into account diverse data, including "non-traditional" information, which today is not taken into account in the approval of credit. This will likely lead to more accurate credit decisions. The process can be carried out in a manner better adapted to customer needs, starting with the offering of credit to those who need it, and including the determination of the terms of the extended credit. It can be broader based, such that credit is extended to a wider range of borrowers, including small and medium enterprises (SMEs).

(3) **Enhancing inclusion and financial accessibility** – In the future, populations currently excluded from the financial sector will be able to use the technology that will be developed as a result of the adoption of AI applications by the financial sector. The adoption of AI technology will likely improve the accessibility of financial products, including credit, insurance and savings and investment capabilities to increasingly larger customer groupings. Investment portfolio management activity, often limited to clients with considerable capital resources, could be executed at lower costs and therefore, could be directed to a wider range of clients.

(4) **Improving compliance** - AI-based systems may help financial entities improve their real-time compliance capabilities. AI systems have the power to process enormous amounts of information instantaneously and to flag suspicions or irregular operations close. These tools can also assist financial regulators identify violations and irregular operations in real time, in a manner that contributes to effective enforcement and deterrence, thereby protecting investors and safeguarding market fairness. In addition, AI-based systems can assist in the real-time detection of criminal operation patterns and the prevention of such operations.

We will now discuss the challenges arising from use of artificial intelligence. The following list comprises two parts. The first focuses on general challenges arising from the deployment of AI systems that are not specifically designed for the financial sector. These challenges are extensively reviewed in the Policy Document. The second part details challenges unique to the financial sector or whose materialization could lead to more severe and significant harm to the financial services relative to other sectors, given the vulnerability and risk associated with financial activity. It should be clarified that this does not constitute a closed list of challenges and risks.[56]

General challenges emanating from the integration of artificial intelligence systems:

(1) **"Black box" and explainability** - The decision-making method of AI systems is often described as a "black box" in the sense that it cannot be traced. That is, it is not possible to know how and why a certain decision was taken as it was taken. The reason for this lies in the complexity of algorithms underlying AI systems and in the amount of information on the basis of which decisions are made. Some argue that there is a trade-off between the degree of explainability and the quality of system performance. That is, there is an inverse relationship between the degree of explainability and the accuracy of system performance.[57] The "black box" problem typifies advanced AI systems as opposed to systems from previous generations, in which it was clear how the algorithm operates and how its products were derived. This opacity could lead to additional problems, such as difficulty in trusting the system, difficulty in identifying and correcting errors, ethical difficulties related to the inability to provide an explanation to the party about whom the decision was made, and more.

(2) **Human involvement** - The question arises to what extent (if at all) human involvement is required at the time decisions are made by AI systems. A central characteristic of advanced AI systems is their ability to operate without human involvement, while the degree of human involvement can be classified as follows: (1) human involvement does not exist at all, and the entire decision-making process is performed by AI systems completely devoid of human involvement; (2) human involvement exists to some degree, so that the decision-making process is performed by AI systems, but a human party is

[56] For further discussion of the issues and challenges arising in connection with artificial intelligence, see the Policy Document, p. 46.

[57] 2021OECD Report, p. 45.

entitled to intervene (to some extent) in the decision made by the system. This involvement can occur at several stages of system activity; in advance, in real time or retrospectively; and (3) human involvement is fully incorporated in the AI system, such that it serves as a "recommending party" aiding human decision-making, to whom discretion in making the final decision is exclusively entrusted.

(3) **Notification and disclosure** - The transparency and disclosure challenge relates to several levels in the activity of AI systems. One question is whether and to what extent a person should be notified that their interlocutor or the party making a decision about them is and AI system rather than a human being. A second question pertains to the scope of transparency required regarding the manner of operation of AI systems. For example, is specific disclosure about the decision-making process required when performed by an AI system. A third question relates to marketing publications concerning AI systems and the necessity that these not mislead the public.

The engagement in the issue of notification and disclosure is largely related to the explainability issue, since notification and disclosure about the use of AI applications can be included as part of explainability obligations (which is, essentially, a certain type of transparency regarding the core of the system and the data feeding it) imposed on AI-system and end-user service providers.

(4) **Privacy** - Artificial intelligence technology is based on the extensive collection and processing of information, including personal information, i.e., information relating to an identified or identifiable person. Advances in technology significantly intensify incentives to collect, store and leverage the use of big data. In addition, the outputs of AI systems may revel personal, individualized insights about a person, including highly sensitive and intimate information. Moreover, AI systems process big data using high-quality, efficient computing operations, which are able to detect, *inter alia*, unexpected connections and interactions between different data sets and to maximize the output and insights that can be derived from them. These and other issues raise considerable challenges and tensions in connection with privacy protection and the safeguarding of personal information throughout the entire life cycle of an AI system.

(5) **Discrimination and biases** - Artificial intelligence is based on information raising concerns that the data on which systems are based will be either intentionally or unconsciously biased. An additional concern is that AI systems will make prohibited use of criteria capable of characterizing certain populations, in a manner that will lead to biased results or prohibited discrimination and could even exacerbate existing discrimination. The existence or absence of discrimination depends largely on the representativeness and relevance of the data feeding the system and on the basis of which the system will make a decision. Hence, the issue of the reliability, quality, accuracy and availability of the data based on the decisions made by the system is crucial.

(6) **Reliability, resilience, security and safety** - To realize the benefits of artificial intelligence, it is essential to ensure the proper, safe and reliable functioning of these systems, which rely on technology

and information that are inherently vulnerable to errors and technical malfunctions. According to the Policy Document, failures that could arise in this context can be segmented into internal failures affecting system reliability (leading to inaccuracy and poor performance) and external failures affecting system resilience in the face of disruptions caused by the exploitation of its vulnerabilities. Reference to internal failures usually focuses on failures related to the quality of information used by the system in decision-making, information validation and system performance testing. Reference to external failures is divided into system resilience and security, focusing primarily on threats arising from external intervention that precipitates damage or misuse of the system, including cybersecurity risks. Reference to system safety relates to both internal and external failures, which could significantly harm individual rights or the public interest, particularly with respect to the safety of users affected by them.[58]

(7) **Liability** - A fundamental issue arising from the use of AI systems relates to the question of which party bears civil or criminal liability with respect to a discriminatory result or system failure. This also raises the question of whether the liable parter (whoever it may be) can undertake actions or implement processes that may discharge their obligation or mitigate the extent of their liability. The liability issue also emerges with respect to outsourcing and the execution of business activities by third parties.

In addition to the challenges articulated above, certain challenges and issues are unique to the financial sector:

(a) **Financial stability risks** - Use of AI systems could increase systemic risks to the stability of the financial system.[59] *Inter alia*, the use of AI systems or similar models by a large number of financial entities could create interconnectedness between financial entities using the same systems, in a manner that could lead to them engaging in extensive identical action and the creation of one-way markets. In times of market stress, this increases concerns about liquidity risk and systemic stability.[60] Similarly, AI systems could increase "herding" behavior and market volatility. An additional risk to financial stability arises from the reliance on third-party, AI system providers, the creation of dependency on them.

An example of a systemic stability risk can be found in the realm of capital market trading. In May 2010, the event that earned the moniker "flash crash" occurred in the United States, during which leading indices on the New York Stock Exchange lost tens of percentage points within a few minutes of trading. As a result, trillions of dollars in value were erased within minutes, precipitating a liquidity crisis across U.S. stock exchanges. The event made waves in the United States and worldwide, and was investigated by two financial regulators in the United States. *Inter alia*, it was found that the interaction between automatic trading systems and algorithmic trading strategies could quickly cause severe harm to market

[58] Policy Document, p. 65.
[59] Danielsson, Jon and Uthemann, Andreas, On the use of artificial intelligence in financial regulations and the impact on financial stability * (June 05, 2024). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4604628.
[60] 2023 OECD Report, p. 23.

liquidity through irregular market activity.[61] Especially during times of market stress, the sudden execution of similar transactions by a large number of market players could lead to sharp fluctuations in the prices of financial instruments, increasing volatility and, in some cases, leading to a breakdown in market stability.

Similarly, a failure in an algorithm or in databases used by several central credit providers could lead to the extension of credit to borrowers lacking adequate repayment capability, which in turn could generate systemic instability and financial crisis. It should be noted that such failure is not necessarily limited to completely automated, autonomous AI systems, since even when systems serve as decision-support tools and the reliance on system recommendations is high, a credit crisis and systemic harm to economic stability could occur.

(b) **Cyber risks, fraud and disinformation** - Artificial intelligence could serve in various ways to harm the investing public and customers, financial entities, and the financial system on the whole, through the spread of disinformation and the execution of financial fraud. This is not a new phenomenon; however, the availability and distribution of AI tools make it easier to expand their scale and tailor them to new recipients, thereby increasing their effectiveness for bad actors.[62]

Particularly, disinformation risk is currently perceived as one of the most salient risks related to the development of artificial intelligence. Disinformation is particularly relevant to the financial sector, where rapid decisions are routinely made based on information. One example of this risk is deepfake content -- material presented as authentic that was generated by combining existing images, videos, voices and other content. The prevalence of deepfake content has grown dramatically in recent years.[63] Financial crime experts regard the phenomenon as one of the greatest risks the financial sector will face in the future.[64] Concerns have been raised that AI systems will render prohibited practices more complex and difficult to detect. For example, the spoofing offense - a situation where buy and sell orders that are not intended to materialize are fed into stock exchange systems solely in order to influence prices. Although this offense is not new, it could intensify and become difficult to detect and prove, especially in situations where different AI systems inadvertently interact.[65]

(c) **Risk to competition** - The risk of undermining competition is not unique to the financial sector, but it could intensify and increase in this sector, in light of the entry of AI technology, particularly in highly

---

[61] Findings Regarding the Market Events of May 6, 2010 – Report of the Staff of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues (September 2010), https://www.sec.gov/files/marketevents-report.pdf.

[62] Roy Goldschmidt, "Generative Artificial Intelligence: Opportunities, Risks, and Regulation," Knesset Research and Information Center (June 2023), p. 2-3, https://fs.knesset.gov.il/globaldocs/MMM/11189446-7fc1-ed11-8159-005056aa4246/2_11189446-7fc1-ed11-8159-005056aa4246_11_20161.pdf?fireglass_rsn=true#fireglass_params&tabid=ae6fbe7252902d0a&start_with_session_counter=4&application_server_address=isagovil-2-me-west1.prod.fire.glass.

[63] https://sumsub.com/newsroom/sumsub-research-global-deepfake-incidents-surge-tenfold-from-2022-to-2023/.

[64] Ibid.

[65] 2021 OECD Report, p. 28.

concentrated financial systems. For example, in a situation where there is a concentrated market with a limited number of AI system providers for the financial sector, the risk of undermining competition increases, adding to the systemic stability risks discussed above. The same is true with respect to the databases underlying AI systems. Concentration of extensive information with a limited number of providers could contribute to the erection of entry barriers that do not necessarily exist in "traditional" financial systems, especially with respect to system providers lacking broad access to information or sufficient resources to compete. In an AI environment, interaction between different AI systems could lead to price coordination or abuse of monopolistic standing, actions generally prohibited under competition law.[66]

(d) **High cost of error** - There are areas of activity where an AI system error could have a severe impact on individuals. In healthcare, for example, the incorrect decoding of a critical test or a mistaken diagnosis of a disease could lead to irreversible outcomes. This holds true for the financial sector as well. An AI-system error, such as a loss of savings resulting from an algorithmic portfolio management error, rejection of a loan application resulting from an AI-induced bias or discrimination, or the rejection of insurance underwriting resulting from the lack of an adequate database, could have a highly negative impact on individuals.

(e) **Confidence in the financial system** - The financial system is based on public confidence. Such confidence is necessary for activity in all financial fields. For example, the public deploys means of payment based on the assumption that these will be honored according to their full value, they deposit money in banks or other financial institutions based on the assumption that these institutions will safeguard and care for their funds; and it invests in securities based on the assumption of fair trading practices and the integrity of the companies issuing the securities. Lack of public confidence in the financial system, unlike other systems in different areas of life, means that the system will not exist at all, with all the implications this entails on the fabric of life familiar to us.

[66] Gal, Michal and Rubinfeld, Daniel L., Algorithms, AI and Mergers (June 5, 2023). Antitrust Law Journal (2023), NYU Law and Economics Research Paper No. 23-36, https://ssrn.com/abstract=4469586 .

F. Artificial Intelligence in Financial Regulation Worldwide

In recent years, the engagement in artificial intelligence, and its many potential applications in financial services, have prompted financial regulators worldwide to address the issue, particularly since the emergence of generative AI. In 2012, the OECD published a comprehensive report on artificial intelligence, machine learning and big data in the financial sector. The report surveys the challenges, benefits and implications of using AI technology in the financial sector and proposes regulatory approaches for policy-makers and regulators to adopt.[67] Subsequently, in 2023, the OECD published a review of the uses of generative AI in the financial sector. The paper describes the progress in the fields of generative AI and its (gradual) integration in financial organizations. It analyzes the potential risks inherent in widespread use of the technology, and recommends steps that policy-makers and regulators should take in this context.[68] Similarly, the International Monetary Fund published a report dealing with opportunities for using artificial intelligence and machine learning in the financial sector the unique challenges arising from this.[69] A subsequent report surveying the challenges in using generative AI in the financial worlds was later issued.[70] The Bank for International Settlements (BIS, the international organization of central banks) published a paper focusing on the function of big data and machine learning in central banks,[71] another paper focusing on principles of AI governance and human involvement,[72] as well as several additional papers. It dedicated a chapter in its 2024 Annual Economic Report to the implications of artificial intelligence on central banks.[73] IOSCO, the international organization of securities authorities, published a report concerning the use of artificial intelligence and machine learning by financial intermediaries and investment managers, surveying the risks associated with the technology and offering guidelines for policy-makers and regulators on issues of AI governance and supervision, algorithm development and monitoring, database quality and biases, transparency and explainability, outsourcing, as well as ethical considerations (hereinafter: "the 2021 IOSCO Report").[74]

---

[67] 2021 OECD Report.
[68] 2023 OECD Report.
[69] IMF, Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance (October 2021), https://www.imf.org/en/publications/departmental-papers-policy-papers/issues/2021/10/21/powering-the-digital-economy-opportunities-and-risks-of-artificial-intelligence-in-finance-494717.
[70] IMF, Generative Artificial Intelligence in Finance: Risk Considerations (August 2023), https://www.imf.org/en/publications/fintech-notes/issues/2023/08/18/generative-artificial-intelligence-in-finance-risk-considerations-537570.
[71] BIS, BIS Working Papers No 930 - Big data and machine learning in central banking (March 2021), https://www.bis.org/publ/work930.htm?force_isolation=true.
[72] FSI, FSI Insights on policy implementation No 35 - Humans keeping AI in check – emerging regulatory expectations in the financial sector (August 2021), https://www.bis.org/fsi/publ/insights35.htm?force_isolation=true.
[73] BIS Annual Economic Report 2024, https://www.bis.org/publ/arpdf/ar2024e.pdf.
[74] International Organization of Securities Commissions (IOSCO), The use of artificial intelligence and machine learning by market intermediaries and asset managers – Final report (September 2021), https://www.iosco.org/library/pubdocs/pdf/IOSCOPD684.pdf?force_isolation=true.

The European Securities Authority (ESMA) also published a paper on use of artificial intelligence in European stock markets, applications in use and the (slow) adoption pace.[75]

In the US, financial regulators addressed various issues in the realm of artificial intelligence. The U.S. Securities and Exchange Commission (SEC) proposed to amend its provisions regarding conflicts of interest arising from the use of predictive data analysis by investment advisors and broker-dealers (including AI-generated analyses) (hereinafter: "the 2023 SEC Proposal").[76] The Consumer Financial Protection Bureau (CFPB), a regulator dealing primarily with consumer protection in the fields of banking and credit, has issued several publications on artificial intelligence. These include guidance on the use of AI applications in customer service,[77] and a circular addressing the requirement for explainability when denying credit based on complex AI models (hereinafter: "the 2023 CFPB Circular").[78] The Federal Trade Commission (FTC) also released several statements, including a joint declaration with other regulators on enforcement actions against discriminatory or biased automated systems,[79] as well as a guidance affirming its authority to take enforcement measures against products and services that employ artificial intelligence.[80]

In the United Kingdom, the Financial Conduct Authority (FCA) and the Bank of England issued several publications on artificial intelligence,[81] including a discussion paper regarding the use of AI applications in the financial sector, and a survey of machine learning-based application usage by financial services in the U.K.[82] Although the papers do not include recommendations, they largely reflect the British government's regulatory approach described above - empowering sectoral regulators to establish policy and regulation in their respective fields.

---

[75] The 2023 ESMA Report.

[76] SEC, Conflicts of Interest Associated with the Use of Predictive Data Analytics by Broker-Dealers and Investment Advisers (July 2023), https://www.sec.gov/rules-regulations/2025/06/s7-12-23.

[77] See *supra* note 45.

[78] CFPB, Consumer Financial Protection Circular 2023-03 (September 2023), https://www.consumerfinance.gov/compliance/circulars/circular-2023-03-adverse-action-notification-requirements-and-the-proper-use-of-the-cfpbs-sample-forms-provided-in-regulation-b/?force_isolation=true.

[79] FTC, Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems (April 2023), https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf?for.

[80] FTC, Resolution Directing Use Of Compulsory Process in a Non-Public Investigation of Products and Services that Use or Are Produced Using Artificial Intelligence or that Purport to Detect the Use of Artificial Intelligence (November 2023), https://www.ftc.gov/system/files/ftc_gov/pdf/232%203087%20AI%20Omnibus%20Resolution.pdf?force_isolat.

[81] Bank of England & FCA, DP5/22 - Artificial Intelligence and Machine Learning: Discussion Paper 5/22 (October 2022), https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence?force_isolation=true.

[82] Bank of England & FCA, Machine Learning in UK Financial Services (October 2022), https://www.bankofengland.co.uk/report/2022/machine-learning-in-uk-financial-services.

Financial regulators in other developed countries such as Singapore,[83] Hong Kong,[84] the Netherlands,[85] Canada (Ontario)[86] have published reports regarding the use of AI systems in the financial sector, most of them dealing with general principles that financial entities should consider when developing or adopting AI systems.

In the time that has elapsed since the publication of the Interim Report for public comment, various regulatory developments have occurred. In December 2024, the US Treasury Department published a report dealing with artificial intelligence in financial services.[87] This report included recommendations directed primarily at government bodies and financial regulators, such as continuing international cooperation to promote consistent and meaningful standards regarding the use of artificial intelligence in the financial sector, promoting a dialogue with the financial services industry regarding existing regulatory gaps that could cause harm to financial services customers; improving existing risk management frameworks, and collaborating with other regulators to clarify oversight expectations in the sector; examining adaptation of AI applications by supervised entities to regulation requirements prior to launching their use and performing periodic examinations system compliance to financial regulation. In June 2025, the SEC decided to withdraw the proposal it published in 2023, which sought to deal with

---

[83] MAS, Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector (February 2019), https://www.mas.gov.sg/-/media/mas/news-and-publications/monographs-and-information-papers/feat-principles-updated-7-feb-19.pdf?force_isolation=true; The Singaporean paper is intended to promote fair, ethical, responsible, and transparent use of artificial intelligence systems and data analytics within financial institutions, through the application of general principles. The objectives of the paper are to provide financial institutions with basic principles for determining when to use artificial intelligence tools; to assist financial institutions in defining and classifying AI governance within their business models; and to promote public confidence and trust in the use of AI tools in the financial sector.

[84] HKMA, High-level Principles on Artificial Intelligence (November 2019), https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2019/20191101e1.pdf?fo; The banking regulator in Hong Kong published a guide for the banking sector on the use of artificial intelligence applications. In a survey of use cases it conducted, the regulator found that many banks have adopted, or plan to adopt, AI applications, including in customer service, marketing, internal and operational processes, risk management, cybersecurity, and more. Similar to the Singaporean paper, the Hong Kong regulator also established a number of general principles that banks are required to take into account in the design, adoption, and use of artificial intelligence applications and big data analytics. Among these principles are the responsibility of the board of directors and senior management for system outputs, the use of high-quality data, a requirement for sufficient explainability, rigorous validation of AI applications, the establishment of procedures and adequate oversight of AI application providers, the implementation of appropriate cybersecurity solutions, and others.

[85] DNB, General principles for the use of Artificial Intelligence in the financial sector (July 2019), https://www.dnb.nl/media/voffsric/general-principles-for-the-use-of-artificial-intelligence-in-the-financial-sector.pdf?force_isolation=true; The central bank of the Netherlands published a set of general principles for the use of artificial intelligence applications in the financial sector. The principles are divided into six main areas: robustness; accountability; fairness; ethics; competence; and transparency.

[86] OSC & EY, Artificial Intelligence in Capital Markets: Exploring Use Cases in Ontario (2023), https://www.osc.ca/sites/default/files/2024-11/pub_20231010_artificial-intelligence-in-capital-markets.pdf; The paper issued by the securities regulator in the Province of Ontario, Canada, reviews the expected benefits and challenges arising from the use of artificial intelligence applications, presents case studies of such uses within financial institutions under its supervision, and maps the risks associated with the main applications. The picture emerging from the report indicates that the adoption of AI applications in Ontario, Canada, is gradually increasing and is currently at an intermediate stage. In addition, the report finds that existing governance frameworks do not provide an adequate response to the challenges posed by AI applications, and that it is necessary to develop and adopt governance rules that emphasize responsible use and explainability of AI applications, alongside appropriate consideration within the risk management frameworks of financial institutions.

[87] Request for Information - https://home.treasury.gov/system/files/136/Treasury-AI-RFI-financial-sector-2024.pdf; Final Report - https://home.treasury.gov/news/press-releases/jy2760.

conflicts of interest in connection with Predictive Data Analytics analysis in broker-dealer and investment advisor operations. This proposal drew criticism, and, as mentioned, the SEC announced that it does not intend to publish new regulation based on it.[88]

In the United Kingdom, the policy is to rely on existing regulation, while granting regulators flexibility in the areas under their purview, and avoiding cross-sectoral AI legislation. This approach was expressed in the FCA's five-year strategic plan for 2025-2030.[89] The British financial regulators (FCA and the Prudential Regulation Authority (PRA)) published an additional usage survey pertaining to the use of artificial intelligence in the financial sector,[90] in which it was found that approximately 75% of financial companies use artificial intelligence, while approximately another 10% intend to deploy AI systems for use in the next three years. These figures are significantly higher than those revealed in the previous usage survey. At the same time, adoption of completely autonomous, automated AI-based solutions still progresses at a slow and measured pace. In only 2% of the cases examined in the usage survey were AI systems found to make decisions independently and conduct fully autonomous operations.[91]

As a complementary move to the FCA's efforts to support adoption of AI systems in the financial sector, a dedicated regulatory AI sandbox was recently established, designed for companies in the early stages of identifying and adopting AI-based solutions.[92] In addition, an AI laboratory was established at the FCA, whose purpose is to enable the regulator to understand the challenges and opportunities involved in using AI systems in the financial sector in depth. Taking a practical, collaborative regulatory approach, it seeks to understand how these systems affect customers and the market overall.[93] In parallel, the FCA announced that it is developing a dedicated statutory code of practice for companies developing and integrating AI systems in the financial sector.[94]

The overall picture emerging from these developments is that globally, there is generally no dedicated regulation of the use of artificial intelligence in the financial sector. Most regulators focus instead on monitoring and reviewing the issue, and on providing guidance concerning the application and implementation of existing regulatory frameworks in the context of AI usage and related issues. It should be noted, however, that in some countries, primarily the European Union member states, cross-sectoral AI legislation has been enacted which generally applies to the financial sector as well. It should also be emphasized that financial regulation tends to follow market developments. Given the currently limited adoption of artificial intelligence in the financial sector, it is not surprising that dedicated regulation has

---

[88] https://www.sec.gov/rules-regulations/2025/06/s7-12-23.
[89] https://www.fca.org.uk/publications/business-plans/annual-work-programme-2025-26?force_solation=true.
[90] This constitutes an update to the survey of uses conducted in 2022; see *supra* note 74.
[91] https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024?force.
[92] https://www.fca.org.uk/news/press-releases/fca-allows-firms-experiment-ai-alongside-nvidia?force.
[93] https://www.fca.org.uk/firms/innovation/ai-lab.
[94] https://www.ftadviser.com/content/dbb2992c-591b-4c65-96b7-302635242a17.

not yet been introduced to address the issue. As the use of artificial intelligence becomes more extensive and more significant, financial regulation can be expected to evolve accordingly.

G. AI Activity in the Financial Sector in Israel

As in other parts of the world, financial entities in Israel have also begun to employ AI-based applications across a range of activities. For example, in the banking sector, AI systems have been used for several years for underwriting and rating borrower creditworthiness. In addition, entities engaged in portfolio management and securities trading use AI systems for algorithmic and high-frequency trading, analyzing business and for economic analysis and forecasting, and the like.

Although the use of AI applications by financial entities in Israel is not new, and in keeping with trends emerging worldwide, the expected change is in the scope of AI use and in its capabilities. Indeed, the development of increasingly sophisticated AI systems, including generative AI, has sparked considerable interest within Israel's financial sector and various potential applications are being vetted. At the same time, the Taskforce's impression is that, for now at least, the pace of adoption remains slow. Key factors contributing to the sluggish pace of adoption include the persistence of legacy systems and the substantial investment required to replace them, concerns regarding the AI system reliability, and the risks associated with integrating such systems into financial operations– for example, with respect to compliance and reputation. The Taskforce observed the emergence of a potential gap in the extent of AI use between large, well-established "traditional" financial entities and smaller new entities The latter appear to be more willing to invest greater efforts to adopt artificial intelligence, *inter alia*, as a means to acquire competitive advantage.

It appears then that we are still some way away from seeing advanced AI applications that can carry out end-to-end core financial entity activities automatically, autonomously, and directly with the customer, devoid of any human involvement. Most activities currently performed by AI systems in Israel's financial sector assist human decision-making when performing complex tasks based on extensive data, or they are operational (back-office) functions designed to streamline processes, detect financial fraud, enhance risk management, and support marketing and sales.

From a usage survey conducted in 2022 by the Capital Market, Insurance and Savings Authority regarding the use of AI applications by institutional bodies representing a variety of financial services companies in the field of insurance and long-term savings, it was found, *inter alia*, that AI applications have been launched at several entities and additional applications are currently in development, such that an upward trajectory in AI use can be expected. The areas of AI application use are numerous and diverse, including analysis of insurance client behavior, claims settlement, customer service and operational streamlining.

A usage survey conducted in 2024 by the Israel Securities Authority regarding AI use in mutual fund management, a similar picture emerged. About half of fund managers testified that they use AI systems, but as mentioned above, this use is limited, concentrating primarily on analysis and research, securities

trading and investment decision-making. In hedged mutual funds, which enjoy more freedom of action, as well as in flexible funds, AI use is somewhat broader relative to its use by other mutual funds.

It should be noted that, in general, it is difficult to accurately assess the pace and scope of advanced artificial intelligence adoption in the financial sector. Surveys on the subject are limited, given the lack of a uniform definition of artificial intelligence activity, and the wide range of potential applications facing financial institutions.

The overall picture regarding the financial sector in Israel is consistent with findings from other countries, indicating that the adoption of artificial intelligence in this sector remains slow and controlled. Nevertheless, the assessment of industry participants is that artificial intelligence will play an increasingly significant role in financial activities, a view reflected in the resources invested in exploring and integrating potential AI applications.

H. The Inter-Ministerial Taskforce for Artificial Intelligence in the Financial Sector

(1) Background to the Taskforce's Establishment

In December 2020, the Economic Department at Consultation and Legislation in the Ministry of Justice published a call for conducting research and a review of common uses of artificial intelligence and big data in the financial sector and a comparative review of how regulators deal with these uses. An interdisciplinary Taskforce from Tel Aviv University, headed by Prof. Assaf Hamdani and Prof. Kobi Kastiel (from the Faculty of Law) and Prof. Dan Amiram and Dr. Ari Ahiezer (from the Faculty of Management), was selected to head the research. The Taskforce mapped common uses of artificial intelligence in the financial sector, presented the fundamental issues arising from the use of AI tools in general (focusing on issues unique to the financial sector), surveyed the regulatory and legal implications of employing artificial intelligence in the financial sector and examined global regulatory trends regarding AI, addressing aspects related specifically to the financial sector (hereinafter: "the Tel Aviv University Report").[95] The Tel Aviv University Report was published in August 2022, and its key recommendations include:

a. Israel's position as an innovation hub should be preserved, while maintaining a balance with the protection of interests, such as consumer protection and systemic financial stability.

b. Future regulation in the field of artificial intelligence, particularly in the financial sector, should be adapted to emerging international standards to prevent the erection of barriers to entry for foreign participants and to encourage the entry of new participants (foreign and local) and new technologies to the Israeli market.

c. At this stage, the establishment of uniform rules for all uses of AI applications solely due to their being based on similar technology should be avoided. Distinctions should be drawn between uses that may also generate various risks. At the same time, however, there is a need for regulatory coordination in the various branches of law and between relevant supervisory entities in cases where different uses share similar characteristics. In either case, government policy must be clear in both design and implementation. The researchers recommended placing such coordination of all AI applications across various sectors in the hands of a single government entity with expertise in the field.

d. At least at this stage, where regulatory uncertainty regarding the use of AI applications prevails, the development of self-regulation based on accepted principles and goals should be encouraged. In

---

[95] Achiaz, Hamdani, Amiram, and Kastiel, Artificial Intelligence in the Financial Sector: Common Applications, Challenges, and a Comparative Survey of Regulatory Approaches (Report submitted to the Economic Department, Counseling and Legislation Division, Ministry of Justice, July 2022), https://www.gov.il/BlobFolder/news/ai_report/he/AI_report.pdf?fireglass_rsn=true.

addition, the report recommends the launch of regulatory sandboxes that can assist both in accelerating the introduction of new technologies to the market, and in enhancing the grasp of the relevant regulators of the technology underlying new services or products and facilitating the development of expertise on the subject (which can subsequently aid the formulation and design of appropriate regulation).

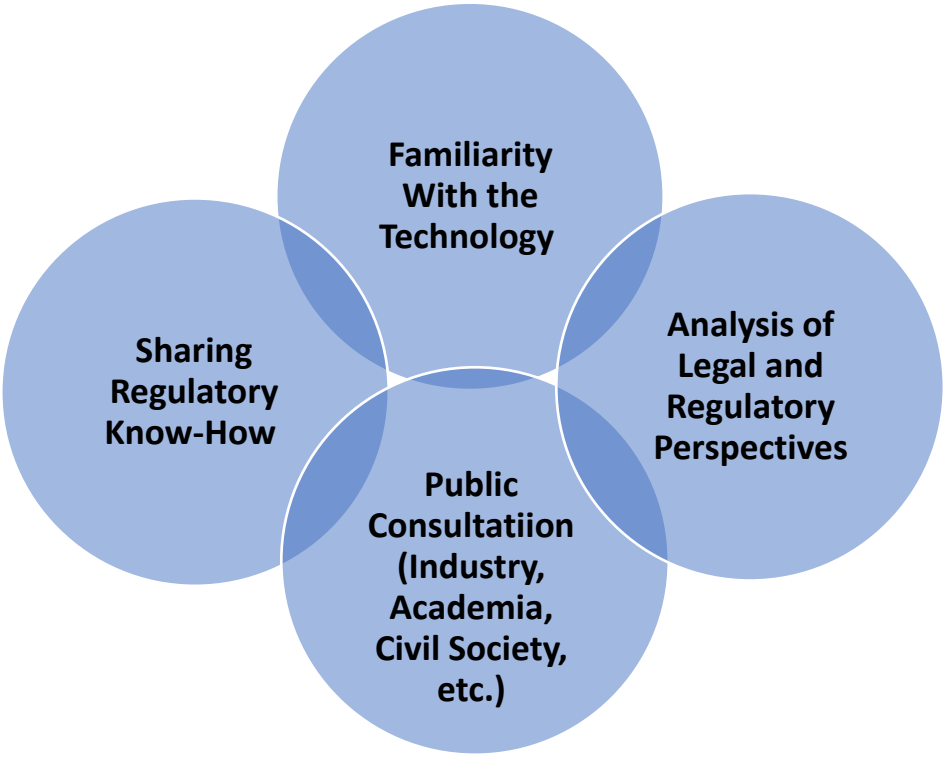
e. The entities engaged in financial oversight should coordinate licensing and liability regimes, including the establishment of uniform arrangements. The report also proposed that entry barriers to the Israeli market and the technological impacts on them be examined, and that licensing procedures in the various fields be adjusted accordingly. Similarly, in light of the fact that there are several financial supervisors operate in the Israeli financial market, it is recommended to focus on unifying the demands of the oversight authorities pertaining to issues with similar technological characteristics to avoid regulatory uncertainty and potential regulatory arbitrage.

f. Among the range of fields reviewed, it was recommended to focus regulation development efforts in the financial sector in the following areas: (a) investment and pension planning advice; (b) credit.

Following this report, against the background of rapid developments in AI technology, and in keeping with the recommendations emerging from the Policy Document with respect to sectoral regulation, the Taskforce for Examining the Impacts of Algorithmic and AI systems in the Financial Sector was established (hereinafter: "the Taskforce") in December 2022.

(2) The Taskforce's Work

The Taskforce held dozens of meetings in which it studied possible financial sector AI applications in depth, analyzed key issues arising in connection with this activity, examined emerging regulation abroad, and discussed recommendations regarding this activity in Israel.



It operated in several areas, including:

a. **Familiarity with the technology underlying AI applications** - The Taskforce met with senior technology leaders, from Israel and abroad, to understand the foundations on which artificial intelligence technology is based.

b. **Sharing regulatory know-how** - Taskforce members shared regulatory and research activities performed by each of the member bodies regarding artificial intelligence.

c. **Analysis of legal and regulatory perspectives of artificial intelligence** - The Taskforce engaged extensively in studying the legal and regulatory issues associated with AI and met with content development experts from academia and the private sector to understand key legal issues pertaining to the use of AI applications in depth .

d. **Call for comments and meetings with various entities from industry, academia, civil society, government entities and foreign regulators** - In April 2023, the Taskforce published a call for comments, in which the general public and stakeholders were requested to express their positions and address various issues, including: the current and future uses of artificial intelligence in the financial sector; barriers (technological, business, legal) facing financial entities and others related to the development, experimentation and use of artificial intelligence tools; benefits and challenges and risks that may arise in the integration of AI tools in financial entities; instruments to deal with challenges and risks in using artificial intelligence in the financial sector; proper AI regulatory policy in the financial sector and more.[96]

Seventeen responses were received to the call from industry entities (including consulting companies, professional associations and international law firms), civil society and academia. The references touched on numerous relevant issues, including insurance, credit, discrimination, money laundering, detecting embezzlement, fraud and capital market anomalies, confidentiality and privacy protection, human capital, responsible AI use and explainability, competition, the various risks associated with AI use, the proper regulatory framework, etc. Appendix X to the final report will include summaries of some of the key comments submitted to the Taskforce.

In addition, the Taskforce met with various stakeholders in the field, including entities that submitted their response to the call as well as other entities from myriad sectors from whom the Taskforce requested their position. For example, the Taskforce met with industry practitioners from banking, insurance, algorithmic trading, technology (including technology giants), anti-money laundering authorities, as well as with members of academia (local and international) specializing in law, public policy and computer science. It also met with representatives from government and regulatory agencies, both local

[96] https://www.new.isa.gov.il/images/Fittings/isa/asset_library_pic/al_lobby/al_lobby-637f62a916518/Artificial_intelligence_team_financial_sector.pdf?fireglass_rsn=true#fireglass_params&tabid=a315d0774bfe64f2&start_with_session_counter=3&application_server_address=isagovil-2-me-west1.prod.fire.glass.

and foreign, including the Israel Privacy Authority, OECD, European Union, ESMA, and regulators in the US, United Kingdom and Singapore, as well as from civil society organizations. The meetings covered a wide range of subjects and were instrumental to the Taskforce in many respects, including in understanding the AI applications currently in use in the market and those expected to be launched in the future. In addition, the Taskforce held workshops dealing with "disruptive" innovative technology and the potential for dealing with such technology and the applications it engenders.

(3) The Interim Report

The Taskforce's Interim Report was published for public comment until November 15, 2024. Additional work was subsequently conducted by the Taskforce, including meetings with numerous entities, seminars, and deliberation of the public comments, toward the submission of its final report and recommendations. 24 public comment documents were received to the Interim Report. A summary table of these comments appears in Appendix D to the report.

Part B

General Approach, Guiding Principles, and the Definition of Artificial Intelligence in the Financial Sector

A. The General Approach – Encouraging Innovation and Mitigating Risks

As the capabilities of artificial intelligence technology and the rapid pace of its development became clearer, global discourse intensified around its potential contributions, the risks it entails, and the appropriate regulatory response to these risks. The suitable regulatory approach with respect to the technology is derived primarily from identifying the potential risks posed by the technology and assessing their severity, alongside evaluating the potential benefits from its use.

Without derogating from the many concerns raised globally regarding the implications of AI technology, the Taskforce believes that with respect to AI activity in the financial sector, the considerations for encouraging such activity are compelling.

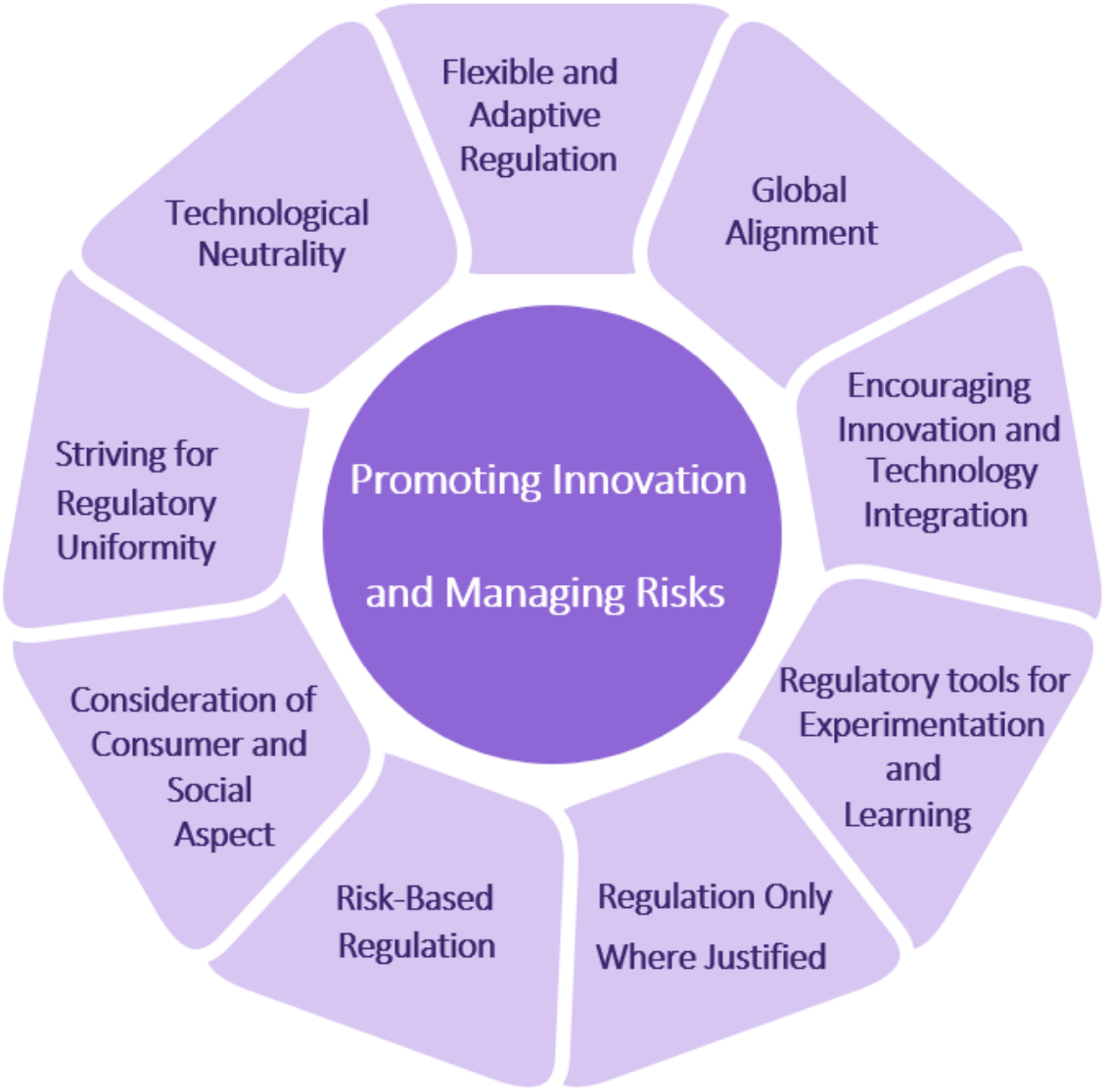
First, the financial sector is already under relatively close supervision, applied to a wide range of activities and entities, facilitating the ability to address the risks arising from AI activity within the framework of regulatory actions and regular oversight. Second, in view of technological and regulatory developments globally, it appears that the broad integration of AI technology in the financial sector is merely a matter of time. In other words, the question is not whether the technology will be integrated, but rather when and in what manner. It appears, therefore, that, under the circumstances, attempting to ignore AI technology (or limit its use) is not the right solution. Third, the potential AI technology engenders to significantly improve financial services, reduce their costs, and enhance competition in the sector is immense. The outputs of AI in the financial sector may have positive implications for a large range of current and future activities.

The above does not ignore the broader discourse on the general risks associated with artificial intelligence, which even if not unique to activity in the financial sector, may nonetheless manifest and take shape in this field. Under these circumstances, there appears to be a need to contend with AI's entry into the financial sector and to prepare for the changes it is expected to precipitate, while encouraging innovation that can benefit consumers and the responsible adoption of the technology by establishing appropriate "rules of the game," and managing the risks involved.

In light of the above, the approach that guided the Taskforce and shaped the recommendations outlined in the report advocates encouraging AI activity in the financial sector alongside establishing tailored regulation where necessary. This approach facilitates better integration of the technology and addresses the risks associated with this activity more effectively. The risks associated with AI's integration into financial activity are significant; however, the Taskforce's position is that these risks that can be adequately mitigated.

This approach aligns with the overall government strategy taking shape with respect to artificial intelligence, as well as with the principles proposed below (for elaboration, see the Policy Document and the discussion in the chapter on AI activity in Israel's financial sector above).

B. Guiding Principles for AI Use in the Financial Sector



1. Flexible and Adaptive Regulation

One of the challenges in designing regulation for technology in general – and artificial intelligence in particular– is the continual progress of the technology both during and after the regulatory framework is formulated. This is especially true with respect to artificial intelligence, where new capabilities emerge on a daily basis, many of which will likely be relevant to the financial sector as well. At the same time, even AI's existing capabilities have not yet been fully deployed or utilized for various reasons. As stated above, this challenge is characteristic of technological development in general, requiring regulators to balance multiple considerations when determining the appropriate timing for regulatory intervention. Foremost among them is the desire not to hinder technological progress while fulfilling the obligation

to protect the public interest. Although this challenge is not a new, it appears to be particularly acute within the context of artificial intelligence.

In light of the above, the ISA favors flexible regulatory tools that be easily updated as the technology develops, and no less importantly, after the regulator has had an opportunity to become familiar with technology, its impact on the area of its oversight, global developments, and the like. Regulatory tools can include directives, regulations (as temporary provisions) and standards. There will also be room to adopt "soft" regulation, such as self-regulation, that leaves a certain degree of discretion in the hands of the supervised entity. In this context, the relationship between technology entities and the supervised parties is important, *inter alia*, in order to learn about the degree of AI application adoption and the internal measures taken by users with respect to AI integration into their operations and the measures required for risk management.

There may, of course, be exceptions to this recommendation; for example, cases in which existing statutory regulation creates a limitation or the legislation itself must be amended. Another exception relates to cases where a special risk is detected or where it is not possible to address the public interest in the development or regulation of the activity without amending legislation.

Thus, the Taskforce's position is that at this stage, a regulatory toolbox should be adopted that will include adaptive and soft regulatory tools. Generally speaking, there is no room at this stage to adopt rigid statutory measures for the regulation of AI use in the financial sector.

2. Alignment with International Regulation

Alongside the desire to maintain the independence of regulation in Israel, the Taskforce sees importance in adopting regulation aligning with the regulation accepted in developed markets abroad. Neither the potential nor the risks associated with artificial intelligence are unique to Israel, and therefore, it is possible in this context to make use of rules or standards as they develop worldwide. It may even be possible to rely on foreign regulation that will be developed, such as regulation on international technology suppliers or on AI-based financial products and services. Second, establishing unique standards in Israel could hinder the financial entities' ability to adopt services and products not specifically tailored to Israeli regulation. For example, requirements that will necessitate technological adaptation that are not required with respect to other large markets could deter technology suppliers from operating in Israel and lead to a situation in which it will not be possible to integrate those products and services. The Israeli public will lose out as a result. Finally, technological and financial "players" operating in Israel also direct their activity to international markets. These entities will have difficulty adapting their products to both unique Israeli regulation and the regulation effective in other markets. This could harm both these players and the financial entities that will struggle to integrate advanced products into their activity. For this reason, as well, then, it is necessary to ensure that the regulation

adopted in Israel aligns with that worldwide and will not create unnecessary barriers to the transfer of products and services between Israel and other countries.

There is currently no global regulatory standard concerning artificial intelligence either in general, or specifically for the financial sector, even though initiatives and emerging regulatory directions abound. Therefore, the recommendations in the report include an overview of developments through the world as well as the aspiration for harmonization with the regulation emerging in developed countries, but do not adopt a specific foreign standard.

3. Promotion of Innovation and Technological Integration

Another challenge accompanying this issue is the lack of clarity regarding the degree of AI technology integration into Israel's financial sector. The Taskforce's impression from meetings held with market participants and from surveys conducted by the entities comprising the Taskforce's membership was that, although the technology arouses significant interest among financial entities, integration within these entities remains nascent and uneven across entities and use. From the meetings, it emerged that "traditional" financial institutions are currently examining possibilities for integrating AI technology into their operations, but appear to be in no rush to do so. AI adoption is being carried out in deliberate and carefully measured steps. Additionally, it appears that extensive technological activity intended for the financial sector, is taking place in Israel, but that the companies engaged in this are directing their efforts at markets abroad. The Taskforce's impression is that the integration of the technology among financial entities is adversely affected by regulatory uncertainty, but it appears that the main reason for this is not rooted in the regulation, but rather in the investment involved in changing legacy systems and the need to build "trust" in the new and highly dynamic technology. This state of affairs, which is seemingly not unique to Israel, enables orderly preparation for the integration of AI technology in financial entities in Israel.

The Taskforce attaches importance to deploying AI applications where the technology holds positive potential for the economy and consumers. Therefore, beyond questions of preparation for the integration of the technology, the recommendations also relate to how to encourage the entry of the technology into financial activity conducted in Israel and to promote competition in the market (including through the removal of barriers, the creation of new opportunities, the elimination of unnecessary restrictions interfering with market development or excessive regulatory costs, etc.). These recommendations reach beyond those concerning the manner in which existing or anticipated activity should be regulated.

4. Experimentation and Learning

Following the above and consistent with the approach of adopting technology in a measured and cautious manner, the Taskforce supports the use of regulatory tools which enable limited experimentation with the technology while providing regulatory certainty, even if only temporary, for the relevant participants.

In this manner, both the financial entity and the regulator will be able to familiarize themselves with the technology and obtain data on how it functions. enabling the subsequent formulation of "smart" and data-based regulation. Several years ago, a bill was prepared for "regulatory experimentation" in the financial sector (the Sandbox Bill). It appears, however, that it is inadequate as written, and modifications should be considered. Additional legal and regulatory tools to encourage experimentation and learning, some of which already exist within the framework of the authorities vested in the financial regulators (interpretative positions, preliminary inquiries, and the like) should also be advanced.[97] The experience accumulated by the member entities of the Taskforce support the notion that recommendations in this area should not be limited to creating a regulatory environment that enables experimentation, but should also address the envelope provided to entities seeking to experiment with the technology, including technological and business shepherding, the provision of business incentives, and similar measures. This approach may also assist the market entry of small and medium-size enterprises who require assistance or relief to cope with Israel's business and financial environment.

5. Circumstances Justifying Regulatory Changes

A key question raised in the Taskforce's discussions – and in meetings held with the private sector representative and researchers in the field – concerned whether regulation or existing rules should be revisited solely because artificial intelligence is being integrated into existing activity, where certain statistical or automated tools are already in use. In other words, should the same activity be regulated differently merely because it is conducted with the assistance of, or through, artificial intelligence? The Taskforce's approach is that the implications of the use of artificial intelligence for the service and for consumers should be assess, and only then should it be determined whether amending the existing regulation is required. That is, we do not assume that the mere use of artificial intelligence in the provision of financial services, necessitates regulatory change. For example, where existing law does not impose an obligation to provide reasoning, the question of whether an explainability requirement towards the customer should be created solely due to the transition to using an AI tool should be examined.

6. Risk-Based Approach

As in the Policy Document, and consistent with the policy currently carried out by financial regulators, the Taskforce's recommendations also adopt a risk-based regulatory approach. That is, the necessity and scope of regulatory change depend on the risk associated with in the specific activity or service, taking into account the availability of tools to address that risk and the impact of artificial intelligence on it. Within this framework, it is essential to assess the specific risk associated with each financial institution

[97] For further discussion, see Part E - Additional Measures to Promote Financial Regulation in the Field of Artificial Intelligence, below.

and each type of financial activity, and to weigh that risk against the potential benefits derived from integrating the technology in these operations. On this matter, see the chapter on comprehensive and risk-based perspective below.

7. Incorporating Consumer Protection, Social and Human Rights Considerations

In addition to the risks confronting financial institutions, artificial intelligence may heighten or create consumer and social risks as well. In line with approaches articulated in regulatory documents assessing AI-related risks to citizens, including the implications of these risks for consumers and the society at large, the adoption of a similar perspective in Israel is warranted. For example, under the EU AI Act, consumer credit is classified as a high-risk area due to its implications on the individual's economic mobility. Additional risks may arise from the use of machine learning techniques, which could make it difficult to interpret system outputs and to assess potential discriminatory effects.

8. Regulatory Consistency

The report and its recommendations cover a wide array of financial services and address various types of financial entities. Some of these entities raise similar questions even when the services differ (for example, the importance of customer service), while others raise similar questions because the financial services share certain characteristics (for example, in the field of advice). The Taskforce's general approach is that regulatory consistency should be pursued for comparable services and risks, even if there may be certain disparities arising from the nature of the financial products subject to regulation. Regulation of a particular financial activity should be grounded in the characteristics and inherent risks of that activity and should not be influenced by the identity of the body serving as regulator. In other words, similar activities bearing similar risks should be subject to similar regulatory treatment in order to prevent regulatory arbitrage. Consequently, the development of a uniform regulatory language is also important – for example, for the purpose of defining the risks associated with the technology or shaping the regulatory response.

The Taskforce's working assumption is that the technology will continue to develop and create opportunities and challenges that require continual cross-sectoral assessments, and therefore, the continuation of the dialogue and coordination that was started in the Taskforce's work is essential.

9. Technological Neutrality

The Taskforce recognizes that the financial services reviewed in its work may be delivered in various ways – through innovative technologies, through traditional methods, through various AI models, and the like. The manner in which the financial service is performed or the choice of a particular technology should not necessarily require different regulatory treatment. Instead, the implications of the technology for the financial service should be assessed and the regulatory response should address the opportunities or challenges arising from its use. For example, traditional portfolio management involves risks

stemming from the transfer of investors' funds to another party authorized to make investment decisions at its discretion. When this activity is conducted using artificial intelligence, the general regulatory framework applicable to traditional portfolio management activity remains relevant, and, in keeping with the principle of technological neutrality, the specific changes stemming from the technology should be identified and addressed. This approach ensures that regulation does not confer an unwarranted advantage of one technology over another, and thereby enables the development of a market in which technologically diverse services can compete under similar conditions.

C. Defining Artificial Intelligence

Defining artificial intelligence is itself a complex task. Globally, there is a wide spectrum of approaches to defining what artificial intelligence entails. The variations among definitions stem, *inter alia*, from the context in they were formulated and the time period in which they were proposed. Some definitions are more theoretical in nature, seeking to pinpoint the fundamental difference between human intelligence and artificial intelligence, while others are more practical, aiming to characterize artificial intelligence for the purpose of determining policy or the law that will govern AI systems. These latter definitions require greater precision. Differences in definitions also stem, as noted above, from the timing of their proposal, reflecting the state of technology and its uses at that time. More recent definitions, for example, refer specifically to more advanced forms of artificial intelligence, such as Generative AI, Artificial General Intelligence, and foundation models. In some cases, separate definitions are given for different types of AI systems.

The definition of artificial intelligence is important to this report not only for the purpose of describing the phenomenon with which we are dealing, and analyzing its potential, risks, the existing and required regulatory response, but also because the definition has serious regulatory implications. The definition we propose for the purposes of financial regulation aims to assist regulators in their regulatory and oversight activities. For example, regulators may make use of the proposed definition for the purpose of drafting provisions for imposing obligations on the supervised entities employing the artificial intelligence. This means that, despite the difficulty inherent in articulating a definition of artificial intelligence, because the definition has practical implications, it must provide sufficient clarity and legal certainty with respect to its application.

What are the main challenges in establishing a definition for artificial intelligence?

A first challenge lies in formulating a definition that can encompass existing and future technologies. Definitions of artificial intelligence generally seek to capture a wide range of technologies – those existing and those yet to emerge. While this objective is understandable, given the pace of technological advancement and the diversity of AI systems, it makes it difficult to formulate a precise definition. This challenge becomes particularly significant when the purpose of the definition is regulatory.

Another challenge stems from the fact that artificial intelligence can be deployed across very diverse contexts. As noted at the outset, artificial intelligence may be used in all areas of life, including throughout the financial sector. It may contribute to compliance with legal requirements, support customer service, or play a role in core financial service operations such as credit underwriting or investment advice. AI systems may serve as a tool to support decision-making or as a decision-making tool. These two challenges highlight a broader tension in defining artificial intelligence for the purposes

of financial regulation. On the one hand, there is a desire to maintain flexibility and technological neutrality, so that that definition accommodates different types of artificial intelligence and keeps up with the dynamic pace with technological change. On the other hand, the need to provide a coherent regulatory response to changes, and ensure legal certainty so that the regulation can be clearly applied and aligned with actual activity. A further tension arises between definitions that are overly broad, which could lead to over-regulation, and definitions that are too narrow, which could lead to insufficient regulatory protection.

A third challenge regarding the definition of artificial intelligence, which arose in the dialogue the Taskforce held with the supervised entities, relates to the distinction between standard and sophisticated computer-based applications. Applications currently perceived as basic and standard, such as Excel software, perform operations that were previously performed manually or assist in making decisions based on statistical calculation. For this reason, the definition of artificial intelligence should be limited to sophisticated computer-based applications that incorporate a machine-learning component. Prominent definitions of artificial intelligence worldwide do not include this element, and instead seek to encompass AI systems that do not incorporate machine learning, while excluding basic and routine applications. This likely reflects the belief that even systems that do not explicitly include machine learning may still be highly significant. We note here that the definitions of artificial intelligence vary in this context, a point which we address this later.

A fourth challenge concerns the question of whether we should rely on an international definition of artificial intelligence. There is an advantage in choosing a definition similar to that accepted abroad. Uniformity of terminology may facilitate AI technology use that is already subject to standards established abroad. At the same time, however, at this stage there is no single agreed-upon definition of artificial intelligence, let alone such a definition with respect to the financial sector.

Regulatory documents address these challenges in different ways. In some cases, multiple definitions appear in the same document – for example, a basic definition of artificial intelligence alongside an additional definition of General-Purpose AI. Similarly, definitions can include additional elements, such as the level of risk associated with using an AI system, or the system's degree of materiality of, which can affect whether or how the regulation is applied. While establishing multiple layers of definitions can be useful, it can create an over-abundance of possible options that adds to an already complex landscape.

Examples of definitions of artificial intelligence

To date, Israel has not adopted a legally binding definition of artificial intelligence. The Policy Document clarifies that, for the purpose of the report, "the field of artificial intelligence is a collective name for development in the field of information technology, communications and data science, which enables decision-making, generating predictions, or performing operations by a computer at a high level of

autonomy, in a manner that simulates or is capable of replacing human intelligence." The Paper further notes that formulating a formal definition of artificial intelligence is a simple task. In particular, it highlights the challenge of drafting a regulatory definition, which captures all computer systems that should be subject to regulation, while excluding simple and everyday systems that do not present any significant risk.[98]

Among the many definitions worldwide, the OECD definition provides a helpful starting point:[99]

> "AI system: An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."

Another definition of importance, which also has binding force (on EU states and parties engaging with EU states on AI), is the one established in the EU AI Act:[100]

> AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments.

In contrast to an earlier version of the definition, which included an annex containing techniques and approaches of artificial intelligence, this definition is general and (intentionally) more similar to the OECD definition. The EU AI Act has also formulated an additional and separate definition for General Purpose AI models.[101] The original definition in the European Union's proposed legislation was written prior to the breakthrough of Large Language Models, and this update is intended to accommodate the technological change. Alongside the additional definition, separate regulatory provisions have also been established.

Similarly, artificial intelligence is defined in the U.S. federal law (15 U.S.C. 9401(3) as follows:[102]

---

[98] Policy Document, p. 18.

[99] OECD (2024), "Explanatory memorandum on the updated OECD definition of an AI system", OECD Artificial Intelligence Papers, No. 8, OECD Publishing, Paris, https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html; This definition replaced the original definition presented by the OECD in 2019.

[100] EU AI Act, Article 3(1), *supra* note 12.

[101] 'general-purpose AI model' means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development and prototyping activities before they are placed on the market. EU AI Act, Article 3(63).

[102] The Federal Law. The federal law included additional relevant definitions such as AI system, generative AI, and machine learning, thereby providing examples of layered definitions relating to artificial intelligence.

> The term "artificial intelligence" means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems use machine- and human-based inputs to: (a) perceive real and virtual environments; (b) abstract such perceptions into models through analysis in an automated manner; and (c) use model inference to formulate options for information or action.

Additional definitions worth review have been issued in Singapore and Canada.[103] These definitions share several common features: in the definition of AI, they avoid reference to any specific technology; they recognize that AI systems can operate at different levels of autonomy and automation; and they refer to the systems' use, capabilities and diverse system outputs, such as predictions, recommendations, decisions or content generation. It is important to note that, due to their general nature, these definitions may also apply in financial contexts, the focus of this report. In the realm of financial regulation, the Bank of England's and the FCA's explicit reference to the term "artificial intelligence" is noteworthy. AI is viewed as a simulation of human intelligence by machines, including the use of computer systems capable of performing tasks that reflect learning, decision-making, problem-solving, and additional tasks that previously required human intelligence. Machine learning is mentioned as one type of artificial intelligence.[104] This general reference bypasses the need for a precise formulation of the definition, but it does not yet serve regulatory purpose, but rather appears in a discussion paper.

The U.S. Securities and Exchange Commission (SEC) proposed a formal definition within the framework of rules it published for public comment regarding potential conflicts of interest of brokers

---

[103] The definition proposed in the Singapore Policy Document of 2020: "AI refers to a set of technologies that seek to simulate human traits such as knowledge, reasoning, problem solving, perception, learning and planning, and, depending on the AI model, produce an output or decision (such as a prediction, recommendation, and/or classification). AI technologies rely on AI algorithms to generate models. The most appropriate model(s) is/are selected and deployed in a production system." Singapore Model Artificial Intelligence Governance Framework (2020), Sections 2.14-2.15, https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf

And the definition in the Canadian bill of 2022: "artificial intelligence system means a technological system that, autonomously or partly autonomously, processes data related to human activities through the use of a generic algorithm, a neural network, machine learning or another technique in order to generate content or make decisions, recommendations or predictions. (système d'intelligence artificielle)".

[104] See *supra* note 74: "What is AI?
2.10 While there is no consensus on a single definition, it is generally accepted that AI is the simulation of human intelligence by machines, including the use of computer systems, which have the ability to perform tasks that demonstrate learning, decision-making, problem solving, and other tasks which previously required human intelligence. Machine learning is a sub-branch of AI. AI, a branch of computer science, is complex and evolving in terms of its precise definition. It is broadly seen as part of a spectrum of computational and mathematical methodologies that include innovative data analytics and data modelling techniques.
2.11 The Bank and FCA previously defined AI in non-legal terms as 'the theory and development of computer systems able to perform tasks which previously required human intelligence'. The term AI system refers to the set of integrated computational elements and microservices that input, output, process, and store data and information. Therefore, one AI system may include multiple AI algorithms, models, and datasets."
https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence.

and investment advisers. The definition aims at incorporating also AI activity technology covered by the rules as well:[105]

> "Covered Technology means an analytical, technological, or computational function, algorithm, model, correlation matrix, or similar method or process that optimizes for, predicts, guides, forecasts, or directs investment-related behaviors or outcomes."

Key issues examined by the Taskforce concerning a possible definition of artificial intelligence:

(1) Is there room to refer in the definition to a particular technology, or to specify several technologies within it?

In light of the pace of change in the development of artificial intelligence technologies, in the Taskforce's opinion, a definition encompassing various technologies and techniques that may develop over time is desirable. This approach is evident in definitions adopted worldwide. In the European Union's original definition of AI, for example, the term "software" appeared together with a list of technologies listed in an annex. The term used in the current definition is "machine-based system" while the original annex, which listed specific techniques, was deleted.

Therefore, it is recommended that specific AI technologies not be referenced in the definition of artificial intelligence. In regulatory documents such as directives or circulars, it will be possible to elaborate on different technologies and techniques in the explanatory notes.

In examining possible terminology, the question as to whether to include the terms "computer" or "machine" in the definition terminology. These terms appear in most definitions, although the SEC definition cited above is broader and includes "method or process" as well. Alternatively, examples or text could be added for the purpose of illustrating the systems to which the definition applies. For example, in definitions worldwide terms such as "algorithm," "model," "machine learning" often appear, usually for the purpose of illustration rather than as an essential component of the definition.

It should be noted that since the Taskforce's work was not specifically directed at preparing legislation on the subject, the engagement with terminology did not reach the stage of precise formulation, but was conducted primarily for the purpose of pinpointing the key essential components of a definition to assist regulators and financial entities that will be required to address this question in the future, at least until a standard global definition is adopted, should this come to pass.

[105] This proposal has been withdrawn by the SEC.

(2) Is there room to refer to the intended use or the actual outputs of artificial intelligence?

As stated above, the purposes of AI system employment and the operational outputs resulting from its use can be diverse – ranging from monitoring tools intended solely for internal purposes, such as detecting anomalies in compliance with anti-money laundering provisions or activities, to the automation of operations that were previously conducted manually, such as back-office or administrative activities like clearing, and including decision-support tools in various financial activities to decision-making tools operating vis-à-vis customers independently.

One issue the Taskforce discussed is whether it is desirable to limit the definition solely to uses relating to interaction with customers. Since the risks inherent in AI systems may also be related to strictly internal systems, it is the Taskforce's opinion that such a limitation is unwarranted. Regulation should seek to address such "back office" risks as well, if these are material to the activity of the supervised entity.

A second issue the Taskforce considered is whether to refer to the system's outputs, such as prediction, recommendations, content-generation, or autonomous decision-making or its function as a supporting tool (similar, for example, to the European Union's definition) as part of the definition. Given that such reference is broadly accepted in regulatory definitions worldwide, reflecting the nature of the technology through its key capabilities and the products generated by AI systems, and since these are not exhaustive, the Taskforce believes that the inclusion of outputs in the definition is desirable.

A third issue is whether it is desirable to refer more specifically to the financial sector within the framework of the definition– for example, to decision-making by supervised financial entities regarding customers or investors. In the Taskforce's opinion, such a reference is beneficial for characterizing the activity addressed in its recommendations, as long as the definition does not limit regulation solely to activity with customers and the activity is indeed material from the perspective of oversight. It should be presumed that regulators may refer to even more specific uses in the future, such as reference to AI systems engaged in credit decisions.

(3) Whether and how is it possible to distinguish between AI systems and "ordinary" computerized systems?

This issue illustrates the intrinsic complexity of defining artificial intelligence and is vital in determining the scope of the definition's application and the regulation based on it. It appears that in leading definitions worldwide, as discussed above, this issue is addressed in two ways. Firsts, they specify that the system is "machine-based," emphasizing that the reliance on computerized rather than human activity, thereby narrowing the scope of the definition. Second, they acknowledge that such systems may operate at varying levels of autonomy, need not necessarily include machine-learning techniques, and

can be described through their capabilities or outputs. These elements broaden the scope of the definition and its application.

In the absence of other solutions, we recommend adhering to the components found in foreign definitions. It is reasonable to assume that, for many AI systems, the applicability of the definition will be clear and uncontested. For other systems, a dialogue between supervised entities and regulators can clarify how the definition should apply, resolving potential ambiguity. It is further proposed that, pursuant to regulatory guidelines, a financial institution may determine that a computer-based system is not, in essence, an AI system. This discretion would allow for the exclusion of systems which are clearly not AI, even though they may formally fall within the definition.

The drafting of several, modular definitions can serve as a response to the difficulties presented above. Such treatment can assist in adapting the regulation to artificial intelligence activity as it develops. The Taskforce therefore finds an advantage in proposing a basic definition, and alongside it, additional definitions, such as these:

- A definition that includes the element of materiality, and therefore applies only to AI systems with material implications for supervised entities' activity (with respect to either the entity itself or its customers). Accordingly, it will be possible, to determine that the regulation, in whole or in part, applies solely to material systems.
- A separate definition of an AI system characterized by machine learning. With respect to a system of this type, the issues of explainability and risk management are more acute. The drafting of a separate definition from which dedicated regulation will be derived can be advantageous in overseeing these issues.

A definition of a generative AI system or general-purpose AI. With respect to a system of this type as well, additional questions may arise, stemming from the vast data repositories it uses, the wide range of diverse content it can generate, or variety of potential uses it may serve. In the public comments submitted on the Interim Report, several responses addressed the difficulties presented above, particularly the concern that the definition of artificial intelligence may be overly broad and encompass systems whose capabilities and roles do not warrant special regulatory treatment. In this context, suggestions were made to narrow the definition or to tailor it to risks related to the customers of financial institutions.

The Taskforce believes that despite these difficulties, and in the absence of "easy" solutions to the issues involved in defining AI, adopting a definition similar to the definitions accepted worldwide is beneficial, as argued above. From the standpoint of risk (which can also be with respect to the financial institution as well as to its customers), the Taskforce believes that the proposed materiality component and the risk-based approach described in this report can adequately address this issue. At the same time, in response to public comments, an amendment was introduced, providing that in accordance with the regulator's

guidelines, a financial institution may determine that a computer-based system does is not, in essence, an AI system. This discretion allows for situations in which it is evident that a particular system in question is not genuinely an AI system, even though it may formally fit the definition.

Proposed Definition of Artificial Intelligence

Below is the Taskforce's proposal for a basic definition of an AI system in the financial sector:

> "Artificial intelligence system" – a machine-based system that operates with varying levels of autonomy, designed to generate outputs such as content, predictions, recommendations, or decisions, that may have an impact on investors, customers, or the activity of the supervised entity.

The regulatory focus can be placed on material AI systems by adding an explicit materiality component to the definition:

> "... that may have a material impact on investors, customers, or the activity of the supervised entity."

Additionally, in pursuant to regulatory guidelines, a financial entity will be able to determine that a computer-based system does not necessarily constitute, in essence, an AI system.

Part C

Cross-Cutting Issues

Artificial Intelligence Challenges and Regulatory Tools

As described in the introductory chapter, the general approach outlined in the government document on artificial intelligence is one of "responsible innovation." This approach also accompanied the writing of this report, and its meaning is that in order to encourage and derive benefits from AI use in the financial sector (and in general), the challenges these such uses raise must be addressed. In the following chapters, we will discuss the main challenges arising from the integration of AI systems in the financial sector, as well as the regulatory tools available to address these challenges. Even though this section distinguishes between various issues, they should be considered comprehensively, since the use of one regulatory tool may affect the necessity of others.

In this section, the cross-cutting issues stemming from the use of AI applications along with the Taskforce's recommendations regarding those issues will be elaborated. Among other things, we will address issues arising from some of the unique characteristics of artificial intelligence – such as the "black box" and explainability issues and the question of human intervention, as well as issues concerning the general law applicable to supervised financial institutions– with respect to discrimination, privacy, AI governance and risk management. We will present tools for addressing all these challenges in an integrated manner, and then turn to issues influencing the market as a whole– namely challenges that are amplified by the introduction of artificial intelligence in areas such as competition, financial stability, and risks related to fraud, cyber security and disinformation. For each issue, we will outline the questions arising from the integration of artificial intelligence, describe approaches taken abroad, discuss possible responses, and provide recommendations.

"Black Box" and Explainability

The term "explainability" refers to the ability to provide an explanation, rationale or justification for the output generated by an artificial intelligence AI system. The focus on the need to be able to "explain" the activity and outputs of an AI system stems from the fact that advanced AI systems are regarded as a "black box." Due to their complexity or the volume of data on which they are based, and the manner in which connections are drawn between the data, it is impossible to track the manner in which the system's outputs were obtained.

In regulatory discourse, the term "explainability" has no fixed definition, and it is sometimes used to express concepts that are similar but distinct from a technological and regulatory standpoint. Among other distinctions, one can differentiate between an explanation of a system operates in general and how a specific decision was made, that is, the available data, the algorithm and other factors that went into the decision-making, which provide an ex-post explanation of the decision. As will be discussed below, there is an extensive ongoing debate regarding the technical implications of a regulatory requirement for explainability and how such a requirement should be implemented at the technical level. In the absence of clarity with respect to a regulatory standard, system developers may interpret it differently, leading to an inconsistent application of the law. For our purposes, and for the sake of clarity, in this paper we will distinguish between general explainability, which pertains to the system's characteristics and mode of operation (referred to in some international documents as transparency), and specific explainability, which refers to an explanation with respect to how a specific decision was reached.

Since the publication of the Interim Report, several developments have occurred in the field of explainability. Large models offer a sort of explanation or a list of main sources referred to in their decision-making process, with the aim of providing users with the ability to at least partially understand the process undertaken in the system when generating the received output. Even if this falls short of full specific explainability, it can aid the comprehension of the model's decision-making process.

As will be outlined below, there are various grounds for imposing a regulatory requirement for explainability – both general and specific – as this serves, among other purposes, regulatory, consumer and broader social interests. At the same time, the explainability requirement could encounter difficulties stemming, for example, from the current maturity level of the technology or from the need to protect trade secrets. Accordingly, any recommendation relating to the explainability must take these considerations into account, including their implications for the supervised entity, for consumers (whether investors or customers of financial institutions), for the regulator, and more broadly, for innovation.[106]

---

[106] Shan Bilal, David Ebert, Beiyu Lin, LLMs for Explainable AI: A Comprehensive Survey (31 March 2025), https://arxiv.org/html/2504.00125v1.

The Explainability Requirement

As stated above, advanced AI systems may be a "black box", lacking transparency. This means that it is difficult for a human to understand how the system operates and why it reached a particular decision. This challenge has led to extensive technological efforts to develop explainable AI systems, as well as regulatory engagement with respect to the characteristics such systems must have prior to deployment in regulated and other domains. These efforts include the integration of requirements of this nature into regulatory documents and proposed regulatory frameworks for artificial intelligence generally and for the financial sector in particular, as detailed below.

Explainability in artificial intelligence:

General Data Protection Regulation (hereinafter: "GDPR") – the European Union regulation dealing, among other things, with automatic processing of personal information. This regulation establishes various protections in cases where decisions regarding a person are made based solely on automated information processing. and impose obligations on system operators accordingly.[107] Among other things, the regulation establishes that with respect to certain decisions, individuals have the right to be provided "meaningful information about the logic involved in the decision-making system."[108] In addition, in certain cases, system operators are required to ensure that the individual has the ability to object to or contest the decision made with respect to him. These sections have sparked academic discussion on the question as to whether the GDPR establishes a right to explanation and what the scope of this right is.[109] Some argue that the meaning of the sections is that an explanation must be provided to the individual at a level that will enable him to address the decision without disclosing trade secrets regarding the system's operations. Others argue that information must be provided to the individual about the decision regarding him in accordance with technological limitations. In any case, it appears that the GDPR recognizes the importance of the need to provide an explanation to individuals regarding an automated decision taken with respect to him.

EU AI Act – The European AI legislation deals with the obligation imposed on high-risk AI functions (including decisions on financial matters such as credit scoring and life and health insurance).[110] According to this obligation, the provider of the system (the party that makes the system available for use) must operate systems that are designed and programmed in such a way as to enable him to

---

[107] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural 107 persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), http://data.europa.eu/eli/reg/2016/679/oj.

[108] GDPR, Articles 13(2)(f), Article 14(2)(g), and Article 15(1)(h).

[109] See, Inter alia - Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. "Why a right to explanation of automated decision making does not exist in the general data protection regulation." International data privacy law 7, no. 2 (2017): 76-99.; Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi.

[110] See *supra* note 11. The EU AI Act uses the term "transparency", and we have therefore adopted it in the description above.

understand the way the system operates, including an understanding of its capabilities and limitations (for example, if there are situations in which human oversight of is required). The law mandates that the system must include a "user manual" detailing the system's operation, and circumstances that may increase its potential risks including those precipitated by the actions of the system provider. Such transparency ensures that the system provider will be able to make informed and safe use of the system, which complies with the regulatory requirements imposed on it. Finally, the system provider must be able to interpret the system's results.[111] The law primarily establishes a general explainability requirement for the system itself (this is in addition to the requirements articulated in the GDPR mentioned above,). Additionally, when the AI system is involved in making specific decisions affecting an individual, that individual has the right to receive an explanation of the AI system's role in the decision-making process.[112] In addition to explainability requirements, a transparency requirement is imposed on AI systems designed to interact with end users. We address this obligation in the chapter dealing with notification and disclosure below.[113]

Additional regulatory documents also focus on the need for explainability. The OECD lists transparency as one of the principles guiding AI-system development.[114] The document notes that transparency is intended to serve several purposes, including providing the party affected by an AI operation the ability to understand its outputs. It also emphasizes that this principle must be implemented within the context of the AI system's activity in keeping technological development.

Explainability in the financial sector

Regulatory documents concerning the financial sector reference the explainability issue where a decision regarding a consumer was made. The European directive governing consumer credit agreements contains specific provisions regarding credit decision-making and pricing conducted by automated systems. The directive includes a disclosure obligation with respect to the involvement of artificial intelligence in the decision-making process and grants consumers the right to request the financial institution providing the credit to provide an explanation regarding the decisions affecting them. Such explanation must include information about the logic guiding the system, particularly how the system's logic was applied specifically to the consumer. In other words, the directive contains both a general explainability obligation and a specific explainability obligation.[115]

---

[111] It should be noted that a provision which explicitly addressed the need to provide an explanation to the individual was removed during the legislative process in Parliament.

[112] EU AI Act, Article 86.

[113] EU AI Act, Article 50.

[114] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[115] See, for example, par. 56 of the Directive - https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32023L2225.

In the United States, legislation imposes a general "reasoning" obligation for certain credit decisions – particularly those adversely affecting a customer– both as a means of preventing discrimination and as a mechanism for shaping conduct.[116] In 2023, the following the increasing reliance on sophisticated algorithms in credit-decisions, the CFPB issued a circular clarifying that that the reasoning obligation applies when the credit decision is generated by an algorithm. It further emphasized that the explanation of the reasoning must be tailored to the customer's specific circumstances; generic information does not suffice:[117]

"A creditor will not be in compliance with the law by disclosing reasons that are overly broad, vague, or otherwise fail to inform the applicant of the specific and principal reason(s) for an adverse action."

Explainability Obligation – Discussion and Recommendations

As the above discussion reflects, there is relatively broad consensus regarding the need for explainability. However, the regulatory requirements concerning explainability differ in scope as well as definition, and, consequently, in the parties to which this requirement applies.

With respect to the scope of the obligation, it is possible to distinguish between general obligations (similar to a disclosure obligation regarding the system's mode of operation) and specific obligations (similar to a reasoning obligation). The general explainability obligation is concerned primarily with providing detailed information about how an AI system operates, the underlying logic on which it is based, the data used, and even the considerations taken in the decision-making process. This general explanation does not provide detailed information regarding the logic and justification of the output that the system produced in a specific matter, but rather relates primarily to the mode of the system's operation and the decision-making process *writ large*. In contrast, specific explainability focuses on an explanation for a specific decision given ex-post, in which the information is not only procedural, but provides the logic for the particular decision well, explaining why this particular decision was made.

Regarding the parties to whom the obligation applies, this can be to the supervised financial entity using the system, to the regulator, and to the public at large. The justification for applying explanation and transparency obligations varies by party. An outline of the rationale with respect to each party follows below:

[116] U.S.C. 1691(d)(2); 12 CFR 1002.9(a)(2)(I).
[117] CFPB Circular (2023), see *supra* note 71.

General explainability regarding how the system operates vis-à-vis the public will enable the system and its mode of operation to be subject to public scrutiny. That is, civil society organizations, academia, and similar actors will be able to examine the characteristics of the system and express their views regarding its	**THE PUBLIC**
Refers to the body of knowledge required for regulation and for carrying out the regulatory function. This knowledge includes data, risk assessment, understanding, and learning. A system that makes use of this body of knowledge can operate in a changing	**THE REGULATOR**
Refers to the body of knowledge required for effective regulatory activity and for achieving regulatory objectives. This includes both existing and emerging information, originating from within and outside the system.	**THE FINANCIAL ENTITY**
Refers to the ability to absorb change, in a broad sense. This includes the ability to adopt changes in a timely manner and in a way that enables adaptation to changes in the environment, by adjusting the available tools and through forward planning.	**THE CONSUMER**

Alongside these justifications, additional considerations must also be taken into account:

Technological limitations: When discussing the question as to whether to apply an explainability obligation, particularly regarding specific explainability, the Taskforce's position is that it is imperative to take technological limitations into account. As of the writing of this report, certain algorithmic systems are indeed opaque black boxes and it is difficult to explain how or why a particular output is generated. These systems, however, may be more accurate precisely because of their size or complexity. In these cases, a specific explainability requirement means, *de facto*, a prohibition on employing the system for an intended use. This could likely lead to the use of less accurate systems and may well slow innovation, since the adoption of such systems will be impossible. As will be detailed below, one of the regulatory challenges facing regulators is to properly define the explainability requirements for each of the parties mentioned above. For example, is general explainability regarding the manner of the system's operation sufficient? Is it possible to make do with partial information regarding the decision-making process with respect to individuals? Can a system provider merely identify the key elements in making the decision, or should they be required to provide a precise explanation regarding the decision, which includes not only the identity, but the weight given to each element.

Materiality of the technological component: The role the technology plays in the overall arrangement must also be taken into account. For example, if the technology is a decision-support tool, while a human who is able to clearly reason his decision is ultimately responsible for the decision, it is

reasonable that a full explainability obligation is not required. In contrast, if the human only "mediates" the decision, it the right of the consumer and the regulator to receive a sufficiently comprehensive explanation for the decision must be ensured.

Legal and contractual constraints: One must also take into account that the explainability requirement, may, in some cases, pose challenges because it could involve the disclosure of trade secrets.

Nature of the decision: As stated above, in the European Union, the explainability requirements applies solely to decisions of a system defined in the Act as high-risk., In the United States, the reasoning requirement applies only within certain financial contexts with respect to decisions that adversely affect the customer. For example, if a customer's application for a loan was accepted rather than denied, a specific explainability requirement may be unnecessary.

Public comments on this chapter referred to both the description of the explainability issue and to the recommendations themselves.

Comments regarding the general issue claimed, among other things, that the clarity of the distinction drawn between "general explainability" and "specific explainability"; is not sharp enough, and that the system's status with respect to the regulatory requirement is not clear enough. In response to these comments, changes were made in the section dealing with general explainability to clarify these points.

With respect to the Taskforce's recommendations, some commented that explainability is not required in every situation and this requirement should be focused only on certain types of systems or decisions (high-risk systems, systems for the purpose of fraud prevention, etc.). Another observation was that the recommendation that explainability should not be deemed necessary when the technology serves solely as a supporting tool should be elaborated and sharpened. It was noted, among other things, that it is essential to ensure that the human agent overseeing the system's operation has a solid understanding of its features and weaknesses. It was also noted that explainability plays a key role in validating the system's outputs, and weight should be given to this in the recommendations as well. Finally, some noted that technological advances currently provide insights into certain aspects of the system's decision-making process, and that this can potentially soften the explainability challenges described in the chapter should be considered.

The Taskforce's impression is that these comments relate primarily to various clarifications that were requested rather than to the essence of the recommendations themselves. Following the comments, several clarifications were indeed made. For example, it was clarified that specific explainability is not required with respect to every use or AI decision, and this requirement should be reserved for solely for material decisions. It should also be noted, following the comments, that often explainability serves the needs of the supervised entity and against this backdrop, technological developments seeking to improve the level of explainability in AI systems can be noted. In this context, the recommendations include providing updates on the issue of providing partial specific explainability.

"Black Box" and Explainability – Recommendations

Pursuant to the considerations presented above the recommendations regarding the imposition of an explainability or transparency obligation include:

1. Specific explainability requirements should be imposed in accordance with existing law:

An explainability requirement should not be established in every case due solely to the fact that the supervised entity employs an AI system in its operations. That is, an assessment should be made as to whether an explanation obligation already exists within the current legal framework and to whom it applies. Where such an obligation does exist, the introduction of an AI system should not necessarily change neither the obligation itself nor its scope. Additionally, there may be room to reconsider the necessity of existing reasoning requirements and possible alternatives to them, since these obligations may constitute a barrier to AI adoption. If no such obligation currently exists under the law, an assessment should be made as to whether the type of system or nature of the decision warrants the imposition of a new obligation (taking into account the considerations listed below).

2. General and specific explainability requirements with respect to material high-risk decisions:

The application of explainability requirements to regulators and consumers with respect to high-risk systems should be considered. The nature of the decision should be vetted, focusing on material decisions and decisions that adversely affect the customer, since, generally speaking, these are the cases in which a customer will most likely seek to contest the decision or gain insight into it. For this purpose, the decision concerning him begs explanation. It should be emphasized that, if the AI system does not play a material role in making the recommendation or decision, and is limited to tasks such as organizing information, even explainability with respect its mode of operation will not be required (similar to EU legislation).

3. "Compensatory" alternatives to specific explainability - it should be examined whether there are alternatives to explainability that will provide an appropriate response to the above challenges:

The imposition of a general explainability requirement together with other compensatory measures, such as human oversight, close monitoring of the system's results, or procedural controls could be considered. These measures may adequately address the challenge posed by the lack of an orderly, structured explanation, and may even suffice with respect to high-risk systems.

Providing service alternatives – where the regulator believes that an explainability requirement can be waived in light of a human service alternative (so that the consumer can choose between an AI-based

service or a human service), it must be ensured that the human alternative is not inferior to the algorithmic one and that the two will be provided under similar conditions, so as not to create an indirect preference for the algorithmic service lacking the ability to provide specific explanations. Making do with partial specific explainability – as stated above, technological developments in the field of artificial intelligence seek to achieve specific explainability, even if it is not complete. For example, an AI system may provide an explanation regarding the weights it assigned to different variables throughout the decision-making process – but will not be able to explain the reasoning behind this. In many cases, even such a partial explanation may have value, such as in checking data errors or for understanding the key data affecting the output. It follows that supervised entities and regulators may adjust the required level of explainability according to the circumstances, and in some cases, may make do with partial or limited specific explainability

4. Specific explainability requirements where the technology plays only a supporting role in decision-making is not necessary.

A distinction should be made between situations in which the technology plays a supporting role in the decision-making process and situations in which the technology is the decision-maker (even if mediated by a human involvement). If the AI system is merely a supporting factor and the decision is ultimately made by a human – who is able to reason it – the explainability requirement should be relaxed. On the other hand, if the human merely "mediates" the decision, the right of consumers and regulators to receive a sufficiently comprehensive explanation for the decision must be ensured.

5. Internal explainability obligation (within the financial entity):

It is important that financial entities operating AI systems understand the system's operations, limitations and other characteristics, in line with EU standards. While this mainly entails a general explainability obligation, a specific explainability obligation should be considered pursuant to the considerations listed above - including the decision risk level, compensatory alternatives to explainability, and the centrality of the technology in decision-making.

6. General explainability obligation to the public:

The general explainability obligation may serve an important social need, but nonetheless must be balanced with other considerations, such as the potential disclosure of trade secrets, competition and the like. It is important to encourage companies and financial entities to disclose certain aspects of the systems they make available to the public, but it appears that requiring them to do so is difficult.

Human in the Loop

Human intervention is one of the regulatory remedies commonly proposed to address the risks and challenges inherent in AI activity. Human supervision over algorithmic decision-making (which will be referred to for the purposes of this report as human intervention) provides greater transparency and control of AI-driven processes.

The human intervention requirement is perceived as an effective regulatory tool, designed to address the potential malfunctions and failures of a wholly algorithmic system. It seeks a balance between the desire to integrate AI systems in various domains and the concerns that these systems are not yet mature enough to handle failures that could harm or endanger individuals or society as a whole.

The premise for the human intervention requirement is that humans can effectively prevent or control algorithmic system failures, at least at the current stage of AI development in which doubt lingers regarding its ability to make accurate decisions which also comply with legal rules and standards. The malfunctions and failures relevant in this context may involve biased system outputs, the dissemination of inaccurate information, erroneous results stemming from operational malfunctions and the like. The more significant the implications of the decision for the individual, the greater the importance of requiring human intervention and oversight. Another rationale raised in connection with the human-oversight requirement is the concern over the loss of individual autonomy. In certain circumstances, this concern may necessitate human involvement in the process – among other things, due to the difficulty of explaining the decisions made by these systems in certain cases, and the concern that this may impair an individual's ability to understand the decision made about him. A decision made by a non-human party may also create a feeling of diminished control over his own fate. Human intervention – both in making the decision and through human oversight – can provide at least a partial response to these concerns. In this context, it has also been argued that there is an intrinsic value in applying human judgement to an individual's case, as human discretion can accommodate complexities that an algorithmic model is not able to incorporate. Finally, there is also a legal rationale for the human intervention requirement. Algorithmic systems are involved in making decisions also in areas where professional licensing is required, such as investment advice, pension advice, legal practice and the like. The absence of human intervention in making decisions in these domains could be perceived as a violation of the law and the licensing requirement therein. This rationale is both formal and substantive. Its proponents argue that it is necessary since licensing requirements aim at ensuring that the fiducial duties of the professional towards the individual are maintained or ensuring a minimal level of competency. The involvement of licensees, therefore, must be preserved.

Despite these justifications, this requirement –particularly the sweeping application of it– has also drawn significant criticism. A key criticism of this requirement relates to the fact that it is uncertain whether human intervention will actually improve the system's output. One of the reasons for this is the tendency

of human beings not to deviate from the system's recommendations, so that if there is indeed a failure in the system, there is no guarantee that human intervention will address the failure. In addition, if the regulatory requirements focus on completely automated decisions, they can be circumvented by introducing limited human intervention that perhaps passes the regulatory threshold but fails to truly fulfill the purposes described above.[118] This critique argues that there is a presumption that human discretion is preferable to algorithmic discretion, even though human beings also discriminate, exercise bias, make mistakes and otherwise act in bad faith. In this context, the critics argue that it is appropriate for the regulator to focus on establishing standards for preventing bias in algorithmic decision-making rather than transferring discretion to a human party, who, as stated, may rely on the system's recommendation or himself engage in discrimination.[119]

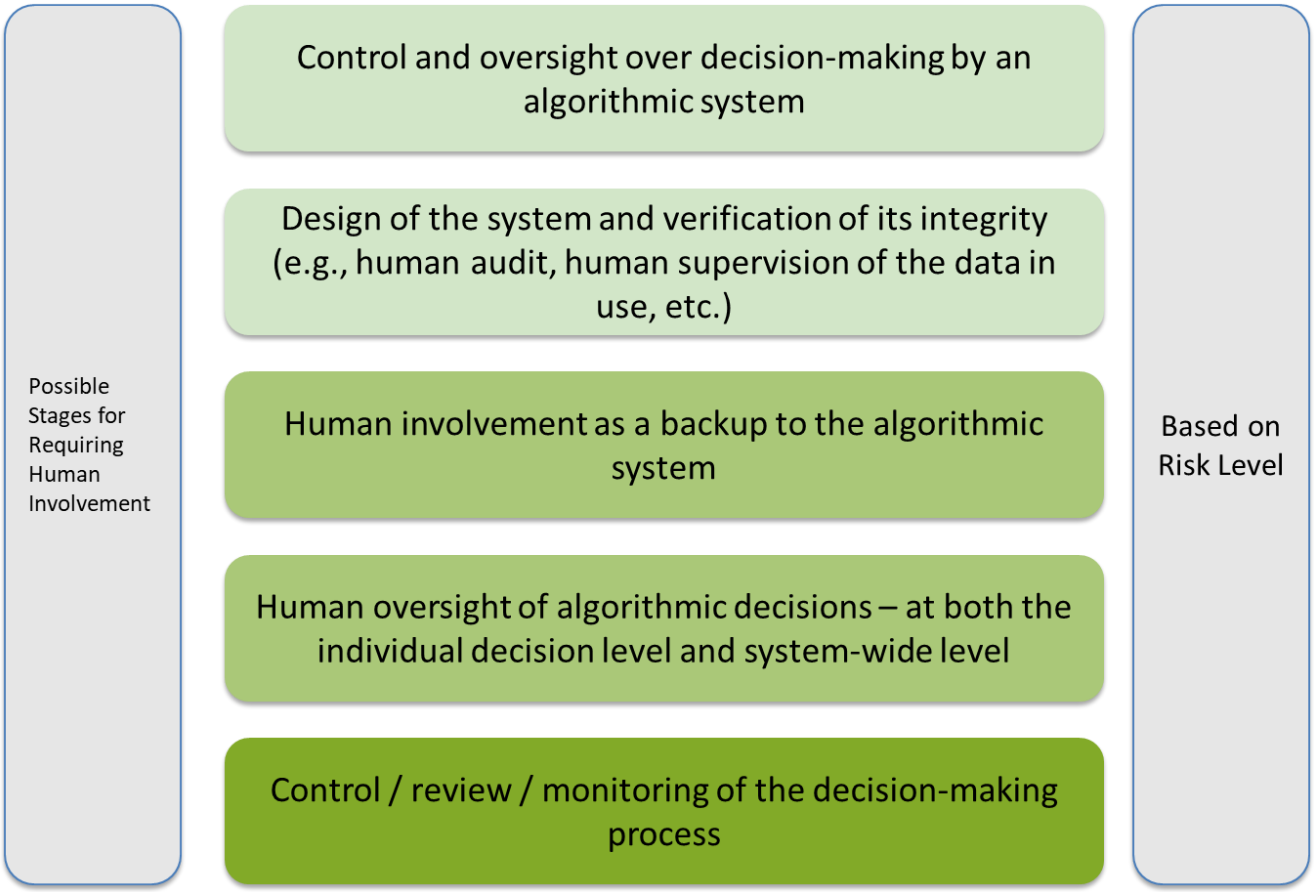
Another criticism pertains to the uncertainty surrounding the division of responsibility between the human and the AI decision-making system. A human intervention requirement without defined rules could blur the lines of responsibility between the human party and the algorithm. For example, will a person tend to intervene in the system's decisions even when it is not clear whether a failure actually occurred, only in order to avoid bearing responsibility for it? Alternatively, will a person avoid intervention due to a lack of adequate understanding of the system or the fear of bearing liability?

Finally, the imposition of mandatory human intervention raises concerns that it will eliminate or reduce the advantages engendered in the integration of AI systems in the organization. Organizations invest in these systems in order to make their work more efficient and to overcome "human" disadvantages such as working hours limitations, faulty discretion, and primarily to exploit the enhanced computing capabilities of systems that are supposed to lead to more accurate decision-making. Leaving the human in the picture could reduce or impair these advantages.

The human intervention and oversight requirement can apply with respect to different stages in an algorithm's life cycle:

[118] Green, Ben, The Flaws of Policies Requiring Human Oversight of Government Algorithms (April 26, 2022). Computer Law & Security Review, Volume 45, 2022, https://ssrn.com/abstract=3921216.

[119] https://www.ecgi.global/publications/blog/regulating-for-humans-in-the-loop.



The first three stages presented in the diagram deal with the process and conditions for integrating an AI system: the decision to employ such a system in financial institution operations, tailoring the system to the institution's needs, and implementing human backup in case of a technical failure. These stages will not be discussed in this section but rather in the chapters dealing with comprehensive and risk-based perspectives and AI governance, since the focus of these stages is placed on the procedural aspects of AI deployment.

In this section, we focus on human intervention in the system's decision-making process, where the human intervention requirement usually aims at oversight over the system's operations - from the standpoint of the decision's content, the information it provides and compliance with the regulatory and legal requirements imposed on the financial entity.

Human intervention requirements in regulatory documents abroad

The requirement for human intervention recurs, in varying scope, in prominent regulatory documents in the field.

- The EU AI Act establishes several obligations of human intervention and oversight, primarily for high-risk systems, which include systems designed for financial decision-making.[120] The AI

[120] EU AI Act, Article 14.

Act provides that high-risk AI systems must be developed and designed in a manner that allows for real-time human system-oversight, to minimize, to the extent possible, the potential risks generated by the system. These include risks arising from sabotage or malfunction, in addition to those arising from "planned" system operations. The Act further provides that the level of human oversight must be adapted to the AI system's risk level the degree of its autonomy and the context in which it is used. It sets a series of conditions including, among other things: (1) the human party must be able to identify anomalies and malfunctions in the system's functioning; (2) the human party must be aware of its tendency to rely on the AI system; (3) it must be capable of interpreting the system's decisions; (4) the system must be designed to grant the human party the technical ability to override or ignore the system's decision; (5) the human part must be able to intervene in the system's activity and to stop its work. The Act also addresses to the requirements and skills required of the human oversight party.[121]

- The GDPR grants the individual the right to demand that any decision that has significant legal and other implications for him (subject to exceptions) not be made based solely on an AI system.[122]

  Following these provisions, in Europe the directive on consumer credit agreements also reiterates the obligation to enable human intervention in credit decisions upon consumer request.[123]

- In Canada, the proposed legislation for regulating artificial intelligence establishes an obligation to conduct risk assessments for all AI systems. With respect to systems with significant implications for the individual (high-impact), human oversight and system monitoring is mandatory. Similar to the European Union, Canada proposes to require that the development and design of AI systems enable human parties to oversee and interpret its activity. In addition, such legislation would impose an obligation to enable the tracking of the system's activity and its output.[124]

- In the United States, the Presidential Order under the Biden administration referred to cases in which human intervention is required. Among other things, the Order established that in certain cases, such as integrating AI systems in government decision-making that can affect certain rights, the system must allow for the examination of the decision by a human party upon the individual's request. It further instructed authorities to establish the cases in which individuals should be entitled to receive service and support from a human party.[125]

---

[121] Ibid.
[122] GDPR, Article 22.
[123] See *supra* note 98.
[124] https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading.
[125] See *supra* note 25.

Reference to human intervention is also expressed in financial sector policy papers. In a policy paper issued by the Bank of England, the importance of human intervention is discussed both during the development of an AI system and in the system's operational stages. With respect to the operational stage, the report refers to the fact that human intervention is an effective means to address problematic system results, since the human party has relevant knowledge that will not necessarily be reflected in the decision-making model (such as the context for making the decision, understanding built-in biases in the information, etc.). Additionally, a human can identify a pattern or specific problematic decisions and refer them for re-examination.[126]

In a CFPB consultation paper on asset valuation for the purpose of granting mortgages and collateral valuation, the limitations of automated models versus human evaluators for these tasks were discussed, leading to the conclusion that human oversight over the model's results is recommended.[127] In addition, in another policy paper on the use of chatbots in the financial sector, the disadvantages of customer service rendered by an automated chat bot versus by a human being were discussed. Here too, the CFPB proposes enabling a transition to a human service provider in certain cases.[128]

The 2023 OECD Report recommends continual human oversight as one of the policy tools to increase control over and accountability of AI systems. The report argues that human oversight should not stop at the system development and design stage but continue throughout the operational stage, among other things, by monitoring results and the model's "behavior" in real time.[129]

Finally, it should be noted that the 2023 ESMA Report found that most companies engaged in the field expressed the position that at this stage, the most effective use of algorithmic systems will be achieved through a combination of algorithmic and human decision-making.[130]

Several comments to the Interim Report argued that human intervention is not preferable to artificial intelligence, and therefore a requirement for human intervention will not necessarily benefit customers and can even be expected to adversely affect AI-system efficiency. Other commenters did not see any problem with the existence of human intervention requirements, but believed that the recommendations in this context should be more concrete, and should present clear criteria for cases in which human intervention will be required and cases exempt from such a requirement. They argued that the recommendation suggesting that human intervention is more necessary with respect to "negative" decisions for the customer is misleading, since seemingly "positive" decisions (for example, a decision to approve a loan request) could also prove harmful to customers. In any case, a licensee should be

---

[126] See *supra* note 74.

[127] https://www.consumerfinance.gov/rules-policy/rules-under-development/quality-control-standards-for-automated-valuation-models/.

[128] See *supra* note 45.

[129] The 2023 OECD Report.

[130] The 2023 ESMA Report.

required to examine positive decisions as well. Another commenter pointed out that benefits and harm notwithstanding, it will not necessarily be possible to require human intervention by financial entities, as this will likely be contingent on the commercial agreement between it and the system developer.

The standards on the question of whether or when human intervention is required are indeed contingent on circumstance. The line presented in this report is that the establishment of regulatory requirements, including a human-intervention requirement, is necessary when dealing with high-risk activity and no other measures are taken to mitigate the risk. Consideration should be given to the potential impact of the system's activity on both the customer and the financial entity, regardless of whether the system's outputs are decisions or relate more to assistance functions, such as conveying information and conducting communication with the individual. Clarifications on this matter were made in the chapter's recommendations. Among other things, the comment that seemingly "positive" decisions could also harm customers was accepted, and therefore, the emphasis has been placed on the materiality of the decision for the customer and the risk of harm to him, regardless if the decision approved or denied the customer's request. In response to other public comments, the report clarified the relationship between the human-intervention requirement and the human-control requirement: the necessity of human control over the system's activity and its results is undeniable, but human control is not a panacea for all cases in which a human-intervention requirement is proposed. The comment suggesting that human intervention does not depend on the financial entity but on the system developer, was rejected, since the liability for introducing the system and its outputs ultimately rests on the financial entity (see also the discussion in the chapter dealing with liability).

Another issue raised in the public comments is the call to require financial entities to provide a human-intervention alternative alongside the AI-based service. Under this approach, a financial entity engaged in providing credit to customers, for example, will be required to provide a credit arrangement based on human decision-making alongside an AI -based credit decision. The Taskforce believes that such a cross-cutting requirement is unwarranted, since it is contingent on questions such as the extent to which the service is essential, potential difficulties in accessing the service through AI systems (similar to the challenges certain populations face when receiving digital rather than in-person services), and the availability of alternative providers with the financial sector. To illustrate, a financial entity may choose to engage exclusively in the provision of AI-based portfolio management services, competitively positioning itself as an entity operating in this manner. Since portfolio management is a voluntary customer activity, and alternative human-based systems are abundant, the establishment of an obligation as proposed appears unwarranted as long the alternative human-based service is available in the market. This issue may reemerge down the road when AI-based services become significantly more prevalent. In the meantime, there is no need to preemptively establish a sweeping obligation to provide a human-intervention alternative. There is, however, room to ensure customers the right to an appeal to a human

party, or require a human-response alternative for customer inquiries. In the future, human alternatives to AI systems may also be required for certain matters.

Human in the Loop – Recommendations:

A. Human intervention in material decisions: While human intervention may be an effective regulatory tool serving multiple purposes, the sweeping application of a human-involvement requirement may generate unnecessary costs, raise other barriers to integrating AI technology in financial operations, and may even impinge on the advantages AI has for consumers. Therefore, and as argued with respect to other tools and similar to arguments appearing in regulatory documents globally, it is the Taskforce's position that to reach an appropriate balance, this requirement should be adopted with respect to decisions that have significant implications for the individual tailed below. In this context, regulators can establish criteria or structure discretionary authority for defining a decision as subject to the requirement, based on the materiality of the decision for the individual and the intrinsic risk of error in that decision (see below on chapter "Connecting the Dots" - A Comprehensive Risk-Based View as well as the chapter entitled Notification and Disclosure).

B. Timing of human intervention: Human intervention regarding a specific decision can be performed at two main points of time: at the time the decision is made or subsequent to it in response to an appeal by the customer. The main disadvantage of early human intervention at the time the decision is made relates to the efficiency considerations discussed above. The main disadvantage of belated human involvement, following a customer appeal, is that it requires the individual to act to activate the human intervention, and it assumes that the individual will have the awareness and ability to appeal a decision made with respect to him. It is also recommended that the distinction articulated in the first recommendation be adopted; that is, that real-time intervention will be required only with respect to material, potentially high-risk decisions for which no compensatory alternatives exist.

C. Compensatory alternatives: The human intervention requirement should be examined within the system of the additional regulatory requirements imposed on the system. Thus, where a system has sufficient explainability capabilities, there is no apparent need to mandate human intervention at the point of decision-making, and it may be sufficient to rely on appeals process or human oversight of the decisions. This implies that even high-risk decisions may not require human intervention, provided that suitable compensatory mechanisms are in place.

D. Distinction between decision-supporting and decision-making systems: If the system is designed such that a human party makes the ultimate decision, then the human-intervention requirement should be reduced solely to general control and oversight of the system's operations. It follows that in decision-support systems, in which the activity is not fully automated, it is possible to reduce human-intervention requirements on specific decisions and focus instead on control over the system's operation and its results.

E. Alternative of human activity and opting out: The Taskforce does not believe that imposing a sweeping obligation to provide human-based activity alongside any activity performed by an AI system is warranted. The need to impose such an obligation depends on multiple factors, including the extent to which the service is essential, difficulties that the entire public may face in receiving the service from AI systems, and the possibility of receiving the same service from other financial entities. At the same time, insofar as there is a possibility to receive a human-based decision or service, this should be brought to the customer's attention so that he may choose whether he is interested in a human or algorithmic alternative, with all their respective features (see Notification and Disclosure chapter). It should be emphasized that opting for the algorithmic alternative does not preclude the consumer's right to appeal to a human party as detailed above.

F. Between human intervention and control over the system's activity: The Taskforce believes that vigorous oversight of the system's operations and output is of vital importance. Such oversight, as opposed to a human-intervention requirement for every specific decision, appears more appropriate at the current stage of technological development, particularly when dealing with a new system for which operational experience is limited. Monitoring the system's activity and results is essential for preventing malfunctions or errors as well as for handling them should they nevertheless occur. It is therefore recommended that a human-oversight requirement be adopted over the system's activity, rather than a requirement for oversight of every individual decision. Such ongoing oversight need not be human at every stage of the process, but must nonetheless include an element of human oversight in it.

G. System design: In view of the recommendations concerning both real time and ex-post oversight, it appears that financial entities will need to implement systems that allow for such oversight in line with the above recommendations. The responsibility for facilitating human control and intervention rests with the financial entity deploying the AI system in its operations.

Notification and Disclosure

One of the regulatory requirements discussed in relation to the deployment of AI systems across services and products is the obligation to disclose that an AI system is being used. This entails informing the public, regulators and other stakeholders that a given service or product relies on a system that can be characterized as an AI system. Other than notification about the use of an AI system, there may be disclosure requirements to an investor or customer concerning the characteristics of that system, its implications, and the manner in which it is integrated into the company's service or product. These disclosure requirements may span a broad spectrum of options, both in terms of the substance of the disclosure (for example, whether it is general or specific content, whether it focuses primarily on risks and warnings or highlights potential benefits as well, and other similar considerations) and in terms of its format (for example, disclosure within the framework of a customer agreement, periodic disclosures, standard vs. prominent disclosure, etc.). These disclosure requirements are closely related to the explainability issue discussed above. Another aspect of disclosure pertains to mitigating the risk of issuing misleading statements about artificial intelligence. New and complex technologies may provide fertile ground for publications that overstate the capabilities of the entities using them. Disclosure obligations can mitigate this risk by requiring entities to publish the truth. We note at this stage that, in the context of Generative AI, disclosure obligations assume heightened importance in light of the unique challenges posed by this technology for issues related the reliability of information, its sources and the like. Regulatory documents, such as the European legislation, already establish distinct disclosure requirements specifically for this technology.

Purpose of Disclosure Obligations

Notification or disclosure obligations designed to enable the public to make informed decisions are familiar features in numerous regulatory regimes. Such obligations narrow information gaps between consumers and service providers by ensuring that all the information relevant to making a decision is made available, allowing consumers to weigh all pertinent considerations – whether the decision concerns investing in a public company subject to disclosure requirements or selecting food products based on nutritional information that must be included on the packaging.[131] This premise is also tied to the outlook that disclosure obligations play a key role in safeguarding an individual's autonomy to control decision-making regarding him. In addition, disclosure and notification obligations are viewed as instruments that facilitate public oversight, promote competition, and advance related objectives. In

[131] Weil, David, Archon Fung, Mary Graham, and Elena Fagotto. “The effectiveness of regulatory disclosure policies”. Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management 25, no. 1 .155-181 :(2006).

the financial sector, disclosure requirements are one of the most common regulatory tools, deployed with respect to virtually every service or product offered to consumers.[132] For example, in Israel it is possible to find comprehensive consumer disclosure obligations within the framework of banking,[133] investment or pension advice,[134] as well as within the framework of indirect or direct investment in securities.[135]

The disclosure obligations proposed in these provisions and implemented in many additional provisions are intended, among other things, to enable the public to understand its financial position, compare different investment and savings alternatives, learn about the entity offering a service or product and its characteristics, identify the commissions or management fees being charged, and assess the risks and limitations associated with a particular service. As noted above, such disclosure requirements serve fundamental objectives: facilitating informed decision-making, reducing information and power gaps between the public and financial institutions, deterring those with positions of authority in financial entities from abusing their power, fostering competition between entities and financial products, and strengthening public confidence in the financial system. In some cases, disclosure requirements also serve broader goals, such as promoting market efficiency and facilitating continuous pricing of traded assets.[136] These assumptions and justifications also serve in the discourse regarding the application of disclosure obligations to AI systems. In this regard, emphasis is placed on the notion that disclosure itself – and the ability to subject services and products incorporating this technology to public scrutiny – enhances their safety, reliability and the level of public confidence in them. As discussed below, the realization of these objectives is linked to the scope of the disclosure obligation with respect to the system's operation.

First, the deployment of AI technology in financial products and institutions is still in its early stages. The efficiency of the technology and the potential risks it engenders have not yet been fully clarified. In this context, there is value in informing the public about changes in the way services are provided, allowing them to choose which service to use (where alternatives exist). It may be possible to reconsider the scope of this requirement once the technology becomes more firmly established in the financial services sector. Second, the disclosure requirement may incentivize improvements in the technology

---

[132] nriques, Luca and Gilotta, Sergio, Disclosure and Financial Market Regulation (April 11, 2014). Chapter in The Oxford 132 Handbook on Financial Regulation, edited by Eilís Ferran, Niamh Moloney, and Jennifer Payne, (Oxford University Press). European Corporate Governance Institute (ECGI) - Law Working Paper No. 252/2014, Oxford Legal Studies Research Paper No. 68/2014, https://ssrn.com/abstract=2423768.

[133] See § 5 of the Banking (Service to Customer) Law, 1981(Israel); and the Banking (Service to Customer) Rules (Proper Disclosure and Document Delivery), 1992 (Israel).

[134] See § 21 of the Supervision of Financial Services Law (Pension Advice, Marketing, and Acquiring Services) 2005 (Israel); §§ 14–16 of The Investment Advice, Investment Marketing, and Investment Portfolio Management Law, 1995 (Israel); Investment Advice, Investment Marketing and Investment Portfolio Management (Reports) Regulations, 2012 (Israel).

[135] See, for example, §§ 16, 17, and 36 of the Securities Law and the regulations issued thereunder; §§ 26, 72, and 73 of the Joint Investments in Trust Law, 1994 (Israel), and the regulations issued thereunder.

[136] See the case law of the Supreme Court regarding the purposes of disclosure in various financial contexts: CA 5320/90 **A.Z. Baranovitz Properties and Leasing Ltd. v. Israel Securities Authority**, IsrSC 46(2) 818 (1992); CA 5893/91 **Tefahot – Israel Mortgage Bank Ltd. v. Tzabah**, IsrSC 48(2) 573 (1994).

while causing minimal harm to efficiency. When customers are informed of the option to transfer to human customer service instead of continuing interaction with an AI agent, their choice will be influenced by the quality of service offered by the financial institution. In this way, the disclosure obligation, combined with the institution's interest in preserving efficiency will encourage the financial institutions to improve their products, thereby reducing instances in which customers are discouraged from using artificial intelligence. Finally, the disclosure requirement is also advantageous to the regulator. As stated above, the integration of the technology in financial products and institutions is still in its nascent stages. In some cases, there is no regulatory restriction placed on the integration of the technology itself. As a result, the regulator may become aware of its adoption only after the fact. A basic disclosure requirement allows the regulator to maintain an up-to-date understanding of how the technology is integrated within financial institutions and to assess its impact (for example, by identifying an increase in customer service complaints).

As discussed below, the attainment of these objectives is related to the scope and format of the disclosure obligation with respect to the AI system's activity.

Disclosure requirements in foreign regulatory documents dealing with artificial intelligence

The vast majority of prominent regulatory and ethical documents in artificial intelligence globally advocate for embedding disclosure, transparency or explainability requirements, positioning them as central pillars of AI regulation.

- One of the OECD principles with respect to the responsible use of AI applications is transparency (i.e., explainability). The objective of transparency, in accordance with these principles, is to: foster a general understanding of AI systems; inform various relevant players about their interface with AI systems; and enable those affected by the system's outputs a deeper understanding of the system as well as their ability to address or appeal its decisions.[137]
- EU AI Act – This legislation lists the principle of transparency as a general principle applied to all AI systems, separate from the question of the risk inherent in these systems. The legislation indeed establishes regulation following a risk-based approach, but the principle of transparency applies to all systems. As part of this obligation, there is a duty to notify individuals when they are interacting with an AI system, in a way that also allows them to understand and respond to the system's outputs. Beyond this general duty, the legislation sets out specific notification requirements depending on the type of system. For example, a notification obligation applies

[137] OECD, Transparency and Explainability (Principle 1.3), https://oecd.ai/en/dashboards/ai-principles/P7.

when certain information is generated by generative AI, and risk assessment reports for high-risk systems must be published. In the European context, the transparency requirement also extends to the obligation to register certain systems in a dedicated database for high-risk AI systems.[138]

- United States – The Biden Presidential Order addressed, among other things, regulators and federal agencies with respect to the use of AI systems by supervised entities and to the agencies' own use of these systems. It recommended that regulators exercise their authority to impose transparency obligations on supervised entities, clarifying that these requirements aim to protect consumers from fraud, breaches of privacy or threats to financial stability. The Order emphasized the importance of transparency in the federal agencies' use of artificial intelligence.[139] Moreover, in the "Bill of Rights" published by the White House, transparency is listed as one of the guiding principles for AI governance. The disclosure obligation also extends to aspects of personal information use.[140]
- The GDPR includes an obligation to disclose to an individual that use was made of information regarding him collected as part of an automated information processing system for decision-making. Individuals are subsequently granted the right to object to such use. This brief review illustrates that disclosure obligations have become established features in leading regulatory documents dealing with AI governance. However, the scope of the disclosure is not necessarily uniform and may vary depending on the type of technology used, the system's risk profile, the type of user and the like.

Discussion on the various dimensions of disclosure

As stated at the outset, it is possible to divide the discussion on disclosure requirements in the financial sector into three dimensions: the first pertaining to the basic notification requirement regarding the very use of an AI system; the second dealing with the format of the disclosure obligation and its content with respect to the characteristics of the AI system and its decisions; and the third addressing misleading marketing disclosures regarding artificial intelligence.

A. The basic notification requirement regarding the deployment of an AI system

The basic disclosure obligation may simply require making it known and clear that a customer's interaction with a bank's customer service department, for example, is not with a human but with an AI-based chatbot, or that an AI system was involved in making a particular decision. In other words, the notification requirement would apply to the fact that artificial intelligence is being used. Such disclosure can motivate the public to take action – for instance, by choosing to seek services from another financial

[138] EU AI Act, Article 50.
[139] See *supra* note. 25.
[140] See *supra* note, the principle of disclosure is anchored as a duty of "notice".

service provider, or by requesting, to transfer to a non-AI system or to a human representative (insofar as the financial institution provides such an alternative). It can also be argued this "basic" notification requirement serves broader purposes as well, such as facilitating public scrutiny of products and services, since the awareness of AI use may lead individuals to respond and even demand to receive additional information about how the system operates.

The primary disadvantage of imposing a sweeping notification requirement is the possible harm it may inflict on the efficiency for which to the decision to integrate artificial intelligence in the financial institution's activity was initially made. Generative models, for example, are designed to give the consumer an almost human "user experience." In this way, a financial institution can provide faster service and save customer service costs. However, imposing a notification obligation could deter consumers and motivate them to request to transfer to human customer service irrespective of the quality of the AI-generated response, thereby compromising the efficiency that the financial institution sought to achieve. Another possible disadvantage of a "basic" notification requirement is the concern that by making do with a basic notification requirement, the regulator will avoid establishing more extensive disclosure requirements when this becomes possible and even necessary for achieving additional objectives, such as scrutiny of the system and the like. Finally, the notification requirement could be interpreted as a negative sign from the financial entity, signaling to the individual that he should avoid such services.

B. Disclosure obligations regarding AI system characteristics and their implications

In the financial sector, beyond the basic notification requirement, numerous disclosure requirements, as described above, are already in place. The question that arises is whether and how to integrate within these obligations disclosures dealing with the AI system's operational characteristics. This question concerns the degree of elaboration that should expand the range of information given to the financial consumer about the system, including details about the system itself and its characteristics. For example, with respect to the scope of disclosure, whether to elaborate on the system's limitations or alternatively, on its advantages as well, or whether to elaborate whether the system is an experimental service. With respect to the disclosure format– whether to convey the disclosure in publications about the product or service, in a customer service agreement, and the degree of prominence, if at all, to require with respect to the disclosure's publication.

To illustrate the need for disclosure regarding an AI system's characteristics and their implications, consider a system that provides securities analysis– that is, investment advice activity regarding the feasibility of investing in various securities. Suppose such an AI system performs its securities analysis using multiple data sources, including corporate reports, relying solely on the corporations' financial statements while ignoring other disclosures. In describing this system, it would be appropriate to clarify the nature of its data inputs and the implications of this for the system's performance. For example, the

system may assess the desirability of investing in a corporation without taking into account a report regarding a material development in its business activity issued after the publication of the company's financial statements. Disclosing the system's limitations may be important for customers using it.

C. Addressing misleading AI marketing disclosures

Misleading marketing related to artificial intelligence – often referred to as "AI washing" – has recently drawn increasing regulatory attention worldwide. The term refers to claims that products or services incorporate AI systems, when in practice no such systems are used at all or their use is marginal and offers little real benefit. For example, the SEC has taken legal action against investment managers who allegedly made false and misleading statements about their use of AI systems.[141] The SEC's 2023 proposal seeks to address this risk as well.[142] The FTC has also issued warnings against exaggerations and unsubstantiated claims in the marketing of products purportedly powered by AI.[143] Misleading marketing disclosures of this nature is not unique to artificial intelligence. Similar activity emerged in the past with respect to phenomena that created interest among the public at large, in cases in which it is impossible to practically establish the credibility of the publications, due to the complexity of the technology or lack of sufficient knowledge on it.

Disclosure obligations in financial regulation contain a negative dimension as well, namely the prohibition against conveying misleading information to the public. Financial entities are not only obligated to convey complete information about their activity but to refrain from disseminating misleading information as well.[144] Addressing inaccuracies in publications across various areas is carried out through supervision and enforcement of the disclosure and the advertising requirements imposed on the supervised entities. A similar approach can be applied with respect to misleading statements concerning artificial intelligence activity.[145] Beyond the prohibition against misleading information, disclosure requirements to explain the nature of the AI use, and where feasible, also address the system's quality (for example, by providing information about its performance results), can help prevent the release of overly "rosy" representations that do not accurately reflect reality. In the public comments submitted on the Interim Report, some argued that notification and disclosure should be required only in cases in which the user of the system interacts directly with the AI system, and that such notification requirements could impair user experience, create information overload with information customers are unlikely to use, and discourage AI-system adoption. The Taskforce believes that these arguments are not

---

[141] https://www.sec.gov/newsroom/press-releases/2024-36.
[142] See *supra* note 69.
[143] FTC Division of Advertising Practices "Keep your AI claims in check" (February 27, 2023).
[144] See, for example, § 44A of the Securities Law, 1968 (Israel), and § 73A of the Joint Investment in Trust Law, 1994 (Israel), which prohibit the inclusion of a "misleading detail" in various disclosure documents.
[145] See enforcement settlements entered into by the SEC in March 2024 with two investment advisory firms regarding misleading representations regarding their alleged use of artificial intelligence, https://www.sec.gov/news/press-release/2024-36.

persuasive. First, notification and disclosure may be important to the customer or investor even when they do not directly interact with the system, but are nevertheless affected by its use. For example, if a portfolio manager uses an AI system for the purpose of making investment decisions on the customer's behalf, the customer himself does not directly engage with the system, yet its use may significantly affect him (as well as his decision to continue to receive services from this portfolio manager). Second, as is the case with existing disclosure requirements, which as stated above, are a key tool in the financial regulatory toolbox, AI-related disclosure requirements can and should be designed in ways that benefit customers and investors. The report includes an example illustrating the need for such disclosure.

Another comment suggested that where the customer is informed about an AI system, an option should be offered for him to use of an alternative non-AI service. The commenter also argued that the right to a human-made decision was already proposed in the literature dealing with AI regulation, and that it will likely advance principles of fairness and justice. These comments extend beyond the subject of notification, addressing instead the "human in the loop" issue, including the question of the need to provide a human-intervention alternative to AI-generated services, and therefore, they were discussed in the previous chapter dedicated to this issue.

Additional comments included claims that the report does not sufficiently clarify which disclosure requirements will be applied in each and every case, and does not present, for example, the disclosure required in situations where systems are operated in an experimental model, or for example specific disclosure requirements and mechanisms to prevent AI washing. However, given that this report addresses diverse financial activities, for which many legal disclosure requirements already exist, there is neither the room nor the practical possibility of addressing such requirements in detail. As stated in the Taskforce's recommendations, disclosure obligations regarding the characteristics of AI systems and their implications should be integrated into existing disclosure requirements with respect to the service or product offered to the customer, and should address the unique characteristics associated with AI technology. In line with this fundamental recommendation, financial regulators will be able to establish specific disclosure requirements in each and every relevant domain whenever this is deemed necessary.

Notification and Disclosure – Recommendations:

A. Notification regarding the use of an AI system -

A fundamental notification obligation regarding the fact AI system use carries public and consumer value that outweighs the drawbacks noted above.[146] Key considerations for implementing the requirement:

- Notification is particularly relevant in the early stage of the technology's adoption, when a parallel financial activity is conducted without the use of artificial intelligence;
- The necessity of notification is more critical when the AI technology has a material impact on the service or product offered to the customer, particularly if this impact is felt relative to a parallel service or product that does not use artificial intelligence;
- It must be ensured that the notification message will be given in neutral and clear language, which does not create an unwarranted preference for one service over another.
- The notification requirement can be relevant when customers mistakenly think that they are interacting with a person. Alongside the above, a notification requirement is not necessary when it is clear that artificial intelligence is being used.

B. Disclosure obligations regarding AI systems -

The Taskforce is of the opinion that disclosure obligations regarding the characteristics of the AI system and their implications should be integrated into existing disclosure requirements concerning the service or product offered consumers. These disclosure obligations should deal with the unique characteristics arising from the AI technology.

Similar to the recommendation on this subject in the chapter *Black Box and Explainability*, the Taskforce holds that enhanced disclosure obligations regarding characteristics of the AI system are warranted when the system has material implications for the institution's activity and customers. Disclosure in this context should focus on the unique characteristics of the activity associated with the AI technology (as opposed to a situation in which the activity is carried out without AI). To illustrate, in the case of an AI-assisted investment portfolio management activity, there is a need to focus on the essence of the service and its unique characteristics, on the potential associated with it, the system's methodology for determining investment policy or making investment decisions.[147]

---

[146] This is also the approach presented in the Policy Document.

[147] See Israel Securities Authority, Directive for License Holders regarding the Provision of Services Using Technological Means (2016), sections 10, 11, 12, 17 and 20 of the Directive. In the explanatory notes to the Directive, it was clarified that this requirement does not include the exposure of trade secrets, and its purpose is to provide a principled explanation, focusing for example on the fact that the algorithm processes the information in light of the client's answers and therefore there is importance in answering questions thoroughly, or for example on the types of assets the algorithm is likely to recommend. See more on this directive in the chapter on Investment Advice and Portfolio Management below.

Format of disclosure – it is required that the disclosure be made prior to the purchase of the financial asset or contracting for the financial service integrating artificial intelligence. Efforts should be made to ensure that the disclosure is clear, concise, simple and easy to understand. Regulators will develop specific disclosure provisions tailored to the nature of the activity and the applicable requirements.

C. Restrictions on misleading marketing claims -

If misleading marketing regarding the use of artificial intelligence (AI washing) emerges, it can be addressed using existing regulatory tools. This includes disclosure obligations imposed on supervised entities, enforcement actions against legal violations involving misleading information provided to investors or customers, and financial education initiatives such as public alerts and explanatory guidance to the public regarding financial products and services, emphasizing the role of artificial intelligence in their provision.

Box 2: What is AI Literacy? And How Can We Achieve It in the Financial Sector?

In the technological and digital environment in which we live today, "literacy" is not limited to the ability to read, write and use language. It encompasses a broader set of skills that enable individuals to engage with advanced technology applications, including the ability to analyze, understand, respond to and critically evaluate the relevant application. The accepted definition of literacy in the academic literature on artificial intelligence refers to a person's ability to understand the mechanism of AI operation, to evaluate it critically, to interact with it, and to make efficient use of it for his needs on the internet, at home and in the workplace.[148]

Despite the increasing use of AI applications, studies indicate that the ability of the public at large to navigate AI technology is highly limited. Most of the population does not understand neither how AI systems operate nor the technology underlying these systems, and is not even aware of the ethical issues and risks associated with artificial intelligence.[149] This is apparently also true for Israel. For example, as part of a discussion that took place in the Knesset in the subcommittee of the Science and Technology Committee on artificial intelligence, the representative of the Bank of Israel noted that, with respect to skills likely required in the future labor market, the level of the Israeli economy's preparedness for adopting artificial intelligence is not sufficiently high.[150]

Regulators and policy-makers worldwide are already addressing AI literacy issues. A prominent example is the EU AI Act which contains a provision dedicated to this subject.[151] The aim of regulation in this context is that companies creating and using AI systems – irrespective of the degree of risk involved in them – ensure that their employees or anyone else using these systems for them or on their behalf possess the appropriate knowledge and orientation abilities, including adequate technical skills, experience and training.

Indeed, one aspect of AI literacy aims to ensure that business entities creating or using AI applications possess the appropriate skills. A second aspect relates to artificial intelligence literacy among public sector employees – both those whose occupation is related specifically to artificial intelligence and the larger group of employees who will use artificial intelligence as part of their job. Both these groups need to acquire AI skills, particularly the ability to make effective use of AI applications and to recognize and address the risks arising from them. In this context, see the Presidential Order established in the United

---

[148] Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376727

[149] Cary Funk, Alec Tyson, Brian Kennedy, How Americans view emerging uses of artificial intelligence, including programs to generate text or art (February 22, 2023) https://pewrsr.ch/3xP4NdO.

[150] Discussion in the Subcommittee of the Science and Technology Committee on Artificial Intelligence matters dated May 15, 2024 - https://main.knesset.gov.il/news/pressreleases/pages/press15052024c.aspx?fireglass_rsn=true#fireglass_params&tabid=a4894da853688af8&start_with_session_counter=3&application_server_address=isagovil-2-me-west1.prod.fire.glass.

[151] EU AI Act, article 4.

States to encourage use of artificial intelligence in government agencies, and the Taskforce's recommendation on this matter. A third aspect is promoting literacy among the general public. In the United States, for example, in December 2023 the AI Literacy Act bill was submitted, which aims at amending the Digital Literacy Act in the United States with the aim of emphasizing the importance of artificial intelligence literacy for the public at large (and not only with respect to producers or users).[152]

Is the AI literacy of the general public important to its financial activity?

The primary academic literature on AI literacy with the labor market and the education system in a cross-cutting rather than sectoral manner. However, given the importance of financial services to individuals, and the anticipated increased usage of AI applications in the financial world – those that interact with customers or make financial decisions on their behalf – render literacy in this field critically important. Financial education has long been recognized as a key challenge for countries seeking to improve the financial well-being of its citizens, encourage participation in the capital market, and mitigate risks associated with uninformed financial activity, including fraud. In Israel, several initiatives have been launched to promote financial literacy among the general public. For example, in 2011 the government adopted a decision to enhance financial education nationwide in Israel, and enacted legislation to establish a government fund for this purpose, although such a fund has not yet been established.[153] Additionally, all the financial regulators in Israel invest resources to advance financial education within their respective fields of oversight. AI literacy in the financial sector is expected to become an integral part of this challenge, forming a key component of the information included in public financial education. Broadly, this involves: understanding the concept and functioning of artificial intelligence; developing the skills to use and interact effectively with it; recognizing the distinction between AI's potential and its actual capabilities (what artificial intelligence can and cannot do); critically evaluating outputs and associated risks of using artificial intelligence; and ensuring responsible AI use, including ethical considerations.

[152] H.R.6791 - Artificial Intelligence Literacy Act of 2023, https://www.congress.gov/bill/118th-congress/house-bill/6791

[153] Resolution No. HC/243 of the Ministerial Committee on Social and Economic Affairs (Social-Economic Cabinet) dated November 28, 2011, https://www.gov.il/he/pages/2011_des3926?fireglass_rsn=true#fireglass_params&tabid=af5424f9d7e12c08&start_with_session_counter=3&application_server_address=isagovil-2-me-west1.prod.fire.glass.

Privacy and Protection of Personal Information

Artificial intelligence technologies are based on the extensive collection and processing of information, including personal information, that is, information pertaining to an identified or identifiable individual. The development of these technologies, whose potential applications extend across diverse contexts and virtually all spheres of life, constitutes a highly significant accelerating force that intensifies the drive to collect, retain, and utilize vast quantities of data in multiple ways. Information serves as the "fuel" or "oxygen" powering artificial intelligence systems, and is the principal and indispensable resource for their development and operation.[154] Therefore, the development of artificial intelligence systems and their use involve an attempt to base themselves on as comprehensive, diverse, and broad information as possible, whether out of an aspiration to produce content, predictions, or recommendations with a high level of accuracy, or because of the desire to develop innovative products that do not necessarily rely on "standard" information typically expected to serve as a benchmark or input for generating a particular forecast.[155]

Moreover, the products of artificial intelligence systems and applications: namely, the predictions, recommendations, and even decisions, are often directed towards humans, and serve as a means of predicting their behavior, identifying patterns of human action, character traits, mental or physical state, etc. The products embody individual and personal insights about a person, sometimes very sensitive and intimate details about various private aspects of their lives.

Artificial intelligence systems process huge amounts of data using efficient and high-quality computational operations, and therefore facilitate the analysis of vast amounts of information that exceed the capabilities of human intelligence. The information processing performed by them is designed to detect, among other things, **unanticipated** connections and interactions between different data and to maximize the products and insights that can be derived from them. This implies that beyond the high level of accuracy and sensitivity of the sensitivity analysis conducted on human behavior, unfortunately at the current stage of development of AI systems, the products and predictions they may produce are not necessarily expected, known or predefined.

These characteristics of AI systems naturally raise significant challenges and tensions in relation to the protection of the right to data privacy and personal information, throughout the life cycle of these systems. This includes collecting information for the purpose of training the algorithm to

---

* This chapter was written by Attorney Liron Mautner-Lugasi and Attorney Shira Gartenberg from Consulting and Legislation (Public-Constitutional).

[154] Amir Kahane and Tehila Schwartz Altshuler **Man, Machine, State: Towards the Regulation of Artificial Intelligence** 53 (Israel Democracy Institute, 2023).

[155] This characteristic likewise has implications with respect to the risk of discrimination arising from such systems, which may rely on "proxy" data that are not discriminatory in themselves but are strongly correlated with protected or discriminatory attributes. For further discussion, see the section on bias and discrimination.

develop the system prior to deployment, as well as the processing of information during the system's implementation and operational phases, in order to achieve its objectives and support its ongoing development and accuracy.

For example, a fundamental and important principle in privacy protection law, the principle of purpose contiguity, states that personal information should not be used except for the purpose for which it was collected. However, as noted with respect to the development stages of AI systems (and possibly even during their deployment), there is not necessarily clarity regarding the nature of the output that will ultimately be generated. This challenge is likewise reflected in the question of the scope and validity of an individual's consent to the processing of personal data in this domain, given the difficulty of specifying in advance the system's objectives, its potential outputs, and the categories of information that will be used for those purposes.

This raises another tension, regarding compliance with the basic requirement in privacy protection laws not to process information beyond what is necessary to achieve the objective of the processing, in accordance with the principle of data minimization. In his article on privacy and artificial intelligence, Prof. Michael Birnhack noted:

> *…when the system operators ask it to detect patterns and connections that the human eye does not recognize,* ***it is difficult to say more than 'we are requesting information in order to learn about the unknown***.*' Without knowing exactly what the purpose is, it is impossible to examine whether the information collected is sufficient, or exceeds what is necessary*.[156]

These and other challenges addressed in this chapter, in the encounter between the inherent characteristics of AI systems and basic rules and principles in privacy protection law, are largely sweeping, and common to various content areas and sectors, beyond financial services. However, in light of the Policy Document's recommendation to regulate the development and use of artificial intelligence on a sectoral basis, and given the establishment at this stage of the inter-ministerial team tasked with examining the use of artificial intelligence in the financial services sector, we considered it appropriate to undertake the legal analysis of this central and sweeping issue at this time, thereby presenting a comprehensive picture regarding the development of artificial intelligence and its employment in the financial services sector and other content areas and sectors as well.

In the context of the financial uses of AI systems, we would first like to note that the Privacy Protection Law, as amended by Amendment 12, classifies information about a person's salary data and financial

---

[156] Michael Birnhack, "Privacy and Artificial Intelligence," Law Society and Culture 8, 139-175, (2025), p. 163. https://law.tau.ac.il/sites/law.tau.ac.il/files/media_server/law_heb/Law_Society_Culture/books/AI/6Michael%20Birnhack.pdf ("Privacy and Artificial Intelligence").

activity as particularly sensitive information.[157] Similarly, information that is subject to legally prescribed confidentiality is defined by law as "particularly sensitive information".[158] This classification carries several implications under the Privacy Protection Law, reflecting the heightened protection afforded to this type of information due to its sensitivity. Among other things, it is manifested in provisions requiring notification to the Privacy Protection Authority for large databases containing sensitive information, the obligation to appoint a privacy protection officer for anyone processing such information on a significant scale, elevated financial penalties imposed for violations of the Law, information security requirements, to mention a few.

In this chapter, we will discuss key challenges at the intersection between the protection of the right to data privacy and personal information in Israel and the features of AI systems at the different stages of their lifecycle. Some of the challenges existed prior to the rise of artificial intelligence technologies, and have only been amplified by them, while some arise specifically from the special characteristics of these technologies. Within this framework, recommendations will also be offered for further examination of the issue, in general and with particular reference to the financial services sector.

As detailed in this report, the use of artificial intelligence raises numerous social and ethical questions. Various approaches are being developed around the world to address abuse and AI-related extreme scenarios, including cognitive manipulation, the exploitation of vulnerabilities, biases, intrusive applications that threaten human autonomy and dignity, and other related risks.[159]

One approach to managing the risks outlined above is to leverage existing legal mechanisms. In this context, the focus may shift to privacy policy principles, such as informed consent, the assessment of reasonable inferences, and the question of whether there is a departure from the data subject's reasonable expectations. This is particularly relevant given the absence of an alternative regulatory framework, the fact that these tools rely on the collection, analysis, and processing of personal information on an unprecedented scale, and their close connection to the potential loss of control individuals have over how their information is used.

In the Team's opinion, privacy protection laws cannot or should not provide a solution to all the challenges arising from the processing of personal information for the purpose of developing and using AI systems. Therefore, the Team believes that, alongside dealing with these challenges in the financial services sector through financial regulation, there is room to re-examine the need to create legal frameworks that will deal with the aforementioned challenges in a comprehensive yet targeted manner, alongside and enhancing privacy protection laws. Such a response will allow the legal and regulatory response to be adapted to specific challenges in an effective and focused manner.

---

[157] Item (10) for the definition of the term.
[158] Item (11) for the definition of the term "particularly sensitive information".
[159] Footnote 11 above, Chapter III.

Review of the main provisions of relevant privacy protection laws in Israel

The right to privacy has been recognized as a fundamental right with constitutional status. It is enshrined in Section 7 of the Basic Law: Human Dignity and Liberty, which states that "Every person has a right to privacy and intimacy in his life." The Protection of Privacy Law, 5741-1981 (hereinafter in this chapter: **the "Privacy Protection Law"** or the "**Law**"**)** establishes torts and offenses related to the violation of a person's privacy as well as provisions governing the information processing of personal information in databases.

Privacy protection laws or the protection of personal information ("data protection") in Israel, as elsewhere in the world, concerns processed **personal information**. According to the current definition in the Privacy Protection Law, "personal information" is:

> *data relating to an identified or identifiable person; for the purpose of this definition, "identifiable person" - someone who can be identified with reasonable effort, directly or indirectly, including by means of an identifier such as a name, an identity number, a biometric identifier, location data, an online identifier, or one or more data relating to his physical condition, health, economic, social or cultural situation.*

The principles of data protection do not apply to anonymous information - information that does not pertain to an identified or identifiable person or information that, after certain actions have been taken on it, the person can no longer be identified by using it.

From the distinction between personal information and anonymous information, the scope of application of the Privacy Protection Law and its regulations (as well as the provisions of the **GDPR -** the European Union Regulation dealing with the information processing of personal information follows. This distinction determines, among other things, the circumstances under which personal information may be processed, the limitations and principles applicable to such processing, and the rights of the **data subject**, namely, the individual to whom the personal information relates or who can be identified from it.

As part of Amendment 13 to Law,[160] basic definitions were either established or updated in the law, including "personal information," database "controller," "information processing"[161] and the categories of personal information considered "particularly sensitive" were also defined, in harmonization with international standards for the protection of personal information, particularly as articulated in the

[160] Law No. 3287 of 14.8.2024, p. 1430.

[161] A controller of a database is defined in Section 1 of the Privacy Protection Law as "One who determines, alone or together with another, the purposes of processing the information in the database or a body that is or an official in it authorized by law to process information in a database;" Section 3 of the Law defines "processing," "use" as "any operation performed on personal information, including receiving it, collecting it, storing it, copying it, viewing it, disclosing it, exposing it, transferring it, delivering it or providing access to it.

GDPR. Alongside these updates, a dedicated chapter on criminal offenses in databases was added to the law, the supervision and enforcement powers of the Privacy Protection Authority were anchored and expanded, and it was granted the authority to impose financial sanctions and additional administrative sanctions for violations of the law, as well as for failure to comply with the information security requirements anchored in Section 17 of the Privacy Protection Law and the Protection of Privacy (Data Security) Regulations, 5777- 2017, enacted under it (hereinafter: the "**Data Security Regulations**") and for violation of the Privacy Protection Regulations (Instructions for Data that was Transferred to Israel from the European Economic Area), 5783-2023 (hereinafter: the "**Mediation Regulations**"). It was also established that a privacy protection officer should be appointed in many entities, including banks and insurance companies.[162]

The Data Security Regulations, which came into effect in 2018, constitute another important pillar of personal information protection laws in Israel. "Information Security" is defined in the Privacy Protection Law as "protection of the integrity of personal information or protection of personal information from processing without lawful authorization".[163] The Regulations establish a detailed and comprehensive framework for the protection of databases, including governance procedures and data security requirements. They apply to databases maintained by both public and private entities. The Regulations classify databases according to the level of security applicable to each and are intended to impose concrete obligations in the field of information security within organizations, in order to safeguard data privacy in the database and to protect the rights of data subjects against misuse of information relating to them, whether by external parties or by the organization's employees.

The responsibility for securing the information in the database is established in Section 17 of the Law, according to which "the controller of a database and the holder of a database, each of them is responsible for the security of the information in the database." Section 17B of the Law establishes the obligation to appoint an Information security officer in various entities, including banks, insurance companies, and a companies engaged in credit rating or evaluation.[164] The information security officer also bears personal responsibility for securing the information in the database.

Pursuant to the Law, and in addition to the arrangements under Israeli law, the analysis set out below will also include a focused reference to the arrangements established under the GDPR (some of which have likewise been addressed in other chapters of this Report). The GDPR constitutes a point of reference

---

[162] Section 17B1(a)(4) of the Law.

[163] Section 3 of the Law. "Data Integrity" is defined in this Section as "the identity of the data in a database to the source from which it was drawn, without being changed, transmitted or destroyed without legal permission." See also Section (f)(1)5 of the GDPR and Section 32.

[164] Section 17B (a)(3) of the Law. See also Section 17B (a)(1), according to which a database controller that is required to register or notify under Section 8A of the Law, or holds five databases as aforesaid, is also required to appoint a data security officer. The appointment of a data security officer obligation also applies to individual arrangements in other legislation, see, for example, Section 29A(c) of the Population Registry Law.

as an international standard as well as a central source of inspiration for data security and personal privacy protection, particularly in light of the recognition of Israel's adequacy status with respect to this standard, which was ratified by the EU in January 2024. In addition, alongside Amendment 13 to the Privacy Protection Law that entered into force in August 2025, the Ministry of Justice intends to promote another amendment to the law, which includes a series of new substantive arrangements, intended, *inter alia*, to harmonize Israeli law with the provisions set forth in the GDPR and other advanced personal information regulation abroad. This is part of the completion of a process that commenced with the adoption of the Mediation Regulations in 2023, which, as of January 2025, also apply to personal information relating to Israelis when such information is held in the same database together with data received from the European Economic Area, and which has continued with the enactment of the legislative amendment in 2025.[165]

Data Privacy and Personal Information Protection in Artificial Intelligence Systems – Challenges and Recommendations

As noted in the introduction, the convergence of the development and deployment of AI systems with data protection and privacy laws gives rise to numerous challenges and tensions. These stem from the characteristics of such systems—both at the development and training stages and at the subsequent operational stage—including issues relating to their accuracy and the outcomes they generate. These challenges are not unique to the use of AI systems in the financial services sector, but the examination must address the sensitivity of financial information, the confidentiality obligations that apply to it, as well as the liabilities of good faith or fairness that apply in relationships between service providers in this sector and consumers, which are relevant, for example, when examining the application of the consent requirement and other rules, as detailed below.

As a principled position, we note that Israel's Privacy Protect Law is an existing law, technologically neutral, and it applies, generally speaking, to public and private entities alike.[166] **As a starting point, it should be clarified that privacy protection laws continue to apply even where**

[165] As mentioned, the European Union Commission's decision of January 2024 confirmed Israel's adequacy status in the field of privacy protection, here. One of the significant steps taken to bring about the ratification of Israel's recognition as a country with compliance status was the enactment of the Privacy Protection Regulations (Instructions for Data that was Transferred to Israel from the European Economic Area), 5783-2023. These regulations were intended to ensure the preservation of the compliance status that was being vetted at the time, by imposing additional liabilities on Israeli entities that receive information from European Economic Area (EEA) countries. During the process of approving the Regulations in the Knesset, the Constitution Committee decided that the Regulations would also apply to "Israeli" information found in the same database, together with information received from the EEA, with all that this implies in terms of the substantive liabilities set forth in these Regulations.

[166] Public bodies are also subject to additional specific requirements under the Privacy Protection Law (for example, the obligation to register databases under Section 8A of the Law), alongside the obligations imposed on them by administrative and constitutional law.

information is processed through intelligent systems, including artificial intelligence,[167] **as well as to financial information and the specific regulatory arrangements governing it.**

Alongside this, it is undeniable that the deployment of AI systems poses challenges and tensions regarding the implementation of privacy protection laws. In numerous reports by various international organizations and countries that have examined aspects of artificial intelligence, the issue of privacy arises in almost all of them.[168] Recent studies show that even among the general public, the issue of privacy and data protection is key for placing confidence in AI systems.[169]

At the current stage, where the AI system technology continues to develop and the possible uses, benefits and risks posed by them are still being studied, the way the world deals with these risks is also developing. In this chapter, we focus on a number of key **challenges**, which are not unique to Israel but arise globally. We will put forward recommendations and proposed lines of analysis for addressing these challenges, reflecting an effort to strike a careful balance among the relevant considerations, while taking into account evolving regulatory trends worldwide.

A. The challenge of distinguishing between personal information and anonymous information

As noted above, privacy protection laws apply to personal information, that is, information that concerns an identified person or a person who can be identified, directly or indirectly with reasonable effort. If the person to whom the information relates cannot be identified – the information is considered anonymous, i.e., unidentified information – privacy protection laws do not apply to it.

Accordingly, the classification of information as personal data is what "opens the door" to the application of data protection laws, including the conditions under which such data may be lawfully processed, the rules governing its processing, and the rights afforded to data subjects.

How is the information tested for identification? As noted above, the identification requirement is fundamental to the question of whether certain information is personal or anonymous. Even prior to the amended definition of personal information pursuant to the enactment of Legislative Amendment 13 to the Law, the test was explained in the Supreme Court ruling in the **Gottesman** case, where it was determined that the identification requirement is met where a reasonable person can link information to a specific person. This applies even in cases where the information is presented in an apparently

[167] It should be noted that the Privacy Protection Authority has published draft guideline on the application of the Privacy Protection Law to Artificial Intelligence Systems (2025) - https://www.gov.il/he/pages/ai_reg.
[168] Michael Birnhak, "Privacy and Artificial Intelligence," Society and Culture Law 8, (2025) p. 142, p. 17 in the article.
[169] See, for example, a large-scale study conducted by the University of Queensland in Australia in 2023 involving 17,000 participants from 17 countries, who ranked the issue of data privacy and information security as a top priority for developing reliable AI systems. p. 42 of the survey findings report, which is available here.

anonymous form, but may still establish a connection to a specific person. Such a link is not necessarily direct and may be obtained by other means, such as by "reverse engineering".[170]

Following the updated definition of personal information in the Law, the Gottesman ruling must be applied in light of the test prescribed in the Law, as specified above. The identifiability test also fits accepted approaches in Regulation worldwide, including in the definition of personal information in Section 4(1) of the GDPR. In the introductory remarks to the GDPR (recitals), it is noted that the question of whether a person can be identified with reasonable effort will be examined based on objective considerations, including costs and the time required for identification, taking into account the technologies available at the time of processing, and emerging technological developments.[171]

Making personal information anonymous, also known as "anonymization," is attained by removing identifying elements from personal information in various ways that reduce the risk of identification or the disclosure of identifiable information, such as: aggregation (data combination), reducing the accuracy of data by using a range instead of a unique value, omitting details, nullification (adding or subtracting a value to data), and so on.

"Pseudo-anonymization" (hereinafter referred to as **coding**) involves separating identifying data from non-identifying data, using a key that is stored separately, and is subject to technical and organizational measures that restricting access. In place of direct identifiers, the data is linked to a unique code, thereby preserving the ability to correlate the relevant characteristics. This method reduces the risk of personal information disclosure, but does not eliminate the possibility of identifying the person, and therefore does not make the information anonymous and does not exempt it from the application of privacy protection laws.

Different anonymization operations result in different levels of risk of re-identifying the information with a specific person. For example, according to the GDPR, information that has been coded (pseudo-anonymized) is still considered personal information, since it can still be attributed to an identifiable

---

[170] A"A 1697/11 **A. Gottesman Architecture Ltd. v. Vardi**, para. 18, 22-21 (23.1.2013) In RAA 7828/17 **Hadassah Medical Association v. John Doe**, para. 14 (9.1.2018) it was determined that "in such cases, in which the public exposed to the medical information cannot link the information presented to a specific person, that person's privacy is not violated, since without the possibility of indicating the patient's identity, no one invades that person's personal space." The medical information remains disconnected from the context and amorphous, and therefore its presentation does not violate medical confidentiality or the patient's right to privacy." See also the ruling of the Tel Aviv District Court (Administrative) 28857-06-17 **Nursing Care Association v. Ministry of Defense**, para. 16.4 of the ruling of Honorable Judge Agmon-Gonen (1.7.2019) (hereinafter: "**The Nursing Care Association Matter**").
In this context, see also the position of the Attorney General submitted in 22141-03-15 **Greenberg v. Cellcom Israel Ltd.**, Section 40, 46 (It is noted that in that case, an amended settlement agreement submitted by the parties was approved, despite the Attorney General's objection, but the court did not rule on the question of whether in that case the subscribers could be identified through the information (Sections 79.4-79.5 of the ruling dated 5.3.2020).

[171] GDPR, Recital 26.

person by means of the key that enables the matching of personal attributes with non-identifying data.[172] Full anonymization is that which no longer allows for re-identification of the data subject.

The Privacy Protection Law does not expressly address data silencing as a mechanism for mitigating the risk of privacy infringements (save for a specific reference in the Mediation Regulations, which likewise apply to information relating to Israelis where such information is contained in a database received from the European Economic Area).[173] however, this possibility may be reflected in legislation regulating the use of personal information for research purposes. With regard to medical information, for example, Section (20A(7)) of the Patient Rights Law, 5756-1996 permits the disclosure of medical information, without the patient's consent, for research purposes**, provided that no identifying details of the patient are disclosed**.[174] Section (1)23 of the Genetic Information Law, 5761-2000 states that genetic information may be disclosed for research purposes approved by law, regulation or publication in a scientific journal, provided that the genetic information is transmitted "**without any identifying details**." Even if the subject gives written consent to disclose the information, the publication will not include the disclosure of any identifying details unless he has given his express prior consent to this in writing.[175]

With regard to the financial services sector, **the Credit Data Law, 5776-2016,** is intended, among other things, to create an unidentified database that will be used by the Bank of Israel to perform its functions[176] and includes a detailed arrangement regarding the use of unidentified information, information security, and a prohibition on performing any action to ascertain a customer's identity based on unidentified information originating from the database established under law.[177] **The Financial Information**

---

[172] Ibid.

[173] Regulation (3C) of the Mediation Regulations, which establishes the obligation of a database owner to delete information at the request of the data subject in certain cases. Where the data subject has requested deletion as stated in the Regulations, and the exceptions stipulated in Regulation (3B) do not apply - the owner must delete the information under its control **or take actions that ensure that it will not be possible, by reasonable means, to identify the data subject**. Similarly, in Regulation 4, which regulates the liabilities of a database owner to ensure that information that is no longer necessary for the purpose for which it was collected or held, or for another permitted purpose, is no longer held in the database and to delete the unnecessary information. It is noted that the aforementioned liability will not apply **if actions have been taken regarding the said information that ensure that it will not be possible, by reasonable means, to identify the data subject** (Regulation (4C)).

[174] See also section 20(B) of this Law, which states that the provision of medical information under Section (20A), which further regulates cases in which medical information is provided even without the patient's consent, shall not be made available except to the extent necessary, while avoiding as much as possible the disclosure of the patient's identity.

[175] Section 23(2) of the Genetic Information Law.

[176] Section 1(A)(4) of the Credit Data Law. In Section 2, the Credit Data Law defines "identifying information" as "information that includes an Identifying Detail of a Customer, or information from which a Customer's Identifying Details have been separated, but the Customer to whom the information refers can be identified with a reasonable effort"; "Nonidentifying information" is defined as information on a customer that is not identifying information.

[177] See sections 16(B)(3)-(4) of this law, which regulate the purposes of the unidentifiable database; sections 44-45, which regulate the use of unidentifiable information and the provision of access to it. Also see section (58B) regarding the prohibition of re-identification of unidentifiable information and the arrangement regarding the database's data security, which stipulates, *inter alia*, that the Bank in Israel shall collect, retain and collect the information in a manner that ensures that the unidentifiable information is retained and disclosed as such (section (60A,(4)); see also sections (60B)-(C).

Services Law, 5772-2021 also stipulates that a financial information service provider may use financial information that it has collected or received for statistical purposes related to providing financial information services to all of its customers, provided that express written consent has been given for this by the customer, and **that the information** presented to customers will be unidentifiable aggregate statistical information.[178]

In the digital world, AI technologies and computational statistical tools are enhancing the capacity to re-identify data that is purportedly anonymous or has undergone full anonymization. These tools enable the association of de-identified data with identifying elements, drawing on big data. The re-identification process is usually based on statistical correlations between unidentifiable information and available identifiable information about that person.

The challenge of re-identification is not unique to AI technologies and has in fact existed with respect to tools that collect, analyze, and cross-reference big data from multiple sources.[179] However, the rise and development of Artificial intelligence systems further enhances the capacity to re-identify data that was presumed to be anonymous.[180] AI systems make it possible to detect unexpected patterns based on the analysis of duplicate data, and in this way the ability to re-identify unidentified individuals is improved and expanded.[181] To this must be added the growing motivation when developing AI technologies to expand the personal information base for the purpose of training AI systems and developing the algorithm, making them as accurate and sophisticated as possible. The higher the level of data abstraction and generalization, the less effective the use of information can be.

The current challenge of anonymization was also addressed by the US Federal Trade Commission (**FTC**) in a July 2022 opinion regarding the unlawful use and sharing of sensitive personal information. *Inter*

---

[178] Section 25(D)(2) of the Financial Information Services Law, and section (27D) of this law. See also the arrangement concerning research by the Central Bureau of Statistics, in section (17(b) of the Statistics Ordinance [New Version], 5732 - 1972, which states that personal information collected from the types specified in the section that no information" shall be so published as to enable the identification of the person to whom it relates".

[179] See Administrative case (Administrative Court Tel Aviv) 28857-06-17 **Nursing Care Association** v. **Ministry of Defense** (July 1, 2019): "In an era of big data and massive information, there is concern that the disclosure of ostensibly 'innocuous' information—such as location data—may ultimately result in an intrusion into the very core of privacy." See also Omar Tana, "Look in the jar and see what's in it: "Communication Data and Personal Information in the Twenty-First Century," **Legal Network: Law and Information Technology**, 287-301 (eds. Niva Elkin-Koren and Michael Birnhack (2011)).

[180] See also the reference in the privacy chapter of the Policy Document, pp. 79-80.

[181] Karl Manheim & Lyric Kaplan, Artificial intelligence: Risks to Privacy & Democracy, 21 Yale J.L. & Tech. 106, 127-129 (2019); Yves-Alexandre de Montjoye, et al., Unique in the shopping mall: On the Reidentifiability of Credit Card Metadata, 347
(2015) 536 .Sci. Birnhack, "Privacy and Artificial Intelligence", p. 21; Avner Levin, "Privacy in the Age of Explosive Technologies", **Explosive Technologies: Challenges in Israeli Law**, 215 (eds. Lior Zemer, Dov Greenbaum, and Aviv Gaon, 2022), 222-223: "In fact, there is currently an arms race between anonymization technologies for data processing that allows for a new day, and the assumption is that as big data technologies and artificial intelligence advance, anonymization (which still preserves some usefulness of the information) will become impossible". See also a study conducted at the request of the European Parliament Research Centre on Artificial Intelligence Giovanni Sartor, "The impact of the General Data Protection Regulation (GDPR)" on the impact of the GDPR on artificial intelligence development (hereinafter: **European Parliament Research**), 36-37 (2020).

alia, it was noted there that claims that information is "anonymous" or anonymized may mislead consumers, as extensive research shows that anonymized information can often be re-identified.[182]

In England, the Information Commissioner's Office (**ICO**) also published a draft guide on anonymization, partial anonymization, and privacy-enhancing technologies (PET), which proposed guidelines for assessing the risk of re-identification, taking into account multiple considerations: the sensitivity of the information that is sought to be anonymized, the circumstances of the information's disclosure, the scope of the disclosure (for example, whether it is made publicly available or transferred to a specific third party), whether there is additional information that could enable identification and its availability, the use of the information after anonymization, and the technologies currently available for re-identification and under development. It was further noted that the assessment of the risk of re-identification should be reviewed periodically, in accordance with available technologies or publicly accessible information in certain areas.[183]

In France, the Data Protection and Digital Freedoms Commission (**CNIL**) published guidelines on the anonymization of personal information, which includes, among other things, a description and detail of various anonymization methods, and an outline of how to structure the anonymization process.[184] These guidelines proposed three criteria that, if met, indicate that the information is indeed complete: individualization, correlation, and inference. Individualization means that it will not be possible to isolate an individual from the database. Correlation requires that it will not be possible to link information to a person by linking it to other data that includes the same information, along with identifying information about it. For example, a database of addresses cannot be considered anonymous if there is a database that contains the addresses with other data that facilitates the identification of the individuals living at the addresses. According to the inference requirement, it will not be possible to deduce new information about a person from exhausted information (for example, if survey results are published in very high-resolution cross-sections, such as vaccination data by building, in a way that allows information about individual people to be discovered through it.)

Where the three aforementioned criteria are not met, according to the French guidelines, a data controller who wishes to use declassified data must demonstrate, through a thorough assessment of the risks of re-

---

[182] https://www.presidency.ucsb.edu/documents/white-house-press-release-location-health-and-other-sensitive-information-ftc-committed. See also the FTC's recent July 2024 statement on the "hashing" technique (as one of the methods of partial anonymization), according to which information that has undergone the process is not yet considered anonymous and cannot be presented as such, given the ability to re-identify a specific person with it: https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/no-hashing-still-doesnt-make-your-data-anonymous.

[183] ICO - Privacy-enhancing technologies (PETs) - dated 16.6.23, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/privacy-enhancing-technologies. Changes may be made to the document in light of the expected entry into force of the Data Access and Use (DUAA) Act 2025.

[184] See CNIL "Anonymization of Personal Data" (19.5.2020) (Available here).

identification, that the risk of re-identification by reasonable means is small. The guidelines in France also refer to the fact that the anonymization operations and the assessment of the risk of re-identification should be based on **periodic monitoring** in light of the development of anonymization methods, alongside the methods for conducting re-identification. The said monitoring must take into account both the available technological means and the other sources of information that can facilitate the identification of previously anonymous information.

In addition to these, it is worth noting the report published by the OECD on **privacy-enhancing technologies**, which describes the latest developments in the field and the main types of technologies available today, along with an assessment of their effectiveness.[185] "Privacy-Enhancing Technologies" (PETs) is an umbrella term that refers to digital technologies, approaches and tools designed to protect the confidentiality of information, and in some cases also its authenticity and access to it, in order to protect the privacy of data subjects as the as well as the commercial interests of database controllers (therefore, the term PETs are broader than anonymization). The report noted that these technologies typically do not stand alone, and should be seen as tools to be used in conjunction other organizational and legal measures to protect data privacy and personal information.[186]

In 2025, the Privacy Protection Authority also published a Guide to Privacy-Enhancing Technologies (PETs), in which various types of existing technologies and methods in this field are extensively reviewed, while considering the advantages and disadvantages of each technology.[187]

Against this backdrop, it seems that the division between identifiable information, to which privacy protection laws apply, and unidentifiable information, which is not subject to them, is becoming less distinct when AI systems are deployed. This blurring of the distinctions may result in significant risks to data privacy and the protection of identifiable information following the re-identification of ostensibly unidentifiable information. Similarly, the act of anonymizing personal information as a possible solution that renders the use of such information permissible, particularly for the purpose of developing and training AI systems, is not an absolute solution that guarantees complete protection of personal information and the privacy of the data subjects. On the other hand, it is clear that anonymization does have significant advantages in protecting the privacy of data subjects, especially where the information was not collected with their express consent. It makes it possible to reduce the harm to privacy resulting from the use of such

[185] OECD, "Emerging privacy-enhancing technologies: Current regulatory and policy approaches", OECD Digital Economy (Papers, No. 351 p. 14-26 (2023) (available here).
[186] Ibid., p. 13.
[187] Privacy Protection Authority "Guide to "Privacy-Enhancing Technologies" (2025), https://www.gov.il/he/pages/guide_enhancing_technologies.

information while preserving the benefits of information processing - among other things, for improving the quality of life, research, promoting innovation, etc.

The conclusion that data anonymization is ineffective due to the growing risk of re-identification in today's technological environment is problematic, as it may discourage entities from investing in anonymization. Conversely, it can be argued that alongside advances in re-identification techniques, appropriate incentives could foster the development of privacy-enhancing technologies and novel anonymization methods that mitigate the risk of re-identification.

Therefore, the development of data anonymization methods should be encouraged to improve their ability to cope with the risk of re-identification, and developers and users of artificial intelligence systems that wish to rely on anonymized information should be encouraged to use anonymization methods that are considered acceptable and advanced, to reduce the risk of re-identification.

In the context of the public comments on the interim report, concerns were raised about costs and the creation of a barrier to the development of technology if too strict an approach is taken with regard to data mining. Another argument in this context was that it is doubtful whether the supervised entities themselves, unlike the developers of AI systems, possess the appropriate tools to examine the effectiveness of data mining. One source commented that there was a need to clarify regulatory expectations regarding data mining criteria, in coordination with the Privacy Protection Authority, while referring to additional tools other than data mining. In contrast, another source recommended that it be sufficient to instruct financial entities to be aware of the risk and manage it. It was also suggested to clarify that the data mining requirement applies only in cases of sharing information with a third party and not in cases of internal use, and that there is no need to impose multiple obligations on information mined to train AI models.

The debate arising from these comments concerns an issue that has also emerged in other contexts, namely the degree of regulatory involvement required in various aspects of artificial intelligence, with some commenters supporting regulatory involvement (for the sake of stronger protection of protected interests or for the sake of increasing regulatory certainty) and others calling for avoiding regulatory involvement (to prevent unnecessary restrictions and allow supervised entities greater discretion). The Team is of the opinion that, given the importance of data protection as a means of protecting privacy, the recommendations proposed in this context are balanced and can help enhance the effectiveness of privacy protection and anonymization, thereby increasing regulatory certainty. In addition, following the comments, it was clarified that with regard to the quality of the required consent or the ability to prove retroactively the existence of one of the protections under Section 18 of the Law, for example for processing personal information for the purpose of developing an AI model, consideration should be

given to the using accepted and advanced anonymization methods, as well as taking additional measures to mitigate the risk of re-identification and the leakage of personal information. This is in addition to considering other factors, such as the purpose of the information processing, the sensitivity of the data, the data subject's reasonable expectations regarding the use of their information, and similar considerations.

Addressing the Distinction Between Personal and Anonymous Information – Recommendations:

1. It is suggested that financial regulators, in consultation with the Privacy Protection Authority, consider directing supervised entities to conduct a risk assessment regarding the risk of re-identification of personal information[188] before using the de-identified information, throughout the life cycle of the AI-based system – for example, through guidelines, opinions and interpretative clarifications. To the extent that the source of information for training the system or its use is within the financial entity, the responsibility for assessing the risk of re-identification lies with the supervised financial entity. This process should integrate expertise from multiple disciplines, including legal and information security domains.

A. Criteria may be defined, or the discretion applied in conducting the risk assessment may be structured, inter alia, by reference to: the nature of the information intended for anonymization and its level of sensitivity; the scope of the data required and the level of granularity necessary for each data field; the circumstances of the data's exposure (including the scope and identity of parties with access and the intended use of the information post-anonymization); the cost and complexity associated with re-identification; the likelihood of re-identification in light of publicly available datasets; the features of the particular anonymization technique and the extent to which it mitigates re-identification risks; and the range of available anonymization methods.[189]

B. Financial entities should employ organizational and contractual tools alongside technological tools to address the risk of re-identification, such as limiting permitted access as well as other privacy-enhancing technologies. To the extent that personal information is transferred to third parties – include a contractual commitment not to use the information for re-identification, alongside measures to monitor violations of this commitment.[190]

C. Due to technological and other developments in both anonymization and re-identification capacity (including the expansion of databases created and made public and the ability to cross-reference them with additional existing information), **periodic monitoring** of the available anonymization methods

---

[188] See a proposal reflecting a similar approach in Dan Hai, "Privacy and Technology in Israel" **Explosive Technologies: Challenges in Israeli Law** 277, 241 (Lior Zemer, Dov Greenbaum and Aviv Gaon, eds., 2022). See also Omer Tana, "Look in the Jar and See What's in It: Communications Data and Personal Information in the Twenty-First Century," Legal Fund: Law and Information Technology, 287, 301 (Niva Elkin Koren and Michael Birnhak, eds., 2011, p. 319).

[189] See also Draft Anonymisation, Pseudonymisation and Privacy Enhancing Technologies guidance, 2021 (October 13-14). Link available here.

[190] Federal Trade Commission Report: Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers 21 (2012). Available here. See also Sharon Bar-Ziv and Tal Zharsky, "Privacy in an Identity Crisis: Strategies of Regulation in the Age of Silence", **Society and Culture Law**, 125, 141 (Michael Birnhak, ed., 2019) (hereinafter: "**Bar-Ziv and Zharsky**").

should be considered and the tools for reducing the risk of re-identification should be assessed accordingly.[191]

D. It is proposed that the risk assessment also encompass a contingency plan in the event of data disclosure, namely, re-identification. Such a plan could, for example, set out the mechanisms for notifying affected individuals of the breach and the nature of the assistance to be provided to them.[192]

2. In light of the difficulty of carrying out "complete" anonymization and in light of the fact that the information held in the financial system is classified as "particularly sensitive" information, supervised entities should undertake data security measures for anonymized information as well.

3. Upon assessing the use of individual identification methods as a means for supervised entities to mitigate re-identification risks, regulators should evaluate such methods in light of practices adopted in developed markets worldwide, particularly within the financial services sector. It is recommended to consult with the Privacy Protection Authority regarding methods that are considered acceptable and advanced and the response they can provide against the risk of re-identification. As mentioned, the Privacy Protection Authority has published a guide to privacy-enhancing technologies (PETs) that can assist financial regulators and the supervised public.[193]

4. With respect to the standard of consent required, or the ability to demonstrate ex post the applicability of one of the defenses under Section 18 of the Law—for example, in the context of processing personal data for model development—consideration should be given to the adoption of recognized and advanced identification methods, together with additional safeguards aimed at mitigating the risk of re-identification and data leakage. This should be undertaken in parallel with an assessment of other relevant factors, including the purpose of the processing, the sensitivity of the data, and the data subject's reasonable expectations regarding the use of their information, among other considerations.[194]

---

[191] Ibid, page 142. See also draft policy of ICO, Draft Anonymisation, Pseudonymisation and Privacy Enhancing Technologies, Guidance 13-14 (October 2021), https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/2023/12/ico-call-for-views-anonymisation-pseudonymisation-and-privacy-enhancing-technologies-guidance/. See also the guide of CNIL, Anonymization of Personal Data (19.5.2020), https://www.cnil.fr/fr/technologies/lanonymisation-de-donnees-personnelles.

[192] Bar-Ziv and Zharsky, "Privacy in an Identity Crisis: Regulatory Strategies in the Age of Silence," Society and Culture Law, 125-141 (Michael Birnhak, ed., 2019, p. 142). See also Section 34 of the GDPR.

[193] In this context, it is also important to note the provision in Section 17I2 of the Privacy Protection Law, added by the most recent Amendment 13, which sets out the mechanism for referrals to the Privacy Protection Authority for a pre-ruling regarding the assessment of a database's compliance with the Law and its regulations.

[194] To dispel any doubt, it is clarified that under existing law, undertaking a de-identification that does not guarantee the impossibility of identifying the customer with reasonable effort, directly or indirectly, does not remove the data from the category of "personal information".

B. The Challenge of Informed Consent for Information Processing in Artificial Intelligence-Based Systems

The Privacy Protection Law states that a person's privacy cannot be violated without his or her consent or unless authorized by law.[195] Accordingly, most of the Information processing of personal information undertaken in Israel involving privacy violations, not by public authorities, is based either on the consent of the data subject, or on the protections set forth in Section 18 of the Law.[196] Consent must be "informed consent"[197] and, specifically, it must be given with respect to all actions that the entity collecting the information sees to take with the information, including the uses to be made of the information and its transfer to third parties.[198] Moreover, pursuant to the principle of purpose limitation, personal information cannot be used contrary to the purpose for which the information was provided, except with the consent of the data subject.

A person's consent is perceived as a means that removes the objectionable from the violation of privacy in a way that expresses the concept of "privacy as control" – privacy as an expression of the agency to decide what information will be disclosed about them, to whom, when and how, if at all. This concept emphasizes not only the person's decision regarding information about them, but that they have the option to decide and choose what to do with that information.[199] Accordingly, when the data subject has

---

[195] See Section 1 of the Law, which states: "No person shall infringe the privacy of another without his consent.," and Section 35 (the non-derogation clause).

[196] With respect to public authorities, the processing of personal information cannot invoke the protections provided under Section 18 of the Law as a basis for a policy that infringes on privacy. See the ruling in RAA 2558/16 **John Doe v. Compensation Officer, Ministry of Defense**, para. 59 (5.11.2017).
The starting point in the regulation of the European Union, the GDPR, in this matter is different, according to which any processing of personal information requires well-founded legal grounds, among those specified in the GDPR. These are referred to as "processing grounds" (Article 6 of the GDPR). Two of the processing grounds are based on the prior consent of the data subject – one on the data subject's consent to the processing of information for a specific purpose, and the other when the processing is essential to perform an agreement to which the data subject is a party.

[197] Section 3 of the Privacy Protection Law defines consent as "informed consent, explicit or implicit." The word "informed" was added to the consent requirement in the law in the 2007 amendment to the Law. The amendment is intended to ensure that the consent given to a violation of privacy will be informed consent, that is: that the person consenting to the violation will have the information "that he reasonably needs to decide whether or not to consent, and the information will be provided to him in an understandable manner" (Explanatory Notes to the Privacy Protection Law (Amendment No. 9, 2005- *Official Gazette*, Law proposal number 230).
As part of the proposed further amendment to the Privacy Protection Law under consideration by the Ministry of Justice, there is an intention, among other measures, to reinforce the conditions under which a person's consent can serve as a legal basis for processing their information, while also introducing additional legal grounds for such processing.

[198] Michael Birnhack **Private Space: The Right to Privacy between Law and Technology** 100 (2010) (hereinafter: **Birnhack "Private Space"**). See also Tel Aviv (Yafo District) 3213/09 **Jane Doe v. John Doe,** para. 58 of the judgment of Justice Kanfi Steinitz (11.10.2011). In an appeal against the judgment, the Supreme Court upheld the District Court's conclusion that there was no consent to the violation of privacy in that case, see A.A. 8954/11 **John Doe v. Jane Doe,** para. 160 of the judgment of the Justice Solberg (22.5.2014) (hereinafter: **Jane Doe case)**.

[199] Michael Birnhack **Constitutional Privacy** 90-89 (2023) See also Appeal No. 25669-04-21 **Aviv v. Facebook** (9.7.2025), paragraph 61 of the decision. See also the draft opinion statement of the Privacy Protection Authority on the subject of consent in privacy protection laws (2025), https://www.gov.il/he/pages/consent2025.

provided informed consent and the use of their information remains consistent with that consent, such use does not constitute a prohibited invasion of privacy.

A practical expression concerning the informed consent requirement is found in Section 11 of the Law, which stipulates a party's duty to inform a data subject when requesting personal information about him. This obligation is intended to ensure the ability of data subjects to control the information about them, and its use.[200] In some cases, additional legal or regulatory obligations are imposed regarding how the notification is implemented in a specific context or by specific entities (such as banks).[201]

Given the centrality of consent to the processing and use of private information, the relationship in which it is given must be examined. Some consents are "inherently suspect" given the power relations and circumstances in which the consent is given. This can affect the reliance on consent as a remedy for the violation of privacy. "Suspicious" consents can be those given in a relationship between parties governed by power imbalances,[202] or when the receipt of an essential service consent is conditional on consent, and particularly, when there is no connection between the information or the intended use of that information and the essential service. In these cases, faced with the possibility of consenting to a violation of privacy, there is no real alternative,[203] and therefore, it is difficult to conclude that consent reflects a genuine choice and an expression of his agency over the information about him.[204]

---

[200] According to this article, a request to a person to receive information or to process it in a database must be accompanied by a notice that includes the following details: whether that person is legally obligated to provide the information, or whether the provision of the information depends on his will and consent, and what the consequence of the non-consent is; the purpose for which the information is requested; the name of the database owner and the methods of contacting him; to whom the information will be provided and the purposes of its provision. Sections 13 and 14 of the Law stipulate the right of data subjects to inspect the personal information and the right to request its correction, respectively.

[201] Or, for example, Section 33 of the Credit Data Law, which stipulates an obligation to inform a customer (in advance and explicitly) where a credit provider intends to ask a credit bureau to receive a credit report, and that in order to receive the aforementioned report, the credit bureau will submit a request to a bank in Israel to receive the credit data regarding him included in the database. See reference in the Privacy Protection Authority's position statement entitled "Obligation to inform with respect to the collection and use of personal information" (18.7.2022) (the document is available here).

[202] For example: the state versus a citizen, a large corporation versus an individual consumer, an employer versus an employee. Birnhack, "**Private Space**,", p. 253. See also the position of the Attorney General submitted in the Court of Appeals 22141-03-15 **Greenberg v. Cellcom Israel Ltd.**, (5.3.2020) paras. 58-54. See HCJ 8298/22 **Public Defender's Office v. Legal Advisor to the Government**, (31.8.2025) para. 42 of the judgment, as well as HCJ 70206-12-14 **Little Brother Ltd. v. Knesset**, (5.10.2025) para. 8 of the judgment.

[203] See also the introductory note to the GDPR below regarding consent given freely.

[204] For example, in the past, receiving unemployment benefits was conditional on registration by providing a fingerprint. The Association for Civil Rights requested to allow alternative channels and further to this request it was clarified that the Bureau will allow those who so desire to continue to appear before an official. See details in Birnhack **"Private Space"**, p.100 (2010), pp. 104-105, 254-253

The consent required for the processing of personal information under the GDPR is free, specific and informed consent, which constitutes an unambiguous indication of the data subject's wishes and is given by a statement or by a clear act of consent (see the definition in Section (11) (4) of the GDPR). To ensure that consent is freely given, it cannot constitute a lawful basis for the processing of personal data under the GDPR in the following circumstances: where a clear imbalance exists between the data subject and the controller (particularly where the controller is a public authority); where separate processing operations involving personal data cannot be consented to individually despite being distinguishable; or where the performance of a contract, including the provision of a service, is made conditional upon consent to the processing of personal data that is not necessary for the performance of that contract (Recital 32).

This latter issue was also discussed **in the financial context** in the Attorney General's request to cancel and change disadvantageous terms in the general terms and conditions agreement for opening a checking and savings account at Bank Leumi. The Standard Contracts Court held that the clause included in the account opening agreement, which provided for a blanket waiver of the confidentiality obligation vis-à-vis companies affiliated with the bank, constitutes a disadvantageous term in a standard contract. This finding was based, among other factors, on the binding nature of the consent linked to both the account opening and the waiver itself, as well as on the customer's inability to provide individual approval for the transfer of information for a specific purpose in order to receive a particular service, and accordingly the Court determined:[205]

> "Another consideration mentioned by the bank in this context lies in the need to transfer required information, for example, to credit card companies or in the transfer of information that takes place between the bank and leasing companies or the transfer of information for the purpose of investigating a customer to determine a credit limit. **In all such cases, the transfer of information serves to advance the customer's interests. Accordingly, the customer may provide specific consent for the transfer of information for a particular purpose to a designated entity. <u>This does not amount to a broad, blanket consent for information transfer in advance</u>.**
>
> **...However, here the bank conditions the opening of the account on a prior waiver of the duty of confidentiality towards the companies affiliated with the banking corporation. Such a waiver is unfair and disadvantages the customers.** If the customer needs credit or any other service from a company related to the banking corporation, and if the customer's financial situation is relevant to the provision of the requested service, then the providers of those services can condition the provision of the services on a waiver of banking confidentiality. There would be no objection to this. However, in the case before us**, a blanket and prior waiver of banking confidentiality is required as a condition for opening the account. Such a waiver is not required, and it constitutes a waiver of an obligation that the bank owes to its customer under law, which establishes a presumption of deprivation under Section 4(6) of the Standard Contracts Law**."

[205] In the case 195/97 **Attorney General v. Bank Leumi**, paras. 282-281 (2004). In the appeal filed against the judgment in 6916/04, this matter was not discussed.

It should be noted that, beyond situations in which, as stated, the consent is "suspicious in nature," and raises doubt as to whether it reflects the owner's free will, the digital environment reveals additional challenges in relying on consent as a legal source for processing personal information, due to the doubt as to whether the data subject's consent is genuinely **informed**. This environment is characterized by large-scale data processing, and by complex, sophisticated processing operations which give rise to doubts as to whether data subjects possess the requisite knowledge or capacity to understand the information provided to them regarding the processing activities and the risks involved.

The notice accompanying the contact with the data subjects in the digital environment in order to obtain their consent to the information processing and its objectives is usually translated into a "privacy policy" – a computerized form that **makes it difficult to provide genuine "informed" consent. In the course of normal daily use of websites and applications, a person is required to review and agree to numerous and lengthy policy documents, which are frequently updated, and which generally employ legalistic and inaccessible language**. Accordingly, recognizing an individual's act of clicking the "I agree" box in such circumstances as constituting informed consent gives rise to difficulty, as it significantly challenges the realization of the fundamental purpose underlying the consent requirement and its accompanying notice mechanism — namely, enabling the data subject to exercise control over personal data concerning him or her and the uses made thereof.[206]

With regard to the challenge of informed consent in the context of individual AI systems, as described in detail in the chapter "Black Box" and Explainability. One of the distinguishing features of certain artificial intelligence systems is their operation as a "black box," meaning that the manner in which the system's outputs are generated cannot readily be traced or explained. In the same chapter, a distinction is made between **general disclosure**, which refers to the characteristics of the system and its mode of operation (sometimes also referred to as transparency), and **specific disclosure**, which refers to explainability regarding the manner in which a specific decision was made.

The aforementioned characteristic of AI systems also affects the implementation of the consent requirement as a condition for processing personal information that involves an invasion of privacy. The difficulties that we have discussed in relying on consent for processing information in the digital environment are therefore compounded by another significant challenge, which concerns the limitation related to the ability to present to the individual the purposes for which his or her personal information

[206] See in detail Michael Birnhack, "Privacy and Artificial Intelligence," pp. 159-160. In his article, Birnhack cites examples of empirical studies that found that privacy policies in practice do not achieve their goals of adequately informing the data subject. Many users assume that a document titled "Privacy Policy" signifies the protection of their privacy on the website they are visiting, when in fact it does not. It has been argued that standard privacy policies are structured in a manner that takes advantage of users' cognitive limitations, and that reviewing privacy policies across various websites would require, on average, approximately 201 hours per year.

will be used, what actions will be taken with the information, and what information about him or her will be used for those purposes.

This difficulty appears to be even more pronounced where information is collected for purposes other than the use of an existing AI system to make an individual decision — circumstances in which it is otherwise possible to present to the data subject, in a clear and detailed manner, the specific purposes of the processing, the categories of data to be used, and the actions to be taken with that data.

In contrast, in a request to receive personal information for the purpose of developing AI systems, or training them, it seems that the layers of **general** information are relevant, and especially, as mentioned, information about how the AI system operates and its goals (which at this stage are not necessarily known or clear to the developer or the party intending to use the system), the logic underlying it, the data that will be used, and the considerations used in decision-making. To this, add the fact that this is a new technology for which the public's understanding, and especially regarding the potential for its various uses, may be limited.

Furthermore, the consent requirement "constrains" the collection of information with respect to the permitted uses of personal information, when the purposes of use and the system's outputs are not necessarily known. Accordingly, where data subjects are asked to provide personal data for the purpose of **developing** artificial intelligence systems, it is often difficult, regrettably, to specify beyond a relatively general level the precise uses that will ultimately be made of the information.[207]

In a position statement published by the Privacy Protection Authority on the subject of the duty to inform in the context of the collection and use of personal information, the Authority noted, with reference to using AI systems information processing, that the request under Section 11 of the Privacy Protection Law should include details about the operation of the aforementioned systems, to the extent that this is relevant to the formation of consent, and to the extent that this detail is legally, technologically and commercially feasible (a finding that is also discussed in the chapter dealing with the "black box" and explanatory matters). The Authority further recommended that the data subject be provided with an explanation of the categories of information that the systems may process in relation to him or her, as well as the sources from which such information is obtained.[208] Further to these remarks, it was noted that "nothing in this document says that the duty to inform requires the disclosure of **all details** and data relating to the collection and use of personal information." There may be situations in which such disclosure is not feasible or contrary to legal purposes or other legitimate interests. **In such situations,**

[207] Ibid., p. 18. And see the examples cited there for privacy policies that include a very general reference to the uses of artificial intelligence.

[208] See the position of the Privacy Protection Authority, "Duty of Information in the Framework of the Collection and Use of Personal Information," (18.7.2022) in section 24 https://www.gov.il/he/pages/duty_to_notify.

the requesting party must balance between the various interests. It must examine how it provides the data subject with the key data relating to the collection and use of the information (such as data without which the data subject cannot give informed consent to the collection), while avoiding disclosing data that the supplier is prohibited from disclosing.

Another issue that arises when personal data is used for the development or training of artificial intelligence systems relates to the use of historical data, i.e., data concerning current or former customers, for decisions affecting them within the context of a service agreement that has expired, or that did not include consent at the time for information processing for the purposes of training or developing AI applications ("the inventory problem").

Whereas Israeli law generally relies on consent (or authorization) in cases of privacy intrusion, the GDPR prohibits the processing of personal data in the absence of a lawful basis, while also providing for additional grounds that may justify such processing. Thus, in addition to the data subject's consent to the requested processing, the GDPR establishes additional grounds that permit the processing of personal information in general, except for "special categories of personal information" (which are generally equivalent to the definition of "particularly sensitive information" in Israeli law) the processing of which is prohibited, unless one of the limited circumstances stipulated the GDPR is met.[209] In addition to the additional grounds for information processing, the GDPR imposes additional obligations on data controllers and owners, while granting significant rights to data subjects. This issue of expanding the grounds for processing, alongside imposing additional obligations and rights, similar to the GDPR, is included in the Ministry of Justice's work in formulating an additional amendment to the Privacy Protection Law, irrespective of the use of AI systems. In addition, given the difficulties and challenges involved in the reliance on "informed consent" when dealing with AI systems, it is necessary to consider additional legal grounds for processing personal information specifically for the purposes of developing and training AI systems. This will facilitate technological development on the one hand, but will provide remedies and protections for data subjects on the other. In this context, the question has arisen as to whether it would be sufficient to rely solely on the additional grounds for processing provided under the GDPR for the development of AI systems, given that their application is complex and raises issues even at the international level.[210] Accordingly, the team sought public feedback on this matter and on whether

[209] Articles 6 and 9 of the GDPR.

[210] See, for example, Dan Solove, "Artificial Intelligence and Privacy", 77 Flo. L. Rev. 30 (Draft, 1.2.2024) (https://papers.ssrn.com/sol3/(papers.cfm?abstract_id=4713111). For example, in the OpenAI case, the issue arose concerning the reliance on the data controller's legitimate interests as a legal basis for processing, in light of the GDPR's additional provisions governing the processing of special categories of data (under Article 9(2)) and against the broader context of web scraping practices. see the initial report published by the EDPB on the subject of the technology's compliance with GDPR, EDPB "Report of the work undertaken by the ChatGPT Taskforce" (23.5.2024) (https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf) and the latest opinion

the creation specific legal grounds for processing personal data tailored to AI systems, alongside the appropriate balances and protections adapted for artificial intelligence is warranted.

As part of the public comments on the interim Report, many comments referred to the distinction between Israeli law and the GDPR in all that concerns the fact that in Israeli law the need for consent or authorization is only relevant to situations in which there is a violation of privacy (rather than in relation to all information processing), while the legal grounds for processing in the GDPR are more numerous than those in Israeli law. It was also specifically argued in the context of AI model training that the consent requirement is inappropriate and not legally mandated, and that introducing a principle of the right to withdraw consent, as under the GDPR, without the availability of the other legal grounds for information processing provided therein, would significantly impede the development and training of AI systems.

Further to the comments and question on this issue that the team referred for public comment, a recommendation was added to expand the grounds for information processing in Israeli law.

Regarding the recommendations that sought to strengthen informed consent with unique requirements that would apply specifically to AI systems, such as a default of consent in the opt-in format, a requirement for explicit and prominent consent, or a mechanism for graduated consent, it was argued that these add non-binding but burdensome layers. Most commenters argued that there is no place for unique rules regarding informed consent for the use of AI technologies. On the other hand, other commenters believe that the use of AI systems raises significant questions regarding the ability of data subjects to understand and provide informed consent to information processing and believe that the strengthening the informed consent mechanism is warranted. The Team believes that it is incorrect assert categorically that the use of artificial intelligence is equivalent to that of any other technology, but rather that AI systems have distinct characteristics that warrant a legal or otherwise tailored approach. However, the report's recommendations took into account the comments received in order to focus on the circumstances in which informed consent is indeed required when the use of artificial intelligence violates privacy.

One of the questions raised by commenters to the Interim Report is the ability of financial entities to use historical data at their disposal as part of AI system operations. This depends, among other things, on the extent of the consent given to the financial entity regarding such information. Their position was that one should not conclude that the mere use of an AI system necessarily requires obtaining renewed consent from each customer. For illustrative purposes, if a credit company has previously used customer

---

of the EDPB on the ChatGPT Directives of December 17, 2024 (https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf).

information in a credit underwriting model based on the customers' consent, it would not necessarily need to obtain new consent to use an AI system for the same purpose. Such a requirement is not required by law, but rather depends, as stated, on the nature of the consent given and the activity proposed to be carried out. The team reviewed the comment and made changes to the report's recommendations with respect to this issue.

Informed Consent Challenge – Recommendations:

1. The greater the risk associated with a proposed AI activity, the more complex its purposes (i.e., the harder it is to understand the processing outcomes or their implications), or the more likely it is to exceed the data subject's reasonable expectations, the more explicit the form of consent should be—such as an opt-in default or consent that is prominently requested.

Another approach is to implement "graded consent" mechanisms, which aim to balance the challenges of providing detailed information at the data collection stage with the obligation to inform, while also helping to enhance customer trust regarding the provision of their data, such as:[211]

a. When the purpose of processing information using artificial intelligence is unknown and more general at the collection stage, state this explicitly.

b. Where possible, specify specific purposes or possible specific uses of the system under development, in a manner that enables the data subject to specify the purposes or uses for which they consent to the processing of their personal information.

2. Making the provision of a financial service conditional on consent to the processing of personal information for the purpose of developing, training and operating the AI system will be possible, provided that the development, training and operation pertains to the system used to provide that service. However, when it comes to **basic financial services**, such a condition will only be possible when it is required for the purpose of providing the service specifically to that data subject.[212]

3. In relation to historical data ("the inventory problem"), the need to obtain renewed consent for the use of AI systems depends on the nature of the consent given and the proposed activity, and it should not be assumed that the mere deployment of an AI system automatically necessitates obtaining renewed consent from each customer. To the extent that the use of artificial intelligence is for the same purpose for which the information was initially provided in a non-AI environment (for example, as part of providing the financial service that the customer requires), then renewed consent will not be required, when using AI systems for the same purpose. However, to the extent that the use of AI systems deviates from the original purpose (for example, when consent was given to process information for the purpose of providing a specific financial service and it is used for training a general AI model or for the purpose of

---

[211] See Birnhack's proposal, "Privacy and Artificial Intelligence", on p.174, as is customary in the field of medical research. See also Floridi, Luciano & Daniel Brent (Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts, 22 Sci. & Eng. Ethics 303, 312 (2016).

[212] See in this context the discussion conducted above in the chapter on human involvement, which indicates that in the future, certain activities may require a human alternative. Such a requirement may naturally also have implications for the requirement of consent to provide information for the purpose of developing and training AI systems. However, as stated, the Team does not believe that there is currently a need to impose a blanket obligation to provide human activity alongside activity carried out by AI systems.

providing a different service than that which the customer requested to receive), then separate consent shall be required for this purpose. Even when the information is used for the same purpose, but the manner of use significantly exceeds the customer's reasonable expectations (for example, using a picture from the customer's ID card for training a facial recognition algorithm to prevent bank fraud), separate consent shall be required from the customer.

In the event that renewed consent is required for the use of information about previous customers or existing customers who did not consent to the use of information about them for the purposes of developing and training AI systems –

(1) The Team recommends freezing the information, subject to guidance from the financial regulator on assessing the risk of re-identification, as a potential solution that balances the desire to use the data with the challenges of contacting customers to obtain renewed consent. The requirement for renewed consent may be waived where there is no reasonable risk of re-identification, provided that the criteria and considerations outlined above are met, and that compensating organizational and legal safeguards are implemented—such as imposing a legal obligation on information recipients not to attempt re-identification.

(2) It is proposed that the financial regulators, in coordination with the Privacy Protection Authority, clarify the threshold for determining when consent or knowledge is sufficient in situations where recognized and advanced identification procedures have been implemented, along with additional safeguards, but a residual risk of re-identification remains at a low level nonetheless. For example, if these conditions are satisfied, it may be possible to consider obtaining consent for the processing of personal information on an opt-out basis should suffice, subject to conditions to be prescribed.

4. It is proposed that financial regulators consider establishing provisions regarding granting customers, the data subjects, the right to withdraw consent to the processing of their personal information in certain circumstances for the purpose of developing and training an AI system, beyond the withdrawal option granted to them under existing law. This right is among those currently afforded to data subjects under the GDPR. It provides that personal data may not be further processed for the relevant purposes on the basis of consent once the data subject has withdrawn that consent. The withdrawal does not affect the lawfulness of processing carried out prior to the notification of withdrawal. According to the GDPR, the information regarding the right to withdraw consent must be provided before consent is given, and the exercise of the right must be accessible in the same way as the granting of consent. It should be clarified that the intention is not to impose an obligation to retrain the model without the information about the data subject who withdraws consent, but rather that the data relating to that data subject should be removed from the database used to train the model, so that it will no longer be used as of that date onward.

5. Similar to the arrangements set out in the GDPR, according to which a condition for processing personal data is the existence of the legal grounds for information processing from among those specified in the regulation. This includes additional grounds beyond the data subject's consent. it is proposed to establish additional legal bases for processing personal data for the purposes of developing and training artificial intelligence systems. This will facilitate the development of technology on the one hand, but will provide remedies and protections for information subjects on the other. It is proposed to enshrine these bases in legislation, while considering whether, in the context of developing AI systems, it would be appropriate to establish a specific legal basis for private information processing, accompanied by tailored safeguards and balancing mechanisms.[213]

C. The challenge of AI system outputs - inferred personal data

Artificial intelligence systems significantly expand the ability to infer new information about data subjects by running their models on personal information. These systems absorb and analyze vast volumes of data, aiming to identify unexpected patterns of behavior or correlations, including those that were previously imperceptible to the human eye. By means of AI models, such systems are used to identify, classify, predict, or evaluate a particular individual. This may occur even where information about that individual does not appear in the raw training data and even where the individual has not personally provided any information.[214] In other words, even if a person refrains from sharing personal data for machine-learning purposes, data supplied by others may be used to discern general patterns that subsequently enable detailed, and at times sensitive, insights about that person (particularly in contexts such as credit or insurance), without requiring direct access to their personal data (or relying only on minimal identifying details).[215]

The manner in which artificial intelligence systems learn from and process training data significantly enhances the quality, precision, and granularity of the information they infer, as well as their capacity to make inferences and draw conclusions from data that appears, at first glance, entirely unrelated to the relevant assessment or outcome. With the development of AI systems, the aforementioned inference

[213] Article (3)7 of the GDPR - The right to withdraw consent forms part of the suite of rights afforded to data subjects and is expected to be incorporated into the additional amendment to the Privacy Protection Law currently being advanced by the Ministry of Justice. At present, this right is also enshrined in various consumer protection contexts. On this matter, see the position statement issued by the Attorney General in Case No. 22141-03-15 **Greenberg v. Cellcom Israel Ltd.**, (5.3.2020) Section 61 A.A. 8954/11 **Jon Doe v. Jane Doe**, Section 160 of the judgment rendered by Justice Solberg (22.5.2014).

[214] For example, in models that provide recommendations based on an analysis of book purchases or viewing habits, artificial intelligence is used to identify reading and viewing patterns, and when it identifies a person, whose information is not included in the raw data but shares a similar behavioral pattern, it assumes that he will be interested in similar products. Birnhack, "Privacy and Artificial Intelligence", pp. 156-157.

[215] Birnhack calls this challenge “the group challenge,” see ibid., pp. 22-23.

ability is constantly being upgraded and it is possible to use it to reach particularly sensitive information about an individual from relatively "distant" data, so that even if a person has control over the information specifically about him, he cannot control the way in which the information will be evaluated and the insights that will be gleaned about him, let alone anticipate the potential uses of such information.[216] For example, systems may generate inferences regarding an individual's sexual orientation based on an analysis of external facial features, or draw conclusions about a person's mental state and personality traits from patterns in smartphone usage or payment history.[217]

Accordingly, the improved computing and processing capabilities, alongside the availability of big data and the development of AI technologies, dramatically improve the ability to find correlations and connections that unlock sensitive and rich insights into individuals with respect to personality and behavioral aspects, ethnicity, sexual preference, and opinions, as well as predictions about their future behavior. The use of these insights carry far reaching and diverse commercial potential, such as the development and marketing of personalized services (see, for example, the chapter on credit in the banking system in relation to credit marketing), behavioral targeting for advertising and marketing purposes, and more. They can also function as a decision-making tool with considerable implications for an individual. In financial contexts, these tools can be instrumental in insurance pricing, credit risk assessment, identifying anomalies in credit card use for fraud detection, personalized investment portfolio management, and the like.

We have discussed above the complexity introduced by AI systems, particularly regarding the potential for unexpected uses and the challenges that can arise in fully achieving their development objectives, especially in relation to fulfilling the obligations to inform and obtain consent for the use of personal data. The present section is devoted to another challenge in privacy and data protection that these capabilities pose. It relates to the issue of how AI **outputs** concerning an identified or identifiable individual should be classified—specifically, whether such outputs of AI algorithms about a person should also be considered "personal information," and whether additional obligations or rights should apply to this type of information, or whether adjustments are needed in the implementation of the relevant legal framework.[218]

---

[216] Alicia Solow-Niederman, "Information Privacy and the Inference Economy", 117 Nw. U. L. Rev. 357 (2022).

[217] Wang, Y. & Kosinski M. "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images", Journal of Personality & Soc. Psychology 114(2), 246 (2018); Stachl, Clemens, et al. "Predicting Personality from Patterns of Behavior Collected with Smartphones", Proceedings of the National Academy of Sciences of the United States of America 117, no. 30 (2020); Ramon Y, Farrokhnia R.A, et al. Explainable AI for Psychological Profiling from Behavioral Data: An Application to Big Five Personality Predictions from Financial Transaction Records, Information 12 (2021).

[218] The question can be illustrated with an example from the analog world, in a case heard by the European Court of Justice (**ECJ**) under the EU Directive on the Processing of Personal Data that preceded the GDPR. In C-434/16 Nowak v. Data Protection Commissioner (2017) (available here), an accounting exam examinee claimed that the examiner's comments on the test he took constitute personal information and therefore he is entitled to request access to them and to correct them since

In Israel, the recently revised definition of "personal information" in the Privacy Protection Law, as "data relating to an identified or identifiable person," implies that even data that is an inference or prediction about a person is data relating to such a person and is therefore considered "personal information," protected under the Privacy Protection Law. To illustrate, a healthcare provider's impression of a person and his assessment of their behavior patterns, or a medical assessment of a person's condition—even if they are subject to subjective interpretation by various experts and even if they were inaccurate or turn out to be so—are personal information relating to that person, to which privacy protection laws apply. In other words, there is no requirement that the information accurately describes the identified or identifiable person in order to extend the protection of the law to them, but it is sufficient that it is data relating to that person.

Furthermore, the Privacy Protection Law, as worded following Amendment 12, states that personal information that is " personality assessment conducted by a professional who, as a matter of practice, gives their opinion on a person's personality, or conducted **by means intended to evaluate** essential personality characteristics, including character traits, intellectual ability, and the ability to function at work or in studies" shall be considered "particularly sensitive information".[219] This definition is designed to categorize as particularly sensitive any information generated by technological tools, including AI applications, that aim to produce assessments or predictions about specific core personality traits. Such classification carries implications for various legal requirements, including obligations to notify the Privacy Protection Authority about databases, the appointment of a privacy officer, penalties for legal violations, mandatory data security standards, and other related provisions. The European data regulation in the **GDPR** also includes special reference to this issue, within the framework of the provisions concerning "profiling," which is defined as follows:[220]

> *... any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to*

---

they are inaccurate. The **ECJ** ruled that the exam notes are also personal information, as even a subjective impression, such as an assessment or opinion, can be information concerning an identified or identifiable person. It was further determined that, as a result, in order to use the examiner's assessment for other purposes outside of the examination process or to transfer it to a third party, the examinee's consent must be obtained, and that the examiner's assessment must also be stored in a secure manner to prevent unauthorized access to it. The court recognized that the rights of data subjects, such as the right to inspect, the right to correct, and the right to delete, also apply to the examiner's assessment, but determined that the right to request correction is in accordance with the purpose for which the information was collected. Therefore, the right to correction does not include the right to correct test answers or the examiner's comments (unless they were recorded incorrectly).

[219] Paragraph (8) of the definition of "particularly sensitive information" in Section 3 of the Privacy Protection Law. See also the Privacy Protection Authority's guideline 2/2012 "Guidelines on the Application of the Protection of Privacy Law on Employee Screening Procedures and Institutions" (2012) (available here). This guideline (based on the previous definition of "information" in the Privacy Protection Law) states that, as a general rule, screening institutes must provide job candidates with the opinion prepared regarding them, as provided to the employer who ordered the tests, including the results of the tests themselves.

[220] Section 4(4) of the GDPR.

analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements."

The emphasis in the definition is on the purpose of the automated processing. Automated processing designed to classify people according to characteristics such as age, height, and gender does not necessarily fall within the definition of profiling – for example, if it is used solely for statistical analysis. However, if the automatic processing is intended for predictions or conclusions about a specific individual, it falls within the definition.[221]

In the introduction to the **GDPR**, it is noted that profiling is subject to GDPR rules regarding the processing of personal information, such as the principles of information processing and the requirement to process information according to one of the legal bases for processing.[222] This is also the guideline issued by the professional committee of privacy protection authorities in the European Union countries (EDPB).[223]

In addition to the scope of the provisions on processing principles and the legal bases for information processing, the **GDPR** addresses the rights of the data subject when profiling is carried out in concrete terms, including imposing a duty of transparency on owners and controllers, and confers the right to inspection and the right to correction on the data subjects. These rights also apply to the of the profile's output rather than just the inputted data. In addition, individuals retain the right to object to profiling, and in particular to profiling for marketing purposes.

This means, for example, that where of personal information processing is based on a customer's consent that the information about him or her will be used for credit scoring purposes, then the output of the algorithmic information processing by which a particular credit score is determined is also considered personal information - the continued processing of which requires specific legal grounds for information processing. Accordingly, the use of the credit score for other purposes (for example, for training a general AI model, or for developing algorithms that will provide other services) is contingent on the data subject's informed consent for this use as well. Otherwise, an additional legal basis, will be required for processing.

In defining personal information, the California Consumer Privacy Act (CCPA) of 2018 includes, among other things, inferences drawn from other types of personal information specified by the law, which are

---

[221] Article 29 Data Protection Working Party "Guidelines on Automated Individual Decision-Making and Profiling for the Purpose of Regulation 2016/679, p. 7 (2018)(hereinafter: "**EDPB Interpretation on Profiling**")(available here).
[222] Recital 72 of the GDPR. EDPB Interpretation on Profiling.
[223] Ibid.

intended to construct a profile of the consumer across various dimensions, as follows:[224] "inferences drawn from any of the information identified in this subdivision to create a profile about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes".[225]

As noted above, artificial intelligence is transforming, the scope and accuracy of inferences and predictions about individuals, at times changing them beyond recognition. These predictions are becoming increasingly detailed, precise, and nuanced, and can draw on both typical and unconventional information to identify core traits and sensitive data about a person, at times even uncovering information the individual may not be consciously aware of, such as behavioral or cognitive patterns, or analyses of the strength of certain relationships. In this context, beyond the generation of new personal information and the general ability to draw conclusions, another aspect of privacy infringement occurs when individuals lose control over how information about them is used to make inferences or draw conclusions regarding their personal characteristics, or over the manner in which such information is applied.[226]

As part of the Interim Report, the Taskforce sought the public's views on the question of whether additional specific obligations or rights should be applied with respect to inferred personal information, or whether adjustments should be made as part of the implementation of the Law on this matter. Comments were also sought on whether, in the context of using AI tools in relationships governed by heightened duties of good faith and fairness – such as in the provision of financial services – individuals should have the right to draw reasonable conclusions regarding their own situation.[227] i.e. one that meets the criteria of Relevance, Reliability and **Acceptability**:

Acceptability – The information used as input for the inference should be of a type deemed acceptable for individuals—for instance, it should exclude protected characteristics such as sexual orientation (see also the recommendations in the chapter on bias and discrimination).[228]

Relevance – The information inferred must be relevant to the decision and acceptable in that context. For example, there is no reason for an application designed to produce a credit-worthiness assessment to convey information about ethnicity.

---

[224] Section 1798.140 (o)(1)(K) 224.
[225] See also the ICO's Artificial Intelligence Guidance, which specifically addresses the use of AI systems to make inferences (link here).
[226] Birnhack, "Privacy and Artificial Intelligence", p. 156 and references therein.
[227] Sandra Wachter & Brent Mittelstadt, "A Right to Reasonable Inference: Re-thinking Data Protection Law in the Age of Big Data and AI", Colum. Bus. L. Rev. 494 (2019).
[228] This last issue stems directly from the existing law on equality and discrimination in Israel, see the chapter on bias and discrimination.

Reliability – The information on which the model is based, including the information on which it was trained, should be statistically reliable and accurate (see also Recommendation 4 in the chapter dealing with bias and discrimination, among other things).

The comments received in this context ranged between two extremes. According to one approach, there can be particular sensitivity in information inferred from sensitive aspects, such as personality characteristics or mental state. This sensitivity, along with the potential privacy concerns stemming from the fact that data subjects neither control nor are aware of the insights generated about them warrant particular regulatory attention. One party proposed requiring explicit consent from data subjects for the derivation and use of such information, subject to the criteria outlined above (acceptability, relevance, and reliability). The adoption of additional measure, such as enhancing transparency toward data subjects and drawing a distinction in financial regulation between raw and processed information. Another respondent suggested that the elements of reasonable inference should form part of the risk assessment process for AI systems, while ensuring that both the information processing and its outcomes comply with the three criteria above.

Another commenter suggested making reference to Recital 71 of the GDPR, which addresses the right not to be subject to decisions based solely on automated processing, including profiling, where such decisions may have a significant impact on the individual—for example, the denial of a credit application. It should be clarified, however, that this Recital relates to Article 22 of the GDPR, which establishes this right and sets out specific exceptions to it, and that these exceptions do not have an equivalent under Israeli law. According an alternative approach, articulated by another party, although inferred information is indeed included in the definition of "personal information," it should not be concluded that special arrangements are necessarily required for it, and that such arrangements should be unique to the financial system. It was further observed that the “importation” of privacy-related arrangements from foreign legal systems should be undertaken through primary legislation, rather than through sector-specific regulation applicable to financial institutions.

After considering the public comments and further reflection on the subject, it is proposed that the inference of personal information by financial entities should be reasonable, i.e. based on standards of acceptability, relevance and reliability. Such a recommendation, even if not fully derived from privacy protection policies, is consistent with the rest of the report's recommendations with respect to managing the risks associated with AI systems, and with the protections required for investors and customers in whose interest AI systems operate, particularly in light of the special relationship existing between financial entities and their customers. It is proposed that financial regulators assess the need to provide further detail and guidance on how these principles should be implemented in the operations of supervised entities that use AI systems to generate sensitive inferences.

Inferred Personal Information – Recommendations:

1. Inferred personal information is included in the definition of "personal information" - the basis for responding to the issue of personal information is found in existing law, which defines "personal information" broadly as "data relating to an identified or identifiable person." The inference that the output of AI systems - predictions, inferences, assessments, classifications made about a specific person - are **personal information** that is protected by privacy laws. This is in addition to the express provision of the law that refers to a personality assessment conducted **using a means designed to assess** material personality characteristics, as **particularly sensitive** material information.

2. Reasonable inference - It is proposed to adopt a rule regarding reasonable inference in the case of an individual, which will apply to high-risk decisions. Reasonable inference is based on the criteria of acceptability, relevance and reliability. It is suggested that financial regulators examine the need for detail and guidance on how to apply this rule regarding the operation of supervised entities in artificial intelligence systems that generate sensitive inferences.

D. Challenges in implementing the principle of relevance

A fundamental principle of privacy and personal data protection law in Israel and worldwide is the principle of purpose limitation, or purpose contiguity, under which personal data may be collected and used only for the specific purposes for which it was originally collected or disclosed.[229] This rule is enshrined in Section (9)2 of the Privacy Protection Law, which states that using information about a person's private affairs for a purpose other than that for which it was disclosed is a violation of privacy (which constitutes a civil wrong and a criminal offense).[230] Purpose contiguity also appears in Chapter B of the Law dealing with databases, in Section (8b), which stipulates that "No person shall process personal information in a database except for the purpose legally determined for the database." Financial sanctions may be imposed on a data base controller or owner[231] for a violation of these provisions. In addition, a judicial order may be issued at the request of the Commissioner of the Privacy Protection

---

[229] See Section (9)2 of the Privacy Protection Law, which states that using information about a person's private affairs or disclosing it to another for a purpose other than that for which it was disclosed is a violation of privacy that constitutes a civil wrong and a criminal offense for which the maximum penalty is 5 years (Sections 4-5 of the Law); Section (8b) states that a person shall not process personal information in a scientific database except for the purpose of the database that was legally determined for it. In the GDPR, the principle of purpose limitation is enshrined as one of the processing principles listed in Section 5. Section 5(b)(1) states that personal information must be collected for specified, explicit and legitimate purposes and that information cannot be processed in a manner incompatible with those purposes.

[230] For the application of this principle in case law, see, for example, A.A. 439/88 **Registrar of Databases v. Ventura**, P.D. 439(3), 808 817 (1994); A.A. (National) 8/90 **Isakov Inbar – State of Israel – Commissioner for the Employment of Women Law**, para. (26d), 41 (8.2.2011).

[231] A database holder is defined by law as an external entity to the controller of the database that processes information for them; (see also the definition of processor in Section 4(8) of the (GDPR).

Authority, to cease information processing activities that violate these provisions or that are likely to lead to their violation.

The principle of relevance of purpose is also expressed in individual arrangements governing transfers of personal information to various entities. For example, the arrangement in Section 29A of the Population Registry Law, 5725-1965, which authorizes the Population Authority to transfer information to bodies which are required to verify an individual's personal identifiers against registration details in the Registry, including banks. The section further reinforces the principle of purpose contiguity, stipulating that the transfer of information is conditional upon a commitment by the receiving party to use the information solely for the purpose for which it was provided.[232] Similarly, Section 25 of the Financial Information Services Law states that a financial information service provider may collect and receive financial information, and use it solely for purposes specified in the law.[233]

In I.D.I. Insurance Company Ltd. v. Registrar of Databases, in which the District Court and later the Supreme Court were required to apply the principle of purpose contiguity, an insurance company to which an order to seize a debtor's assets was transferred, continued to retain information even after the procedure that required the information contained in the order was completed. When the debtor later sought to purchase insurance from the insurance company, he was refused, based on the information contain in the foreclosure order.[234]

The District Court ruled that the principle of purpose contiguity in that matter prohibits the insurance company from using information about foreclosures that it receives to facilitate debt collection by creditors, for the purpose of classifying existing customers or potential customers. In an appeal filed against the verdict, the Supreme Court upheld these findings.[235]

---

[232] Section 29A(a) of the Population Registry Law, and see also Sections 29A(d)-(f).

[233] The Financial Information Services Law further stipulates that financial information service provider may use financial information it has collected or received for statistical purposes related to provision of the financial information service to all of its customers, provided that express written consent has been given for this by the customer, and that the information presented to customers will be aggregate statistical information which is unidentified information (see Section (25D(2)) of this Law). See also Sections 25 and (58-A) of the Credit Data Law; Sections 52 and 136 of the Insolvency and Financial Rehabilitation Law, 5778-2018.

[234] The procedure reached the courts following a complaint filed with the Privacy Protection Authority. The Authority subsequently, issued Directive 1/2011 " Use of Data Regarding the Imposition of Foreclosure" dated September 2011 (link here). The guidelines stated that the violation of privacy resulting from the use of information regarding the imposition of a foreclosure can be legalized by obtaining the **prior** informed consent of the debtors, in accordance with Section 1 of the Law. This means that such use will be possible in practice only in cases where the debtor is a customer of the third party (see the Authority's position statement regarding the circumstances in which such a provision in a standard contract of a service provider will be considered a disadvantageous condition, Sections 3.1.6-3.1.7 of these guidelines). In addition, the guidelines stipulate that a third party that has made a decision concerning the customer based on information regarding a foreclosure imposed on it, must explicitly notify the customer to avoid the damage that may be caused to the party whose assets were seized pursuant to a mistaken or false claim (Section 3.1.9).

[235] ITA (Administrative Tel Aviv) **24867-02-11 I.D.I. Insurance Company Ltd. v. Registrar of Databases, Authority for Law, Technology and Information at the Ministry of Justice**, pp. 22-26 (5.8.2012) for appeal in the Supreme Court –

The principle of purpose **limitation is also recognized as one of the basic principles for information processing in** the **GDPR**, which states that personal information will be collected for a specific, explicit and legitimate purpose and will not be processed in a manner incompatible with those purposes.[236] The GDPR stipulates that, under certain safeguards against misuse, information processing for archiving purposes in the public interest, for scientific or historical research purposes, or for statistical purposes will be considered compatible with the original purpose.[237]

The assumption in this section of the **GDPR** is that the use of information for statistical purposes is This provision of the GDPR proceeds on the assumption that the use of data for statistical purposes is compatible with the originally specified purpose and, as such, will generally be permissible, subject to the implementation of appropriate additional safeguards.[238] However, for information deemed sensitive under the GDPR, processing for statistical purposes is allowed only if it is necessary to achieve those purposes and is grounded in Union or Member State law that is proportionate to the processing objective, respects the fundamental right to personal data protection, and incorporates safeguards for the rights and interests of the data subjects.[239] The preamble to the GDPR specifies and clarifies the processing compatible with statistical purposes for the purpose of this principle, according to which it concerns any act of collecting and processing personal information required for statistical surveys (similar to the CBS) or for producing statistical results. These results can be used for various purposes, such as scientific research. According to the preamble, the statistical purpose means that the result of the statistical processing is not personal information but rather aggregated data, **and that the personal information collected <u>or the result of the processing</u> is not used to support measures or decisions concerning a specific person**.[240]

In light of the above, it seems that a statistical purpose that would be deemed compatible for the principle of purpose relevance under the GDPR cannot bear an interpretation according to which the use of personal information for developing AI tools and algorithms also falls within the scope of "statistical purposes", insofar as these are intended to support measures or decision-making related to specific individuals. To the best of our understanding, a significant portion of AI systems are such. In any case, such an exception to the principle of purpose contiguity does not currently exist in Israeli law.

A.A.M. 7043/12 **I.D.I. Insurance Company Ltd. v. Ministry of Justice, Authority for Law, Technology and Information – Registrar of Databases** (15.1.2014) (hereinafter: "**IDI Appeal**").

[236] GDPR, Article 5 (b)(1).

[237] See Article 89 (1) of the GDPR- which requires, among other things, the adoption of technical and organizational measures to implement the principle of minimization, such as partial anonymization, if the goals can be fulfilled this way. Where statistical purposes can be fulfilled by processing that does not or no longer allows the identification of the data subject, they must be they must be carried out in that manner. See also the introductory note to Recital 156 of the GDPR.

[238] Recital 50 of the GDPR, Recital 238.

[239] Recital 9(j)(2) of the GDPR.

[240] Recital 162 of the GDPR.

Applying the principle of purpose limitation in the context of processing personal information for development of AI-based systems is a key issue because building big data banks often involves combining data that was collected separately for different purposes and processing that data for purposes that were unimaginable at the time the data was collected. In other words, these systems often seek to build upon and even use data that was not traditionally used for certain purposes based on the discovery of unexpected correlations between the aforementioned data.

In this context, it is proposed to focus on three challenges for implementing the principle of purpose limitation: **The first** concerns meeting the requirement that the purpose of information processing be clear, detailed and explicit, in order to ensure compliance with this principle. As noted in the discussion of consent presented above, sometimes during the development phase of AI tools, the purposes of use and the system's output are not known to the developer or the entity that intends to use the system, and it is difficult to specify, beyond a fairly general level, how the system will utilize the information. A **second** challenge concerns the issue of "inventory," i.e. a scenario in which information collected from customers for a specific purpose is intended to be repurposed for the development of AI tools. In the discussion on consent, we considered a potential approach to the issues arising in this context, involving either clarification or renewed communication with customers, in line with the framework described above. **Third**, pursuant to the growing motivation to expand the data set on which AI systems are based, various entities harvest information from the Internet (scraping) for the purposes of developing AI systems (for example, images or other information from social networks). Although this information is ostensibly available on public websites, collecting it from these platforms and using it to develop AI systems for financial services (or for any other purpose) may go beyond the scope of consent or the reasonable expectations of users when they posted the information, and would violate the principle of purpose relevance.[241]

As part of the public comments on the Interim Report, one commenter proposed the following measures to address the issue of purpose contiguity: tailoring the level of detail required about the purpose of information processing to the assessed risks posed by the system (i.e., the greater the use or the more it exceeds a customer's reasonable expectations, the more detailed and explicit the purpose description should be); aligning new purposes with existing purposes based on a risk assessment of the system; and amending the legislation to permit the limited use of data for developing artificial intelligence systems, subject to appropriate safeguards. Another party contended that the concept under Israeli law, which prohibits the use of information for any additional purpose without the express consent of the information’s owner, is neither necessary nor analogous to the GDPR, which addresses purpose

[241] Michael Birnhack, "Privacy and Artificial Intelligence", p. 164-165.

limitation rather than purpose contiguity. Accordingly, it was proposed to consider extending the principle of purpose contiguity through a purposive interpretation of Section 9(2) of the Privacy Protection Law, as reflected in Section 18 of that Law.

A third party reiterated his argument, which also addresses other matters, according to which the issue of the relevance of the purpose is not unique to financial institutions or AI systems, and no special provisions should be imposed. The same source also noted that the use of information for the purposes of training AI systems is within the reasonable expectation of the customer and therefore there is no justification for requiring additional consent.

The recommendations regarding the challenge to the principle of congruence of purpose remained essentially as they were in the Interim Report. Following public comments and modifications made to the recommendations regarding the inventory problem with respect to the consent challenge, a change was also made to the recommendation regarding the use of personal information for the development of AI systems, clarifying that the change only applies to the development of a model that is not for the original purpose for which the information was originally provided. For example, if the purpose of using the information was, *inter alia*, to improve a financial entity's credit underwriting activity, then conduct of this activity within the framework of an AI system would be included within the scope of the original purpose and would not constitute a violation of privacy necessitating renewed consent.

One of the recommendations dealt with the issue of scraping. Following the publication of the Interim Report, the Privacy Protection Authority published a draft guideline on the application of the provisions of the Privacy Protection Law to AI systems,[242] which stated that mining personal information from the Internet for the purpose of processing it in an AI system, including for the purpose of training the algorithm, involves a prohibited violation of privacy, if the informed consent of the data subject has not been given. The team left the recommendation that the financial regulators, together with the Ministry of Justice, examine the possibility of establishing provisions governing the conditions and circumstances under which information may be collected by means of scraping in a manner that does not violate the provisions of the Privacy Protection Law or any other law.

[242] Draft Privacy Protection Authority Directive: Application of the Privacy Protection Law to Artificial Intelligence Systems (2025), published for public comment (https://www.gov.il/he/pages/ai_reg).

Principle of Purpose Contiguity Challenge – Recommendations:

1. It is proposed that the level of detail required regarding the purpose of processing personal information and its use in the development of AI systems be calibrated to reflect the level of risk posed by the system that the financial entity seeks to develop. The more complex the purposes for which personal information is used—or the further they deviate from the customer's reasonable expectations or from the service for which the customer contracted with the financial service provider—the greater the level of **clarity** and **detail** required in the explanation provided to the customer. In this context, see the above proposal concerning a tiered consent mechanism as a means of addressing the difficulty in specifying the purposes of processing at the development stage, where the processing of information is based on consent. However, in our view, given the challenges posed by AI systems in terms of an individual's control over information concerning him and the application of the fundamental principles governing data processing, there is no justification for treating the development of AI systems for purposes other than those for which the information was originally collected (for example, the development of a general AI model or of a model for a service different from the one sought by the customer) as a use compatible with the original purpose of collection, such as the provision of the financial service. In any event, as noted above, the principle of purpose contiguity in Israeli law does not presently include authorization to process personal information for so-called "compatible" purposes.[243]

2. On the other hand, where the use of personal information is for the purpose of **using** an AI system that has already been activated, or where the requested use is for the purpose of **training** an existing system, the purposes of use must be detailed and clear. The assumption is that at this later stage, the initial difficulty in specifying and explaining the purposes of processing personal information that characterizes the system development stage exists to a lesser extent.

3. It is proposed to consider legislation that would facilitate the application of the principle of purpose contiguity for the development of AI systems, based on data collected for other purposes, subject to certain restrictions and control measures. The EU-AI Act establishes such an arrangement, subject to, among other things, the following conditions: the use of information may only be for the purpose of an AI system **intended to protect vital public interests** such as the prevention and investigation of criminal offences, public safety and health, and environmental protection; **personal information is required and information that has been anonymized cannot be used**; information processed within the sandbox **will neither be transferred or made accessible to third parties, and will not be used in**

[243] To illustrate, the use of information provided for the purpose of rendering financial services by a supervised financial entity can also be for the purpose of training the AI systems that will be deployed in providing the services. This is similar to the use that the same entity would make of information for improving systems, managing risks, and the like, without an AI system. On the other hand, if it concerns, for example, the sale of the information to another entity for the purpose of independently developing an AI system for other purposes, the principle of contiguity of purpose is obviously not met.

or influenced by decisions made regarding the data subjects; the information shall be stored in a secure, isolated and controlled environment that can be accessed only by authorized persons; personal information processing operations shall be fully documented; rules shall be set governing the deletion of information and the period of time it can be stored.[244]

4. It is proposed that the financial regulators, together with the Ministry of Justice, examine the possibility of establishing provisions governing the conditions and circumstances under which information may be collected by means of scraping the Internet in a manner that does not violate the provisions of the Privacy Protection Law and any other law, for the purposes of developing AI tools by or used by financial entities (for example, images or other information taken from social networks). This information is ostensibly available on public websites, but a question arises whether extracting it from those platforms and using it[245] for the development and training of AI systems in the financial services sector, or for making a specific decision in relation to an individual in the context of providing a financial service, exceeds the scope of the user's consent or reasonable expectations at the time of publication, and thereby contravenes the principle of relevance.

E. Challenges in implementing the principle of data minimization in Artificial Intelligence-based systems

An additional core principle is that of data minimization, or the restriction on the use of excessive data in the processing of personal information. Under this principle, only data that are relevant and necessary for the purpose of the processing may be collected, and the collection or processing of excessive data, namely, information that goes beyond what is required to achieve the original purpose, is prohibited.[246]

Excess information can be created at the initial information collection stage or can be rendered excessive during its retention over time, subsequent to the completion of the purpose or following the collection of additional information, etc. Accordingly, the obligation to minimize excess information applies both

[244] EU AI Act - Sections 53-55.

[245] See sections 104-106 if the European Data Protection Board (EDPB) - Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, of December 17, 2024, which proposes to take various measures in the context of dealing with the risk that the use of web scraping creates - for example, restricting the collection of personal information and the sources from which it is obtained where such collection may pose risks to particular individuals or groups (including potential victims of harm, discrimination, or violence); excluding certain categories of data, for example due to their sensitivity; honoring the objections of websites that employ technological measures to prevent data scraping; and establishing an "opt-out" list enabling data subjects to object to the collection of their information from specific websites.

[246] See the ruling in A.M. 7043/12 **IDI Appeal,** in which the Supreme Court also addressed the issue of retaining excess information and its use. The ruling determined that "information that has reached an entity such as an insurance company or bank regarding a particular matter must not be used for purposes outside of the original permitted use, **and it must be removed from their institutional memory for any purpose other than the original one**".

at the initial collection stage and throughout the period of the information's retention in the database.[247] Collecting and using personal information when it is no longer necessary for the purpose for which it was provided, or for the purposes of the database, violates the principle of purpose contiguity as detailed above, creating unnecessary security risks, which may constitute a violation of the obligation to ensure information security.[248] The more personal information stored in the database, the greater the risk of privacy infringement as a result of leakage and unintended disclosure of the information.[249]

The principle of data minimization is expressed, among other things, in Regulation (2c) of the Data Security Regulations, which requires that database owners examine at least once a year whether he holds excessive information in a database that is no longer needed for its original purposes.[250] In addition, Regulation 6 of the Privacy Protection Regulations (Conditions for Retaining and Preservation of Information and Procedures for Transferring Information Between Public Bodies), 5746-1986, which deals with the procedure for transferring information between public bodies under Chapter D of the Law, requires the body receiving the information to separate the excess data upon receipt and to delete it immediately.[251]

In the financial context, see also the arrangement in Section 27 of the Financial Information Service Law, which stipulates provisions limiting the duration of the period for which a financial information service provider retains financial information that it has collected or that has been transferred to it.[252] This law also stipulates that the service provider must delete the financial information at the end of a period prescribed by law.[253] The Credit Data Law also stipulates that a credit bureau that has received credit data shall retain it "for the minimum period required" for the purpose of providing a credit data service or a service based on credit data, and for setting a ceiling on the data retention period.[254]

---

[247] See the Privacy Protection Authority's policy statement on data minimization, a draft published for public comments in March 2021 (here). An expression of the principle of data minimization already at the initial stage of data collection can be seen in the Gush Dan Congestion Charge Law, 5782-2021, which stipulates that vehicle footage captured by cameras will not be viewable and will not be stored except in a manner that does not allow the identification of the passengers in the vehicle being filmed or other passersby (Section (28a) of the Law).

[248] Section 17 of the Privacy Protection Law.

[249] Policy document on data reduction by the Privacy Protection Authority, draft published for public comment in March 2021 https://www.gov.il/.BlobFolder/rfp/data_minimization_public_hearing/he/data_minimization_public_hearing.pdf.

[250] For the sanction prescribed for violation of this provision within the framework of Amendment No. 13 to the Privacy Protection Law, see item (3) of the Third Schedule.

[251] See Section 23E(b) of the Privacy Protection Law. See also Regulation 4 of the Mediation Regulations, which requires the owner of a database to operate an organizational, technological or other mechanism, the purpose of which is to ensure that the database does not hold information that is no longer necessary for the purpose for which it was originally collected or held or for another purpose for which it is permitted to be held under any law, and to delete information found to be unnecessary as promptly as circumstances allow. Starting in January 2025, the regulations also apply to information about Israelis held in a database in Israel containing information transferred from Europe.

[252] Section 27(A)(1) of this law. See also section 27(A)(2) which states that the Minister may shorten the periods.

[253] Section (27B) of the Financial Information Services Law, except in the case of retaining account information for the purpose of a legal proceeding or an internal audit or supervision proceeding by law, which is regulated in Section 27(C), including the various conditions applicable to retaining account information for these purposes.

[254] Section (49A) of the Credit Data Law.

The principle of data minimization is also one of the principles of Information Processing set out in the GDPR.[255] In the introduction to the GDPR, it is noted that the meaning of this principle is that it is necessary to ensure that the period during which personal information is stored is strictly **limited to the** minimum possible, and that in order to ensure this, the controller must set aside periods of time during which personal information is deleted or conduct periodic monitoring of the stored information. It was further determined that this principle implies that personal information should be processed only if the purpose of the processing cannot reasonably be achieved by other means.[256]

An inherent tension exists between the implementation of the data minimization principle and the development of AI-based system. While data minimization seeks to confine the collection and use of information to what is relevant and necessary for the purpose for which it was provided, AI-based systems typically depend on processing data at the broadest possible scale in order to draw unforeseen insights, and to train and refine their outputs. As mentioned, during the model development stage, it is not necessarily possible to know what insights and benefits the system will derive from the information and what correlations will be found useful for predicting or achieving the processing goals. Moreover, due to the lack of clarity that can exist in the context of the intended uses of an artificial intelligence system, it may be difficult to characterize certain data as excessive. Moreover, implementing the obligation to delete personal data may present practical challenges once the algorithm has already been trained on that data. It may also impair the ability to trace the model's development and to monitor or audit the outcomes generated by the algorithm.

As part of the public comments on the Interim Report, a comment was received that likewise underscored the challenges inherent in predicting in advance which information may prove to be relevant. The commenter proposed, inter alia, examining the development of specific minimization indicators, incorporating an explicit reference to the difficulty of minimizing unstructured data, and establishing mechanisms that would permit deviation from the minimization principle where justified. Since the report's recommendations already set out several possible approaches to implementing the data minimization principle—each of which requires further consideration and development—the additional alternatives raised in general terms in the comment were not incorporated.

Another recommendation advanced in the public comments, and subsequently adopted, was to create a separation between data collected at the various stages of the system's life cycle, to store such data separately, and to determine retention periods, access authorizations, and related controls in accordance with the needs of each stage.

---

[255] GDPR Article 5(1)(c).
[256] GDPR, Recital 39.

Data Minimization Challenge – Recommendations:

1. Structuring the data minimization requirement with a focus on the personal nature of the information, namely reducing as much as possible the ability to link the data to the individuals to whom it relates, rather than concentrating primarily on reducing the volume of data collected (which is essential for model development). See the discussion above in the chapter on data minimization and below regarding privacy-enhancing technologies (PETs).

2. Taking organizational and technological measures to review and document the data required for the development of the AI system throughout the development process, and cleaning and deleting data that appears to be irrelevant at this stage.

3. Financial entities should assess, having regard to the relevant costs and benefits, whether in certain circumstances it is appropriate to segregate the information collected and processed according to the various stages of the system's life cycle (development and training; deployment and use; maintenance and model improvement), to store it separately, and to define retention periods, access rights, and related controls in accordance with the evolving needs of each stage.

4. Restrictions on the use of certain types of information for the purpose of developing AI systems in certain areas (for example, a restriction on the use of facial images) should be examined in advance.

Additional General Recommendations:

In view of the various challenges outlined above, it is clear that the development of AI tools gives rise to substantial and multifaceted risks with respect to safeguarding customers' data privacy and their personal information". Against this backdrop, and in particular where the assessment of the risks of developing a certain type of artificial intelligence system leads to the conclusion that it should be classified as a high-risk system, it is proposed that financial regulators consider directing entities under their supervision, through guidelines, opinions or recommendations as to best practices and the adoption of organizational tools that will address the privacy risks involved in deploying such tools, including:

1. Privacy Impact Assessment: a methodological process designed to assist the organization in identifying, assessing and managing privacy risks in projects or other business and organizational activities that involve the processing of personal information. There is currently no such obligation in Israeli law, but it is proposed to recommend adopting this mechanism as a tool for managing and assessing the risk in the use of AI systems that generate predictions and assessments regarding people

for the purpose of receiving financial services.[257] It is proposed to guide the review in the early stages of initiating a project that involves the collection or processing of personal information, in parallel with the process and planning of the system – and prior to the commencement of information processing operations.[258]

2. Implementation and assimilation of Privacy Design (Privacy by Design) and Privacy by Default (Privacy by Default) concepts (hereinafter collectively referred to as "Privacy Design") – The privacy design concept is intended to serve the organization's purposes (business or other), while minimizing privacy risks and investing resources in appropriate protective measures against the risks identified in the organization's personal information processing process, whether as part of a privacy impact assessment or otherwise. According to this concept, starting from the early planning stage and throughout the entire life cycle of information collection and use, as well as the available and accepted technological alternatives, information systems will be designed, and additional organizational measures implemented, with a view to maximizing the protection of data subjects' privacy and safeguarding their rights under applicable privacy protection laws. How the privacy-by-design principle is applied in each case depends on the concrete circumstances of the information-processing operation and the organization, as well as the state of technological development at the given time. Without detracting from the generality of the above, privacy design can be implemented, among other things, by reducing the collection and processing of information to the essential minimum, anonymizing or pseudo-anonymizing personal information, using synthetic information rather than personal information, employing differential privacy techniques (encryption), reducing access permissions, establishing work procedures, improving transparency towards data subjects, including in the form of consent, and a long list of other privacy-enhancing technologies.[259] These recommendations are in addition to the obligation to appoint a **privacy protection officer** pursuant to Amendment 13 to the Privacy Protection Law. The Law requires many organizations to appoint such an officer, including any organization in which the principal activity of a database controller or owner involves the large-scale processing of particularly sensitive information. In this context, it expressly enumerates examples of entities falling within this category, including **financial entities**, such as banking corporations as defined in the Banking Law (Service to Customers), 5741-1981, and insurers as defined in the Financial Services Supervision Law (Insurance), 5741-1981.[260] The role of the privacy protection officer is to ensure the financial entity's

[257] See Article 35(a)(3) of the GDPR.

[258] See the methodological guide for conducting a privacy impact assessment published by the Privacy Protection Authority, (updated December 7, 2025), available here.

[259] OECD, "Emerging privacy-enhancing technologies: Current regulatory and policy approaches", OECD Digital Economy Papers, No. 351 (2023). https://www.oecd.org/en/publications/emerging-privacy-enhancing-technologies_bf121be4-.en.html.

[260] For more information regarding privacy-enhancing technologies, see a comprehensive guide published by the Privacy Protection Authority on the subject, above. In addition, see section 17B1(A)(4) of the Law.

compliance with the provisions of the Law and to promote the safeguarding of the privacy and the security of information held in its databases.[261]

[261] See the details of his duties in Section 17B2 of the Law. For more information, see the draft statement published by the Privacy Protection Authority on the subject of appointing a privacy protection officer in organizations according to the requirements of Amendment 13 to the Law (https://www.gov.il/he/pages/dpo25_1).

Bias and Discrimination

A key challenge associated with AI-based systems is the risk of discrimination and bias. Although algorithmic decision-making may reduce the impact of unconscious or intentional human biases, AI systems are also exposed to the risk of biased or discriminatory output for various reasons outlined below. Moreover, the extensive scale at which these systems operate greatly increases the likelihood that such issues will emerge on a much larger scale, having a far larger impact than would a single human decision. One of the primary causes of bias and discrimination in AI systems relates to the databases used to train them. When the system is trained on data that are not sufficiently representative – meaning they lack information about all populations relevant to the system's activity– it may generate biased or discriminatory results. For example, in the medical field, an AI system trained solely on data relating to men may perform sub-optimally and unfairly when assessing medical cases involving women.[262] In addition to issues of underrepresentation, AI training data may contain existing societal biases and discriminatory patterns. Thus, if certain demographic groups have faced bias in employment opportunities or access to credit, an AI system trained on such data may perpetuate and reinforce the biased outcomes.[263] Additional causes of bias and discrimination relate to the manner in which the algorithm is designed and how the variables used by the model for decision-making are selected. In this context, even if the database on which an algorithm is based is representative, AI models may lead to prohibited discrimination due to the consideration of "protected" group affiliation data, such as nationality, gender, race, sex and more. In this regard, one of the key challenges of AI-system design is the inability to completely and consistently block the system from considering discriminatory group affiliation data. This is because AI systems may make use of "proxy" data, i.e., data that is not inherently discriminatory (such as institution of higher education, zip code, browsing history, etc.) but is strongly correlated with discriminatory variables (such as race, nationality, etc.). A leading study in the United States found that an AI system used to identify candidates for specialized treatment programs incorporated data regarding the scope of an individual's medical expenses. Although such data is not discriminatory in itself, its use ultimately resulted in fewer black patients being identified relative to white patients.[264]

---

[262] Isabel Straw, The Automation of Bias in Medical Artificial Intelligence (AI): Decoding the Past to Create a Better Future, 110 ARTIFICIAL INTELLIGENCE MED. 101965, 101966 (2019).
[263] Michael A. Griffith, AI Lending and ECOA: Avoiding Accidental Discrimination, 27 N.C. Banking Inst. 349 (2023), https://scholarship.law.unc.edu/ncbi/vol27/iss1/16.
[264] Ziad Obermeyer et al., Dissecting racial bias in an algorithm used to manage the health of populations, Science, 366, 447-453 (2019), https://www.science.org/doi/10.1126/science.aax2342.

Alongside the above factors that heighten the risk of bias and discrimination, "feedback loops" can precipitate the reproduction and amplification of bias or discrimination by the system. If an AI system generates biased outputs or results in biased or discriminatory consequences in practice, it may subsequently learn from those very outcomes, thereby reinforcing and perpetuating its biases over time.[265] In the context of the financial sector, the concern for AI-based bias and discrimination may arise in numerous areas of activity. One of the these, which has also received considerable attention in global regulation, pertains to the extension of credit. Indeed, models used for risk assessment in the context of extending credit enable more accurate and efficient assessments, and consequently, may expand credit accessibility to various populations due to more accurate predictive power based on richer information,[266] thereby reducing the risk of discrimination and bias for a specific individual human decision. However, the use of algorithmic models can also raise the risk of bias or discrimination due to their training on databases containing insufficient or biased information with respect to certain populations or due to the system's use of proxy data.[267]

Claims of AI-based discrimination in the extension of credit have arisen in recent years in various contexts in the United States. For example, in 2019, a claim posted on social networks regarding discrimination by Apple's credit application, operated in cooperation with Goldman Sachs Bank, as it set lower credit limits for women relative to men. The New York State Department of Financial Services (DFS), the regulator of financial services and insurance in New York State, examined these claims and found no gender-based discrimination arising from the credit application.[268] In another case, a 2020 report published by the "Student Borrower Protection Center," identified significant disparities in fees and interest rates for student loans requested by applicants with otherwise identical data—differences stemming from the institution of higher education where they studied, which is correlated with racial characteristics.[269] The risk of bias or discrimination may arise in other financial contexts as well. In the field of insurance, for example, regarding the pricing of an insurance policy, while the use

---

[265] Emilio Ferrara, The Butterfly Effect in Artificial Intelligence Systems: Implications for AI Bias and Fairness, Machine Learning with Applications, Volume 15, 2024, https://doi.org/10.48550/arXiv.2307.05842?force_isolation=true.

[266] In this context, it should be noted that one of the declared objectives of the Credit Data Law, under which the credit data register at the Bank of Israel began operating, is expanding access to credit and reducing discrimination in granting credit.

[267] For example, the system's reliance on a datum such as higher education, which does not constitute a protected identity criterion in itself - may lead in practice to result-based discrimination in the provision of credit to certain population groups. For an analysis of the reasons for discrimination in this context, see: Ruth Plato-Shinar and Maayan Perl, "Consumer Credit Underwriting Based on Algorithmic Data Processing – Is There Reason to Fear Unlawful Discrimination?" Iyunei Mishpat 45, 553 (2022).

[268] New York State Department of Financial **Services Report on Apple Card Investigation** March 2021, https://www.dfs.ny.gov/system/files/documents/2021/03/rpt_202103_apple_card_investigation.pdf.

[269] Student Borrower Protection Center, Educational Redlining (February 5, 2020), https://protectborrowers.org/wp-content/uploads/2020/02/Education-Redlining-Report.pdf.

of AI enables precise pricing of risk for each individual insured, there is a concurrent risk that the system will take protected identity characteristics, such as nationality or race, into consideration and do so in a discriminatory manner.[270] Also, in the field of investments, a system based on not sufficiently representative data may produce biased or sub-optimal results for certain segments of the population.[271]

The Law Concerning Discrimination in Israel

In Israel, the duty to act without discrimination applies to public bodies or entities performing public functions under administrative law.[272] In parallel, regarding private entities, the Prohibition of Discrimination in Products, Services and Entry to Places of Entertainment and Public Places Law, 5761-2000 (hereinafter: "the Prohibition of Discrimination Law"), establishes a sweeping ban on discrimination in the provision of a public product or service or the operation of a public place. Accordingly, Section 3(a) of the Law states as follows:

> 3(a). A person whose occupation is the provision of a public product or service or the operation of a public place shall not discriminate in the provision of the public product or service, in granting entry to the public place or in providing service in the public place, on the grounds of race, religion or religious affiliation, nationality, country of origin, sex, sexual orientation, viewpoint, party affiliation, age, personal status, parenthood, or wearing uniforms or symbols of the security and rescue forces.

This section establishes a ban on discrimination based on a list of grounds as enumerated in the section, relating to groups for which there is a history of deprivation and discrimination.[273] In addition to these grounds, Section 3(a1) of the Law also establishes a prohibition on discrimination on grounds of place of residence.

---

[270] For illustration, in an investigation on behalf of "Pro Publica", from 2017, it was found that residents in minority neighborhoods in various states in the US paid a higher car insurance premium than in other areas - https://www.propublica.org/article/minority-neighborhoods-higher-car-insurance-premiums-white-areas-same-risk , and see also a reference to this example in a report published in England, on behalf of the Center for Data Ethics and Innovation, titled "Review into bias in algorithmic decision-making", p. 56, https://assets.publishing.service.gov.uk/media/60142096d3bf7f70ba377b20/Review_into_bias_in_algorithmic_decision-making.pdf .

[271] For example, models based on the client's investment history may lead to less favorable results regarding young clients for whom there is insufficient information. See IOSCO report from 2021, at pp. 10-11, and in particular in *supra* note 9.

[272] Itzhak Zamir, Administrative Authority, Vol. 5 - Grounds for Judicial Review, p. 3970 (2020). In this context, regarding banking corporations, "the position has been expressed more than once that banking corporations are hybrid bodies (dual-nature bodies), and as such they owe duties from public law, including fairness, reasonableness, equality and the rules of natural justice" (CA 1691/11 **Bank Leumi Le-Israel Mortgage Bank Ltd. v. Tsuberi**, paragraph B of the opinion of Deputy President E. Rubinstein (15.12.2015), and see also LCA 9065/23 **S.A. Top Investments Ltd. v. Mercantile Discount Bank Ltd**., paragraph 33 and the references there (3.1.2024).

[273] See explanatory notes to § 3 in the Prohibition of Discrimination Bill, Bill 2871 (May 15, 2000).

The Prohibition of Discrimination Law is a sweeping law, but it expressly clarifies that it also applies to a wide range of financial services. Section 2 of the Law (the definitions section) defines "public service" to include "financial services", which are further defined in the that section as "banking services, extension of credit and insurance."

Alongside the prohibition on discrimination on the grounds stipulated in the Prohibition of Discrimination Law, exceptions to the prohibition are specified in Section 3(d) of the Law. Section 3(d)(1) of the Prohibition of Discrimination Law stipulates that discrimination is not deemed to exist "when this is made necessary by the nature or essence of the product, the public service or the public place." In other words, differentiation based on the grounds listed in Section 3 will not be considered discrimination when there is a **substantive reason** to discriminate pursuant to the nature or essence of the product.

In the realm of financial services, such as extending credit or underwriting insurance, the question of what constitutes permissible differentiation sheltered under this exception is a key issue that has been examined from time to time in court rulings.[274] In this context, considering certain personal characteristics of a customer (for example, age) may be essential for providing individually tailored service, for example, when assessing a customer's credit risk or determining an insurance premium—and may, therefore, not be considered prohibited discrimination. It should be emphasized that this report does not seek to take a position regarding the boundary between permissible differentiation and prohibited discrimination. This is a broad legal question that is not unique to cases of AI system deployment, and its boundaries continue to take shape on a case-by-case basis in judicial rulings.[275] At the same time, it should be noted that this issue may raise a particular challenge in relation to AI systems, as such systems may struggle to provide a clear, reasoned explanation for the system's choice to use a certain criterion.

A further challenge in applying the Prohibition of Discrimination Law to AI systems concerns how discrimination on one of the grounds listed in the Law can be assessed. Although the Law specifies a set of identity-based characteristics for which discrimination is prohibited, AI systems, may make it difficult to determine which criterion actually influenced the decision. This difficulty may be due to

---

[274] See, for example, HCJ 4948/03 **Elhanati v. Minister of Finance** (2008); Labor Appeal (National) 26328-09-14 **Harel Insurance Company Ltd. v. Ashbel** (3.12.2019) and also Class Action (Central) 8214-05-14 **Merav v. IDI Insurance Company Ltd.** (23.8.2018).

[275] For example, in the field of insurance, see CA 7308/19 **Tamir v. Bank Hapoalim Ltd**. (20.10.2022) (hereinafter – the Tamir case). Among other things, see the remark of Justice Y. Amit in par. 9 regarding consideration of the age criterion in collecting insurance premiums, and the remark of Justice O. Grosskopf (in the minority in the judgment) in paragraph 32, regarding a theoretical distinction in the pricing of life insurance policies between different population groups based on origin, for example.

the systems' complex mode of operation, the use of "proxy" variables, or due to the opacity of the "black box" phenomenon. These characteristics of AI systems may hinder the detection of prohibited criteria–, whether by customers seeking to determine whether they were discriminated against, by financial institutions seeking to comply with anti-discrimination requirements, or by regulators conducting oversight.

Alongside the sweeping application of the Prohibition of Discrimination Law, various financial statutes have also established specific provisions addressing discrimination in specific contexts. For example:

Part H of the **Equal Rights for Persons with Disabilities Law**, 5758-1998, sets out provisions addressing discrimination on the basis of disability in insurance contracts. These include, for example, prohibitions against refusing to insure a person with a disability, or providing terms so markedly different that, in effect, no insurance contract is concluded.[276] In this context, Section 19II(a) of the Law establishes an exception, under which differential treatment in an insurance contract will not be considered discriminatory "if it is based on actuarial data, statistical data, medical information, or any other information that is reliable and relevant to estimate certain insurance risk, and the treatment is reasonable under the circumstances."[277]

In the **Credit Data Law**, 5776–2016, one of the purposes of sharing credit data, enumerated in Section 1 of the Law, is "reducing of discrimination in the granting of credit and the reducing of economic gaps". In fact, one of the fundamental principles underlying the Law is that sharing credit data helps reduce reliance on discriminatory characteristics, which increasingly lose their importance as more personal information about the customer becomes available.[278] In addition, with respect to the activity of credit bureaus, Section 51 of the Law stipulates that "…a Credit Bureau shall not take into account data on the Customer's gender, age, sexual orientation, race, religion, country of origin, nationality, place of residence and family or health status" for the statistical model required for providing services

[276] See Commission for Equal Rights of Persons with Disabilities "Prohibition of Discrimination against a Person with Disabilities in an Insurance Contract - Information for the Public" (2.8.2018) for elaboration on the arrangement, https://www.gov.il/he/pages/nondiscrimination_in_insurance.

[277] It should be noted that the Equal Rights for Persons with Disabilities Law establishes a unique duty of reasoning in this context, regarding the reasons for the decision. Thus, in § 19le(g) it was determined that if "an insurer decided to give an insured person different treatment or an insurer refused to insure a person, the insurer shall deliver to him a reasoned notice in writing with a summary of the database, if the insurer relied on them in making the decision, and a summary of the information on which he relied...".

[278] In the explanatory notes to the Bill it was written that "in the absence of information regarding individuals, credit providers tend to use statistical data based on population characteristics, such as sex, religion and race. Sharing credit data allows reducing the use of these characteristics, as they lose their importance the more personal information exists regarding the client. In addition, collecting information regarding individuals will allow policymakers to examine, by comparing credit terms received by different segments of the population, whether credit discrimination actually exists, and to conduct control over the risk levels of the different borrowers. Without collecting systematic information on household debts, it is not possible to estimate the volume of credit received by different groups in the population".

based on credit data and credit worthiness assessments, for purposes of providing a credit worthiness assessment regarding a customer and for determining his credit score. Under Section 2(a) of the Banking (Service to Customer) Law, 5741-1981, banking corporations shall not "unreasonably refuse" to provide various services, such as opening a checking account under certain conditions or selling bank checks.[279] Although the Law does not expressly use the term "prohibition of discrimination", the restriction placed on "unreasonably" refusing may, in certain cases, form part of the relevant analytical framework.[280]

Discrimination in Comparative Perspective

In the international arena, concerns regarding discrimination and bias by AI systems have been addressed both within general regulatory frameworks and in sector-specific arrangements, including the financial services sector.

Within the framework of the EU AI Act, issues of bias are addressed primarily in the provisions dealing with **data and data governance**, which apply to AI systems classified under the Act as high-risk systems.[281] These provisions require that training, validation and testing the datasets of high-risk AI systems comply with AI governance and data governance practices. These practices include examining potential biases that could, *inter alia*, infringe on fundamental rights or lead to prohibited discrimination under EU law – particularly in cases where the data output feeds back into future inputs s (i.e., a feedback loop) – and taking appropriate measures to detect, prevent and minimize any identified biases.

In addition, with respect to the **databases** used for training, validation and testing, the EU AI Act stipulates, *inter alia*, that these must be relevant, sufficiently representative, and to the extent possible, also accurate and complete with respect to their intended purpose.[282] The Act further stipulates that the databases must possess appropriate statistical properties in relation to the group or individual for whom

[279] See in this context, Barak Medina "Prohibition of Discrimination in the Private Sector from the Perspective of Economic Theory" Alei Mishpat Vol. 3.

[280] See, for example, the Respondents' claim in the Tamir case, and the discussion in CC (Tel Aviv) 29979-08-14 **Renaissance School Ltd. v. Massad Bank Ltd.**, paragraph 35 of the judgment (2.8.2017).

[281] EU AI Act, Article 10 – Data and data governance.

[282] Article 10.3. Also, in recital 67 of the recitals to the Act the importance of databases was explained: High-quality data and access to high-quality data plays a vital role in providing structure and in ensuring the performance of many AI systems, especially when techniques involving the training of models are used, with a view to ensure that the high-risk AI system performs as intended and safely and it does not become a source of discrimination prohibited by Union law. High-quality data sets for training, validation and testing require the implementation of appropriate data governance and management practices. Data sets for training, validation and testing, including the labels, should be relevant, sufficiently representative, and to the best extent possible free of errors and complete in view of the intended purpose of the system.

the high-risk system is intended.[283] It also sets out rules governing the types of data that may be used to detect and correct bias in the system, including personal data, where its use is absolutely essential.[284] The issue of bias also received attention in the section of the Act dealing with accuracy, robustness and cybersecurity[285] – where it was stipulated that, as a general rule, high-risk AI systems must be designed and developed to achieve an appropriate level of accuracy, robustness and cybersecurity, and must maintain consistent performance in these respects throughout their entire lifecycle.

Although the EU AI Act provides a general legislative framework, in the preamble, it explicitly explains that AI systems used for **credit scoring assessments or vetting credit eligibility**, which are classified as high-risk systems, may generate discriminatory results. Such systems may differentiate between individuals or groups and perpetuate historical patterns of discrimination, including those based on racial or ethnic origin, gender, disability, age or sexual orientation, and may also create discriminatory effects.[286]

The 2021 OECD Report devoted extensive attention to the issue of bias and discrimination in financial services. Within the framework of the report, it was stipulated that policymakers may need to sharpen their existing toolkit for addressing risks arising from, or exacerbated by, the use of AI in several domains. With respect to discrimination and bias, these include the following recommendation:[287]

> "Promote practices that will help overcome the risk of unintended bias and discrimination. In addition to efforts around data quality, safeguards could be put in place to provide assurance about the robustness of the model when it comes to avoiding potential biases. Appropriate sense checking of model results against baseline datasets and other tests based on whether protected classes can be inferred from other attributes in the data are two examples of best practices to mitigate risks of discrimination. The validation of the appropriateness of variables used by the model could reduce a source of potential biases."

In the 2021 IOSCO Report, one of the recommendations pertaining to discrimination is that regulators consider requiring that companies maintain appropriate control mechanisms to ensure that the data on

---

[283] EU AI Act, Article 10.3.
[284] EU AI Act, Article 10.5.
[285] EU AI Act, Article 15.
[286] Para. 58 to the recitals: "AI systems used to evaluate the credit score or creditworthiness of natural persons should be classified as high-risk AI systems, since they determine those persons' access to financial resources or essential services such as housing, electricity, and telecommunication services. AI systems used for those purposes may lead to discrimination between persons or groups and may perpetuate historical patterns of discrimination, such as that based on racial or ethnic origins, gender, disabilities, age or sexual orientation, or may create new forms of **discriminatory impacts**.".
[287] 2021 OECD Report, pp. 59-60.

which their systems rely is of sufficient quality to minimize bias and is sufficiently broad for their robust implementation. The report further recommends that financial institutions ensure that both databases and system outputs are analyzed for discrimination risks. Additionally, companies must have appropriate control processes to identify and remove unjustified biases from the databases used. The report also recommends considering holding specific training courses for relevant personnel (data scientists, and/or other relevant staff) to raise awareness of the potential for system-induced bias.[288]

In the United States, the legal infrastructure in the field of discrimination in the financial sector is extensive, with several laws addressing discrimination specifically in financial contexts. In this framework, the FTC published a statement entitled "Aiming for Truth, Fairness, and Equity in Your Company's Use of AI",[289] clarifying that existing laws relating, *inter alia*, to equality and the prevention of discrimination also apply to developers and users of AI.

The statement refers to the provisions of the following laws as having significance for developers and users of AI systems: Section 5 of the Federal Trade Commission Act prohibiting deceptive or unfair practices; the Fair Credit Reporting Act (FCRA); and the arrangement anchored in the Equal Credit Opportunity Act (ECOA), under which a biased algorithm that generates discrimination in credit based on race, color, religion, etc., may not be used.[290]

Alongside a review of laws that may apply to discrimination by financial-services AI systems, the statement includes numerous recommendations regarding the avoidance of bias or other unfair outcomes. For example, regarding data, the statement recommends considering methods for improving datasets, and also proposes limiting the use of models in certain cases. In addition, it recommends remaining vigilant regarding potential discriminatory outcomes, by examining algorithms both prior to their deployment and periodically thereafter to ensure they do not discriminate based on protected characteristics. Furthermore, the statement recommends the promotion of transparency and independence by publishing the results of independent audits or making data or code available for external review.

---

[288] IOSCO Report 2021.

[289] https://privacysecurityacademy.com/wp-content/uploads/2021/04/Aiming-for-truth-fairness-and-equity-in-your-companys-use-of-AI.pdf

[290] **The Equal Credit Opportunity Act** (ECOA), enacted in 1974, lists a series of actions that constitute discrimination regarding credit transactions, based on race, color, religion, national origin, sex, marital status, age, and the fact that the applicant's income derives from a public assistance program. Alongside this, the law lists a series of actions that will not constitute discrimination, such as checking marital status or age, if the inquiry relates to various required aspects detailed in the law, and not in order to discriminate.

An additional FTC statement entitled *Using Artificial Intelligence and Algorithms*,[291] clarified, *inter alia*, that to detect illegal discrimination, the FTC examines the data fed into the model (inputs), including through proxies, as well as the results generated by the model. In this context, the focus is placed on whether the model generates discriminatory outcomes in practice and whether a model deemed neutral has nonetheless discriminatory implications for protected groups. If discriminatory outcomes are identified, the financial institution's justification for using this model is assessed, along with whether a less discriminatory alternative could have achieved the same results.

It should be noted that the emphasis placed on examining the **results** reached by the AI system in the context of financial law in the United States (as well as in the 2021 OECD Report) constitutes a certain response to the challenge described above regarding the difficulty in identifying an AI system's actual use of protected identity characteristics due to its complex mode of operation (such as the use of "proxy" variables or occurrence of the "black box" phenomenon). Under certain legal doctrines, such as indirect discrimination or disparate impact,[292] which are well-established in various areas of American law and have also been applied in various contexts in Israeli law,[293] the analysis centers on whether the outcomes of a practice result in *de facto* prohibited discrimination against protected groups.[294]

In this regard, see the question posed to the public in Appendix A concerning the effectiveness and feasibility of applying such a doctrine to AI in Israel's financial sector. In the United States, the CFPB issued a circular regarding credit requiring supervised entities to explain to customers the precise reasons for adverse decisions against them in accordance with ECOA, regardless of the technology employed in the decision-making. The circular further clarifies that complex algorithmic models (for

---

[291] Andrew Smith, Director, FTC Bureau of Consumer Protection, Using Artificial Intelligence and Algorithms (April 8, 2020).

[292] See in this context Talia B. Gillis, The Input Fallacy, 106 Minn. L. Rev. 1175 (2022), https://scholarship.law.columbia.edu/faculty_scholarship/3312. Gillis demonstrates how examination

[293] On the distinction between indirect discrimination and direct discrimination, see Talia Steiner**, Breaking Inequality – Coping with Discrimination against Arabs in the Israeli Labor Market**, Policy Study 97, The Israel Democracy Institute (2013), (https://www.idi.org.il/media/3434/pp97.pdf; and see CA 7308/19 **Tamir v. Bank Hapoalim Ltd.**., in paragraph 12 of the judgment of Justice Amit and the references there (hereinafter: "the Tamir case").

[294] In this context, it should be noted that the Prohibition of Discrimination Law lists several scenarios, upon the fulfillment of which the burden of proof is transferred to the defendant, who is required to prove that he did not act in violation of § 3 within the framework of civil proceedings. Thus, for example, in a case where the defendant refused to supply a product or public service to a person belonging to a suspect group, and did not refuse in the same circumstances to one who does not belong to such a group – he will have to prove that he did not violate the provisions of § 3.

example, those including a black box) may not be used if they prevent the entity from conveying the precise reasons for a credit denial or any other adverse action.[295]

Alongside these clarifications, the CFPB published a joint clarification document together with other federal agencies regarding enforcement efforts with respect to discrimination or bias in automated systems, in which it was clarified that existing legal authority also applies with respect to these systems.[296]

In the United Kingdom, as well, several reports and guidance documents addressing, *inter alia*, AI-generated discrimination and bias have been issued. Thus, the Information Commissioner's Office (ICO) released a guide, most recently updated in March 2023, providing extensive information on the legal requirements and challenges associated with AI systems, including bias and discrimination issues.[297] In September 2021, the Department for Digital, Culture, Media and Sport published a report entitled *Data: A New Direction*, which dealt extensively with AI fairness across three dimensions: actual discrimination, bias, and broader fairness considerations, such as procedural fairness and an individual's ability to understand or challenge decisions made with respect to them.[298]

Tools for Dealing with Discrimination

While it can be noted that some of the recommendations made by the various international organizations cited above outline principled frameworks for assessing and monitoring AI systems to address concerns of bias and discrimination, a range of entities and bodies are seeking to establish best practices for tackling this challenge, which may assist, *inter alia*, developers and implementers of AI systems.

For example, the National Institute of Standards and Technology in the United States (NIST) published an initial socio-technical framework addressing bias in AI, encompassing key conceptual foundations, an outline of the primary challenges, and directions for future guidelines. The document identifies core areas in which discrimination concerns arise and proposes corresponding guidelines: databases; test evaluation, validation, and verification (TEVV); and human factors (social and historical biases; a

---

[295] Consumer Financial Protection Circular 2022-03, https://www.consumerfinance.gov/compliance/circulars/circular-2022-03-adverse-action-notification-requirements-in-connection-with-credit-decisions-based-on-complex-algorithms, and see also Consumer Financial Protection Circular 2023-03, regarding the use of forms provided by the CFPB in this context.

[296] See *supra* note 72.

[297] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/.

[298] Data: a new direction - government response to consultation, https://www.gov.uk/government/consultations/data-a-new-direction.

human-centered approach; and human involvement).[299] As discussed below, NIST has since issued several additional guides on similar topics. Another example can be found in Singapore, where "AI Verify", a framework of principles and a technical toolkit for testing AI systems, was published. This framework allows organizations to independently test the performance of AI systems, to verify that they meet recognized international standards, including with respect to fairness and the avoidance of discrimination.[300] The toolkit is a downloadable software package that can be installed and used to generate test reports for AI systems.

Control tools (audit tools) are being developed worldwide—including those based on AI technologies to systematically and consistently examine whether AI systems exhibit bias or discrimination. Such technological tools can support the control processes of developers or financial institutions deploying these systems, as well as assist regulators in their supervisory activities. An accessible compilation of relevant tools and metrics is available on the OECD website. The organization has published a catalogue containing an array of resources designed to assist stakeholders in developing and implementing trustworthy AI systems. The catalogue consolidates a wide variety of technical, procedural and educational tools, including codes and software (*inter alia* for identifying bias), best practice guidelines, risk management tools, training programs and other educational materials.[301]

In the public comments on the Interim Report, one view emphasized the challenge of distinguishing between data that may be used legitimately and data that amounts to prohibited discrimination, highlighting the need for oversight and transparency as proposed in the report. It was also noted that the development of case law on this matter will likely lag behind AI development, making it essential for the regulator to play a leading role in shaping the public discourse. The Taskforce agrees with this approach, as expressed in its recommendations concerning the application of oversight and control measures throughout the system's life cycle, as well as in its call for regulatory action to guide supervised entities.

An alternative approach suggested in the public comments was that bias and discrimination risks are not unique to the financial system, and therefore no additional AI-specific regulation is required. According to this view, since the report does not clearly delineate neither the scope nor necessity of controls, its implementation could entail significant costs and other challenges for entities seeking to

---

[299] NIST Special Publication 1270 Towards a Standard for Identifying and Managing Bias in Artificial Intelligence, https://www.nist.gov/publications/towards-standard-identifying-and-managing-bias-artificial-intelligence.

[300] "Singapore's A.I. Verify builds trust through transparency", and see explanation on the OECD website https://aiverifyfoundation.sg, https://oecd.ai/en/wonk/singapore-ai-verify.

[301] OECD, Catalogue of Tools & Metrics for Trustworthy AI, https://oecd.ai/en/.

deploy AI systems. In the Taskforce's view, however, the report does set out the unique problems posed by AI with respect to bias and discrimination, such as a greater reliance on non-representative databases, the manner in which algorithms are designed, and the use of "proxy" data, which, in itself may not be discriminatory, but may be correlated with discriminatory variables. Accordingly, adopting additional measures as proposed in the Interim Report is warranted, where needed and pursuant to existing law.

Another comment concerned the fact that the legal framework itself is not always clear—for example in distinguishing between permissible differentiation and prohibited discrimination in the context of car insurance. The Taskforce agrees with this observation, and indeed considers there may be merit in addressing such basic questions to help clarifying the law.

Following the question raised by the Taskforce in the Interim Report regarding the effectiveness and feasibility of applying a doctrine of indirect discrimination or disparate impact, responses were received regarding the potential value such a doctrine may have for both financial institutions and their customers. Given the sweeping implications of adopting such a doctrine beyond the financial sector, the report does not address it directly but refers instead to outcome-based oversight as a response to challenges that may arise in identifying the improper use of a protected criterion (for example, due to the use of "proxy" variables or due to the "black box" issue in certain systems).

Bias and Discrimination – Recommendations:

1. Application of discrimination laws – The law on discrimination in Israel is established law, and with respect to private entities, the principle of equality and the prohibition of discrimination are regulated primarily in the Prohibition of Discrimination Law, which explicitly refers to financial services (banking, extension of credit and insurance) as public services, as well as in additional laws, as discussed above. **The actions and decisions of the parties are subject to the prohibition of discrimination irrespective of the technology they use, and therefore it is understood that this is also the case in the use of AI systems**.
2. Taking actions to direct supervised entities – The use of AI systems may generate inequality risks, *inter alia* because infringements of equality may be latent, unconscious or otherwise difficult to detect. As noted above, the operational features of AI systems may make it difficult to identify instances in which improper use is made of a protected characteristic, for example due to the use of "proxy" variables or the "black box" issue. Therefore, special caution and oversight is required in these areas when employing AI. Against this backdrop, and in line with practices abroad, it is

proposed that financial regulators consider **measures to guide supervised entities** and impose appropriate obligations within the scope of their authority—for example through directives, statements and clarifications (including following inquiries, public consultations, rulings and directives on these matters). This applies generally and is particularly relevant for financial operations that have a significant impact on individuals, such as credit scoring and certain types of insurance underwriting. The aim is to reduce the risk of discrimination (whether direct, indirect, intentional or unconscious), to promote a safe, fair and responsible environment for the deployment of AI systems, and to improve regulatory certainty. Key issues to be addressed under this framework include **oversight and measures** to identify and prevent bias throughout the system's entire life cycle (see the following recommendation), particularly with respect to the **outcomes** generated by the system.

3. Oversight and taking measures throughout the system's entire life cycle – Given the propensity of AI systems to exhibit discrimination and bias, enhanced, documented and transparent oversight is required throughout the system's entire life cycle, including during the implementation and operation stages, in order to mitigate the risk of bias and discrimination. It is recommended that supervised entities be directed and instructed to focus on the following key areas, to monitor on a regular basis, document their findings and to take corrective action when defects are discovered:

   a. Databases – It must be ensured that the databases used for training, validation and testing of the systems are representative and diverse databases, free from bias, containing relevant information about the populations on which the systems will make decisions.

   b. System mode of operation – The system's operations must be regularly reviewed, including during the implementation and deployment stages, to avoid reliance on protected characteristics that could result in discrimination, including through proxy variables.

   c. System results – Review of bias and discrimination in the model's results, for example by examining whether discriminatory treatment is given in practice to different populations with similar data.

4. Outcome-based control – Pursuant to the previous recommendation, because AI systems sometimes lack clear specific explainability, or there are concerns that the use of "proxy" variables will not be identified, it is proposed to adopt outcome-based control tools, *inter alia*, for examining claims of discrimination (or its absence). These tools can help both the supervised entity and the regulator detect discrimination in systems that may appear neutral an do not explicitly use

discriminatory data. This approach is not intended to replace the other control measures noted above, such as ensuring database representativeness or preventing the system from using discriminatory data, but rather to complement them.

5. Explainability – Explainability, and primarily specific explainability regarding a particular decision, is one possible tool for dealing with discrimination risk. It can verify that the system was not based on a protected identity characteristic in a discriminatory manner, and facilitate external oversight by the public, consumers or the regulator with respect to the system's activity (see the chapter dealing with "black box" and explainability).

Liability and Accountability

The question of liability for AI systems is a key issue. It encompasses not only the identity of who bears legal responsibility when harm is caused to a user, but also relates to the broader question of who is the accountable for an AI system. In other words, the debate extends beyond civil or criminal liability, incorporating considerations of who should bear responsibility in regulatory contexts.

It should be noted at the outset that this chapter will not address the application of general tort law in AI contexts. Rather, it seeks to focus on responsibility for compliance with the regulatory obligations imposed on AI systems, as well as the regulatory duties within the financial sector (in another chapter, we address the question of liability for discrimination). The rationale for this distinction stems, *inter alia*, from the fact that tort and criminal law apply across a wide range of AI-related activity and reflect considerations broader than those specific to the financial sector. The need to update and interpret these laws will certainly be discussed by the legislature and policymakers within the broader framework of AI regulation and will, undoubtably apply to the financial sector as well, falling outside the narrow scope of this report.

The issue of liability becomes acute in AI systems first and foremost due to the absence of human discretion, or at least the downgrading of the human role in the decision-making process. Traditional legal regulation relies on a human actor who is responsible for fulfilling obligations, while liability is assessed by examining that actor's conduct and discretion in cases of alleged violations. In the age of AI, the human role is diminished and at times is not clearly defined. In the absence of human involvement in part of or the entire decision-making process, the established legal framework for assigning liability, which has been historically premised on the presence of a natural person or corporation at the center of the event and accountable for their actions, is disrupted.[302]

The question of liability becomes acute given the automated and autonomous nature of AI systems, as well as the fact that multiple systems and factors may contribute to the output on which the public ultimately relies. For example, questions arise as to the behavioral standard expected of an AI system and whether knowledge or belief can be attributed to a computer system, among other issues.

These matters also find expression in financial contexts. For example, Section 2 of the Regulation of Investment Advice, Investment Marketing and Investment Portfolio Management Law, 5755-1995

[302] Policy document, p. 62

(hereinafter: the Investment Advice Law), stipulates that only a duly licensed person may engage in providing investment advice, investment marketing and portfolio management services to the public, while Sections 11 and 20 impose the fiduciary duties of trust and care (respectively) on licensees. Human characteristics, such as performing actions "in good faith and diligently" or the "the reasonable licensee" test are required for the fulfillment of these duties. A similar provision was established in the **Supervision of Financial Services (Provident Funds) Law**, 5765-2005 (hereinafter: "the Provident Fund Supervision Law"), regarding the engagement in providing pension advice and with respect to the duties of trust and care (Sections 15 and 16 of the Law).

Possible Approaches to Assigning Liability and Accountability

Despite the general consensus regarding the need to define the areas of responsibility of the relevant parties involved in the "life cycle" of AI systems, there different approaches regarding the manner of allocating responsibility among the various parties abound. We will briefly outline several key actors on whom responsibility may be placed and the justifications for doing so:

The algorithm developer (provider) – This refers to the system developer, who is responsible for writing the code or algorithm. AI systems are marked by a high degree of technological complexity. The system developer has the deepest understanding of the system's operation, its decision-making mechanisms, and its vulnerabilities, and is, therefore, viewed as the actor best positioned to assume responsibility for its functioning. In addition, the developer is best able to anticipate the key risks that may result from the system's use and mechanisms in advance and to mitigate or prevent their manifestation.

However, cases can be cited in which the developer's connection to the system is diminished, making the imposition of liability on the developer problematic. For example, the more autonomous a system is and the more it relies on self-learning capabilities, the less able the developer is to anticipate the risks associated with its use. A machine learning-based AI system developed for stock trading may adopt illegal trading techniques without the developer having any real ability to anticipate this. Similarly, the broader and more general the system's applications are, the harder it is for the developer to anticipate emerging risks in specific use cases, including cases where the system was adjusted for a specific use. In addition, system developers may be unable to address problematic content fed into the AI system. These systems operate largely on the basis of the data on which they are trained, thus the representativeness, quality and reliability of the data are critically important. Where the developer does not control the data feeding the system (for example, when users provide additional data), it becomes difficult to impose responsibility on the developer, given his limited ability to anticipate, prevent or

mitigate damages. In such circumstances the developer is no longer the most efficient party to control the system.

The system operator (deployer) – This refers to an entity deploying AI its business, whether for the corporation's internal needs or for customer service. The business entity is not merely a passive party providing access to an AI-based system; rather, it actively chooses to use AI, selects the specific system itself, determines the nature of its use, and often even makes (or requests) modifications to tailor it to its needs. When a business entity deploys an AI-based system to perform tasks or make decisions, including providing decision-making support (particularly when changes were made to the system during corporate implementation), the system operator can be considered the party best positioned to control the system and anticipate potential harm and to act to prevent or reduce their likelihood. Additionally, imposing liability on system operators may create incentives for them to ensure proper implementation, use and oversight of AI systems, including conducting sufficient due diligence, maintaining monitoring and supervision processes, and upholding good corporate governance.

At the same time, in cases where system operators use AI systems without making any changes to them, certainly not any material changes, and particularly where the systems are not fed from databases belonging to those business entities, it can be argued that in such cases, the system operators are not the party best positioned to prevent damage. They too may claim that they cannot be responsible for the actions taken by either developers or end users.

The end user – Under certain circumstances, the end user can be assigned responsibility for damages incurred by him. This is especially true in cases where the end user has "contributory fault" in causing the damage caused to him or in cases of irresponsible system use, acting in bad faith, concealing or withholding vital information, and similar conduct.[303] In addition, one might envision a legal doctrine whereby, by virtue of his awareness of the risks associated with using AI systems, the end user is deemed to assume those risks as well as their consequences, for better or worse, in a manner analogous to the “user beware” principle, paralleling the “buyer beware” doctrine.

This approach is not without its difficulties, since the ability of an end user to understand the full risks of the system, to ask its developers questions, to understand its technological complexity and, if necessary, to conduct tests on it is very limited, if not impossible or impractical. In most cases, the end user will not have access to the database feeding the system. Adopting a doctrine which imposes

[303] Tel Aviv University Report, pp. 103-104.

liability on the end user may inhibit the entry of new technologies to the market and stifle innovation. In addition, imposing responsibility on the end user may lead to the introduction of AI systems to the market in an irresponsible, inappropriate and unfair manner, when precisely those who developed or implemented these systems in the market are not held accountable for this.

Human "closest to the center of occurrence" – Another approach is to assign liability to a human party based on the extent of their involvement in the decision-making process, with the assessment varying from case to case. However, there are considerable difficulties involved in holding the human party who made the decision in practice, or who is "closest" to the point of occurrence. First, identifying the human party "closest to the center of occurrence" may be complex, and the absence of clear criteria may give rise to uncertainty. Second, even if such a party can be identified, questions arise as to their ability to control the system's recommendations, to reject them or to intervene in its operation. This approach also fails to address situations where AI-based systems make decisions entirely automatically and autonomously, without any human involvement. Moreover, it generates significant uncertainty as to who bears responsibility and may create a "chilling effect" on the deployment of such systems into the financial entity's operations.

Separate legal entity – Another alternative is to recognize AI-based systems as independent entities, separate from the parties that develop, or deploy them. This option requires use of a legal framework that may raise logical and practical challenges and currently lacks foundation.[304] Attempts worldwide to recognize AI systems as separate entities for the purpose of imposing tort liability have so far come to naught.[305]

Additional parties on whom responsibility for AI systems can be imposed are other parties involved in the various stages of AI system operations. For example, responsibility can be imposed on the importer of the system or its distributors if they did not perform the tests required of them prior to the system's market entry or its distribution to users. It can also be assigned to third parties who had contact or involvement with the systems. Another approach would remove responsibility from each one of the actors involved in AI systems. At this stage, and in light of the discussion below, there is no need to consider every one of these options in depth.

[304] Policy document, pp. 65-66.
[305] https://www.washingtonpost.com/travel/2024/02/18/air-canada-airline-chatbot-ruling/?force_isolation=true.

Liability and Accountability for AI Systems – International Survey

The European Union's approach,[306] as reflected in the EU AI Act, establishes a primary liability regime that applies to AI system developers when these systems are classified as high-risk. While the law uses the term "provider", its definition of this party deals primarily with party responsible for developing the AI system.[307] Article 8.2 of the Act outlines requirements applicable to High-Risk AI systems:

> "Where a product contains an AI system, to which the requirements of this Regulation as well as requirements of the Union harmonisation legislation listed in Section A of Annex I apply, providers shall be responsible for ensuring that their product is fully compliant with all applicable requirements under applicable Union harmonisation legislation."

Section 2 of Chapter V of the Act impose comprehensive obligations on the "providers" (developers) of general-purpose AI systems.

The European law distinguishes between obligations that apply to actors involved in AI systems prior to their release to the market, and obligations applying to actors involved after the systems are deployed in the market.

Article 16 of the European law details the obligations applying to the providers of high-risk AI systems, *inter alia*, after they are already operating in the market. These include obligations to ensure compliance with relevant laws and requirements to conduct quality control of the system, maintain orderly documentation, perform conformity assessments to check compliance with the law, and to register themselves and the system in a dedicated EU database.

Alongside the primary approach, under which responsibility for AI-based systems is assigned to developers, certain other parties involved with AI systems (distributors, importers, operators and third parties) may, under certain circumstances, also be considered system developers. This includes cases

---

[306] In this chapter we will refer to the definition of responsibility regarding the fulfillment of regulatory obligations included in the EU AI Act. In tort contexts, the European Union promotes a separate directive intended to harmonize the laws applicable in the Union member states – https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0496.

[307] Definition of provider in EU AI Act: 'provider' means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places them on the market or puts the system into service under its own name or trademark, whether for payment or free of charge;

where they made material modifications to high-risk AI systems (Article 25 of the EU AI Act). Article 26 of the Act imposes distinct obligations on system deployers, covering areas such as human oversight, the vetting of the input data feeding the system (to the extent that the deployer exercises control over it). It is important to note that this division of responsibility described above refers exclusively to responsibility for regulatory obligations arising from the specific legislation and does not refer to general tort law. As mentioned above, no similar sweeping AI legislation has been established in the United States, as of the writing of these lines, nor have federal laws addressing the liability issue been enacted. However, assigning liability to various actors has been proposed within the framework of individual state laws.

In April 2021, the FTC issued a statement providing guidelines for assessing the use of AI systems under the existing legal framework, emphasizing that the obligations currently imposed on business entities will continue to apply when providing services using AI-based systems.[308]

In addition, in 2023, a group of U.S. regulators, including the FTC, issued a joint statement emphasizing that their enforcement authority extends to "automated systems", which encompasses AI systems, and that existing regulation applies to these systems to the same extent as they apply to other technologies. The statement also asserted that oversight covers both the development and the deployment of automated systems, while promoting responsible innovation. Furthermore, the signatory agencies affirmed their intention to exercise their authority to protect individual rights regardless of whether violations arise from using traditional methods or advanced technologies.[309] In a policy paper published by the British government relating to the use of generative AI by government entities, it was stipulated that the responsibility for any output of an AI-based system will be imposed on the organization that chose to implement it:

Ultimately, responsibility for any output or decision made or supported by an AI system always rests with the public organisation. Where generative AI is bought commercially, ensure that vendors understand their responsibilities and liabilities, put the required risk mitigations in place and share all relevant information.[310]

---

[308] See *supra* note 266. For example, the FTC Act also applies to companies producing consumer reports, and including for example, companies producing reports for the purpose of credit scoring for loans. In the published statement it was clarified that the FTC Act prohibits discrimination against credit applicants (the law applies in retail contexts, financial companies and credit companies that are not supervised by another authority). The statement determines that when the use of an algorithm leads to a discriminatory result, this will constitute a violation of the Law.

[309] See supra note 72.

[310] HM Government "Generative AI framework for HM Government V.1", https://www.gov.uk/government/publications/generative-ai-framework-for-hmg.

Liability and Accountability in the Financial Sector in Israel

The starting point in regulation for Israel's financial sector is that as a rule, responsibility rests with the supervised financial institution under the regulatory framework applicable to it. This applies regardless of whether the institution developed the service independently or whether it relies on outsourcing or third-party providers. Regulation in this context focuses on risk management within financial institutions and includes provisions regarding legal liability arising, *inter alia*, from the relationships and duties owed by the financial entity to its customers (including duty of care, heightened trust obligations and similar fiduciary duties). This approach is also reflected in Israeli court rulings. Courts have repeatedly emphasized the significant power imbalance between financial institutions and their customers, as well as the customer's trust and reliance on the institution's advice. However, the liability imposed on financial entities toward their customers is neither absolute nor unilateral, and in some cases, customers also bear obligations toward the financial entity, which may impact the scope of the entity's obligations towards them. This approach is reflected in judicial rulings. Courts have consistently highlighted the significant power imbalance between financial institutions and their customers, as well as the customer's reliance on the institution's advice.[311]

Provisions governing the liability of financial entities appear across the frameworks of all financial regulators. For example:

Israel Securities Authority - The Directive to Licensees Regarding Provision of Services Using Technological Means, stipulates that investment advice, investment marketing or portfolio management licensees bear liability toward their customers for damage resulting from the use of technological means in the provision of its services to the customer. This provision relates to both the stage of onboarding and entering into an agreement, including clarifying the customer's needs, determining investment policy and investment strategy and to the later stage of providing ongoing service. In this context, the Directive clarifies that conducting this process through technological means is a risk that the financial service provider must assume. Accordingly, the provider is responsible for

[311] For example, in case law, duties of a customer towards the bank have been recognized, and in a number of cases, a customer's non-compliance with his duties was recognized as grounds for reducing the financial entity's liability. Also, in the framework of insurance laws for example, it is accepted that a duty of care applies to an insured to behave in the same manner as if he were not insured as a condition for the applicability of the insurance, and a broad duty of disclosure applies towards the insurer; Yaron Elias, Insurance Law (Second Edition, 2009), p. 331 et seq. and p. 449 et seq.

any resulting damage, and he cannot direct the customer to seek recourse from a third party involved in delivering the service.[312]

Capital Market, Insurance and Savings Authority - In the explanatory notes to Section 2 of the Authority's circular on "Outsourcing in Institutional Bodies," which applies to insurance companies and provident funds, it is stated that:

"…the legal provisions applying to an institutional body define, *inter alia*, the obligations and liability imposed on it and on its officeholders. This provision seeks to clarify that transferring activities to a service provider through outsourcing does not allow the institutional body or its officeholders to transfer or delegate these obligations and liability established by law."

This principle is also reflected in additional sections of the circular: Sections 6(b)(2)(c), 6(b)(2)(f) and 8, which seek to clarify and ensure that the institutional body remains responsible and liable for any activity outsourced to a service provider and must act to ensure that the service provider meets the requirements imposed on the institutional body.[313]

Banking Supervision **Department** - In the *Proper Conduct of Banking Business Directive* No. A359 entitled "Outsourcing", Section 6 stipulates that outsourcing activity does not derogate from the banking corporation's responsibility for complying with all laws and directives that apply to it. The provision includes a series of obligations imposed on a banking corporation in the field of corporate governance.[314]

Accordingly, in various areas of financial regulation, the prevailing principle is that obligations relating to a service or financial product offered to the public are imposed on the supervised

---

[312] Directive for License Holders regarding the Provision of Services Using Technological Means, Directive under § 28(b) of the Regulation of Investment Advice, Investment Marketing and Investment Portfolio Management Law, 1995 (Israel); For example, also regarding members of the Stock Exchange: In Chapter D1 of the First Part of the Stock Exchange Rules, duties of Stock Exchange members who are not banks (Non-Banking Members - NBMs) are established regarding receiving outsourcing services within the framework of information technology management. § 24jd(h) of the First Part of the Stock Exchange Rules states that "receipt of outsourcing services by the NBM does not derogate from the NBM's duty to comply with all the instructions under the Rules and/or the provisions of any law, and does not derogate from the NBM's responsibility that the services provided to him via outsourcing, meet the conditions established under the Rules".

[313] Circular dated 3.6.2018, amended on 31.12.2018.

[314] Such as: In § 10 of the Proper Conduct of Banking Business Directive it was determined that "the Board of Directors bears the overall responsibility for outsourcing and must ensure that outsourcing an activity does not derogate from the fulfillment of the banking corporation's duties and responsibility to comply with any provision of law including the instructions of the Supervisor of Banks and in particular: (a) its responsibility towards its clients, including transparency and proper disclosure, providing telephone support for questions and clarifications and handling complaints; (b) its responsibility to ensure that the ability of the Supervisor of Banks to receive information and exercise his powers will not be impaired" and internal audit duties. In addition, the banking corporation is required to conduct due diligence on the service provider in outsourcing, and specific instructions were established regarding the manner of engagement with the service provider via an outsourcing contract. The directive also established duties regarding risk management and reporting duties to the Supervisor of Banks.

financial entity, and liability for that service or product rests with it as well. In terms of the alternatives presented above, this reflects the model in which liability is assigned to the system deployer. The involvement of an external party or system does not remove statutory liability towards customers from the financial entity. While it may contractually provide for indemnification from third parties in cases where damage occurs (such as harm resulting from a malfunction or the failure of computer systems purchased from a third party), its legal liability toward customers remains intact.

Liability and Accountability in the Financial Sector – International Review

As noted in the introduction, in recent years, international organizations involved in shaping economic and financial regulation have issued reports pertaining to AI, which also address the question of liability. The 2021 OECD Report stipulates that responsibility is placed on the shoulders of financial institutions' boards of directors and management:

> "Ultimate responsibility, oversight and accountability over AI-based systems lies by definition with executive and board-level of management of the financial services provider, who have to establish an organization-wide approach to model risk management, and ensure that the level of model risk is within their tolerance."[315]

Similar statements are echoed in the 2023 OECD Report:

> "In alignment with their fiduciary duties, financial service providers have the legal responsibility to act in the best interest of the clients, which also means protecting them from risks of deceptive outputs, misinformation or other risks to financial consumers related to GenAI tools in particular."[316]

It is further noted that in the context of AI-based systems, which are not necessarily generative AI:

> "Similar to any type of model, the use of AI-based models by financial market participants would need to comply with their existing model governance frameworks and oversight arrangements. This includes clear lines of responsibility for the development and overseeing of AI-based systems throughout their lifecycle, from development to deployment, and explicit designation of accountability for any adverse outcome produced by the model."[317]

[315] 2021 OECD Report, p. 50.
[316] 2023 OECD Report, p. 10.
[317] 2023 OECD Report, p. 24.

The 2021 IOSCO Report proposes a series of measures for authorities and policymakers to adopt with respect to AI adoption. The first proposed measure refers to imposing liability on officeholders serving in financial intermediaries and asset managers using AI-based systems:

> "Measure 1: Regulators should consider requiring firms to have designated senior management responsible for the oversight of the development, testing, deployment, monitoring and controls of AI and ML. This includes a documented internal governance framework, with clear lines of accountability.[318]"

Similar recommendations can also be found in the reports of various national financial regulators. For example, in a joint report issued by the British financial regulators:

> "The supervisory authorities' existing rules and guidance, in particular, those implementing the SM&CR, emphasise senior management accountability and responsibility and are relevant to the use of AI. Currently, technology systems are the responsibility of the SMF24 (Chief Operations function ...). [...] PRA-authorised SM&CR banking and insurance firms and FCA-authorised enhanced scope SM&CR firms must ensure that one or more of their SMF managers have overall responsibility for each of the activities, business areas, and management functions of the firm. That means any use of AI in relation to an activity, business area, or management function of firm would fall within the scope of a SMF manager's responsibilities. For banks, the General Organisational Requirements (part 2.1) states that 'firms must have robust governance arrangements, which include a clear organisational structure with well defined, transparent, and consistent lines of responsibility and effective control and safeguard arrangements for information processing systems."[319]

The Singaporean regulator also referred to imposing liability on the financial entity deploying the AI-based systems, regardless of whether that system was developed by the financial entity or through a third-party supplier:

> "It is important that there is clear responsibility for and ownership of AIDA-driven decisions within an AIDA firm, with appropriate internal approving authorities for the use of AIDA. Such accountability applies to all uses of AIDA, whether internally developed or externally sourced."[320]

---

[318] IOSCO Report 2021, p. 17.
[319] See *supra* note 74.
[320] See *supra* note 76.

Recommendations with respect to Liability and Accountability

As stated above, the recommendations of this part refer to the question of liability for the regulatory obligations applicable to financial institutions in cases when these entities are assisted by or make use of AI systems. At present, there are no specific obligations or regulations governing the development and deployment of AI systems (as, for example, under the EU AI Act). However, should such obligations be established in the future, these recommendations may need to be reconsidered in light of the nature of obligations that will be imposed in the context of AI system development and use.

In the public comments submitted on the Interim Report, some expressed support for the approach presented, while others claimed that the report does not provide a clear answer regarding legal liability for damage resulting from the use of AI systems, and emphasized the need to be aligned with internationally accepted liability regimes. As noted above, the Taskforce's principled position, as articulated in the Interim Report, is that there should be no deviation from the existing framework, under which supervised financial entities bear liability for regulatory obligations. This is the established approach in Israel's financial sector, and the Taskforce sees no reason to change it when various processes are carried out using AI systems. This applies, *inter alia*, to outsourcing, the use of systems developed by third parties, and to contractual arrangements between financial entities and their service providers

The Taskforce also understands that this is also the prevailing approach abroad. Additional proposals raised in the public comments, such as introducing criteria for imposing liability, defining the duties of care and oversight for each party involved in the system's development and use, or addressing cases where external suppliers deviate from the supervised entity's instructions, are considered unnecessary at this stage. It is not feasible to anticipate every possible circumstance in engagements between system developers and operators, and similar issues already arise under the existing liability regime regarding the relationships between supervised entities and third-party service providers.

Another comment addressed the recommendation to prohibit limitations on liability, arguing that such a prohibition would impose an excessively broad scope of financial institution liability. In this regard, it should be emphasized that the restriction on limiting liability is a direct extension of the existing liability framework applicable to supervised entities and is already embedded for other areas of their activity. For example, Section 13(e) of the Advice Law, provides that any clause in an agreement that exempts a licensee from liability imposed on it under that law or any other law, whether concerning how it fulfills its role or limiting its liability—is null and void.

Nothing in this context is intended to preclude a supervised entity from providing clarifications regarding reasonable limitations of an AI system. For example, when using an AI system that analyzes publicly traded companies, the entity may inform customers that the system does not incorporate information published after the companies' most recent financial statements were file, just as a comparable disclosure could be made in relation to a system not based on AI. Also, nothing in this context is intended to harm the establishment of specific arrangements between the financial entity and third parties with respect to the AI systems (suppliers, developers, etc.), provided that the financial entity's liability in the relationship between it and the customer is not limited.

In light of the above, and in order to ensure alignment with the liability framework established in existing regulation, the Taskforce recommends the following:

Liability and Accountability – Recommendations:

1. Maintaining the principle of supervised-entity liability

The prevailing approach in Israel's financial sector is that regulatory obligations apply to the financial entities supervised by the relevant regulator— regardless of whether the service is provided directly by the entity itself or through outsourcing or any other reliance on third-party providers. Consequently, legal liability is also imposed on the supervised entity. We see no reason to deviate from this principle, whether from a normative perspective or a practical one. **The adoption of AI-based systems by financial institutions should not diminish their regulatory obligations and or their liability toward customers. These obligations are key for safeguarding the public interest and ensuring that the public has a clear point of contact to turn to**. Moreover, the rationale imposing liability on financial institutions remains unchanged with the introduction of AI-based systems. An end customer who suffers damages as a result of his dealings with a financial entity cannot be expected to be required to pursue a third party to seek redress. Customers are at a disadvantage in terms of knowledge and access vis-à-vis the financial entities, which are subject to a range of regulatory requirements, including those related to technical matters, decision-making processes and AI governance. Finally, the financial entity continues to serve as the most effective point of deterrence, since lability in this field is still subject to the regulatory obligations applicable to the entity.

2. Outsourcing

The use of outsourcing by financial institutions is a well-established and regulated practice (for example, acquiring software rather than developing it in-house). Outsourcing offers significant advantages in conducting financial activity, and is, in practice essential for most financial institutions. Financial regulators have long recognized that many activities may be outsourced, while making clear that outsourcing does not relieve financial institutions of their responsibility to meet all the regulatory obligations imposed on them under law. **In this context, we propose viewing a financial entity's engagement with a third party for AI-based systems as a form of outsourcing, and we recommend maintaining the existing regulatory approach to outsourcing, without introducing AI-specific deviations**.

3. Maintaining liability rules in the financial institutions themselves

In certain circumstances, financial regulation imposes specific liability on organs and position holders within the financial entity, while overall responsibility for the actions of the supervised entity rests with the board of directors. In keeping with the approach presented above, and with regulatory

recommendations in other jurisdictions, we propose that these liability rules should continue to apply with respect to AI systems both in terms of the systems' integration and their operation in the financial entity, including the implementation of appropriate risk-management policies. This does not preclude the possibility that additional AI-specific obligations and liabilities may also apply (as discussed in the AI governance chapter below).

4. Developer liability as a consideration for operation of AI systems **in the future**

As noted above, the regulation of AI systems is developing worldwide. The most prominent to date is the EU AI Act, which imposes liability and compliance obligations on AI system providers. These obligations and the attendant liability may also be relevant in the context of financial supervision.

In light of the above, it may be appropriate in the future to consider the regulatory framework applicable to AI developers, including the liability imposed on them through regulation, as a factor or potential condition for the operation of AI systems. For example, regulations can stipulate that only AI-based systems subject to recognized foreign regulation and proof of compliance may be deployed in the Israeli market. Alternatively, compliance with such regulation could constitute a consideration when assessing the liability of supervised entities in Israel in cases where injury results from the use of these systems.

5. Prohibition of limiting liability toward the customer

It is proposed that a financial entity be prohibited from disclaiming or limiting its liability regarding services delivered to customers through AI systems. For example, issuing a disclaimer in relation to AI-based services would neither absolve nor reduce the financial entity's liability toward its customers. This does not preclude customer liability where, for example the customer knowingly takes an investment risk that materialized in the course of an AI-based service, or where he engaged in malicious use or conduct that exceeds the reasonable use of the system. Nor is this intended to prevent the supervised entity from providing clarifications regarding reasonable intrinsic limitations of the AI system.

6. Establishing AI governance provisions

As noted at the beginning of this chapter, the issue of liability also relates to AI governance, which is designed, *inter alia*, to set out detailed arrangements for implementing the financial entity's responsibilities. For further discussion, see the chapter on AI governance below.

AI Governance

In contrast to the challenges described thus far, AI governance may be viewed as a framework designed to reduce the challenges and mitigate the risks associated with AI use. Its purpose is to establish appropriate management and control mechanisms within an organization, which determine and regulate decision-making, risk management, internal control and oversight of AI system operations. Principles-based governance serves as a means of corporate risk management and supports the organization's overall resilience and financial stability. In essence, AI governance is intended to "translate" the risks and challenges inherent in AI systems use into methods, processes and practical measures that enable responsible and fair use of these systems, the input data feeding them and the outputs they generate. There is currently no universally accepted definition of AI governance, although the academic literature on the subject, which is expanding alongside the broader engagement with AI, offers various formulations.

Nonetheless, most attempts to define AI governance emphasize the need for an extensive set of tools, practices, processes, policies and standards that support the responsible and appropriate development, management and use of AI systems and applications.[321] A comprehensive AI governance framework spans all stages of an AI system's life cycle, starting from the initial decision to adopt AI, through algorithm development, system integration in the organization, system use, operation, oversight and control, as well as the collection and management of the data feeding the system and beyond. International organizations have justifiably emphasized the importance of establishing governance rules for entities and organizations that use AI systems. For example, Chapter VII of the EU AI Act is devoted entirely to AI governance in the European Union. Among other measures, it provides for the establishment of a European Artificial Intelligence Board to advise the European Commission on matters within its mandate, the formation of an advisory forum to provide technical expertise, and the creation of a scientific panel of independent experts to advise the Commission on implementation, enforcement, robustness, development, monitoring, etc. The Act also establishes cooperation mechanisms among the competent authorities, and mandates the creation of a database containing information on high-risk AI systems.

The OECD together with the G7 organization, have also emphasized the importance of AI governance rules:

[321] Mäntymäki, M., Minkkinen, M., Birkstedt, T. et al. Defining organizational AI governance. AI Ethics 2, 603–609 (2022). https://doi.org/10.1007/s43681-022-00143-x.

> "The 'responsible' use of generative AI technologies was also viewed as the most 'important' priority for policy, followed by governance and by addressing disinformation."[322]

The G7 Hiroshima Summit (which dealt with the global response to risks associated with AI, and particularly with generative AI called, *inter alia*, on all stakeholders and interested parties to develop and implement AI governance rules:

> "Develop, implement and disclose AI governance and risk management policies, grounded in a risk-based approach – including privacy policies, and mitigation measures, in particular for organizations developing advanced AI systems."[323]

In the UK, the government published an AI strategy paper. Principles of proper and appropriate governance were articulated in this report, while in a statement, the British government emphasized their importance:

> "Governance measures should be in place to ensure effective oversight of the supply and use of AI systems, with clear lines of accountability established across the AI life cycle. AI life cycle actors should take steps to consider, incorporate and adhere to the principles and introduce measures necessary for the effective implementation of the principles at all stages of the AI life cycle. Regulators will need to look for ways to ensure that clear expectations for regulatory compliance and good practice are placed on appropriate actors in the AI supply chain, and may need to encourage the use of governance procedures that reliably ensure these expectations are met."[324]

In the United States, the White House Office of Science and Technology Policy (OSTP) issued its *Blueprint for an AI Bill of Rights*, relating to the importance of governance rules for entities using and entities developing these tools:

> "Entities responsible for the development or use of automated systems should lay out clear governance structures and procedures. This includes clearly-stated governance procedures before deploying the system, as well as responsibility of specific individuals or entities to oversee ongoing assessment and mitigation."[325]

---

[322] OECD (2023), G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI, OECD Publishing, Paris, https://doi.org/10.1787/bf3c0c60-en.

[323] Hiroshima Process International Guiding Principles for All AI Actors, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document03_en.pdf?force_isolation=true.

[324] A pro-innovation approach to AI regulation (August 2023), https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper#executive-summary.

[325] See *supra* note 24.

In addition, the Blueprint also addresses the need to examine systems prior to their deployment, identify risks associated with using AI systems and mitigate them, monitor system activity on a regular basis, provide oversight and more. Although the Blueprint does not focus on the financial sector, it appears that its principles apply to all industries and areas of life in which AI systems operate.

AI Governance in the Financial Sector

The EU AI Act refers to specifically to AI governance in financial entities. See, for example, Recital 158:

> "Union financial services law includes internal governance and risk-management rules and requirements which are applicable to regulated financial institutions in the course of provision of those services, including when they make use of AI systems…"

The EU AI Act also establishes provisions regarding operating quality control systems,[326] maintaining documentation,[327] and conducting both prospective and retrospective monitoring of AI systems,[328] among other requirements.

Furthermore, Article 17 of the EU AI Act stipulates that providers of AI systems that are financial institutions, will be deemed to satisfy certain requirements for maintaining quality control systems if they comply with internal governance arrangements under Union financial legislation. A similar provision included in Article 26(6) of the Act with respect to deployers that are financial institutions, regarding the implementation of technical and organizational measures to ensure compliance with the requirements associated with the AI system.

In the 2021 IOSCO Report, which, as mentioned above, focuses on AI and machine learning applications for financial intermediaries and asset managers, governance and supervision have been identified as requiring particular attention due to the increasing use of AI systems in financial operations. All measures proposed by IOSCO to address this risk are, in one way or another, tied to AI governance principles. The report underscores the need for financial entities to adapt their governance principles to the age of AI. In fact, IOSCO calls on national regulators to require supervised financial entities to operate in accordance with specific guidelines that, while not

[326] EU AI Act, Article 17(3).
[327] For example, Articles 18 and 19 of the EU AI Act.
[328] For example, Article 72 of the EU AI Act.

necessarily at the core of "conventional" governance frameworks, are expected to be embedded within entities through appropriate governance rules: [329]

(1) Management accountability: to appoint a party from senior management who will be responsible for overseeing, developing, testing, using and monitoring AI policies and practices. The appointed party must possess technological knowledge and skills.

(2) Prior examination and supervision: to perform tests of the algorithm and validate AI system results prior to the system's deployment. Once AI systems are deployed, supervised entities must be required to perform ongoing monitoring of the systems, noting that the algorithm may change over time as a result of its self-learning capabilities and autonomy.

(3) Skills and expertise: to employ professionals with skills, knowledge, expertise and experience in developing, testing, operating, monitoring and supervising AI systems. In addition, persons engaged in compliance and risk-management functions must be sufficiently skilled to examine AI applications developed by third parties that the financial entity seeks to use.

(4) Regulating relationships with third parties: to enter into dedicated agreements with third parties providing the organization with AI systems.

(5) Data quality: to operate control and oversight measures regarding the databases underlying AI systems, to ensure that the information and data are of sufficient quality and representativeness.

(6) Transparency: to establish disclosure obligations toward customers regarding the use of algorithms.

The OECD also focused on the governance issue in the its report on AI in finance. The report discussed the importance of governance rules particularly regarding material decisions that may have a significant impact on the individual (for example in the field of credit or portfolio management), and clarified that the impact on consumers should be reflected in the governance principles:[330]

> "Solid governance arrangements and clear accountability mechanisms are fundamentally important as AI models are deployed in high-value decision-making use-cases (e.g. in determining who gets access to credit or how investment portfolio allocation is decided) [...] Internal governance frameworks could include minimum standards or best practice guidelines and approaches for the implementation of such

[329] IOSCO Report 2021.
[330] 2021 OECD Report, p. 49.

guidelines [...] Importantly, intended outcomes for consumers would need to be incorporated in any governance framework, together with an assessment of whether and how such outcomes are reached using AI technologies."

Similar statements in this spirit were also echoed in the 2023 OECD Report on generative AI:[331]

"Strengthen model governance and promote accountability mechanisms: Currently applicable frameworks for model governance in finance may need to be enhanced or adjusted to address incremental risks emerging from advances in AI. Solid governance arrangements and clear accountability mechanisms are fundamental in AI models deployed in high-value use-cases (e.g. in determining access to credit or investment advice)."

The 2021 OECD Report also addresses governance issues, such as the monitoring and supervision of AI systems, outsourcing and third-party suppliers, and the assessment and verification of systems and their output, and more.

In a comprehensive report, the British financial regulators (FCA and PRA) examined the principles for using AI in financial entities, emphasizing the critical role of governance in ensuring the safe and responsible adoption of the technology. In this context, the regulators discussed numerous aspects of sound corporate governance, including organizational structure, transparency and clarity in accountability, risk management, board composition, internal audit functions, and related matters.[332]

Corporate Governance

The broad concept of "AI governance" can also encompass incorporate corporate governance principles, which refer to the system of laws, rules, practices, processes and controls through which organizations are directed and managed. A key focus of corporate governance is the conduct of the board of directors and its committees, although it also addresses other organizational components and the relationships between the company, its shareholders and other stakeholders. Effective corporate governance takes into account the interests of all organization stakeholders, including shareholders, senior management, employees, customers, suppliers, creditors, regulators and the wider community in which the organization operates. The underlying concept is that sound corporate governance

[331] 2021 OECD Report, p. 30.
[332] See supra note 74.

provides the infrastructure for the organization to operate responsibly and appropriately, starting with the board of directors, through senior management and down to junior employees.

As part of AI governance, one measure to address the intrinsic challenges of AI systems is the adoption and implementation of corporate governance rules, adapted and clarified to meet the needs of the AI age.[333] To date, there are no widely recognized or established corporate governance standards set specifically for AI activity within financial entities, which generally rely on existing corporate governance frameworks. This approach is understandable given the principle that technological innovations should be treated neutrally, without the need to treat each new technology differently from its predecessor, and given the currently limited scope of AI-system deployment.[334] This approach is expected to change, however, as AI-system deployment expands, and given the risks and challenges (or exacerbation of existing risks) unique to AI systems, which did not exist with respect to previous technological innovations or were not necessarily reflected in existing corporate governance rules. These risks can include risks to the entities themselves as well as to consumers arising from the information and data underlying the system (for example, inadequate data representativeness or the use of outdated data that may lead to the introduction of new biases or amplify existing ones, misleading information resulting in flawed system outputs, etc.), risks associated with the systems themselves (for example, autonomous changes of the algorithm driven by self-learning that may materially alter in system activity beyond its original design, the need for frequent system maintenance, and the lack of effective examination monitoring and control tools); or risks engendered by the absence of human involvement throughout all decision-making stages.[335]

Corporate governance provisions are already in place with each of the financial regulators, both in statutes and in regulations established by the competent authorities. For instance, the Securities Law, 5728-1968 (hereinafter: the "**Securities Law**") defines the corporate governance rules required of stock exchanges and clearing houses. Chapter D1 of the Investment Advice Law (see below) also deals with corporate governance and establishes explicit provisions for licensed portfolio managers, investment advisors and investment marketers.

The Directive to Licensees for the Provision of Services Using Technological Means, also includes corporate governance provisions, covering areas such: as acquiring relevant knowledge, skills and

[333] Activity of the Banking Supervision Department in the field of corporate governance in the banking system 14.5.2018, https://boi.org.il/media/0zaj2atu/%D7%AA%D7%99%D7%91%D7%94-%D7%A2%D7%9C-%D7%9E%D7%9E%D7%A9%D7%9C-%D7%AA%D7%90%D7%92%D7%99%D7%93%D7%99.pdf?force_isolation=true.

[334] IOSCO Report 2021, p. 9.

[335] NIST AI 100-1, Artificial Intelligence Risk Management Framework (AI RMF 1.0)

understanding regarding the technological means; implementing controls to detect and prevent failures; ensuring board of directors' involvement in approving the technology and its use; appointing a responsible licensee; and maintaining proper documentation, among other requirements.[336]

In banking, the Proper Conduct of Banking Business Directive No. 310 entitled "Risk Management" deals, *inter alia*, with corporate governance rules required of a banking corporation by the Banking Supervision Department. According to this Directive, corporate governance is considered one of the key mechanisms for risk management in a banking corporation, structured around three "lines of defense":[337]

(1) Business line management (business unit management) - These bear responsibility for identifying, assessing, measuring, monitoring, reducing and reporting on risks embedded in products, activities, processes and systems under their responsibility, and for managing an appropriate risk management control environment;

(2) Independent risk management function - Its role is to complete, as a second line of defense, the business line's risk management activities. This function is responsible for the planning, maintenance and ongoing development of the bank's risk management work plan. A main role of this function is to challenge the appropriateness of business-line inputs for risk management, risk measurement and the banking corporation's reporting systems, as well as the appropriateness of the outputs received. Other compliance, monitoring and control functions – such as the compliance and anti-money laundering functions, the chief accountant and financial reporting oversight – also constitute part of the second line of defense;

(3) Internal auditor – The internal auditor conducts an independent review of and challenges the bank's internal controls, processes and risk-management systems. Its responsibilities are detailed in the Proper Conduct of Banking Business Directive No. 307. Alongside this provision, the corporate governance provisions stipulated in the Directive deal, *inter alia*, with supervision and control, risk management, compliance and organizational culture. The principles of corporate governance are reflected primarily in the effective functioning of the corporation's governing bodies, underscoring the roles of: the board of directors, the independent auditor, the internal auditor, the risk manager, legal counsel, the compliance officer and the chief financial officer.[338] As evident from the above

[336] See *infra* note 387.
[337] Section 4 of Proper Conduct of Banking Business Directive No. 310.
[338] Activity of the Banking Supervision Department in the field of corporate governance in the banking system 14.5.2018 https://boi.org.il/media/0zaj2atu/%D7%AA%D7%99%D7%91%D7%94-%D7%A2%D7%9C-

review, corporate governance rules form a core component of risk management for financial sector corporations, and receive considerable attention within the framework of financial regulation.

The Taskforce considers corporate governance mechanisms to be a potential cornerstone in how financial corporations handle the issues and challenges stemming from the integration of AI systems. In the public comments on the Interim Report, some respondents argued that a structured process should be established regarding AI uses, or that a regulated licensing regime should be introduced. While these comments have merit, and while establishing a process within supervised entities for considering the adoption of AI systems is warranted, the Taskforce takes the view that it is difficult to prescribe in the Interim Report a mandatory, uniform process for all AI applications. The reason for this is that AI uses may vary widely, and imposing mandatory procedures could result in the duplication of existing processes (for example, licensing for financial activity, which will also be conducted with the assistance of AI systems). Nevertheless, financial regulators may ultimately decide to do so, either for certain areas of activity or more broadly.

Another issue raised in the public comments– apparently due to a misunderstanding of the Interim Report's recommendations– was the concern that certain requirements may be overly burdensome. The Interim Report clarified that the AI-governance measures it presents comprise a toolbox from which appropriate measures can be drawn based on the circumstances. This clarification was also added to the final report. By way of illustration, within the discussion of decision-making processes and accountability, the report described a possible mechanism whereby an AI system could be subject to board approval. This should not be understood as a recommendation that every AI system requires board approval. Whether approval of by the board or any other corporate organ is required depends on the specific circumstances, the financial entity's risk management considerations, the materiality of the relevant product or service, and to the board's own policy.

%D7%9E%D7%9E%D7%A9%D7%9C-
%D7%AA%D7%90%D7%92%D7%99%D7%93%D7%99.pdf?force_isolation=true.

AI Governance – Recommendations:

Set out below are several measures that may assist participants in the financial sector in developing appropriate governance mechanisms within financial entities that deploy AI systems. This list is neither mandatory (until determined otherwise by financial regulators) nor exhaustive. Its purpose is to provide regulators and supervised entities with a toolbox with which they can design suitable governance arrangements. The measures set out below are articulated in operational terms and illustrate a broad range of potential governance practices drawn from emerging international approaches. They can be applied, in whole or in part, depending on the relevant considerations and circumstances, such as the nature and materiality of the AI-assisted service provided, the potential harm to individuals and similar factors. Any such application should, of course, follow a review of existing governance arrangements and an assessment of whether these measures are needed or should be integrated into existing governance frameworks. Even in the absence of such requirements, this set of measures can assist financial entities that choose to establish. AI-adapted governance practices on a voluntary basis.

Procedures, policy documents, processes, and practices	Establishing policy documents, processes, procedures (including technical procedures) and practices (including self-practices) that support the mapping, assessment and management of risks associated with the use of AI systems. These policy processes, procedures and practices must be documented and disclosed, and they should operate alongside existing corporate governance requirements, which do not necessarily address AI-related issues. The policy documents, processes, procedures and practices should be updated in line with technological developments and with the scope and materiality of AI use in financial entities. It must also be ensured that they are clear, understood and transparent to all relevant parties in the organization. The policy documents, processes, procedures and practices must apply across the entire lifecycle of the AI system, including pre-deployment review, system deployment, ongoing monitoring and supervision, and taking appropriate action when defects are identified and their consequences addressed. Employees of the financial entity involved in AI-system activity or their results must undergo training and workshops to acquire sufficient knowledge to fulfill their role satisfactorily. As part of the ongoing monitoring and system oversight, consideration must be given to feedback on the system's operation from the various parties involved in the AI system's activity (e.g., system users), and not

	solely from those who developed the systems or those responsible for their implementation within the financial entity. The adoption of AI systems and their usage shall be approved by the financial entity's board of directors, as part of the entity's overall risk-management policy and risk appetite. The financial entity's board of directors and senior management shall be the parties accountable for its activity, including the adoption, supervision, monitoring and use of AI systems.
Decision-making processes and accountability	The financial entity, through one of its organs, must acquire knowledge, understanding and skills with respect to AI systems prior to their deployment and throughout the entire period of their use. The financial entity shall appoint an AI officer, whose role is to know and understand the methodology underlying the AI systems used by the organization. The officer must have the knowledge, skills and experience to make decisions regarding AI systems, provide explanations of their modes of operation and perform testing, monitoring, supervision and validation of existing or new systems.[339] The said AI officer can also be an existing officer (for example, chief technology officer, chief risk officer and so forth), who will perform this role parallel to his other roles.
Monitoring, supervision, validation	The entity should establish processes for the assessment and validation of new and existing AI systems. Monitoring and supervision must begin prior to the AI system deployment, including validation that the system results are proper, and continue on a regular basis using the appropriate audit tools,

[339] The appointment of an AI Officer function corresponds with various provisions in financial regulation dealing with the appointment of office holders entrusted with areas of activity or issues the corporation deals with. Thus, for example, see the duty to appoint a Chief Risk Officer in institutional entities (see review on the subject 'Risk Management in Institutional Entities – A Horizontal Review in Comparative Perspective' (https://www.gov.il/BlobFolder/generalpage/information-entities-corporate-governance/he/nihul-siconim_em31082016.pdf?force_isolation=true) and in banking corporations (https://www.boi.org.il/media/lajbhthh/200944.pdf), as well as the appointment of a licensed individual in charge in financial entities providing portfolio management, investment advice and marketing services [Directive to License Holders regarding the Provision of Services Using Technological Means, Directive under § 28(b) of the Regulation of Investment Advice, Investment Marketing and Investment Portfolio Management Law, 1995 (Israel). The AI Officer shall be subordinate to the management of the financial entity and report to the management and the Board. Furthermore, such an appointment also corresponds with the duty to appoint a Chief Compliance Officer in accordance with the Banking Supervision instructions (https://boi.org.il/media/1xdf0b3e/308.pdf?force_isolation=true) which is also subordinate to the banking corporation's Board, and with the duty to establish effective enforcement programs in accordance with the Companies Law and the Securities Law (in this context see Amendment No. 16 to the Companies Law and Amendment No. 47 to the Securities Law. In a similar context, see the duty of federal agencies in the United States to appoint a Chief AI Officer (CAIO) (https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf?force_isolation=true) following the Executive Order. The instruction, which was also issued on behalf of the US Administration, establishes the roles and authorities of the AI Officer, the scope of their responsibility, and their status in the organizational hierarchy.

	encompassing outcome-based reviews as well as auditing the algorithms themselves on a regular basis. The financial entity must conduct a periodic review of AI system results to retrospectively examine their accuracy (by the entity or by an external party). To dispel any doubt, this review is not intended to replace any regular review of the systems themselves or the assessment conducted prior to the systems' deployment.
Emergencies and cessation of activity	A financial entity should operate appropriate control and oversight measures for managing AI systems. As part of control and oversight measures it must have the ability to halt AI-system activity if a material fault is found requiring the immediate cessation of that activity. This measure, acting as a "kill switch", must be incorporated into the financial entity's procedures, specifying the circumstances under which it may be activated and the organizational position-holders authorized to activate it. The kill switch must also be tested periodically to verify its functionality, even in a system failure and various emergency scenarios. This mechanism is analogous to "circuit-breaker" trading halts in stock-exchange trading, as set forth in the exchange's rules and regulations.[340] The financial entity must maintain appropriate backup for its AI systems in case the kill switch is activated or a there is an AI systems failure barring access to it.
Use of outsourcing and third-party suppliers	Financial entities must ensure the maintenance of their policy documents, processes and procedures addressing risks arising from the use of AI systems provided by third parties or where service is provided through outsourcing. Inter alia, these documents, processes and procedures must include business continuity mechanisms, the management of system or information failures, and related considerations. Information and data received from third parties or in outsourcing must be validated as being representative, non-discriminatory and not introduce bias. This validation must be conducted prior to using the data, and continued on an ongoing basis thereafter to confirm their relevance and accuracy.

[340] For example only, see the circuit breaker mechanism in the Tel Aviv Stock Exchange Trading Guide, Guidelines under the Third Part of the Rules.

Data Governance

Data governance is a core issue that is often associated with AI governance and is of particular importance for AI systems whose functioning and outputs rely extensively on the data used for training or operation. The well-worn phrase, "garbage in, garbage out" illustrates in a figurative way the impact that low-quality information and data can have on the results generated by AI systems. Article 10 of the EU AI Act focuses on data and data governance. It specifically addresses high-risk systems and discusses the requirements applied to information and data in the AI system's training, validation and testing stages, including that rules governing the process of collecting information and its sources, the relevance and representativeness of the data, measures to detect, prevent and minimize discrimination and bias, gaps or deficiencies in the data, and similar considerations.

The 2021 OECD Report, also addresses the subject of data governance and puts forth proposals for policymakers to follow in this context:[341]

Policy makers could consider the introduction of specific requirements or best practices for data management in AI-based techniques. These could touch upon data quality, adequacy of the dataset used, depending on the intended use of the AI model, and safeguards that provide assurance about the robustness of the model when it comes to avoiding potential biases. Appropriate sense checking of model results against baseline datasets and other tests based on whether protected classes can be inferred from other attributes in the data are two examples of best practices to mitigate risks of discrimination. The validation of the appropriateness of variables used by the model could reduce a source of potential biases. Tools could be developed and used to monitor and correct for conceptual drifts.

In the 2023 OECD Report's recommendations, data governance is identified as a critical issue, one that assumes heightened importance:

Strengthen data governance practices by model developers and deployers: The importance of data is undisputed when it comes to the training of GenAI models and their usage by financial market participants. Best practices for data management and governance practices may be considered to ensure data quality, data adequacy depending on the intended use, data privacy when financial consumer data is fed into the model, and data authenticity and appropriate source attribution/copyrighting when applicable. This could include increased transparency, model documentation and reporting about the data used to train the model and any other data introduced into the model, including their location,

[341] 2021 OECD Report, p. 56.

origin and source attribution for copyrighted data used [...] This would need to include any data inputted in the model through prompts or otherwise, and the output of the model itself, given feedback loops for its self-training.[342]

Data Governance – Recommendations:

Several measures which may assist financial regulators in determining appropriate data-governance rules for financial entities using AI systems are presented below. The list is not exhaustive and is intended as a toolbox for regulators formulating suitable data governance arrangements:

Sources of information

- Financial entities should develop policy documents, processes, procedures (including technical procedures) and practices for assessing information and data sources, and for conducting due diligence on the representativeness and accuracy of the data, the manner in which they were obtained and their sensitivity, when such data are received or purchased from third parties. Regarding reliance on third parties that are critical to the financial entity's activity, including with respect to information and data, see above.
- Consideration should be given to appointing a data and information officer, parallel or subordinate to appointment of an AI officer.

Use of information

- Financial entities should develop policy documents, processes, procedures (including technical procedures) and practices (including self-practices) specific to data governance. These documents must address, *inter alia*, the use of sensitive information and data and/or data that may generate heightened risk. In this context, policy documents, processes, procedures and practices must give expression to issues of model training, discrimination and bias, representativeness and validation, privacy, relevance and the original context of data.

[342] 2023 OECD Report, p. 28.

"Connecting the Dots" - A Comprehensive & Risk-Based View

Risk management in relation to AI systems should be approached holistically, as the various risks may be intertwined and the measures required to address them correspondingly interrelated. For example, the "black box" nature and limited explainability of an AI system used to support insurance underwriting decisions may heighten the risk of discriminatory outcomes (given the difficulty in detecting the discrimination) or increase the risk exposure to liability for the system's performance (since the developer or deployer may struggle to explain or adjust the system). The means to mitigate these risks can also be diverse and interrelated. Human oversight requirements, for example, may reduce the risks associated with the system, as may notification and disclosure requirements. Adopting a comprehensive perspective, therefore, enables more effective identification of the relevant risks and tailoring of the means to address them. The management of AI systems should also be grounded in a risk-based approach. Risk-based regulation has emerged as a common theme in AI-related policy frameworks both globally and in Israel, and its benefits are clear: it enhances regulatory effectiveness and by prioritizing material risks, promotes efficient resource allocation by both regulators and the regulated entities. It provides flexibility to tailor regulatory measures to evolving and context-specific risks and it encourages business activity and innovation, assuming that the regulatory burden remains proportionate to the level of risk and does not over-burden the regulated entities.

What is the practical meaning of risk-based regulation?

The EU AI Act is a salient example of risk-based AI regulation. The Act outlines four categories of AI-system risks. The restrictions and rules applicable to them under the Act were derived accordingly: Unacceptable risk – The law prohibits certain AI-assisted activities. Systems at this risk level include, *inter alia*, systems using manipulative techniques to distort behavior and impair informed decision-making; systems exploiting vulnerabilities related to age, disability or socio-economic circumstances to manipulate behavior; systems that engage in social scoring, leading to the adverse treatment of individuals or groups based on such evaluations or classifications; systems used for emotion recognition in workplaces or educational institutions, except for medical or safety purposes; systems building facial recognition databases through the uncontrolled collection of data from the

Internet or security cameras as well as certain biometric systems.[343] As a rule, AI systems classified as engendering "unacceptable risks" have been prohibited for use, except for certain specific exceptions.

High risk – The Act sets criteria for classifying AI systems as high risk and lists specific examples in its annexes. These include, *inter alia*: systems engaging in biometric identification that is not prohibited; systems serving as safety components in critical infrastructure; systems influencing admission to educational institutions or student assessment; systems serving in employment-related decisions; systems affecting access to essential public or private services – including credit scoring and risk assessment for life or health insurance pricing; and certain law enforcement applications. Additionally, systems are considered high-risk if they engage in profiling, i.e., an automated analysis of personal data to assess aspects of a person's life, such as work performance, financial or health status, preferences or behavior. The risk level depends on the extent of the system's involvement in decision-making. A system performing a minimal procedural role in these contexts is not necessarily deemed high-risk under the Act.[344] High-risk AI systems are subject to regulatory oversight and must comply with requirements throughout the system's entire lifecycle, including risk management, data governance requirements, documentation, transparency, human oversight, registration in a dedicated database and more. Most obligations fall on AI-system providers, though certain requirements also apply to system deployers.

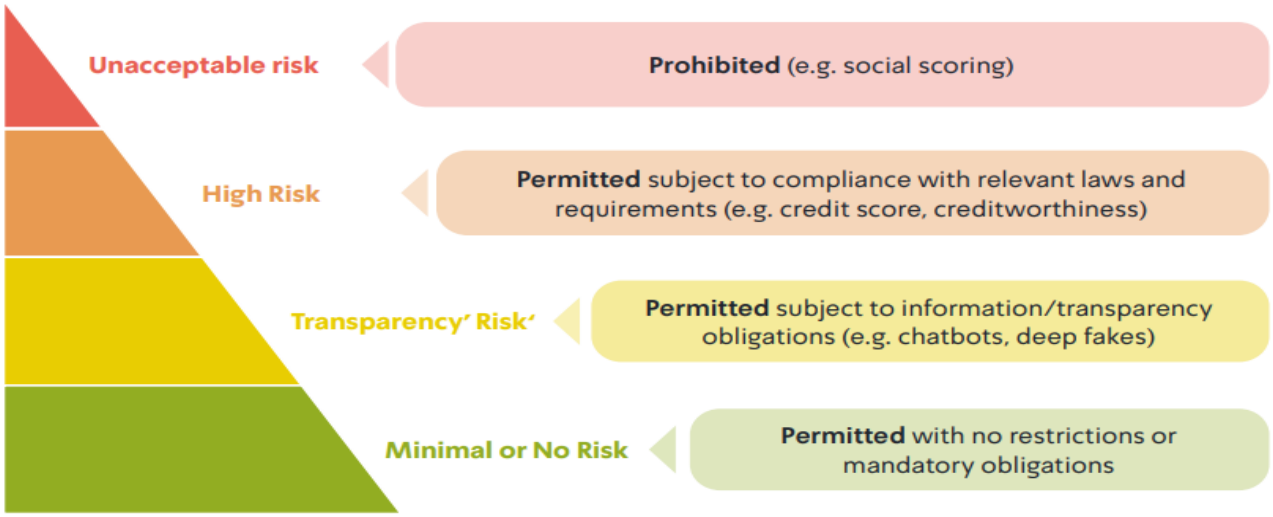
Systems engendering transparency risk – Certain systems – including those that simultaneously may be considered high-risk, but are not categorically so – are required to meet transparency requirements. These systems are intended to interact directly with human users and are therefore subject to transparency obligations. Such obligations comprise, *inter alia*, informing users that they are interacting with an AI system (unless this is evident from the context), labeling content that is AI-generated or AI-modified—including audio-visual deepfakes—and applying watermarks to AI-generated content, subject to applicable exceptions.

Minimal or no risk – Systems at this very low-risk level, such as systems in video games or content filtering systems (spam filters), are not subject to regulation.

In addition to the aforementioned risk levels, the Act imposes regulatory requirements on a new category of AI – general-purpose AI.[345] Developers of such systems are required to provide

[343] EU AI Act, Article 5.
[344] EU AI Act, Article 6(3).
[345] This term refers to AI systems that can be used in a variety of contexts and trained for a variety of uses, as opposed to systems developed for a specific use. See EU AI Act, Article 50.

documentation, usage guidelines, a compliance policy addressing copyright regulations, and a disclosure explaining the content used in model training. If the system is defined as posing a systemic risk (for example, as a result of its computing power), additional obligations apply, including ongoing reporting and cybersecurity risk management.



As stated above, a risk-based approach is not unique to the European Union and can be found in regulatory and policy documents across other jurisdictions as well – generally with respect to AI, but also regarding specific use cases.[346] A distinctive feature of the European legislation at this stage is its specification of the uses and systems included in each risk category.

Multi-tiered risk grading is a useful tool for classifying the level of risk assigned to financial-sector AI systems and for determining the appropriate measures to address each level of risk. For example, the classification of AI systems can be based on a risk continuum – low-medium and high.

Risk classification in financial institutions should be conducted along two dimensions: first, the risk level for customers, investors or the general public, and second, the risk level for the financial institution. For example, an AI system that has a material impact on customers or investors, such as one assisting in credit decision-making, may expose customers to harm by limiting their access to credit or by increasing its cost. In contrast, an AI system used to review anti-money laundering (AML) compliance processes may not necessarily directly harm the financial institution's customers, but its failure could harm the financial institution (through sanctions, reputational damage, and related consequences). As part of this assessment, the extent of the system's involvement in decision-making

[346] See *supra* note 301, sections 1.2, 3.2.2.

or task performance must also be evaluated, akin to the exclusion of systems with only marginal involvement under the European legislation. Another consideration in risk classification is the regulatory framework applicable to the system and the parties involved in its activity – for example, regulation imposed on AI-system providers, which may affect the assessment of the risks inherent in the system. The measures applied to an AI system should correspond to its risk level. High-risk systems may warrant more robust governance requirements, while systems classified as low-risk may require only minimal safeguards. These measures operate alongside any mandatory requirements imposed under applicable law and the regulatory provisions governing that activity. To illustrate, the tools detailed in this report are presented below according to a graduated risk scale. This is not intended as a closed or exhaustive list, but rather as a framework for categorizing AI systems by risk level, facilitating the characterization of a system's risk profile and the corresponding adaptation of the regulatory requirements (for example, regarding explainability, human oversight and more) applicable to it. As noted above, when tailoring this adaptation, it is important to consider that a single mitigation measure may serve multiple purposes, and the combined effect of the various measures should be assessed holistically. The concrete examples presented alongside each risk level are solely for illustration. A specific assessment of the system in use is required.

Risk level	Low-Medium Risk	High Risk
Examples and issues for examination	AI systems serving certain back-office tasks, such as performing routine data processing operations in various workflows, assisting in protecting financial institutions (for example identifying and deleting spam), chatbots serving simple operations not involving significant customer decision-making (for example providing general information about the financial institution).	AI systems engaging, *inter alia*, in analyzing customers' personal information and affecting the service provided to them. For example, a system serving to analyze clients' risk profiles and constructing a managed investment portfolio for them accordingly; or a system serving to identify fraud risks whose result may be preventing customer access to banking activity; or a system serving to price life insurance based on the customers' personal information.
"Black Box" and Explainability	General and specific explainability. General explainability - regarding the manner of algorithm operation, as distinct from specific explanation of how a certain decision was reached. Specific explainability - the regulated	Beyond specific explainability for the financial institution's internal control purposes and vis-à-vis the regulator, the regulated entity should have the capability to explain to the customer the specific decision that was made

	entity must understand the algorithm's functioning, have access to operating documentation and so forth.	regarding them. Specific explainability is required, unless compensatory measures offsetting its absence are in place.
Human Oversight	Human control over the system's activity (as distinct from control over each of its decisions). Retrospective human oversight over the system's decisions, in scope adapted to low and medium risk activity.	Human control over the system's activity. Real-time and retrospective human oversight over the system's decisions in scope adapted to high- risk activity.
Prevention of Discrimination	Oversight and taking measures to prevent discrimination in focused aspects of the system's activity (customer onboarding in the above example) and in focused aspects throughout the system's lifecycle.	Oversight and taking measures to prevent discrimination and bias throughout the system's entire lifecycle, particularly in the following: databases; manner of system operation; system outputs
Liability and Accountability	Liability of the regulated entity; Prohibition of limiting liability towards customers	Liability of the regulated entity; Prohibition of limiting liability towards customers A liability of the developer ("provider") as a consideration or a future condition with respect to operation of AI systems or being subject to common foreign regulation.
Disclosure and Notification	Notification requirement regarding the very use of an AI system unless this is clearly made evident by the circumstances; Disclosure – depending on circumstances (if it is a system for internal purposes not affecting activity vis-à-vis customers, there is no need for notification or disclosure)	Notification requirement of the very use of an AI system; specific disclosure obligations regarding the system and unique characteristics arising from the technology; disclosure in clear, concise and neutral language; restriction of misleading marketing publications.
AI Governance	Limited framework of AI governance measures such as: activity pursuant to policy documents and internal procedures of the regulated entity concerning information systems; activity pursuant to policy documents and internal procedures of the regulated entity concerning outsourcing. Insofar as these are systems involving greater	Comprehensive framework of AI governance measures such as: appointment of a designated AI officer; approval of the AI system by the regulated entity's board of directors; robust oversight before deploying the system and during its operation;

risk, controls are also required prior to system deployment and during its operation; acquisition of knowledge, understanding and deeper proficiency regarding the system; system backup and ability to replace it in case of material failure.	acquisition of knowledge, understanding and deeper proficiency regarding the system; System backup and ability to replace it in case of material failure; due diligence regarding third-party providers, including with respect to dependence on such third-party and data used.

In the public comments submitted on the Interim Report, responses were received pertaining, *inter alia*, to the importance of risk management and the determination of risk levels. Some commenters argued that, given the new and evolving nature of risk, building expertise is crucial; however, flexibility as to how to implement such practices should be retained at this time. Others took the opposite view, claiming that the regulator should prescribe a clear risk management process as well as a clearly defined list of risks. In this context, and with respect to the AI-governance tools proposed in the report, some commenters maintained that it is appropriate, for example, to set rigid requirements regarding the appointment of a designated individual within the organization responsible for the AI-related risks. The Taskforce requested, *inter alia*, the public's response to the classification proposed in the Interim Report, into categories of low, medium and high-risk activities. One commenter argued that the focus of regulation should be placed on applications with high-risk potential. Another claimed that the definition of risk levels proposed in the report is not sufficiently clear, lacking quantitative and qualitative criteria, and that criteria, such as the level of possible impact on the individual or the quantity and nature of personal data processed in the system, should be set.

These comments essentially concern the desired level of detail and the rigidity of the report's recommendations pertaining to the risk-based approach. Considering that the adoption of AI systems is still in its infancy, the potential diversity of such systems, and that in many financial-sector entities risk-management frameworks are already in place, the Taskforce believes that establishing a highly rigid and detailed prescriptive framework within the report's recommendations would be premature at this point of time, although such an approach should be revisited in the future.

With respect to the division into risk categories, the report has been revised to propose two categories: low-medium risk and high risk. This adjustment is intended to facilitate implementation, taking into account the fact that the three-tier division of risk levels – low, medium and high – does not align with the division adopted in the European Union legislation. In the EU framework, the two highest risk levels were characterized as specified above, as "unacceptable risk" and "high risk". Since the

"unacceptable risk" category are in fact practices deemed unlawful, the Taskforce did not consider it necessary to dedicate a separate risk category to it. Additionally, the risk categories articulated in European Union legislation – transparency risk and General-Purpose AI risk – can, in the Taskforce's opinion, be reflected across varying levels of intensity, and therefore, they defy segmentation as set out by the EU. Consequently, the Taskforce ultimately rejected the EU's division methodology and it has not been adopted in the final report. The Taskforce further believes that the flexibility afforded in addressing different risk levels supports a simpler two-tier division, as this allows measures to be tailored to the specific risks of each system. At the same time, the comment suggesting that systems below the high-risk level should fall outside the scope of this report – and financial regulation more broadly– was not accepted. The report does address systems at these risk categories, but proposes the regulatory treatment be proportionate to their risk level.

In response to the Taskforce's question whether the risk classification should be conducted by the regulated entities or the regulator, conflicting comments were received. One commenter argued that the regulated entities are the most efficient party for determining risk level, while others called for more prescriptive regulation, claiming that such regulation would provide clearer guidance regarding risk classification. The Taskforce is of the opinion that, at this stage of technology adoption and given AI's broad range of potential applications, the risk-level classification of AI systems should be conducted primarily by the regulated entities. The financial regulator will be able to establish specific rules in the future, according to which, systems with certain functions or in certain domains could be deemed high-risk. A recommendation on this matter was added to the report.

Another comment that was not accepted by the Taskforce was that the report should be dedicated solely to AI systems intended for the external activity of financial institutions, exempting those systems used for internal operations. According to the risk-based approach presented above, the AI risks associated with financial institutions may be limited to customers and investors, but nonetheless affect the financial institution itself and must be assessed and addressed. Accordingly, limiting the focus solely to systems intended for external use—particularly those involving customer interaction—is unjustified. AI systems developed for internal purposes may present relevant risks, which are likewise addressed in this report.

Financial Stability

Addressing risks to financial stability is a central regulatory priority, intended to protect the public's savings, ensure the sound functioning of the financial system, and prevent financial crises that could affect the entire economy. The importance of systemic risks to financial stability became evident following the global financial crisis of 2008, which illustrated the need to assess the financial system as a whole rather than focusing solely on specific components or institutions, and to identify trends and risks that could materially affect its functioning. In the aftermath of the crisis, oversight of systemic risks was strengthened worldwide, including in Israel, where the Financial Stability Committee was established to help safeguard the stability and sound functioning of the financial system. Established in 2018 under the Bank of Israel Law, the Committee operates through coordination, cooperation and information-sharing among all key stakeholders in the financial stability issue: the Bank of Israel, the Ministry of Finance and the financial regulators.[347] Could the entry of AI to the financial sector bring financial stability risks with it? The fundamental answer to this is yes.

Prudential risks resulting from the entry of AI are attributed primarily to the following:

Herding and volatility – The use of AI models by a large number of players in the financial system could lead to herd behavior, i.e., massive similar activity by numerous participants. This could increase risk, especially during times of crisis. For example, a fall in securities markets may provoke a broad-based sell-off, thereby exacerbating both market declines and volatility.[348] It should be noted that even in the absence of an immediate crisis, repeated use of the same models and information for decision-making could undermine the diversity of opinions and approaches within the financial system. The former chair of the U.S. Securities and Exchange Commission, Gary Gensler, warned about the risk of convergence and the emergence of "monocultures" in the financial system if prompt measures are not taken to address the use of AI.[349] Another concern raised is that AI systems trained primarily on data relating to routine activity may make flawed decision during non-routine market events, increasing the likelihood of a financial crisis.

Concentration and third-party dependency – a concentration of financial activity by a small number of parties possessing the data and technological capabilities for using AI raises concerns not

[347] For elaboration see Bank of Israel Law (Amendment No. 7), 2018 (Israel), which was approved on 26.11.2018 and the Bill dated 25.1.2017.

[348] 2021 OECD Report, p. 8.

[349] See remarks by Prof. Gensler dated 17.7.2023 - https://www.sec.gov/news/speech/gensler-isaac-newton-ai-remarks-07-17-2023; and his remarks in an interview dated 15.10.2023: https://www.ft.com/content/8227636f-e819-443a-aeba-c8237f0ec1ac.

only regarding market competition, but regarding systemic financial stability as well. Concentration can increase risks precipitated by an operational or other failure, due to their potential impact on larger monetary volumes and a greater number of stakeholders. In the financial sector, the concentration risk associated with AI is closely linked to the anticipated dependence on a small number of AI providers. When AI applications rely on or are supplied by only a few entities, the functioning of these systems depends in effect on those third parties. This presents a regulatory concern because these providers are usually not subject to financial regulation, and fall outside the scope of direct supervision.

Interconnectedness and contagion – These risks refer to the concern that shocks affecting one institution or segment of the financial system will quickly spread to others. Interconnectedness and contagion stem from direct and indirect linkages between financial institutions and markets, as well as from the speed, and at times, the unpredictability with which correlations or dependencies within the system manifest. In the context of AI, interconnectedness and contagion may occur, *inter alia*, when market dynamics are heavily influenced by AI-driven decision-making or when external suppliers of AI systems serve multiple institutions, such that a disruption in their services affects the operations of many financial actors.

Operational risks and cybersecurity risks – As the value of AI use grows and the technology becomes embedded across a wide range of internal processes within the financial system, the operational risks and cybersecurity risks associated with it may increase as well. The implications of these risks for financial stability are, in turn, closely linked to the other dimensions mentioned above.

The risks noted above are compounded by the characteristics of AI systems discussed in this report. For example, the "black box" nature of certain models complicates any analysis of the system's behavior and, as a consequence, increases its inherent risk. Additional concerns include the possibility of over-reliance on AI systems and given the potential sweeping deployment of them relative to traditional systems. It is also important to consider that AI systems can execute a very large number of decisions in a short period of time – potentially accelerating the onset and spread of crises.[350] Several examples of extreme situations and the potential risks resulting from the expanding use of AI in the financial sector follow:

[350] European Central Bank, Financial Stability Review (May 2024), https://www.ecb.europa.eu/press/financial-stability-publications/fsr/special/html/ecb.fsrart202405_02_58c3ce5246.en.html

1. Sales of financial assets accelerating a "chain reaction" – a situation in which several leading financial institutions use the same model to execute automated sales on a significant scale simultaneously, accelerating a fall in financial asset prices.
2. Overpricing in the capital market – uniformity in investment strategies employed by financial institutions raises concerns that certain or similar financial assets will be favored by AI. As a result, heightened demand leading to overpricing could be created for financial assets, in a way that overstates their real value, leaving then vulnerable to sudden declines.
3. Cyber attack – exploitation of vulnerabilities in AI systems serving numerous parties, and the ability of attackers to affect numerous players operating in the same manner or by means of the same systems.
4. Model errors and over-reliance by financial institutions – where numerous institutions train their models on the same datasets, concern arises regarding their exposure to the same biases and weaknesses embodied in those datasets. In extreme circumstances, systemic failure could occur if the datasets contain errors or are not current.
5. Intensification of procyclicality – alongside the concerns arising from the use of identical or similar models across financial institutions, an additional concern arises that other models with machine learning capabilities will identify a trend that was apparently created by other models, will learn from these models and execute similar operations, potentially amplifying those actions beyond what the underlying data would justify.

   It is possible to envisage several approaches to mitigating these risks, including requirements pertaining to engagement with AI system providers to monitor and mitigate risks stemming from possible potential dependency on these systems. At this time, however, these remain potential risks, and therefore it appears that, for now, countries worldwide are focusing on monitoring and assessing the implications of AI's future deployment for financial stability.[351]

[351] See for example the references of the Prudential Regulation Authority (PRA) and the Financial Policy Committee (FPC) in the UK on the subject: The Bank and the PRA's response to DSIT/HMT: update on our approach to AI (April 22, 2024), section 12, at https://www.bankofengland.co.uk/prudential-regulation/letter/2024/artificial-intelligence-and-machine-learning-letter; Financial Policy Summary and Record of the Financial Policy Committee meeting on 21 November (December 6, 2023), sections 89- 94 at https://www.bankofengland.co.uk/-/media/boe/files/financial-policy-summary-and-record/2023/fpc-summary-and-record-december-2023.pdf.

Financial Stability – Recommendations:

As noted at the outset, Israel's Financial Stability Committee operates as the forum through which the implications of AI use for the financial system's resilience, and its capacity to contend with systemic risks to financial stability can continue to be assessed. At this stage, given the still limited adoption of AI applications and the lack of concrete measures taken in other jurisdictions, the Taskforce considers it necessary to continually track this issue, while reassessing the need for action at a later point. Potential measures may include: mapping activities that could pose significant risk; reviewing the current framework for addressing operational and cybersecurity risks in financial institutions; and developing a strategy to mitigate systemic risks arising from the reliance on third parties critical to the functioning of the financial system.

Box 3: What is a Flash Crash Resulting from Algorithmic Trading Activity?

The year is 2029. Panic erupts in capital markets worldwide as a result of massive sell-off of financial assets by algorithmic trading systems, causing a "flash crash" resulting from the uncontrolled use by banks, investment houses and funds of identical AI models. The models submit orders to execute sales at the same time, and these massive sales lead to an uncontrolled chain reaction, a market crash and a large-scale credit crisis. The value of some of the financial assets recovers that same day, but the economic damages are estimated at trillions of dollars. A fictional scenario? Not really. The (former) SEC Chair, Professor Gary Gensler, warned that the growing use of AI systems will almost certainly cause a collapse of financial markets at some point in the coming decade.[352]

If such a scenario materializes, it will not be the first time that algorithmic trading precipitates a crash in financial markets. The event, called the Flash Crash of 2010, occurred on May 6, 2010, when within a few minutes, the Dow Jones index fell by approximately 900 points, which, at the time, comprised approximately 9% of the index. Several minutes later, the index "corrected" a lion's share of the decline.

On that day, trading in the U.S. derivatives market was marked by negative sentiment, unusual volatility and little liquidity. Investigation of the event pointed to the execution of a very large-scale futures contract sales program executed by algorithmic trading. As emerges from a report by U.S. financial regulators, high- frequency trading (HFT) accelerated the effect and intensified the decline. The liquidity crisis that originated in the derivatives market also affected stock trading, when automated mechanisms built into automated trading systems used by liquidity providers in the stock market led to temporary trading halts. The report concluded, *inter alia*, that the collapse occurred due to the interaction between the built-in automated mechanisms in the trading systems and the strategies employed by algorithmic trading systems.[353]

Since 2010, the concern over market crashes precipitated by algorithmic activity has taken on greater force. With the expected penetration of AI technology into financial markets, there is a growing concern that the concentration of models and AI system providers on one hand and their increasing use on the other will increase the probability that such events reoccur.

---

[352] Gensler, G. and L. Bailey (2020), "Deep Learning and Financial Stability", SSRN Electronic Journal, http://dx.doi.org/10.2139/ssrn.3723132 ; In this context see also Financial Stability Board (2017), Artificial intelligence and machine learning in financial services Market developments and financial stability implications, https://www.fsb.org/uploads/P011117.pdf

[353] See SEC and CFTC, Findings Regarding The Market Events Of May 6, 2010 (September 30 ,2010), at https://www.sec.gov/files/marketevents-report.pdf; Gitit Gur Gershgoren, Doron Greenberg, "High Frequency Trading – Its Effects on the Capital Market and Regulatory Effects on It", Israel Securities Authority (October 2015), https://www.new.isa.gov.il/images/Fittings/isa/asset_library_pic///mechkar.pdf.

Prof. Gensler argued that the solution to the concern should be found in risk management. While existing risk management models examine risk at the firm level, a method should be adopted that will allow the assessment of macroeconomic risk. According to him, regulators must quickly identify tools to manage the risks associated with the growing use of AI in the financial sector, in order to reduce threats to financial stability.[354]

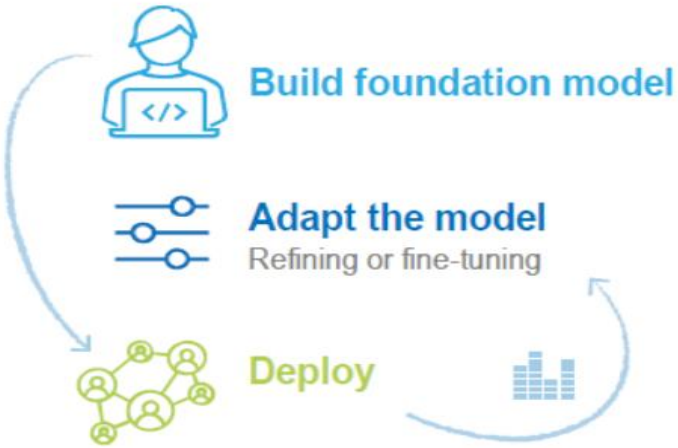
[354] See remarks by Prof. Gensler, *supra* note 326, and the European Central Bank (ECB) review from May 2024, *supra* note 327.

Competition

Artificial Intelligence has the potential to boost productivity in the financial services industry, foster innovation and promote the creation of new products and services that benefit the public. Given the anticipated importance of AI in the coming years, competition in this field also becomes crucial. To advance the technology, integrate it into markets and ensure that consumers realize its benefits, effective and dynamic competition in AI services is essential.[355]

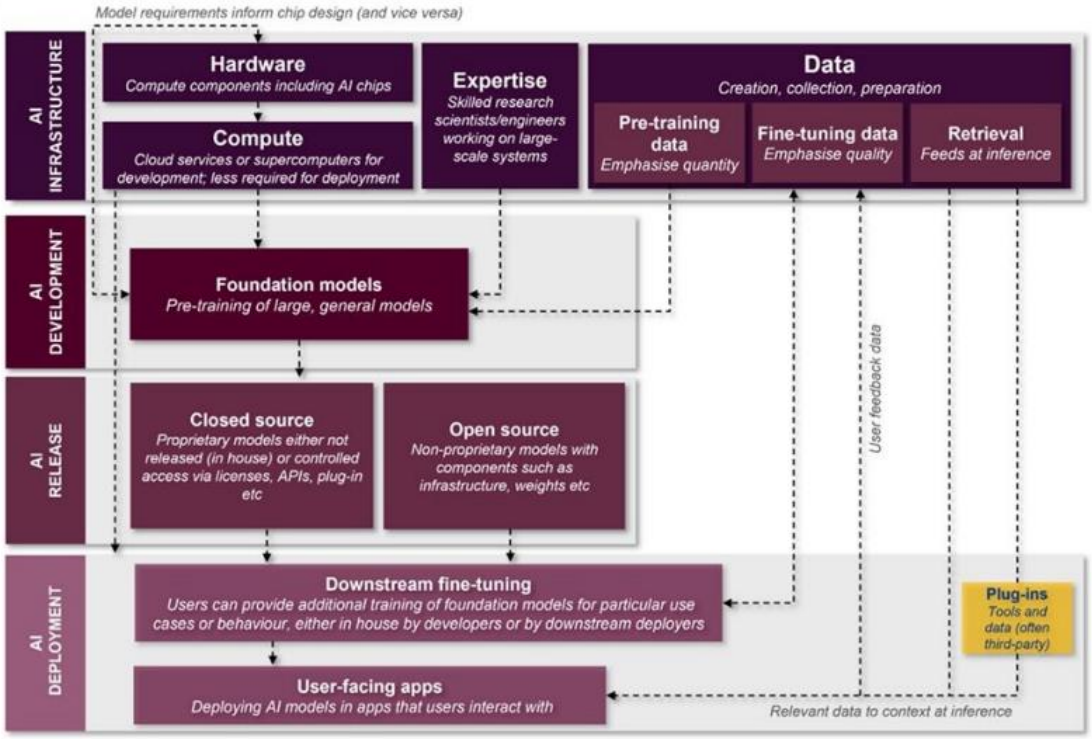
Stages in developing AI models

To analyze the AI sector from a competitive perspective, it is first necessary to understand the key stages in developing an AI model and the inputs they require. A 2024 document published by the OECD Competition Committee outlines these stages as follows:[356]



In 2024, the UK Competition and Markets Authority (CMA) presented a graphic breakdown of inputs required for each stage of AI model development, as follows:[357]

[355] OECD (2024), "Artificial intelligence, data and competition", OECD Artificial Intelligence Papers, No. 18, OECD Publishing, Paris, https://doi.org/10.1787/e7e88884-en, (hereinafter in this chapter: "OECD Competition Committee Report").
[356] OECD Competition Committee Report, p.15.
[357] CMA AI strategic update (29 April 2024), https://www.gov.uk/government/publications/cma-ai-strategic-update/cma-ai-strategic-update, (hereinafter in this chapter: "CMA Report from 2024").



As evidenced above, the main development stages are:

a. Creating the infrastructure for model development – for this purpose, access to three main inputs is required:

- First, computational capacity. AI models use a large number of mathematical computations in order to develop the model and generate its output.[358] The cost of obtaining computational capacity for developing a model varies according to the size of the model and the amount of data needed for its development, but in any case, significant cost is involved. The main options for obtaining computational capacity are building an independent data center; using a publicly available supercomputer; or through one of the cloud service providers.
- Second, access to expertise, knowledge and development capabilities – for this purpose, skilled human resources are required.
- Third, access to data. The data must be generated, collected and prepared for use. As will be detailed below, the data required is both public data and proprietary data.

b. Developing the foundation model – at this stage the model's "knowledge" is primarily developed. This stage involves using data for the purpose of "pre-training" the model, with an emphasis

[358] UK Competition and Markets Authority (CMA), AI Foundation Models Initial Report (18 September 2023) (hereinafter in this chapter: "CMA Report from 2023"), p. 12.

is on the volume of data rather than its quality. As a result, public data is typically used more extensively than proprietary data at this stage.

c. Releasing the model to market –

- Open-source models – these are models open for public use, without cost, subject to their license terms. In many cases, open-source models allow use of a foundation model and enable the user to perform fine-tuning of the model for their required needs.
- Closed-source models – these are models developed by specific companies, intended for their needs and not open to the public. Sometimes companies integrate these models into their existing products, sometimes they create new services or products through, and sometimes they market the AI model itself as a service to third parties who will integrate it within their business.

d. Model deployment and use –

- Model fine-tuning – at this stage a process of fine-tuning a pre-trained model is performed for the purpose of a specific task or operation. Now the emphasis is placed on data quality rather than data volume, and in many cases, proprietary rather than public data are used.
- User applications – this is the stage where users try out the model. As a rule, an AI model continues to develop and improve during its use as well. The use of the model generates additional data by means of which it is possible to reperform the fine-tuning stage and improve the model's performance. Use of AI models is generally divided into use of open-source models and use of closed-source models.[359]

Competitive characteristics of the AI field

As specified below, most competitive concerns in the field of developing AI models relate to barriers to access to required inputs, concern over reinforcement of the dominant position of large technology companies through AI models, and concern over the development of markets marked by "winner-takes-all" or "winner-takes-most" dynamics.[360]

a. Advantages for first-movers in the field

In the AI market, the first-mover advantage appears to be alive and well. However, since competitive dynamics in this field are still evolving, it is not possible at this stage to determine how significant these advantages will be over time.

[359] Ibid, p. 14.
[360] OECD Competition Committee Report, p. 27.

- Network effects and "feedback loops" – the more customers use a certain model, the more knowledge feeding the model accumulates, enabling it to continually improve. As the model improves, incentive for more customers to use it grows. When a competitor enters the market early and acquires customers, it can develop a high-quality model, tailor that model to its customers' needs and foster a tendency for customers to remain with it due to the initial advantage.[361] This, in turn, can create a barrier for later entrants seeking to enter the market.
- Use of imperfect models – at the initial stage of a technology's market entry, users tend to be more tolerant of errors or deficiencies in models. As more users engage with the model, enabling further data collection and additional training, these deficiencies and errors can be addressed. However, as the technology develops, users become increasingly less tolerant of imperfect models and tend to prefer established veteran models. A new company seeking to enter the market at a later stage will need to offer a model that operates without significant errors or shortcomings.[362]
- Entry costs – The cost of entering an industry are generally lower at an early stage than at a later stage. For example, computing and data costs will likely increase as more companies enter the market and drive demand for the same inputs.[363]
- Brand recognition and customer loyalty – companies that enter first have the ability to acquire customer loyalty and position themselves as market leaders.[364]

As stated above, since the AI field is still in its infancy, the competitive dynamics in this field are not sufficiently clear. It is possible that first movers will face disadvantages. Initial AI models may prove to be less accurate or less efficient than models developed at later stages, after gleaning insights from the initial models. It is also possible that foundation models built in an open model will enable those

[361] Ibid, p. 28.
[362] CMA Report from 2023, p. 48.
[363] Ibid.
[364] Ibid.

entering at a later stage to close the gap rapidly and potentially in a more tailored and effective manner than early entrants, and do so without investing in the development of foundation models.

b. Advantages for large players

- Economies of scale – the initial development costs of a model are high. The ability to use software, algorithms, data, model training techniques and similar tools, in a large number of models and for a large number of customers, enables large companies to exploit economies of scale, developing multiple models at a lower cost.[365]
- Use of existing customer base – the more customers a company has, the more significantly network effects and "feedback loops" can be exploited and the model continuously improved.[366]
- Economies of scope – companies active in many markets can use a model, development methods, algorithms or inputs for products and services in other markets and achieve economies of scope accordingly. For example, data may confer economies of scope on firms where data collected in one market enables the development of more effective AI models in other markets.[367]
- Large technology companies – one of the main challenges in the AI field is that large technology companies, which have market power in multiple technology markets, may develop the field and the markets affected by it in a way that could undermine effective competition. Some of these technology companies possess substantial power in securing the inputs needed to develop AI models and have can significantly the ability of AI model developers to penetrate markets. Moreover, their strong position in existing technology markets– which may at some point be competitively threatened by AI models– creates an incentive to act to restrict competition in the AI sector.[368]
- Switching barriers – AI models can learn user preferences and capabilities in a manner that makes switching between competitors difficult. For example, if the model can learn the financial consumer's preferences, there will be a cost

[365] Ibid, p. 49.
[366] Ibid.
[367] OECD Competition Committee Report, p. 27.
[368] CMA Report from 2024, p. 12.

for that consumer to switch to another competitor, since it will take time until the new model learns those same preferences as well.

Competitive concerns in the AI field

In light of the foregoing, the main competitive concerns arising in the context of the AI field include:

a. Limited access to inputs (computing power, data, expertise) – Companies holding the main inputs for developing AI models may prevent third-party access to them in order to fortify their competitive standing.

b. Reduction of customer choice– There is concern that large technology companies will reduce the choice of actors in AI services. For example, this encompasses the initial decision on how to deploy an AI model, the ability to switch between models or between different providers over time, and other related considerations.

c. Fortification of dominant position – Partnerships between key players in the field can reinforce the position of companies with market power across the supply chain.[369]

d. 'Winner-takes-all' – There is concern that AI-based markets will develop "winner-takes-all" dynamics. The question whether markets will develop in this direction depends on a series of secondary questions: the extent to which the markets are characterized by economies of scale from the standpoint of data; the extent to which is necessary for the service provider to provide personalized service for each customer; and the extent to which a network effect prevails in the relevant market.

e. Adoption of anti-competitive practices– Examples of anti-competitive practices include conduct such as excluding competitors, exploiting customers, or discriminating among customers. In addition, AI models may, autonomously and without human intervention detect and adopt coordination strategies between competitors, potentially resulting in *de facto* cartels.

f. Alignment through use of the same AI model – When competitors rely on the same AI model to inform their separate business decisions, there is a risk that the decisions of all competitors will become uniform, giving rise to concerns about alignment among market participants, even in the absence of explicit coordination between them.

Competition and AI in Israel's financial sector

a. The place of financial players in Israel in the AI field

[369] Ibid, p. 13.

The OECD divides the AI field supply chain into three main groups: providers of knowledge and inputs for developing AI models (data, code building, chips, server farms and more); model developers and designers; business users who embed the models in their work processes, products and services.[370]

Financial institutions in Israel typically rely on third parties to provide the inputs needed to operate AI systems. However, because these possess extensive customer data, some may develop their own models or collaborate with technology-sector partners, who supply the additional inputs needed for in-house model development.[371] Financial services in Israel are also expected to integrate AI models within the existing services they offer customers. It is possible that AI models will enable these institutions to offer new financial services, and collaborations may even enable them to act as providers of AI-based services to other financial institutions.

b. Competitive characteristics of Israel's financial market

Despite various regulatory measures undertaken in recent years aimed at enhancing competition and facilitating the entry of new players into Israel's financial sector, certain areas of the sector remain highly concentrated, *inter alia*, due to considerable entry and switching barriers. In the banking sector, the five largest banking groups collectively hold approximately 98% of the total financial assets. The sector is marked by high entry and switching barriers and a limited competition, with banks generally refraining from competing aggressively for customers of other banks or from offering "disruptive" products to individuals who do not already have accounts with them. As a result, most consumers purchase a range of banking services from the bank holding their checking account, without engaging in comparative benchmarking with other banks or non-bank institutions. Consequently, the banking system exhibits conditions conducive to a coordinated equilibrium, where the various players are incentivized to preserve these characteristics, granting each bank substantial market power vis-à-vis its customers.

Generally speaking, the insurance industry is more competitive, although certain segments are highly concentrated such as health insurance, life insurance and certain specialized niches. At the same time, the level of competition experienced by consumers is probably lower, due to significant structural issues related to the operations and incentives of insurance agents. These issues often result in competition among insurance companies being focused on agents, rather than trickling down to benefit the insured public. As a rule, financial services are marked by

[370] OECD Competition Committee Report.
[371] Ibid.

pronounced economies of scale, while personalizing the service for the customer and their individual characteristics is highly important.

c. Competitive concerns from deployment of AI models in Israel's financial services sector

The above raises concern over the development of market power in the provision of AI services to the financial sector in Israel, namely that Israeli financial institutions' use of AI will serve to aggravate existing concentration in the financial services, as outlined below:

1. Concern over the concentration of market power of AI service providers vis-à-vis financial institutions in Israel – Due to economies of scale, network effects and the importance of established reputation in the provision of AI services, combined with the potential for AI markets in various areas to evolve into "winner-takes-all" (or "winner takes most") markets, there is concern about the emergence and intensification of market power among AI service providers, such as large technology companies, vis-à-vis the financial institutions in Israel..
2. Limited access to data – As stated above, limited access to inputs is one of the most significant competitive concerns with respect to AI. Established financial institutions have access to extensive proprietary data accumulated over years of activity, continues to grow and be updated continuously This data enables the development of AI models tailored to the diverse needs of each financial institution, and adaptable to the fixed and evolving needs of individual customers. While developing an AI model also requires additional inputs, such as computational resources and expertise, existing financial institutions can overcome this by partnering with technology companies to develop models for their exclusive use. In contrast, new entrants, lacking a comparable volume of data may face significant barriers to entry barriers due to their inability to develop and train models on a similar scale.
3. Existing financial institutions expected to be first-to-market – As described above, there may be competitive advantages for early entrants to the AI field. The fact that existing financial institutions will likely be the first to deploy AI models may grant them market advantages such as creating network

effects, strengthening brand loyalty, or low entry costs, which new entrants will find difficult to bridge at a later stage.

4. Economies of scale and scope – As noted above, the AI market is characterized by economies of scale, economies of scope and the ability to leverage an existing customer base to enhance model performance. Because established financial institutions are large, financially robust entities, many of which operate in multiple markets, they are able to develop AI models and deploy them across a wide range of financial services, while also using their existing customer base to improve such models quickly and efficiently. These advantages may raise barriers to entry or expansion for other players in the market.
5. Anti-competitive practices – As discussed above, AI models may directly or indirectly adopt practices of self-preferencing, customer discrimination or tying of services. The more personalized the service and pricing offered to each customer, and the more heavily pricing relies on individual customer information, the greater the potential for differential treatment, tying or self-preferencing. In the financial sector, where service is tailored to individual customers and incumbent financial institutions hold extensive information about them, there is concern that AI models can leverage this information to discriminate between customers in ways that impair competition in financial services. Furthermore, AI models may enable large financial institutions to implement tying strategies between services more effectively by identifying customer preferences, needs and behavior patterns, AI models could bundle services in a way that maximizes the financial institution's profits.[372] While some of these practices may sometimes benefit consumers, may also, under certain circumstance, cause harm to competition and the public. Another concern is that financial institutions possessing the data required for developing AI models might restrict or otherwise impair access to that data and unreasonably refuse to make it available to their competitors.[373]

[372] Ibid, p. 45.
[373] Ibid

d. Concern over implicit coordination and market alignment – This concern stems from use of the same AI model by numerous financial institutions. The greater the propensity for different financial institutions to use the same model for the purpose of business decision-making, such as investment decisions, decisions on credit extension and pricing, the greater the concern for formation of implicit coordination or alignment between them. The higher the concentration among existing players in the financial market, the greater the concern. It should be emphasized that this concern may affect both competition and financial stability. In the public comments on the Interim Report, it was argued, *inter alia*, that caution is needed to avoid the erroneous application of competition law. Commentors contended that viewing the financial sector as uniformly highly concentrated is not necessarily accurate across all segments, that AI can be advantageous for competition, and that competitive concern should focus primarily on the power of technology companies. The Taskforce believes that this report is not the proper forum to discuss the competitive structure of the financial sector in detail, however, it is clear that certain activities in this sector do suffer from over-concentration. The report outlines the potential advantages of AI (as reflected at the start of this chapter), and acknowledges the power of large technology companies. This, however, does not negate the need to evaluate and provide recommendations regarding the conduct of local regulated entities.
Additional comments suggested distinguishing between different levels of competition: the development of AI tools, which is expected to include mainly foreign companies; markets for creating AI-based applications, also expected to be largely international; and the provision of application usage services, which is marked primarily by local competition. The Taskforce accepts this segmentation, as reflected in the report and follows the OECD's analysis on the subject. Based on this, it was proposed to examine specific issues, such as exclusivity or priority agreements for AI tool usage, or rules regarding data sharing. Given the uncertain pace of AI system adoption described in the report, it appears that taking immediate concrete steps on these issues may be premature at this time, although they may be necessary in the future. Additional comments addressed, on one hand, concerns regarding data coordination (for example price coordination or conversely trade activity coordination) and, on the other hand, the need for data sharing between traditional financial institutions and new entrants to facilitate competition. The Taskforce agrees on the importance of these factors, as reflected in the first recommendation in the chapter, which establishes the principle that competition law applies to AI activities, and in the third recommendation, which addresses access to aggregated and anonymized data to create conditions conducive to competitive development.

Competition – Recommendations:

Based on the market analysis and competitive concerns outlined above, the following responses and recommendations are presented with the aim of mitigating these competition-related concerns.

(1) Application of competition law to AI activity –

It is important to underscore that competition law remains applicable even where anti-competitive conduct is implemented through AI models or arises from their operation. For example, a restrictive arrangement implemented through AI may constitute a violation of Section 2 of the Economic Competition Law, 5748-1988 (hereinafter: "Competition Law"), with liability potentially extending to both the model's users and the developers. Similarly, if a monopolist blocks its competitors' access to key inputs, such as data, needed to develop AI models, this could constitute a violation of Sections 29 and 29A of the Competition Law. It is also true for cases in which access is blocked to the AI models themselves in certain circumstances. Likewise, if a monopolist engages in tying between financial services or discriminates between customers, these actions could violate Section 29A of the Competition Law, even if they are executed by means of AI models and occur unintentionally.

(2) Examination of concerns regarding formation of market power or barriers to competition within merger reviews –

As part of merger reviews brought before the Commissioner of Competition, the Competition Authority will consider, *inter alia*, concerns related to the consolidation of market power or barriers to competition in the contexts described above, as well as concerns regarding the potential for enabling practices that could impair competition,[374] including through the use of AI models.

(3) Access to aggregated and anonymous data –

[374] An example of this is the Authority's objection in early 2024 to the merger between the credit card company Isracard and the insurance company Harel. The examination found that the merger raises a reasonable concern of significant harm to competition in the field of private health insurance, in which Harel is a leading company. It was found that if the merger is approved, there is a concern that the merged company could push its competitors out of the market or harm their competitive ability and their ability to restrain its market power. The concern focused on harm to competition that would be caused by using Isracard's data in addition to Harel's data to determine a personalized price for customers (among other things, based on the customer's willingness to pay) and thereby rake in more profit from a given customer base, in a manner that would allow Harel to attract insurance agents to it. See Objection of the Director General of Competition to the merger between Harel Insurance Investments and Financial Services Ltd. and Isracard Ltd., 2024, https://www.gov.il/he/pages/cases2023-008372.

In order to create conditions conducive to the development of competition in the provision of AI services to Israel's financial sector, taking into account the concerns outlined above, it is proposed to enable financial entities access to aggregated, anonymized data from the credit data repository at the Bank of Israel, as well as leveraging existing means and infrastructure under Open Banking to expand the data accessible to various financial players, enabling them those entities to develop AI models using that data. Currently, Open Banking allows bank customers, credit card companies and other entities to share their financial information with third parties. This enables financial entities, with the customer's consent, to access bank account information, and offer their customers banking services adapted to their needs. However, developing AI models requires access to extensive data collected over a long time-horizon and data on a single customer or data collected through customer consents is insufficient. The Bank of Israel dissents from this recommendation.

(4) Examination of ways to prevent the development of market power –

It is proposed not to be content with simply addressing entry and switching barriers to the provision of AI services to Israel's financial sector, problematic mergers and violations of the Competition Law, but rather to preemptively preclude the development of market power in this field. In the hands of an AI service provider, market power can have multiple effects, including: enabling the provider to exercise market power vis-à-vis financial institutions; increasing the likelihood of alignment or coordination (implicit or explicit) between different players in the market resulting from of the use of the same model; enhancing the provider's bargaining power with financial institutions and regulators; and eroding systemic stability due to dependency on a single dominant entity. Various measures could be considered to prevent the emergence of such market power. For example, regulation could limit a single AI service provider from serving clients that collectively hold more than a specified market share in a particular market in Israel. Alternatively, financial players in a given market could be required to engage with multiple AI service providers. Each approach has potential advantages and disadvantages, and it is recommended that any measure be assessed in light of current market conditions and the operational practices of the regulated entities.

Operational Risks Including Cybersecurity Risks, Third-Party Risks, Fraud and Disinformation

Operational risks cover a wide range of potential failures that could disrupt an organization's operations, arising from either internal deficiencies or external events.[375] Every financial institution is required to implement effective operational risk management, which involves the identification, assessment, control and mitigation of these risks by a variety of tools and processes. The development and availability of AI tools contribute to the increase in operational risks threatening financial institutions, including cybersecurity risks, third-party risks, fraud risks and disinformation. Technological innovation lowers the barrier to entry for attackers and creates new techniques for the dissemination and automation of online attacks. Generative tools further amplify these potential risks. They enable attackers to develop and deploy effective malicious software at low cost, expanding the scope of attackers and threats. The reliance of AI systems on datasets and third parties amplifies existing challenges and requires attention not only within financial institutions, but also with respect to external partners involved in AI systems development and the broader supply chain.

Cybersecurity threats

AI tools may assist attackers, *inter alia* in the following respects:

1. Disinformation: AI technology is used to create diverse content that may appear authentic, and thus serve parties seeking to harm the financial system as a whole, specific institutions or investors and customers. See elaboration on this risk and its implications in this chapter below.
2. Malicious computer code: Generative AI has been found useful in code development. Potential attackers may, therefore, develop new malicious code or a new version of existing malware more rapidly, and in a manner that will be more difficult to detect.
3. Identifying system vulnerabilities: Attackers can use advanced AI-based tools to identify weak points and vulnerabilities in the attacked party's cyber defense system.

   Cybersecurity attacks may harm the systems of financial institutions systems in several ways, including:

[375] See for example the definition of "Operational Risk" in Proper Conduct of Banking Business Directive 350 of the Banking Supervision Department as "the risk of loss resulting from inadequacy or failure of internal processes, people and systems or from external events. This definition includes legal risk, but does not include strategic risk and reputational risk" (in Section 5 of the Directive).

- data poisoning – tampering with or corrupting training data used or the model itself to skew the algorithm toward a desired outcome;
- data leakage – crafting queries that cause the model to "get confused" by misinterpreting instructions or revealing information it should not disclose;
- model exposure – for example, by using repeated queries to infer the model's structure or parameters; and
- manipulating model outputs—feeding inputs designed to generate a predetermined desired result.

Third-party risks

Whether a financial institution decides to develop an AI solution independently or seeks to purchase it through a third-party supplier, the resource requirements of AI systems will generally increase its dependence on third-party computing infrastructure and data. Therefore, financial institutions must assess and manage the risks associated with an extended supply chain, including potential threats to databases and the data- processing activity of a wide range of suppliers, intermediaries and infrastructure providers. Third parties that are related to the data and its handling can expose institutions to financial, legal and cybersecurity risks. To a large extent, reliance on third parties creates a need for financial institutions to examine how to address the risks discussed in this report regarding activity conducted on their behalf by third parties. This need already exists today in relation to activities undertaken at financial institutions that rely on third parties or outsourcing arrangements, such as cloud data storage and backup. However, as AI usage becomes more widespread, the need to address dependence on third parties will increase, including in areas that are not strictly technical in nature.

Fraud

AI tools may be leveraged for fraudulent purposes in a variety of ways. For example, they may allow fraudsters to impersonate employees or customers of financial institutions more easily than in the past, through voice, image or video imitation, as well as by mimicking other behavioral indicators used to verify identity. Additional potential means include using AI to crack passwords and gain access to customer accounts, or using AI to identify personal information about customers from the web or databases in order to facilitate such access.[376]

[376] As detailed below, fraud risks may also materialize through disinformation, but there is no overlap between the things. There may be fraud attempts using artificial intelligence that are not by way of disinformation, and there may be disinformation activities that are not for fraud purposes (but for example to harm systems by elements hostile to the state).

Although the existing regulatory framework regarding operational risks does not explicitly address AI, its underlying principles can support the safe and responsible adoption of the technology. At the same time, the process of learning and evaluating AI technology progresses, amendments and clarifications to existing regulation may be required if gaps are identified.

Disinformation

Within the broader category of operational risks, the Taskforce found it appropriate to devote particular attention to disinformation risks, which currently give rise to significant concerns, and to examine their potential implications for the financial sector in greater detail.

Manipulation and the systematic undermining of information, often also referred to as "disinformation" or "false information," are widespread phenomena in the digital age, the purpose of which is to influence content disseminated online or through other digital means, and, influence decisions made based on that information accordingly. Even prior to the emergence of AI technology, various interested parties operated to spread disinformation by means of creating false reports and disseminating them through fictitious entities, deliberately mixing authentic with fictitious documents, and directing journalists and public figures to that distorted information.[377]

Although disinformation is not a new phenomenon, the availability and proliferation of AI tools facilitate its creation, increase its scope and its adaptation to recipients, and, as a consequence, its effectiveness in achieving the deceiver's objectives.[378] Generative AI applications enable the creation of new content – texts, images, videos, music and more– based only on a short query, thereby facilitating the creation of synthetic, fictitious content that is difficult and sometimes nearly impossible to distinguish from authentic content. Other AI tools further expand the ways in which such content is distributed.[379] The risk of disinformation is perceived as one of the key risks associated with AI development. In the risk survey conducted by the World Economic Forum for 2023-2024, this risk was ranked particularly high,[380] and there are already illustrations from the past and present of its manifestation in relation to election systems in democratic countries, as well as in the context of political or international conflict.[381]

---

[377] Roy Goldschmidt "The use of artificial intelligence to create and disseminate fake news" Knesset Research and Information Center, p. 2 (21.12.2023).

[378] See *supra* note 55.

[379] See *supra* note 354.

[380] World Economic Forum, Global Risks Report 2023 (January 10, 2024), https://www.weforum.org/publications/global-risks-report-2024/in-full ; It should be noted that this high ranking stemmed, among other things, from concerns about the realization of disinformation risks during important election campaigns scheduled to take place, inter alia in the US and the UK.

[381] For a discussion of recent examples of this phenomenon: Hanna Murphy "The rising threat to democracy of AI-powered disinformation" Financial Times (January 11, 2024), https://www.ft.com/content/16f23c01-fa51-408e-acf5-0d30a5a1ebf2;

Deepfakes, which represent a significant component of the disinformation associated with AI technology, have been identified by financial crime experts as one of the major threats the financial sector may face in the coming years.[382] The term "deepfake" refers to AI-generated content that manipulates images, voices or videos, *inter alia* to falsify reality, including the imitation of a person's conduct or statements, that did not actually occur. Numerous examples of such forgeries exist, and advances in AI, are making them increasingly sophisticated and difficult to detect. At the same time, their creation has become more accessible, easy to execute and cost-effective.[383] Such content could be produced outside the country's borders by criminal or hostile actors.

To what extent is disinformation risk relevant to the financial sector and how can it be addressed?

The decisions of the general public's regarding their financial activity, as well as market trading and the operations of large financial institutions, rely on access to reliable information in large volumes and at high frequency, alongside public trust in institutions and the financial system. Consequently, these areas (like other activities based on the collection and processing of extensive information) are particularly vulnerable to disinformation. The realization of large-scale disinformation risk could not only result in financial losses for investors and customers, but could also pose a threat to systemic stability. In recent years, there have been instances of AI being used to generate content that caused significant damage in the financial sector. A well-known example occurred in 2013, when a fake post on the social network X (Twitter at the time) claimed that an explosion occurred at the White House, leading to a sharp drop in U.S. stock market indices. The post triggered massive automated securities sales through algorithmic trading systems, resulting in losses of billions of dollars.[384] Similarly, in May 2023, fabricated images of smoke rising near the Pentagon circulated on X. In the minutes after the information spread on social networks (and in some cases, even through traditional media), the U.S. stock markets experienced

Tate Ryan-Mosley "How Generative AI is boosting the spread of disinformation and propaganda" MIT Technology Review (October 4, 2023), https://www.technologyreview.com/2023/10/04/1080801/generative-ai-boosting-disinformation-and-propaganda-freedom-house/.

382 https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime.

383 Examples of Deepfakes made using artificial intelligence include recordings or photos allegedly of leaders and other well-known people while falsifying the content. See elaboration in this context: Meredith Somers "Deepfakes, explained" MIT (July 21, 2020), https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained; Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab, Publications Office of the European Union, Europol (13 March 2024), https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes.

384 Peter Foster, 'Bogus' AP Tweet about Explosion at the White House Wipes Billions Off US Markets, The Telegraph (23.04.2013), https://www.telegraph.co.uk/finance/markets/10013768/Bogus-AP-tweet-about-explosion-at-the-White-House-wipes-billions-off-US-markets.html.

notable declines.[385] It should be noted that disinformation does not necessarily entail the dissemination of false and fake information. In some cases, attempts have been made to present accurate information in a way that AI systems interpret differently than expected. For example, publicly traded companies have sometimes avoided using words that AI-based algorithmic trading systems relying on natural language processing might interpret as negative, in order to improve their perceived status assessment in the eyes of the market and traders.[386] As stated, although disinformation risk does not stem exclusively from AI, advances in AI technology has greatly enhanced the ability to easily generate false content, at a quality level that makes it very difficult to detect it as such. However, the severity of disinformation risk is influenced by additional factors, such as the extent of social network usage and the ease with which false content can be disseminated through them, the public's reliance on these platforms, the spillover of information to other channels such as traditional media, and the level of trust the public places in this content. Hence, addressing disinformation risk may require taking measures at multiple levels, such as: obligating social networks to monitor content; granting authorities the power to remove harmful content; and, in certain case, requiring disclosure of information sources.

Addressing disinformation risk in the financial field

As noted above, addressing disinformation risks is partly integrated into the financial regulators' strategies for addressing cybersecurity risks. For example, a circular issued by the Capital Market Authority defines "cyber risk" to include unauthorized use of identity (impersonation),[387] while a "cyber event" is defined as an attack on systems or other technological assets belonging to an institutional body that could compromise, *inter alia*, the integrity of its information. These risks often involve, to one degree or another, disinformation operations by attackers aimed at harming the institution – such as impersonating a customer or supplier or manipulating existing information in its systems using cyber tools to disrupt operations.

The circular outlines various mechanisms to address these risks, including setting procedures for user and authorization management, and implementing controls and monitoring operations and queries in systems that manage sensitive customer information or that are highly exposed to unauthorized activity, with the goal of detecting impersonation-based breaches. Institutional bodies are also required to maintain cyber-event preparedness and management plans, covering unauthorized access, information

[385] https://edition.cnn.com/2023/05/22/tech/twitter-fake-image-pentagon-explosion/index.html?force_isolation=true.
[386] Tel Aviv University Report, pp. 21-22.
[387] Institutional Bodies Circular 2016-9-14: Cyber Risk Management in Institutional Bodies, Capital Market, Insurance and Savings Authority (31.8.2016).

leakage, impersonation and fraud, and to conduct periodic penetration tests, including social engineering tests, impersonation and phishing exercises.

Similar principles and mechanisms for addressing disinformation and cyber risks are reflected in a proper conduct of banking business directive issued by the Supervisor of Banks concerning cyber defense management.[388] Importantly, AI can also enhance cybersecurity by providing advanced defense capabilities, such as detecting connections, linking sequences of operations into one sequence, developing more secure software, and accurately and rapidly identifying anomalies in processes – capabilities that should be considered for integration into the aforementioned mechanisms.

Accordingly, addressing disinformation risks can, at least in part, be managed within existing frameworks for dealing with cybersecurity risks of all types, or through separate regulation.[389] However, even in the cybersecurity domain, AI fundamentally changes the rules of the game. AI accelerates and amplifies attack processes, more quickly identifies system vulnerabilities, and through them, systematically disrupts information, simulates human behavior or impersonates targets in ways that appear more credible than current methods. In effect, AI increases both the rate at which threats emerge and the urgency with which they must be addressed, while also enhancing automated attack capabilities.[390] The capabilities of AI and its rapid development, therefore, call for a reassessment of the effectiveness of existing mechanisms for addressing cyber attacks. Additionally, traditional approaches to cybersecurity risks focus primarily on potential attacks against financial institutions themselves, rather than on risks stemming from the dissemination of market disinformation, which could have an indirect impact on financial activity. For this reason, as well, existing mechanisms for managing these risks should be reexamined. At the core of AI technology lies vast amounts of information, and damage to this data (such as the dissemination of market disinformation or deliberate action undermining data reliability) will inevitably affect the outputs generated by AI systems. Among possible disinformation risks regarding financial activities, the Taskforce believes that the following issues warrant special regulatory attention:

---

[388] Proper Conduct of Banking Business (9/21): Cyber Defense Management, Supervisor of Banks (30.09.2021).

[389] See, for example, the Digital Services Act legislation in the European Union which is intended to deal with the dissemination of disinformation on various platforms, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), https://eur-lex.europa.eu/eli/reg/2022/2065/oj.

[390] See remarks of the Head of the National Cyber Directorate at the Cybertech Next conference "Artificial Intelligence Changes the Rules of the Game in Cyber", National Cyber Directorate website (24.11.2019), https://www.gov.il/he/pages/nextec2019?fireglass_rsn=true#fireglass_params&tabid=aab55ae619fd94f7&start_with_session_counter=9&application_server_address=isagovil-2-me-west1.prod.fire.glass.

- Financial activities involving rapid decision-making by the general public. These activities are characterized by the ability to quickly withdraw or divert funds, and are often conducted without the guidance of a professional financial intermediation (who might otherwise "cool down" hasty or ill-informed actions). Examples include investing in securities or mutual funds without advice and depositing money in savings, current, or payment accounts. The associated risk increases when the public relies on information from unreliable sources, such as social media or unverified websites. In this context, AI could be used to bias decisions toward investment channels that deliberately favor certain market participants or to deliberately cause harm to investors and the financial system.[391]
- Financial activities driven by continuous information-based pricing. The most prominent example of this is stock exchange activity, where securities are traded continuously based on information disclosed by the traded companies. Disinformation about a certain company or an external event with ostensible macroeconomic implications could influence investor behavior and move securities prices (see, for example, the previously mentioned cases of fake White House and Pentagon images). Another example is broker-dealer activity, which also relies on pricing that is highly sensitive to information.
- Financial institutions exposed to interconnectedness and contagion risks. The global financial crisis in 2007-2009 exposed the risks for individual financial institutions and the financial system as a whole resulting from direct and indirect linkages between numerous entities. In addition, this crisis highlighted the risks stemming from information uncertainty regarding financial institutions. For example, the collapse of the investment bank, Lehman Brothers, which occurred shortly after other financial institutions refused to continue dealings with it, is attributed *inter alia* to a lack of complete information and the inability to assess the risks associated with its operations. Disinformation could amplify contagion and interconnectedness risks, as it affects the information held by financial institutions about other institutions and assets. Furthermore, reliance by financial institutions on identical AI systems using similar algorithms and drawing similar (mistaken) information could lead to a chain effect that could undermine the integrity and stability of the entire market. As previously stated, this could also occur as a result of human error, entirely unrelated to AI use. However, the speed that essentially defines automated mechanisms operating on external information increases both the likelihood of such events and amplifies their potential impact, since

[391] Tel Aviv University Report, at p. 22.

it is difficult to interrupt the chain of events in time. In this context, widespread use of identical or similar AI-based models by traders could create a "feedback loop", potentially resulting in extreme market volatility or even flash crashes.[392]

- Financial institutions using AI systems for rapid financial operations. Similar to the risk described above concerning the general public's financial activities, a comparable risk may arise regarding the use of AI systems by financial institutions. Examples include algorithmic securities trading by institutional investors or portfolio managers using AI to make investment decisions on behalf of their clients. The speed, automation and reliance on external information in these operations contribute to the heightened risk described earlier.
- Use of AI to circumvent security mechanisms. AI could facilitate impersonation employed to gain unauthorized access to customer accounts, circumvent anti-money laundering regulations that rely on verifying identification, or avoid other financial supervision rules. Such AI-driven impersonation attempts may target both the general public for fraudulent purposes and financial institutions.[393] In this context, AI contributes to an increase in cybersecurity risks.

Since disinformation is not new, and financial regulators have previously addressed it through various means.

For example, in the context of securities trading on public markets, several tools are already in place to address the effects of unreported or unreliable information that may trigger "rumor mills" and sharp fluctuations in securities prices. One such tool is a temporary trading halt that can be imposed by the Israel Securities Authority under the Securities Law, or by the stock exchange itself by under the exchange's rules and regulations. Trading may be halted in cases such as a material breach of a company's reporting obligations, the publication of erroneous information concerning the price of a security or group of securities, the dissemination in any manner of material information relating to a company including its controlling shareholders or its securities, or the occurrence of a material macroeconomic event that could affect trading in multiple securities.[394]

---

[392] Tel Aviv University Report, at p. 22, see also: Ghiath Shabsigh and El Bachir Boukherouaa Generative Artificial Intelligence in Finance: Risk Considerations (International Monetary Fund Fintech Notes (August 2023), https://www.imf.org/en/publications/fintech-notes/issues/2023/08/18/generative-artificial-intelligence-in-finance-risk-considerations-537570.

[393] See regarding a police investigation in Hong Kong on suspicion that a financial entity employee fell victim to fraud in which he transferred a sum of 25 million dollars, after joining a video conference call, in which his co-workers allegedly participated, but their figures were deepfake forgeries. Heather Chen and Kathleen Magramo "Finance worker pays out $25 million after video call with deepfake 'chief financial officer' "CNN (February 4, 2024), https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html.

[394] § 38A of the Securities Law and the Fourth Part of the Stock Exchange Rules.

Another tool is a mechanism that suspends trading in a security for a set period, allowing the market to assess the information and its pricing implications.[395] An additional layer of response lies in enforcement, which already enables the prosecution of securities trading offenses, including those committed using AI technology. The primary difficulty in this context concerns investigative powers, which currently fail to provide an adequate response to computer-related offenses. The Criminal Procedure Bill (Enforcement Powers – Service, Search and Seizure), 5783-2023, which seeks to update the search and seizure powers conferred on enforcement authorities in criminal proceedings, may address this gap.

Another means of addressing unreliable information is the publication of warnings and clarifications by regulators and financial institutions. For example, publication of warnings by financial institutions to their customers (such as warnings by credit card companies regarding misuse of personal information) or the publication of clarifications by regulators regarding market conditions or information concerning them.

Since the publication of the Interim Report, incidents of financial fraud involving AI-generated voice and visual impersonation using AI tools increased both globally and in Israel. *Inter alia*, there have been attempts to contact bank customers while falsifying the caller's identity (posing as a bank, credit card company or governmental authority) as well as posts on social media featuring fake videos of public figures intended to prompt investments in certain securities. In response, financial regulators have issued warnings and recommended protective measures against such criminal activity.[396]

In the public comments submitted on the Interim Report, it was suggested that the discussion of cybersecurity risks should be further expanded, and that legislative changes to Know-Your-Customer (KYC) procedures should be considered, since these are particularly vulnerable to deepfake-based fraud. The Taskforce agrees that cybersecurity risks have become increasingly significant in light of AI development and recognizes their relevance to KYC processes. However, it does not consider an amendment of the report on this matter necessary, since the relevant chapter already reflects the heightened importance of cybersecurity risks, and KYC processes are specifically identified elsewhere in the report as an area where focused regulatory attention is required.

[395] Guidelines under the Third Part of the Stock Exchange Rules.

[396] See for example the warning of the American organization FINRA against investor fraud using artificial intelligence: FINRA "Artificial Intelligence (AI) and Investment Fraud" (January 25, 2024), https://www.finra.org/investors/insights/artificial-intelligence-and-investment-fraud.

Operational Risks Including Cybersecurity Risks, Third-Party Risks, Fraud and Disinformation – Recommendations:

The Taskforce's recommendation to financial regulators is to map activities that have particular exposure to disinformation risk and examine whether existing tools for addressing this risk are sufficient or whether refining them is warranted. The regulatory tools proposed for consideration in this context include *inter alia*:

- Directing regulated entities to evaluate disinformation risk, as a component of broader cybersecurity risk, and to review whether the existing tools and controls at their disposal are sufficient to address it.
- Directing regulated entities to establish procedures for preventive activity to address disinformation events, as a component of cybersecurity risk, including through monitoring as early as possible, as well as establishing response procedures for when such events occur (as part of advance preparedness for addressing cyber events and/or emergency situations).
- Determining mechanisms or reviewing existing mechanisms for addressing disinformation risk in relevant institutions, such as directives and procedures dealing with cybersecurity risk among regulated entities, or the trading-halt mechanisms discussed above regarding stock exchange trading.
- Increasing public awareness of disinformation risk, through publishing explanations and warnings, as well as other financial education initiatives.
- Establishing early preparedness measures among financial regulators for addressing disinformation events (similar to early preparedness for conduct in emergency situations). For example, early monitoring of disinformation spread (such as unusual activity on social networks), establishing procedures and work processes concerning the issuance of clarifications over the media by regulators, requiring regulated entities to publish information in response to false information, and so forth.
- Incorporating disinformation risk scenarios into cybersecurity response procedures and/or emergency preparedness procedures, as well as into drills conducted by financial regulators and the financial system to practice emergency response.
- Improving investigation capabilities regarding disinformation resulting from AI activity. Particular importance is placed on this with respect to securities trading activity that may

be particularly vulnerable to disinformation events. In this context, there is scope to further advance current legislative proposals aimed at strengthening the enforcement of offenses under the Penal Law and the Computers Law within the securities sector.

- Reviewing the allocation of powers under general law to remove false content and the appropriate referral to the competent authority in such cases.

Box 4: Securities Trading Offenses Using AI

Is it possible to commit securities trading offenses using AI systems? And can these offenses be enforced in Israel today?

Stock exchange securities trading is highly complex. The exchange allows a virtually unlimited number of participants to operate simultaneously by submitting buy and sell orders for securities. Trading is conducted across a wide range of securities, based on information published by companies. Pricing is continuous and transparent, reflecting securities transactions executed in real time, while clearing—namely, the transfer of securities against monetary consideration—is carried out shortly after each transaction is executed.

AI technology can be exploited by parties to commit trading offenses. For instance, publishing false information generated by AI in a way that appears credible could move securities prices, allowing the party responsible to benefit from the resulting price fluctuations. They may, for example, sell shares of a company at an elevated price, after the stock price reacted positively to false information about a positive development in the company's business. This activity could harm not only specific investors, but the overall integrity of trading on the exchange and public confidence in it as well.

Section 54 of the Securities Law prohibits fraudulent activity in securities. It covers two types of prohibited conduct: inducing the purchase or sale of securities through fraud, and manipulating securities prices by fraudulent means.

The inducement offense prohibits creating a false representation by providing details or concealing material facts, with the intent of inducing investors to engage in securities transactions. The trading fraud offense prohibits actions that undermine the pricing mechanism of securities traded on the exchange. Frequently both types of conduct occur simultaneously. For example, a person who publishes false information may induce investors to purchase or sell securities (violating the first prohibition against inducement), while also influencing securities prices in a fraudulent manner (violating the second prohibition against securities fraud).

Courts have emphasized that the fraud offense may encompass a wide and evolving range of activities. For example, the Supreme Court determined with respect to trading fraud that "life is richer than imagination. It is not appropriate to establish in advance a 'closed list' of situations prohibited under the provisions of Section 54(a)(2) of the Securities Law. Comprehensive definitions should be avoided, which upon their formulation immediately find the next case that does not fall within their scope, but whose conduct is also prohibited. In this living and vibrant field of securities market trading, an inductive

approach is preferable, moving from case to case, and based on accumulated life experience expanding and refining its principled concepts."[397]

Indeed, fraud offenses involving disinformation have been discussed and adjudicated in prior case law. In the *Podim* case, a defendant was convicted of securities fraud for posting false messages on a website's bulletin board under multiple fictitious identities. In the *Dahan* case, five defendants were convicted after purchasing thinly traded shares and subsequently publishing false reports on forums and investment site comments to inflate their prices and sell at a profit Similarly, in the *Holtzman* case, a defendant was convicted for posting a document resembling an online article from the *Globes* newspaper, which contained favorable information about securities he owned, on investor forums with the intent of inducing investors to purchase them.[398]

Similarly, in the U.S., notable securities fraud cases involve disinformation. For example, in the *Burns* case, that dealt with uploading a false report to the U.S. Securities and Exchange Commission's corporate filings website, following which the price jumped and options were exercised at a profit. In the *Craig* case, false "tweets" were posted on fake Twitter accounts impersonating securities firms, with the intent of manipulating the stock price for financial gain.[399]

Accordingly, the fraud offense also extends to operations carried out through AI. The main challenge in this context lies in the investigative process. These offenses can be committed on multiple internet platforms, where vast volumes of content are generated, and can also be committed remotely, including from computers located outside of Israel. As discussed in the chapter dealing with AI-driven disinformation, an additional difficulty arises from the quality of the forgery, the ease with which it can be produced, and the challenge of identifying its source.

In this context it is crucial to update investigation powers regarding computer materials to ensure effective enforcement in the future, the Criminal Procedure Bill (Enforcement Powers – Service, Search and Seizure), 5783-2023, is under parliamentary consideration. This bill contains, *inter alia*, search and access powers for computer materials that do not necessarily exist at this time. This is designed to enable investigative bodies to address complex crimes, while providing a response to dynamic technological developments. For example, the bill provides for covert remote searches of computer materials, when necessary to address the speed and sophistication of offenses committed using sophisticated technological

[397] Criminal Case 5052/95 **Vaknin v. State of Israel**, PD 50(2) 642, in paragraph 29 of the judgment.

[398] Criminal Case (Tel Aviv) 70571/01 State of Israel v. Podim Tal (Nevo, 23.3.2002); Criminal Case (Tel Aviv) 6952/06 State of Israel v. Avi Dahan (22.22.2010); Criminal Case (Tel Aviv) 31515-05-16 State of Israel v. Sharon Holtzman (11.12.2016).

[399] https://www.sec.gov/litigation/litreleases/lr-24557 ; https://www.sec.gov/litigation/litreleases/lr-24481.

means, including AI. Given these considerations, granting such powers is essential to ensure effective enforcement in the field of securities offenses.

Part D

Specific Areas Examined by the Taskforce

This chapter focuses on three specific areas of activity examined by the Taskforce: investment advice and portfolio management, credit, and insurance underwriting. For each area, the financial regulators assess how they anticipate AI applications may be used within their respective domains, describe what they consider the relevant regulatory framework to be, highlight issues requiring attention, and provide preliminary recommendations.

Investment Advice and Portfolio Management

Potential Uses of Artificial Intelligence in Investment Advice and Portfolio Management

Artificial Intelligence may serve a wide range of functions in investment advice and portfolio management. Its potential applications span virtually all professional activities related to providing advice or managing client assets. For example, artificial intelligence can assist in the process of clarifying client needs, *inter alia* through online questionnaires and analysis of the responses; it can support the design and presentation of investment tracks or investment recommendations tailored to the client. In portfolio management, AI systems can directly participate in portfolio selection and the management of the client's funds. In this context, artificial intelligence may contribute to core functions, such as selecting investments and, in some cases, even executing them.

Beyond these activities, artificial intelligence may also serve other investment licensees' business functions, such as marketing activity and customer relations (for example, identifying prospective clients and proposing suitable services), as well as compliance functions (such as compliance with anti-money laundering obligations or supporting risk-management processes). Similar to other financial activities, the extent to which artificial intelligence is used may differ from case to case, and the level of human involvement will vary accordingly. For example, algorithms used in investment selection may function merely as a decision-support mechanism for a portfolio manager, or may be employed as the primary mechanism for making investment decisions.

This chapter covers both long-standing, relatively simple uses of artificial intelligence and automated decision-making, such as customer service algorithms that generated predetermined alternatives, as well as more advanced emerging applications, such as machine-learning-based applications and Large Language Models (LLM) used to assess client needs. As becomes evident throughout this chapter, the regulatory challenge lies in addressing this broad spectrum of uses - both basic and complex - alongside the current transition between different generations of technological systems. The new generation of systems is expected to greatly expand potential applications, addressing issues that emerged in earlier technologies, while raising new challenges that require a regulatory response.

Below is a detailed description of potential applications of artificial intelligence in investment advice and investment portfolio management, based on surveys and examples from other jurisdictions.[400] Potential applications of artificial intelligence in investment advice and portfolio management activities can occur throughout the entire chain of activity:

Potential applications of artificial intelligence in investment advice and portfolio management activities can take place throughout the entire lifecycle of the services, for example	
Advertising and market	identifying prospective clients, generating interest among clients, personalizing marketing, and reaching a larger customer base
Client onboarding process	introducing new clients to service providers' systems, familiarizing clients with these various services, while ensuring compliance with legal requirements including KYC requirements, and signing the agreement for providing services
Remote client identification	assisting in the transition from "face-to-face" identification to remote identification through facial recognition systems and additional identification markers
Client needs-clarification process	conducting the process that forms the basis of investment advice and portfolio management activities - clarifying with clients their investment objectives, financial situation, preferences, and other relevant circumstances, for the purpose of tailoring investment activity to their needs and instructions. Artificial intelligence makes it possible, *inter alia*, to collect, aggregate and analyze far more information about a client, such as their purchasing history and financial behavior. As a result, the assessment of client needs can be carried out through data-driven insights rather than just through the client's direct responses to questions
Investment analysis and selection	scanning and analyzing multiple investment options and choosing among them, as well as determining the optimal way to carry out ancillary operations (currency conversion, broker selection)

[400] See Tel Aviv University Report, at pp. 17–19; SEC Proposal from 2023, at pp. 15-21; OECD 2023 Report, at pp. 13–18; IOSCO Report from 2021, at pp. 6-8; Blake Schmidt and Amanda Albright, AI Is Coming to Wealth Management. Here's What That Means, Bloomberg Markets (April 21, 2023), https://www.wealthmanagement.com/technology/ai-coming-wealth-management-here-s-what-means.

Investment recommendations	recommending an investment portfolio tailored to client needs, based on weighing the information that emerged during the needs-clarification process (investment scope, investment objectives, risk profile, and so forth) alongside the available investment options
Investment decisions	providing decision support or decision-making functions n portfolio management, as well as executing transactions. In particular, these systems can autonomously rebalance a client's portfolio, issue alerts regarding the investment portfolio, make adjustments for new market conditions, and more
Documentation, reasoning, updates and client reports	documenting advice and investment operations, conversations between the client and service provider, the reasoning behind the actions performed by the service provider, client consent for certain actions (such as executing special-risk transactions), and regular disclosures to clients. Beyond compliance with legal requirements, AI systems can be employed to conduct additional client outreach, provide updates on their status not necessarily required by law as well as reviews of their investments
Technical support and customer service	automizing and providing rapid responses to client inquiries, especially for routine matters; supporting personalized client service by reviewing client details, their previous inquiries, and account status; assisting human staff by quickly presenting the resources and information needed to resolve client problems; facilitating referrals to the appropriate party while saving time for the service provider and client; and other similar functions
Compliance and risk management mechanisms	detecting fraud and unusual activity, ensuring compliance with legal requirements, and implementing money-laundering controls

The General Legal Framework

The Investment Advice Law establishes a regulatory and oversight regime for the three types of services mentioned. At the basis of this regime stands a licensing requirement imposed on those seeking to engage in these activities, followed by a series of additional obligations applied to the licensee to ensure client

protection, such as the fiduciary duties of trust and care, an obligation to prepare a written agreement, and an obligation to adapt the service to client needs after clarifying them.[401]

The Investment Advice Law, along with its regulations and the Authority's directives under it comprises a comprehensive regulatory framework. The primary objectives of this regulation are to ensure the quality and reliability of services provided to the investing public, to regulate the operations of licensees and their interactions with clients, to mitigate potential risks, including those arising from conflicts of interest; to create opportunities for generating returns on investors' funds while tailoring investments to each investor's needs, and to broaden public participation in the capital market reflecting the state's economic interest in the development of capital market activity.[402]

The key distinction between investment advice and portfolio management lies in the degree of decision-making authority held by the client or portfolio manager respectively. In investment advice, the advisor provides recommendations, while the client decides whether to act on them. In contrast, in investment portfolio management, the portfolio manager has discretion to execute transactions in the securities and financial assets held in the client's account, on behalf of the client through a power of attorney to act in the account, in accordance with the investment policy agreed upon with the client. Accordingly, the Law imposes additional obligations on portfolio management related, *inter alia*, to requirements for the separate holding and management of client assets, and prohibitions of using client assets for any purpose other than executing transactions on the client's behalf in accordance with the agreed-upon terms.[403]

Reflecting this distinction, Section 1 of the Investment Advice Law defines investment advice activity as "the provision of advice to others with respect to the profitability of an investment, a holding, a purchase or sale of securities or financial assets." Investment portfolio management, on the other hand, is defined as "the discretionary execution of transactions, in the accounts of others." [404]

Regulatory Engagement to Date

The extensive regulation applicable to investment advice and portfolio management has evolved over decades, yet it has yet to explicitly reference artificial intelligence. Largely written in the 1990s, the provisions of the Investment Advice Law, presume human engagement, typically through in-person

[401] Sections 11–21 of The Investment Advice Law.

[402] See for example The Investment Advice Law, Investment Marketing and Investment Portfolio Management Bill, 5755-1995, Bill 2320, p. 92; HCJ 1715/97 **Investment Managers Association v. Minister of Finance**, PD 51(4) 367 (1997), at p. 391; Class Action 47119-12-15 **Efrimov v. USG Capital Ltd. et al.** (June 24, 2019), in sections 26–27 (hereinafter: the "**Efrimov case**").

[403] Sections 22–24 of The Investment Advice Law.

[404] Section 1 of The Investment Advice Law. Activity of "Investment Marketing", which we will not deal with in this report, is also defined in Section 1 of the Law as "providing advice to others regarding the worthwhileness of investment, holding, purchase or sale of securities or of financial assets, when the advice provider has an affinity to the financial asset".

meetings between licensees and their clients, needs assessments conducted either face-to-face or by telephone, and the provision of investment recommendations and portfolio management in a traditional manner. In recent years, the Israel Securities Authority has advanced legal adjustments to permit the delivery of investment advice and investment portfolio management services through sophisticated technological means, reflecting the substantial changes of how the public consumes financial services in light of technological developments.[405] Among these measures, the Israel Securities Authority established the Directive to Licensees for Providing Services Using Technological Means, which also addresses artificial intelligence activity, which we will expand upon below.

(1) Unsupervised Computerized Portfolio Management Activity

The Israel Securities Authority's first significant engagement with artificial intelligence in the context of investment advice and portfolio management arose in response to a growing problem. Starting around 2014, against the backdrop of a low -interest-rate environment and widespread online access to clients, the practice of offering unregulated investments to the public at large gained traction. These offers, appearing across various financial channels, also included a category of computerized portfolio management services often referred to as "robotic trading" or "algo-trading." In this context, unsupervised entities approached the public, promoting investment of their funds through algorithms, and in their presentations to investors, claimed that the investment would be managed in a sophisticated, secure manner and would yield exceptional returns.

The Israel Securities Authority's activity regarding such unsupervised investments included, *inter alia*, issuing warnings to the public and promoting financial education, identifying unsupervised activities that violated the law, demanding that the responsible parties cease such activities, taking enforcement actions, and enacting legislation to regulate previously unsupervised activities Specifically, in the context of "robotic trading" or "algo-trading", the Authority took action in 2015 and 2016 against a number of ventures operating without oversight.[406]

---

[405] See for example the Bill for Encouragement of Activity in the Capital Market (Legislative Amendments), 2024 (June 26, 2024), regarding general investment advice.

[406] See Israel Securities Authority "Warning to the Public regarding Computerized Investment Management ("Robotic Trading", "Algo Trading" etc.) (2.8.2015). The most well-known case among these ventures was that of Utrade which raised funds from hundreds of investors after offering them a computerized trading service based on software including an algorithm for selecting investments, which focused on trading in contracts for difference on a variety of underlying assets. The Israel Securities Authority stopped the company's activity, which was done without supervision, and opened a criminal investigation into its matter. The proceeding ended with the conviction of the company's owner and CEO, inter alia, for false representations to the client, theft of funds, investment portfolio management without a license, and obstruction of justice. See also Criminal Case 60588-12-18 **State of Israel v. Talmor et al**. (13.9.2022) (hereinafter: "the Talmor case").

One of the arguments raised by these ventures, and rejected by the Israel Securities Authority, was that they were not subject to the Investment Advice Law, because a computer, rather than a human, exercised the discretion and executed the transactions. The Authority countered that the fact that investment decisions were made by computer software does not alter the fact that transactions were carried out at the discretion of the service provider, through that software, and without the discretionary involvement of the client. Therefore, such activity is subject to licensing and the other requirements set in the Law with respect to portfolio management. Another argument raised by ventures at that time was that the activity engaged in was not portfolio management, but rather trading platform activity, for which, at the time, pending regulation had not become effective. In response to this claim, the Authority published a legal position clarifying when and under what conditions the use of algorithmic trading software in a trading on a dealer-operated trading platform is to be regarded as providing a platform for independent trading by platform clients, as opposed to portfolio management or investment advice.[407]

The Israel Securities Authority's position that the Investment Advice Law applies even to computerized portfolio management activity was accepted in the courts in both criminal and civil proceedings conducted against unsupervised ventures. For example, in the *USG Capital* case regarding the issue of computer discretion in securities trading, the ruling stated:[408]

> "Indeed, the question of exercising 'discretion' in executing investment decisions is more complex when computer-executed transactions are involved. Executing transactions through an algorithm is a technological innovation of recent years. 'Classic' portfolio management is conducted by depositing the client's funds with another party who exercises discretion and makes independent decisions about where and how to invest the funds. The Investment Advice Law enacted in 1995 envisioned the traditional trading that took place directly by human traders. At the time the law was enacted, it was not possible to anticipate all the technological developments that have occurred in recent years, enabling technological solutions in financial services as well.
>
> […]

---

[407] Israel Securities Authority "Legal Staff Position: Offering Trading Services Using an Algorithm on a Trading Platform" (14.9.2016). The regulation applicable to trading platforms restricts them from performing investment advice and investment portfolio management activities, and therefore the need arose to clarify what use they can make of algorithmic trading software without violating these restrictions. See Regulation 10(a) and 10(c) of the Securities Regulations (Trading Platform for its Own Account), 5775-2014, and the conditions established in the said position.

[408] The Efrimov case, in sections 30 and 33 of the judgment. See also the Talmor case in sections 320-323, 328-332 of the judgment.

I believe that the discretion regarding the choice between different transactions must necessarily lie with some human agent– the client who owns the money or the company through which they invest it and which provided them with the software that makes the investment decisions. Therefore, if the discretion is not the client's, then it necessarily lies with the company and vice versa. Any other determination would create a 'vacuum,' and lead to a situation where the discretion ostensibly belongs to no human agent - and therefore no party bears responsibility for it.

I also do not accept the argument that the use of the term 'discretion' requires human intervention before each concrete decision. First, even when computer software is involved, that software was designed by some human agent, which is the one that 'instructed' it in the way of making decisions. The computerized algorithm 'makes' the investment decision based on examination and analysis of data fed into it. However, even if the concrete investment decision is made automatically and presumably 'without the human touch,' in practice, the decision is made according to a decision-making process shaped in advance and based on predetermined parameters. Moreover, as clarified above, the choice of the algorithm that will make the decisions is also a human decision, for which 'discretion' is required. Therefore, the question is not whether human discretion was exercised, but whether the discretion should be attributed to the client or to the company."

The phenomenon of unsupervised ventures offering computerized portfolio management services largely subsided in 2016, yielding several key lessons with respect to the use of artificial intelligence.

One lesson highlights the importance of the principle of "technological neutrality," which means, *inter alia*, that the circumvention of legal provisions solely because a certain regulated activity is conducted using different technology is unwarranted. Although it will often be appropriate to make adjustments to legal provisions to promote technological innovation or address it specifically, the general presumption is that the interests relating to protecting the public's welfare should be preserved regardless of the technology through which the regulated activity is conducted. This principle of neutrality has additional importance, since formulating regulation according to the existing technology at a given time may incentivize the creation of technological workarounds, and thereby also grant certain companies conducting the same activity an advantage solely due to technological differences, which do not affect the nature of the activity.

An additional lesson relates to the potential misuse of new technology for marketing and "branding" purposes, when in practice, such advertising lacks substance. This lesson is relevant to supervised and unsupervised activity alike, and from it, arises the need to clarify to the client in a reliable manner what the technology being used is, including all its advantages and disadvantages.[409]

(2) Online Services Directive

Alongside its efforts to address the phenomenon of unsupervised investment advice and management services, the Israel Securities Authority increasingly recognized the need for adjustments and clarifications to the legal framework are needed in order to facilitate the use of technological means in the provision of investment advice and investment portfolio management services. The Authority viewed the use of technology means as having many potential advantages, such as enhancing service quality, efficiency, and cost-effectiveness, as well as expanding access to a wider segment of the public, including those able to invest only modest amounts, who licensees may otherwise be reluctant to serve for economic reasons. These adjustments and clarifications were reflected in the "**Directive to Licensees Regarding the Provision of Services Using Technological Means**" issued by the Israel Securities Authority in 2016 (hereinafter: "**the Online Services Directive**," or, in this chapter "**the Directive**").[410] Although the Directive does not explicitly refer to artificial intelligence, its broadly framed provisions allow for its application in certain contexts involving the of use of AI technology. The Directive essentially addresses two interrelated topics: The first concerns processes carried out by licensees without human involvement. For example, the clarification of client needs may be conducted through an online questionnaire rather than through conversation. This shift to digital processes raises questions related to the transition to digital activity; for example, whether the client fully understands the implications of the information they provide in the online questionnaire, or whether they may act hastily on an investment recommendation without adequately considering its significance. The second area the Directive addresses relates to the use of algorithmic applications by licensees in various aspects of their operations. In this context, questions arise that are directly relevant to artificial intelligence, such as the licensee's responsibility when relying on an algorithm to make investment decisions on behalf of the client, and under which circumstances human involvement by the licensee is required. Viewed in this

[409] See in this regard the sub-chapter dealing with misleading marketing disclosure regarding artificial intelligence in the chapter on Notification and Disclosure above.

[410] Israel Securities Authority, Directive to License Holders for Providing Services Using Technological Means (23.8.2016). For an updated version of the Directive see https://www.new.isa.gov.il/images/Fittings/isa/asset_library_pic/al_lobby/al_lobby-6283705a39f5d/horaa7822.pdf.

light, the Directive may be regarded as a pioneering step in Israeli financial regulation in addressing some of the core issues examined in this report.

The Directive focuses on the three principal stages in the process of providing investment advice and portfolio management services.[411] It proceeds from the premise that these stages may be conducted through advanced technological means, subject to addressing the issues arising from such use. The Directive identifies the following three stages:

1. Clarifying client needs and entering into an agreement - conducting the process online, without a face-to-face meeting and without full interaction between a human actor, who is a licensee and the client.
2. Determining the investment policy and client's risk level - processing the data received from the client, for the purpose of determining the risk level and investment policy, through an algorithmic application.
3. Providing ongoing service - generating investment recommendations and delivering them to the client (in cases of investment advice), and making investment decisions (in cases of portfolio management), while using an algorithmic application, and maintaining ongoing contact with the client online including updating client needs from time to time.

The Directive is structured modularly, so that a licensee can use technological means in one or more of the aforementioned stages in the service provided by them. For example, a licensee may choose to conduct the contracting and needs assessment stage through a traditional face-to-face meeting with the client, and conduct the following two stages through technological means. Moreover, the nature of AI use may vary from case to case, ranging from basic or highly sophisticated uses. "Providing services using a technological means" was defined generally as "providing services online, or using an algorithm for the purpose of making decisions, without interaction with a human agent" in each of the aforementioned stages.[412]

The Directive adopts a principles-based approach presenting a series of obligations, but leaving discretion regarding how to implement them to the licensee. These are, however, mandatory

[411] In addition, the Directive on Online Services deals with two services which, due to their nature, are provided online: independent trading advice service (formerly called "signals service") within which investment recommendations are dispatched to clients on an ongoing basis, and social trading services where a client can view transactions executed by another person on a social trading platform or automatically follow that person's investment portfolio in managing their funds. In both said types of services, use can be made of artificial intelligence, but the Directive does not deal specifically with this possibility but rather regulates other aspects regarding these services.

[412] Section 1 of the Directive on Online Services.

requirements providing for the proper use of technological means by licensee holders, and not merely recommendations.

The key issues addressed in the Directive and the obligations it establishes include:

Proficiency - The Directive establishes that, in light of the fiduciary duties of trust and care applicable to licensees, they may employ technological means only after acquiring sufficient proficiency and understanding of those means and only after determining that their use allows the services to be provided in a manner that will neither harm their clients nor compromise their obligations under the law.[413] The explanatory notes to the Directive clarify that this proficiency requirement applies already at the stage of the licensee's decision to adopt, acquire or develop a technological means, and that it becomes particularly significant in the second and third stages of service provision, where the technological means is also used for decision-making purposes.

Investment advice or portfolio management licensees may be either entities or individuals. When an entity is involved, there is an obligation to appoint an individual licensee to be in charge of providing services as stated (hereinafter: "**Appointed Licensee**"). Beyond the control and liability aspects, this requirement specifically establishes that with respect to the determination of investment policy and risk level as well as to the provision of service on an on-going basis (the second and third stages), the Appointed Licensee must be familiar with and understand the methodology underlying the algorithm so as to be able to make informed decisions regarding its use and the explanations given to the client regarding is operations, and to test whether the algorithm is operating as expected.[414]

Controls - The obligation to maintain controls ensures that the appropriateness of using the technological means, including detecting and preventing potential failures is also presented as part of the licensee's fiduciary duties. Beyond this general obligation, the Directive specifically stipulates the need to take all necessary measures with respect to information security, privacy protection, client identity verification and the detection of failures in the needs assessment process.[415] Alongside the Appointed Licensee mentioned above, the entity's board of directors is required to approve the technological means and its use as well.[416]

---

[413] Section 2.1 of the Directive.

[414] Sections 15–16, 19 of the Directive. In the explanatory notes to the Directive, it is noted that all licensees involved in providing the service are required to know and understand the methodology at a level adapted to the stage of service provision in which they are involved, although this is not required to be the enhanced level of understanding of the Licensee in Charge.

[415] Sections 2.3, 14 of the Directive.

[416] Section 6 of the Directive.

Liability - The Directive emphasizes that using a technological means does not derogate from the licensee's liability. This liability is not only reflected in the proficiency requirement described above, but in explicit provisions prescribing that the licensee bears liability toward a client for damage resulting from the use of a technological means.[417] This provision prevents the licensee from claiming that a third party, rather than he, is responsible for damages caused to them, ensuring customers a clear and unequivocal "address" for recourse.[418] In addition to this, the licensee is required to ensure that their insurance coverage includes the insurer's liability for the licensee's use of the technological means.[419]

Explanation to the client - The Directive places special emphasis on the level of explanations owed the client, against the backdrop of the transition from consuming services in a traditional manner to consuming services through technological means. The obligation to provide explanations to clients, in a clear manner, is also intended to prevent the use of false representations of technological capabilities for marketing purposes. The Directive includes requirements to provide explanations to the client before initiating service provision, including incorporating specific reference to this matter in the service agreement, and in each one of the stages of service provision. Here as well, the obligation is framed in general terms. The Directive clarifies that the explanations should cover the nature of the service and its special characteristics, the potential risks arising from it, and the methodology of the technological means used to determine investment policy and make investment decisions.[420] It is further clarified that this requirement does not extend to the disclosure of trade secrets. Its purpose, rather, is to provide a high-level principled explanation, focusing for example, on the fact that the algorithm processes information based on the client's responses and therefore, requires careful and complete answers to the questionnaire, or that it may generate recommendations limited to only certain assets or asset classes.[421]

Human involvement - The Directive addresses situations in which involvement of a human agent is mandatory. For example, human involvement is required in any situation in which a client requests a response from a human agent, or in any circumstance in which a client requests to deviate from an investment policy other than the one determined for them by the technological means.[422] In such cases,

---

[417] Section 4 of the Directive.
[418] The licensee may contractually establish liability arrangements with a third party providing them with a technological system, and thus obtain indemnification in case of damage, but such arrangements do not affect the extent of their direct liability towards the client.
[419] Section 5 of the Directive.
[420] Sections 10, 11, 12, 17 and 20 of the Directive.
[421] Explanatory notes to the Directive.
[422] Sections 3, 18 of the Directive.

the licensee must make both technical and professional human support available to the client, within a reasonable timeframe and in a reasonable manner, at a predetermined hour.[423]

Broad implications of using the algorithm - As a rule, the Directive does not address the broad implications that may result from the public's transition to online services. There is one exception to this in the Directive, reflecting the risk that portfolio management may be carried out in an identical manner for a large number of clients through the use of algorithmic application. In such circumstances, portfolio management activity may in essence resemble mutual fund management, which involves the collective investment of the investing public's funds under the management of the fund manager. In mutual fund activity, investors' funds are deposited in a trust account of the fund and are invested collectively, as a single pool, in accordance with the fund's investment policy.

The concern arising in this context is that such activity could lead to the circumvention of restrictions established under the *Bachar Reform*, in which banks were obligated divest ownership of the mutual fund firms they owned. Since the employment of an algorithmic application to conduct portfolio management may turn the activity from a personal service to an off-the-shelf product equivalent to mutual funds, the Directive established that portfolio management by a company related to a banking corporation, is limited to investments in mutual fund units, deposits and cash.[424]

Implementation of the Directive and Artificial Intelligence Uses in Investment Advice and Portfolio Management in Israel

Since its adoption, the Online Services Directive has been implemented by supervised entities. In practice, it has proven useful, mainly with respect to the first two stages described above - the onboarding stage of needs assessment and the stage of determining the client's investment policy and risk profile. *Inter alia*, supervised entities make use of online questionnaires and chatbots to conduct client onboarding and to integrate them into the supervised entity's systems.

The use online services in the third stage, the provision of ongoing service, is currently still limited. In practice, supervised entities use algorithmic applications primarily as decision-support tools, while preserving human involvement. In contrast, cases in which licensees perform the core activity of generating investment recommendations and delivering them to the client (in the case of investment advice) and making investment decisions (in investment portfolio management), exclusively through AI applications, are rare.

[423] Section 3 of the Directive.
[424] Section 21 of the Directive and explanatory notes to the Directive.

The impression is that both licensees and clients find value in personal interaction and are not willing to give it up completely. For licensees, personal interaction is a means to learn more about their clients, create a relationship and build trust with them, and assess what investment needs can be addressed. From the clients' perspective as well, a fully computerized response may seem lacking, while personal interaction can add value in terms of convenience and trust. It may, therefore, be assumed that considerations of expertise and cost, together with the inclination of both licensees and their clients to retain a degree of personal interaction, are what drive the popularity of hybrid models at this stage.

It can therefore be concluded that, to date, the primary use of artificial intelligence and automated decision-making in investment advice s in Israel has been in conducting the needs assessment and in determining the investment policy and the client risk profile. While some licensees also employ artificial intelligence in the actual provision of investment recommendations or portfolio management, such use remains relatively uncommon at this time. While some licensees also employ artificial intelligence in the actual provision of investment recommendations or in portfolio management, such use still remains relatively uncommon.[425]

The limited scope of use in Israel to date is consistent with findings reported internationally. The 2023 ESMA report notes that, while the potential use of artificial intelligence are numerous and promising, the extent of actual implementation varies across sectors, and widespread use of artificial intelligence in core business processes in the financial industry has not yet come to fruition. Overall, despite the increasing use of artificial intelligence to support and optimize certain activities, there is currently no indication of a rapid or disruptive transformation of business processes.[426] With respect to automated investment advice, in practice, it is largely based on relatively simple algorithms and limited data, while with respect to portfolio management, although industry interest has been identified, no significant change in investment practices is evident.[427]

Similarly, the 2023 OECD Report notes that advanced applications of artificial intelligence involving full automation are still, if present at all, only in development stages, and the pace of adoption of AI systems remains slow. The report points to several contributing factors, including the regulatory environment, the allocation of liability to financial entities, various inherent risks associated with these

---

[425] See, for example, the response given in October 2022 by the Israel Securities Authority in response to an inquiry from a company holding an investment advice license to produce digital analyses based on algorithms it developed, in accordance with the Directive on Online Services: https://app.new.isa.gov.il/public/download/media/12766.json

[426] ESMA Report 2023, pp. 3, 19. NLP activity is also noted as becoming more widespread in situations that require processing large amounts of text and identifying information within unstructured data.

[427] ESMA Report 2023, pp. 5–10.

systems, such as the accuracy of outputs, data protection and privacy concerns, and the complexity involved in replacing existing systems.[428]

Regulation of Artificial Intelligence in Investment Advice and Portfolio Management: A Comparative Perspective

Similar to Israel's Online Services Directive, arrangements have been established in other countries abroad seeking to regulate digital investment advice and management activity. This activity has been called by names such as digital advice, robo-advice or automated advice, and is characterized by the fact that operations are conducted online and with only limited, if any, human involvement. The purpose of these regulatory arrangements is not only to protect the investing public receiving these services, but also to make investment advice and portfolio management services accessible to the public at large.

The contours of the arrangements established in various countries are fundamentally similar. They establish that, pursuant to the principle of technological neutrality and consistent with the objective of providing investor protection, the obligations applicable to "regular" licensees apply to AI-conducted services as well. These arrangements outline how these obligations should be implemented when activities are automated and conducted online, and they present the various adjustments necessary for conducting such activity. These include, for example, provisions relating to proficiency and comprehension obligations in connection with technological means, controls over technological means and outputs of the activity, disclosures to be given to the client regarding the characteristics of the activity, documentation, data protection and business continuity requirements, and rules addressing outsourcing through which the activity is conducted. A similar approach was also highlighted in a publication by ESMA, following the publication of the Interim Report, which explained that, to comply with existing legal requirements (under the European MiFID II Directive), supervised entities using artificial intelligence technologies must act in their clients' best interests, disclose to them the role these technologies play in the services provided, and establish organizational arrangements covering governance, risk management, employee knowledge and proficiency.[429]

It should be noted that certain foreign regulatory frameworks are more detailed, and their scope of application is not identical to that of the Israeli Directive, as these also address additional activities, such as advice relating to additional financial products or brokerage services. Nevertheless, it is apparent that

---

[428] 2023 OECD Report, pp. 10–12.

[429] See ESMA, Public Statement On the Use of Artificial Intelligence (AI) in the provision of retail investment services (30 May 2024), https://www.esma.europa.eu/sites/default/files/2024-05/ESMA35-335435667-5924_Public_Statement_on_AI_and_investment_services.pdf .

these frameworks reflect standards similar to those developed in Israel.[430] At the same time, it is evident that both the foreign and Israeli arrangements, were primarily designed to address the shift toward digital and automated investment advice services. Recent technological breakthroughs in artificial intelligence indicate the potential for more advanced and sophisticated applications, which were not taken into account in the existing regulatory framework.

Alongside the said technological advances, the SEC has shown growing concern around automated services (robo-advising) and the incorporation of artificial intelligence in investment advice and portfolio management. This concern is reflected, *inter alia*, in audit findings that pointed to deficiencies in the activity of investment advisers, including with respect to compliance programs, portfolio management conducted in a manner consistent with the duty to act in the clients' best interests, and marketing practices;[431] in statements by the SEC Chair regarding the risks inherent in artificial intelligence, *inter alia*, the potential for harm to financial stability arising from uniform decision-making based on shared AI models;[432] and in the SEC's 2023 proposal concerning conflicts of interest associate with the use of Predictive Data Analytics in the work of broker-dealers and investment advisers (which has later been withdrawn).[433]

The proposed rule on Predictive Data Analytics was intended to apply to a range of technologies included under the term "Covered Technologies".[434] Under the proposal, broker-dealers and investment advisers would have been required to evaluate the predictive data analytics technologies they use, and to identify, and either eliminate or neutralize conflicts of interest that could result in the supervised entity's interest prevailing over those of their clients. To this end, supervised entities would have been required to adopt and implement policies and mechanisms for this purpose, and to document the controls they implemented. The concern underlying the proposal was that applications of artificial intelligence could exacerbate traditional conflicts of interest, such as the provision of biased investment advice or investment management designed to generate higher management fees or trading commissions. This risk arises, *inter alia*, because conflicts of interest may be more difficult to detect and mitigate, and because

---

[430] For elaboration on the regulation applying to Robo Advising abroad, see Tel Aviv University Report, at pp. 24–27.

[431] SEC Risk Alert, Observations from Examinations of Advisers that Provide Electronic Investment Advice (November 9, 2021), https://www.sec.gov/files/exams-eia-risk-alert.pdf.

[432] SEC Chair Gary Gansler "Remarks before the Financial Stability Oversight Council: 2023 Annual Report" (December 14, 2023), https://www.sec.gov/news/speech/gensler-remarks-fsoc-121423.

[433] SEC Proposal 2023.

[434] Covered technology means an analytical, technological, or computational function, algorithm, model, correlation matrix, or similar method or process that optimizes for, predicts, guides, forecasts, or directs investment-related behaviors or outcomes.

the potential scope of harm from biased outcomes may be greater relative to human-provided investment advice or portfolio management.[435]

Considerable criticism has been leveled at this rule proposal which was withdrawn in June 2025 following the change in administration in the United States, and is no longer being pursued. Some argue that if it had been approved in the proposed or similar format, it may greatly hinder the use of AI applications. For example, there may be difficulty in accurately identifying conflicts of interest and exercising effective oversight of over the outputs of AI systems, particularly given their limited explainability. The proposal also reflected an approach to address conflicts of interest by neutralizing them, rather than by disclosing or mitigating them.[436] Following the governmental change in the U.S., the SEC announced on June 2025 that it withdraws this proposal.[437]

Key Advantages and Risks of Using Artificial Intelligence in Investment Advice and Portfolio Management

Throughout the report, various benefits and risks associated with the use of artificial intelligence in the financial sector have been examined. These include, *inter alia*, risks related to data privacy, the reliance on unreliable or partial information, and reduced explainability. At this stage, the focus to the principal advantages and risks that are more specifically relevant to investment advice and portfolio management activities.

Key Advantages

The main advantages of deploying artificial intelligence into investment advice and portfolio management activity include the ability to serve a significantly larger client base; anticipated savings in service provision costs; and the potential to enhance both service quality of and the quality of supporting processes at service providers.

(1) Broadening the Scope of Activity and the Client Base for Investment Advice and Portfolio Management Services

Investment advice and portfolio management services can play a significant role in helping the general public manage their funds effectively. Investment advisers provide objective guidance, based on a

[435] SEC Proposal 2023, pp. 6–33.

[436] See for example Daniel Michael, David Simon, Merin Cherian "Understanding SEC's Focus Amid Lack of Final AI Rules" Skadden Insights (February 23, 2024), https://www.skadden.com/insights/publications/2024/02/understanding-sec-focus-amid-lack-of-final-ai-rules; James Deekan "SEC Predictive Tech Ban Causes Confusion for Investment Advisers" Bloomberg Law (January 8, 2024), https://news.bloomberglaw.com/us-law-week/sec-predictive-tech-ban-causes-confusion-for-investment-advisers.

[437] https://www.sec.gov/rules-regulations/2025/06/s7-12-23.

careful assessment of the investor's goals and risk tolerance. They advise clients on how to construct their investment portfolios, and assist in selecting specific investments. Likewise, portfolio managers are authorized to manage their clients' securities accounts, construct investment portfolios, in accordance with a predefined risk profile and investment preferences, and execute investment transactions on the client's behalf.

Broadening access to investment advice or portfolio management services provided by a regulated licensee serves the clear interest of both the public and the state. Individuals who lack financial literacy and do not receive professional intermediation services, are likely to remain outside the circle of investors or turn to unlicensed providers, an action that may lead to financial losses, including those resulting from fraud or deception.

In spite of the above, in recent years, there has been a noticeable and continuous decline in the number of licensees in Israel, in addition to a decline in the number of clients benefiting from professional investment advice and portfolio management services.[438] The primary reasons for this decline are the costs associated with providing these services, which lead licensees to focus primarily on clients with a high minimum investment threshold, while the general public turns to unsupervised channels, whether operating legally or illegally. The implementation of AI systems could help reverse this trend by making such services accessible to a wider audience.

(2) Cost Savings

A possible shift toward automation and a reduced reliance on human capital in both core service provision and ancillary activities, such as back-office operations or compliance, is expected to lower operational costs. These savings could enable the realization of the previously mentioned benefit: expanding investment advice and portfolio management services to a broader client base, while potentially reducing service costs to clients as well.

(3) Improving Service Quality

The various potential applications for artificial intelligence in investment advice and portfolio management were outlined above. AI systems may enhance service quality, for instance, by improving response availability, the scope of responses, and customization of services to clients' needs. It should be noted, however, that while, the entry of artificial intelligence is expected to improve certain aspects of these services, there is debate over its effectiveness (as opposed to human-delivered services) in the core functions of investment advice and portfolio management services – namely, clarifying client needs,

[438] See the data below in this chapter.

tailoring activity to those needs, and providing investment recommendations or making investment decisions. For example, studies of automated investment advice services (robo-advising) have raised concerns about service quality, particularly the ability of computerized systems to understand the client's situation holistically.[439] Even if advanced AI improves some aspects of service, this report highlights potential risks that may arise, such as limited explainability, conflicts of interest, and other factors that could adversely impact service quality.

Key Risks

(1) Failure to Meet Fiduciary Duties and Materialization of Potential Conflicts of Interest

A key criticism raised internationally with respect to robo-advisory services concerns the ability of computerized systems to satisfy the fiduciary duties of trust and care imposed on investment advice and portfolio management licensees. With respect to the duty of care, it has been argued, *inter alia*, that reliance on digital questionnaires may fail to capture material information about the client, that a human perspective is essential to properly understand certain situations, and that automated processes, while efficient, may also be overly simplistic.[440] AI systems based on LLM technology or other advanced technologies are expected to be more sophisticated and to enhance the understanding of client needs and the tailoring of services accordingly. However, at least in the current transitional stage, such systems present other challenges, including generic responses that are insufficiently customized, imbalanced outputs reflecting partial or unreliable information and even so-called "hallucinations," i.e., the generation of false content.[441]

With respect to the duty of trust, it has been argued that AI systems may take into account extraneous considerations and, in doing so, prioritize the service provider's interest over those of the client. Moreover, due to the nature of these systems, and the limited transparency surrounding their operations, the emergence of such conflicts of interest may be more difficult detect than is human-provided investment advice or portfolio management services. This concern regarding the potential

[439] See Tel Aviv University Report, at p. 20 and the references mentioned there, and later in this chapter.

[440] Ibid, pp. 19-21 and the references mentioned there; Itamar Caspi, Sarit S. Felber, Talia B. Gillis, Generative AI and the Future of Financial Advice Regulation, Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 2021, https://blog.genlaw.org/CameraReady/19.pdf.

[441] Ibid, in section 4. On the other hand, it can be argued that advice by humans is also not free of problems and may also suffer from biases, conflicts of interest or failure to consider relevant information. This argument naturally arises in other contexts as well, in the transition from human activity to artificial intelligence.

materialization of conflicts of interest stood, as noted above, was at the basis of the SEC's rule proposal to tighten regulation on AI systems in the said fields.[442]

Generative AI technology may also pose challenges the fiduciary obligations of licensees, as it complicates the establishment of clear operational boundaries. This difficulty arises, for example, in portfolio management activity that relies on adaptive learning algorithms, or in interactions between clients and a licensee's chatbot, where communications involve content generation and are not limited to a predefined set of responses.

(2) Gamification Risks

In recent years, the phenomenon of gamification in financial activity has expanded across global capital markets. Gamification refers to the use of technological tools designed to make the financial activity engaged in by clients resemble game-like experiences. This includes user-friendly mobile applications, intuitive customer interfaces, designs that encourage the execution of trades and investments, the use of "push" notifications to deliver information and offers, and the offer of symbolic or tangible "rewards" for carrying out transactions. These tools primarily target unsophisticated retail investors, particularly young investors.[443]

The gamification phenomenon has intensified not only as a result of technological developments, but also in the context of other factors that have increased public participation in self-directed securities investment, such as the ability to trade on trading platforms in small amounts and without commissions. While it primarily characterizes securities trading activity, it is also relevant to related activities such as investment advice. In the United States, for example, trading platforms such as Robinhood have become very popular among investors.

Although gamification in the financial sector offers certain benefits, including stimulating public interest, increasing accessibility to investment activity, and supporting financial education, it also has given rise to serious among regulators and others. These concerns include, *inter alia*: the potential exploitation of human behavioral biases to prompt unnecessary or inappropriate transactions; the diversion of investors from rational decision-making toward executing specific transactions, such as

---

[442] See Tel Aviv University Report, pp. 21-23 and the references mentioned there; Nicole G. Iannarone, Computer as Confidant: Digital Investment Advice and the Fiduciary Standard, 93 Chi.-Kent L. Rev., 141 (2018), https://scholarship.kentlaw.iit.edu/cklawreview/vol93/iss1/5 ; Sophia Duffy and Steve Parrish, You Say Fiduciary, I Say Binary: A Review and Recommendation of Robo-Advisors and the Fiduciary and Best Interest Standards, 17 Hastings Bus. L.J. 3 (2021), https://repository.uclawsf.edu/hastings_business_law_journal/vol17/iss1/3 ; SEC Proposal 2023.

[443] James Fallow Tierney, Investment Games, 72 Duke Law Journal 353-446 (2022), https://scholarship.law.duke.edu/dlj/vol72/iss2/2; Nolan Schloneger "A Case for Regulating Gamified Investing" 56 Ind. L. Rev. 175 (2022), https://doi.org/10.18060/27147.

purchasing certain designated stocks; and encouraging herd behavior, such as inducing investors to purchase stocks after their prices have risen significantly. The GameStop affair, which involved widespread speculative trading resulting in a sharp spike and subsequent collapse in the company's stock price, illustrates the risks inherent in investment divorced from economic fundamentals, to both individual investors and market integrity overall, as well as the dangers of uncontrolled encouragement of trading through platforms and social networks.[444]

Given the scope of the phenomenon in the United States, regulatory authorities have also taken action in this area. In 2021, the SEC issued a request for public input on the use of broker-dealer and investment-adviser Digital Engagement Practices when interacting with clients.[445] These practices encompass behavioral prompts, personalized and differential marketing activity, game-like features, as well as other mechanisms designed to stimulate client activity. Subsequently, the 2023 SEC Proposal described above was published, addressing conflicts of interest stemming from the use of predictive data analytics by broker-dealers and investment advisers. This proposal, however, was ultimately withdrawn.

s artificial intelligence becomes more prevalent and advanced, it may further amplify the use of gamification in financial activities, along with the risks inherent in such practices. *Inter alia*, algorithms can tailor investment advice and trading platforms to the preferences (and potential vulnerabilities) of individual clients, conduct in-depth analyses of client behavior and use such analysis to incentivize investment activity, and provide automated advisory responses or trading prompts to clients' actions. These adverse effects are compounded by the challenges of explainability and potential conflicts of interest described above.

(3) Harm to Service Quality

As stated above, the concern regarding the impairment to the quality of customer service has been discussed globally in light of the expansion of automated investment-advice services. Question marks regarding service quality continue to hover over more sophisticated artificial intelligence activity as well. With respect to automated investment advice, concerns have been raised that the process of clarifying client needs may be overly simplistic, that the client's situation may not be fully or accurately understood,

[444] See more on the GameStop affair:
Game Stopped? Who Wins and Loses When Short Sellers, Social Media, and Retail Investors Collide: Hearing Before the H. Comm. on Fin. Servs., 117th Cong. (2021), https://www.govinfo.gov/content/pkg/CHRG-117hhrg44837/pdf/CHRG-117hhrg44837.pdf; Jill E. Fisch "GameStop and the Reemergence of the Retail Investor" 102 B.U.L. Rev. 1799 (2022), https://scholarship.law.upenn.edu/faculty_scholarship/2817/.

[445] Request for Information and Comments on Broker-Dealer and Investment Adviser Digital Engagement Practices, Related Tools and Methods, and Regulatory Considerations and Potential Approaches; Information and Comments on Investment Adviser Use of Technology To Develop and Provide Investment Advice, Securities Exchange Act Release No. 92766 (August 27, 2021), 86 Fed Reg. 49067 (Sept. 1, 2021), available at https://www.federalregister.gov/d/2021-18901.

that resulting investment recommendations would be generic and insufficiently tailored to individual client needs and that ultimately, the benefits of remote accessibility and cost efficiency may come at the expense of the quality of the advice.

The recent leap in the capabilities of AI systems carries with it a promise for substantial improvements over the automated investment advice available to date. Advances in natural language comprehension and human-like interaction, the enhanced capacity for information analysis and the presentation of multiple alternatives, and the improved ability to achieve meaningful personalization in the assessment of client needs as well as the tailoring of service to those individualized needs - all these are expected to enhance the scope and quality of digital services beyond what has been previously possible. At the same time, at least at the current stage of technological development, reliance on general LLM models may still yield mixed results, including incomplete or unbalanced information, recommendations that are generic, insufficiently current or not subject to adequate controls. The development of dedicated AI systems is expected to improve the quality of outputs, and many anticipate that such improvements will materialize in the near future, across multiple areas of life, including in the financial sector. However, during the interim period preceding the launch and deployment of such systems, the risk to service quality remains.

(4) Concentration and Third-Party Dependency

A final distinct risk associated with the use of artificial intelligence in investment advice and portfolio management arises from that the potential that a small number of technological systems may serve numerous licensees and a large client base. Concerns in this context include: the potential for reduced investment diversity and increased correlation across the investment portfolios of many clients; the possibility that a malfunction or system error could have an impact on the entire market; and the creation of dependency on a small number of providers, including reliance on outsourcing to a small group of dominant developers.[446] With respect to investment advice and portfolio management activity in Israel, this risk currently appears to be of limited material relevance, but it should be considered as patterns of activity evolve.

Unique Advantages and Risks to Artificial Intelligence Activity in Investment Advice and Portfolio Management:

[446] See Tel Aviv University Report, pp. 23–24; 2023 OECD Report, pp. 24–25. See further discussion of these risks in the chapters dealing with financial stability and competition.

Advantages	Risks
Expanding the scope of activity and client base benefiting from investment advice and portfolio management services	Failure to meet fiduciary duties and materialization of potential conflicts of interest
Cost efficiency	Gamification risks
Improving service quality	Compromised service quality
	Concentration and third-party dependency

Investment Advice and Portfolio Management – Recommendations:

The transition to automated investment advice and investment portfolio management began gaining traction globally approximately a decade ago. In Israel, the transition remains partial, with services not fully automated and still relying on human involvement in core functions, such as generating recommendations (in advisory services) or making investment decisions (in portfolio management). Recent advancements in artificial intelligence are expected to accelerate the transition, both by expanding the number of participating entities and clients and by increasing the scope of the automated activity.

The shift toward automation may carry substantial implications. On one hand, it offers notable advantages, including broader accessibility of services, more efficient processes, and cost reductions for investors and licensees alike. On the other hand, an unsuccessful transition may give rise to material risks, such as the clients' insufficient understanding of the service, the provision of recommendations or execution of decisions that are misaligned with client needs (for example with their risk profile), and technological or operational risks associated with system operations.

(1) Active Encouragement of Integrating Artificial Intelligence Technology Among Investment Advice and Portfolio Management Licensees to Increase the Number of Clients Receiving These Services

Data from recent years indicate a decline in the scope of regulated investment advice and portfolio management activity. In the field of investment advice, this is reflected in a decline in the number of investment advice licensees, in the number of clients receiving advice, and in high costs associated with

investment advice.[447] For example, the number of active advisors in banking corporations declined from 1,513 in 2020 to 1,200 at the end of 2023, the number of advised accounts declined from 453 thousand in 2018 to 323 thousand at the end of 2023, and the average market capitalization of an advised account rose from 656 thousand shekels in 2018 to more than one million shekels in 2023. In the field of portfolio management, this is reflected in the stagnation of the number of portfolio management companies, in the number of clients and in the value of assets under management. For example, the number of portfolio management companies stood at 124 and 130 in 2018 and 2023, respectively; the number of clients stood at approximately 83 and approximately 85.5 thousand in 2018 and 2023, respectively; and the value of assets under management stood at approximately 271 and approximately 324 billion shekels in 2018 and 2023, respectively.[448]

The weakness of the said activity as reflected in the data is particularly prominent when one considers that in those five years, the scope of economic activity and the scope of public savings in the economy grew. The reasons for the observed decline in investment advice and portfolio management activity may be varied. Nevertheless, it is recognized that this reduction in the scope of these services has adverse implications for public access to professional financial services and for the quality of the management of public savings. As noted above, broadening the client base is clearly in the interest of both the public and the state.

Artificial intelligence technologies have the potential to reverse the current trend in the scope of these services, to the benefit of the public. The automation of a series of processes is expected to reduce costs for supervised entities and lower barriers to entry for the public. In addition, artificial intelligence offers the capacity to analyze individual investor needs and to tailor recommendations or portfolio management accordingly. As a result, the public– including populations that do not currently benefit from investment advice or portfolio management services– may gain access to more affordable professional services, thereby reducing reliance on problematic alternatives, such as unprofessional self-management, unsupervised channels, or generic advice that is not tailored to individual needs.

In view of this trend and the potential offered by artificial intelligence technologies, it is proposed to take proactive steps to promote their adoption among licensees. The regulatory review did not identify any clear legal barriers preventing this, and the limited level of implementation to date may also be

[447] Israel Securities Authority, Activity of Advisory Arrays in Banks 2023, https://www.new.isa.gov.il/images/Fittings/isa-be/asset_library_pic/al_lobby/al_lobby-6283705a39f5d/Consulting2023.pdf.

[448] Israel Securities Authority, Activity of Portfolio Management Licensee Companies, Data as of 31.12.2022, https://www.new.isa.gov.il/images/Fittings/isa-be/asset_library_pic/al_lobby/al_lobby-637b90c7be39b/NihulTikim2023.pdf.

attributed to technological or business constraints. Nevertheless, experience indicates that active engagement, in cooperation with financial market participants can foster the development of new activity. In addition to the measures outlined in this report to promote regulatory certainty and innovation (see Part E - Additional Actions to Promote Financial Regulation in the Field of Artificial Intelligence), it is proposed to engage directly with existing and prospective market participants, and to consider publishing a user-friendly guidance document explaining potential service models and practical approaches to compliance with the applicable legal obligations. By way of comparison, see a guidance publication published in Australia regarding robo-advisory activity, which includes tips and explanations on how to comply with legal requirements. Information of this type can assist both supervised entities and those interested in developing systems for them.[449]

(2) Updating the Online Services Directive

The Online Services Directive provides a regulatory framework for the use of artificial intelligence while safeguarding the public interest, and is broadly consistent with comparable legal arrangements in other jurisdictions. However, as the Directive was adopted in 2016, prior to recent advances in AI capabilities, it is proposed to examine potential updates to the Directive, both from a technical and substantive perspective.

Technically, it is desirable to consider updating references and terminology, so they correspond with current AI regulation. For example, the Directive does not currently explicitly refer to artificial intelligence and does not include established terminology such as "explainability" and "generative AI". In this context, it is proposed to consider adding a specific chapter to the Directive addressing current AI technology.[450]

Substantively, it is proposed to further clarify the roles of the person at the licensee responsible for providing technological services (the "Appointed Licensee").[451] This mechanism has also been adopted in other jurisdictions and plays a significant role in ensuring licensee compliance with legal requirements. For example, it is proposed to clarify what the roles of the Appointed Licensee with respect to reviewing and overseeing system outputs

[449] ASIC, Regulatory Guide 255 – Providing digital financial product advice to retail clients (August 2016), https://asic.gov.au/regulatory-resources/find-a-document/regulatory-guides/rg-255-providing-digital-financial-product-advice-to-retail-clients.

[450] Currently the Directive includes chapters specifically addressing social trading services and signals. In light of its expected importance in the coming years, it is proposed to consider dedicating a more detailed reference to it.

[451] Section 7 of the Online Services Directive.

Additionally, consideration should be given to providing regulatory guidance, within the framework of the Directive or elsewhere, to AI activity that does not involve direct client interfaces. As outlined in this report, artificial intelligence is expected to be deployed across a range of additional functions within supervised entities. In this context, an appropriate regulatory response may take the form of a dedicated directive or interpretive position clarifying how existing legal obligations apply to the use of AI systems. Additional matters warranting specific consideration in the course of updating the Directive include the use of chatbots in client-facing interactions (see below), as well as conflicts of interest that are at the focus of regulatory engagement in the United States.

(3) Examining Chatbot Activity and the Regulation That Should Apply to It

The recent development that has raised the profile of artificial intelligence is the emergence of large language models (LLM). These models constitute a category of generative AI, where the content created by it is language/text. They are capable of generating new content instantaneously through interactive dialogue with human users.[452] The capabilities of these models, including ChatGPT, which was the first model widely available to the general public, have attracted considerable attention for several reasons, including their ability to address a immense range of topics, their exceptional capacity to absorb content and identify the essence of a query, the high quality of their language output, the diversity of possible responses in terms of length, style, and format; and the immediacy of their responses.

The use of chatbots predates the development of LLM-based AI systems. The term "chatbot" refers to any application that conducts a dialogue with a user in natural language including systems based on relatively simple models. For example, rule-based chatbots operate by matching a user's query question and a predefined response selected from a set of possible responses. Such chatbots do not generate new content.[453] In many areas, including the financial sector, the use of chatbots have been widely adopted, primarily for customer support operations or self-service functions.[454]

Advances in LLM-based models, together with the user-friendly interfaces of ChatGPT and similar applications. following its launch, have significantly increased the usability of chatbots. As a result, chatbots are expected to be integrated into a broader range of service provider operations, including functions at the core of investment advice and portfolio management services, such as clarifying client needs and tailoring services accordingly. Beyond the expansion of chatbot use, the enhanced capabilities

[452] See 2023 OECD Report, at pp. 8–9, 12–13, and also in the chapter on AI Applications in the Financial Sector above.
[453] On types of chatbots see: Eleni Adamopoulou, Lefteris Moussiades "Chatbots: History, technology, and applications" Machine Learning with Applications 2 (2020) 100006, https://www.sciencedirect.com/science/article/pii/S2666827020300062.
[454] See Tel Aviv University Report, pp. 28–30.

of AI-based chatbots are expected to lead to greater client interaction. Such chatbots may operate more proactively (as distinct from reactively), learn from client's inquiries, adapt responses to individual users, ostensibly display emotions, foster engagement and increase trust.

The expansion of chatbot use also gives rise to regulatory challenges. For example, in 2023, the U.S. CFPB published a report examining the use of chatbots in financial services. The report noted the rapid expansion of chatbot adoption: the ten largest banks in the U.S. have deployed chatbots as part of their customer service operations; approximately 37% of the U.S. population interacted with a bank chatbot in 2022 a figure that is expected to increase. A shift is underway as banks transition from simple chatbots to more sophisticated chatbots relying on LLM-based technologies. The report further describes customer complaints received by the CFPB in connection with chatbot use and highlights, *inter alia* the following issues: the limited ability of chatbots to address complex issues; constraints on customer accessibility to timely human assistance; technical shortcomings and related data protection risks; and failures to comply with applicable legal requirements.[455]

In Israel, the regulation applicable to investment advice and portfolio management, including the Online Services Directive, does not specifically address chatbots, let alone chatbots based on advanced AI technologies. The Taskforce proposes to monitor this activity, track developments in comparative law, and if necessary, update Israeli regulation in the future.[456] Such updates could be implemented by revising the Online Services Directive, as stated above, or by providing clarifications regarding duties of care and trust when using chatbots. Regulatory rules in this are expected to address issues discussed in this report, including explainability, reliability, information privacy, as well as matters more specific to chatbots in investment advice and portfolio management. These include training algorithms on relevant topics and controlling their content, managing sensitive operations, such as disputes between service providers and clients, and ensuring the availability of human support alongside chatbot interactions.

(4) Clarifying the Law Regarding General-Purpose AI Systems

he deployment of general-purpose AI systems, including ChatGPT and Claude, capable of responding on virtually any subject, raises questions regarding content quality and legal applicability. With respect to content quality, the key question is how accurate and useful the responses are to clients' questions on

[455] See *supra* note 45.

[456] See regarding regulatory frameworks applying to chatbots in the Tel Aviv University Report; and frameworks in other contexts: World Economic Forum "Chatbots RESET – A Framework for Governing Responsible Use of Conversational AI in Healthcare" (December 2020), https://www.weforum.org/publications/chatbots-reset-a-framework-for-governing-responsible-use-of-conversational-ai-in-healthcare/.

investment topics. With respect to the applicability of law, the question is whether the Investment Advice Law applies to systems of this type when they provide the user investment advice.

Views on content quality are divided. Some argue that investment recommendations generated by artificial intelligence tend to be generic, lack depth and at times are not even up-to-date. Others contend that AI-systems already provide reasonably effective analytical tools that can be further refined and enhanced.[457] In any event, these assessments relate to systems that have only been launched recently, and it can be assumed that their outputs will become increasingly sophisticated and improved over time.

With respect to the applicability of law, relevant characteristics may include: the range of topics addressed by the AI system, whether use of the system involves consideration, and the identity of the entity making the system available. It should be noted that a related question arose several years ago regarding algorithmic portfolio management activity. Dealer platforms, which were prohibited from providing investment advice and executing discretionary transactions on behalf of a third party, claimed they were merely making "off-the-shelf software" available to clients enabling them to trade securities, and that, therefore, the law was not applicable to this activity. Accordingly, the Israel Securities Authority set out conditions under which the Authority would not regard a client's ability to operate trading software as part of its activity on the platform as constituting investment advice or portfolio management.[458] These conditions include, *inter alia*, a prohibition of the platform marketing algorithmic trading software or encouraging clients to trade through it; a prohibition on any connection between the platform and the developer or designer of the trading software; and the absence of control over the operation of the algorithmic trading software.[459] A similar approach was adopted by the court, which held that making such software available to clients by a local party is not merely a technical act, but falls within the scope of the law's applicability.[460]

With respect to response quality, it is proposed to consider publishing materials as part of financial education efforts, informing the public about the limitations of general AI systems and highlighting licensed human financial services as the appropriate alternative.

With respect to the applicability of law, it is proposed to consider issuing a clarification regarding the prohibition of making general AI systems accessible.

---

[457] See for example Dowling, Michael M. and Lucey, Brian M., ChatGPT for (Finance) Research: The Bananarama Conjecture (January 11, 2023), https://ssrn.com/abstract=4322651 ; and also, Kim, Alex G. and Muhn, Maximilian and Nikolaev, Valeri V., Financial Statement Analysis with Large Language Models (May 20, 2024). Chicago Booth Research Paper, Fama-Miller Working Paper, https://arxiv.org/abs/2407.17866.

[458] Regulations 10(a) and 10(c) of the Securities Regulations (Trading Platform for its Own Account), 2014.

[459] Israel Securities Authority, Legal Staff Position: Offering Trading Services Using an Algorithm on a Trading Platform (September 14, 2016).

[460] See Section 34 in the Efrimov case, and the reference to the Authority's position in Sections 50–52 of the judgment.

(5) Promoting Research Examining Customer Interaction with Computerized Systems

For the purpose of establishing optimal regulation, it is essential to have a deep understanding of user behavior with respect to AI activity and automated decision-making. For example, it is important to examine whether investors prefer to use automated systems or, alternatively, choose to avoid them, and the reasons underlying these preferences; how investors respond to information and recommendations given to them by automated systems; whether investors are unduly influenced by such information and recommendations by automated systems, and whether their behavior differs from that of clients receiving traditional investment advice or portfolio management services; the extent to which investors assign weight, if any, to warnings presented by automated systems; and more.[461]

These and other issues pose interesting questions for research, and indeed engagement with them is growing globally. The answers to these questions, depend, *inter alia*, on the pace of technological adoption among the public and may differ across countries, cultures, time periods, age groups and other characteristics.

In Israel, research on these issues is lacking and there is scope to promote further study. From a regulatory perspective, conducting research in this area is important both for market development and for understanding the impact of technology on investor behavior. Such research would support the assessment of existing regulation and the identification of any regulatory gaps. It may be conducted by the regulators themselves or through collaboration between regulators, academia and financial entities. In addition, regulators can encourage research through various means, such as providing financial support, facilitating access to relevant data, or issuing calls for proposals to stimulate research on the topic.

(6) Regulation of Brokerage Services

The Taskforce was instructed to focus specifically on the use of AI technology in investment advice and portfolio management. In regulatory discussions worldwide, the close link between these activities and brokerage services is particularly prominent. Brokerage services include trading platforms that allow the public to trade securities on the stock exchange, as well as brokers and dealers that execute trading operations in securities transactions on behalf of clients. In recent years, the brokerage services sector

[461] See Caspi, Felber, Gillis, *supra* note 415; Belanche, Daniel, Luis V. Casaló and Carlos Flavián "Artificial Intelligence in FinTech: understanding robo-advisors adoption among customers" Industrial Management & Data Systems 119.7 (2019) 1411, https://www.researchgate.net/publication/344743097_Artificial_Intelligence_in_FinTech_understanding_robo-advisors_adoption_among_customers#read; World Bank, Facundo Abraha, Sergio L. Schmukler and Tessada, Jose "Robo-Advisors : Investing through Machines" Research & Policy Briefs no. 21 (2019), https://documents1.worldbank.org/curated/en/275041551196836758/pdf/Robo-Advisors-Investing-through-Machines.pdf.

has undergone substantial changes, with the launch of digital platforms enabling the general public to execute transactions in securities with ease and with the expansion of direct public participation in securities trading, particularly during the COVID-19 pandemic period.

In view of these developments, the brokerage services sector presents significant regulatory challenges. For example, in the United States there has been extensive focus in recent years on risks associated with the gamification of trading activity, as described above, including the encouragement of retail investment without sufficient understanding of the associated risks, conflicts of interest in platform operations, and the potential harm to market integrity.

As explained above, AI technology can amplify these risks. For example, it may increase the capacity to influence investor behavior on digital trading platforms (through offering services to clients, or encouraging them to execute transaction,). Similarly, it may increase the risk of herd behavior among investors trading securities on these platforms.

In Israel, the brokerage services sector is not directly supervised by the Israel Securities Authority Oversight is indirect and partial, conducted mainly through the Tel Aviv Stock Exchange's supervision over non-bank Exchange members, through whom trading on the Exchange is conducted under provisions established in the Exchange's Rules and Regulations.[462] This supervisory gap is unusual by global standards and has long been recognized warranting attention.[463] The rationale for filling the gap includes protecting the investing public, supporting broader public participation in securities trading, accommodating the entry of new market players while fostering competition and innovation, and ensuring that entities operating in the sector adhere to established regulatory standards.

These considerations exceed the scope of this report. In the context of artificial intelligence, as noted above the link between these fields is particularly significant, highlighting the need to regulate the brokerage services sector to avoid regulatory gaps. To illustrate: investment advice provided through artificial intelligence carries the risk of conflicts of interest, potentially leading clients to execute transactions in ways that increase trading commissions. This risk can be mitigated through regulation

[462] The Stock Exchange supervises all Stock Exchange members in aspects of compliance and adherence to trading rules. Stock Exchange members that are banking corporations are also supervised by the Supervisor of Banks at the Bank of Israel.

[463] In 2010 a first attempt was made to regulate the brokerage sector in Israel, similar to regulation in Western countries. In 2020 a Memorandum of Law for Securities (Regulation of Broker Dealer Activity), 5781-2020 was published, but it was not promoted further. See also the Israel Securities Authority and the Competition Authority regarding the importance of promoting said regulation "Competition in the Retail Brokerage Market" (August 2019), at pp. 27–29, https://www.gov.il/BlobFolder/reports/marketresearch-brokeragefinal/he/marketresearch_%D7%AA%D7%97%D7%A8%D7%95%D7%AA%20%D7%91%D7%A9%D7%95%D7%A7%20%D7%94%D7%91%D7%A8%D7%95%D7%A7%D7%A8%D7%90%D7%96'%20%D7%94%D7%A7%D7%9E%D7%A2%D7%95%D7%A0%D7%90%D7%99%200819.pdf.

governing investment advice. However, a similar risk arises with respect to trading platforms, where the use of artificial intelligence could create conflicts of interest and influence clients to trade in ways that generate higher trading commissions. Effective regulation should address this risk in both areas.

In light of the above, it is important to complete this segment of oversight, enabling AI applications related to securities activity to be addressed in a consistent and comparable manner.

In the public comments to the Interim Report, several comments were submitted regarding the trends and risks described above. The Taskforce believes that these comments do not alter the content of the report, including the recommendations pertaining to this sector of activity.

Summary of Recommendations – Investment Advice and Portfolio Management Activity:

- Active encouragement of integrating artificial intelligence technology among investment advice and portfolio management licensees with the aim of increasing the number of clients accessing these services. This can be achieved, *inter alia* through direct engagement with existing and prospective market participants, publishing a guide that encourages and explains how to implement the service, and taking additional measures to promote regulatory certainty and foster innovation.
- Updating the Online Services Directive, including adding a chapter to the Directive specifically addressing current AI technology, clarifying the role of the Appointed Licensee (the person charged with providing the technological services), considering directives that address AI systems that are not used in client interfaces, as well as directives regarding the use of chatbots.
- Assessment of chatbot activity in the provision of investment advice and portfolio management as well as the regulation applied.
- Clarification of the law with respect to general-purpose AI systems.
- Promotion of research examining customer interaction with computerized systems. Research pertaining to the interaction of customers using AI systems and automated decision-making is necessary to develop the market, understand the effects of technology on investor behavior and assess existing and required regulation regarding this activity.
- Regulation of brokerage services to close regulatory gaps in this field. Given the close connection between brokerage and investment advice and portfolio management activities, as well as the potential extensive use of AI in brokerage activity, the regulation of this segment of the securities industry is essential for its development and ensuring consistent comprehensive regulation of it.

Credit in the Banking System

Within the financial sector, the credit sector presents significant potential for the deployment of artificial intelligence applications. Artificial intelligence can be applied across a wide range of functions, including identifying client needs, credit underwriting, credit scoring and ongoing credit management. It is assumed that the adoption of artificial intelligence represents a natural progression from the statistical models currently used in the credit sector.

In this chapter, we will provide a general overview of the credit sector within the banking system and the regulation governing it, outline potential applications of artificial intelligence in this sector in Israel, discuss the applicability of existing regulation to these uses, assess current trends in this area, and finally, address the issues arising from these applications and offer related recommendations.

Background – Credit

Credit plays an important role in creating conditions for economic growth, *inter alia*, because credit increases the ability to establish new businesses, and to expand and invest in existing businesses. Today the general public obtains loans from a range of financial entities, including banks, institutional lenders and credit card companies.

The credit sector is typically divided into several segments:

Type of lender:

A common distinction is made between bank credit and non-bank credit, which determines the type of lender involved:

Bank credit – credit provided by a banking entity (banks and credit card companies) and subject to directives of the Supervisor of Banks covering the rules and restrictions on extending credit, the supervision and monitoring required during the life of the loan as well as the handling of cases in which the borrower has difficulty meeting their obligations.

Non-bank credit – credit provided by entities supervised by the Capital Market, Insurance and Savings Authority.

Type of borrower:

A distinction is also drawn between consumer credit and business credit, which determines the type of borrower:

Consumer credit – refers to extending credit to private clients. Consumer credit is characterized by loans in relatively small amounts to private individuals for the benefit of private consumption. This segment of the credit services industry is controlled almost entirely by banks and credit card companies, although, in recent years, alternatives have emerged in the form of so-called people-to-people (P2P) platforms for borrowing money directly from other individuals, as well as non-bank entities that extend consumer credit. Until the end of the first decade of the twenty-first century, business credit was also concentrated within the banking system; however, following various shifts in financial markets, business credit has become more dispersed between the banking system and institutional entities.

Business credit – extending credit to business clients for the benefit of establishing or developing an existing or new business. In these cases, the borrower is in many cases the business itself rather than a private individual.

Credit/loan purposes

A final distinction is drawn between the purposes of the credit/loan. For example, housing credit is provided after a dedicated underwriting process and after the lending entity examines the borrower's repayment capacity and the value of the collateral, since, in most cases, credit of this type is long-term and secured by a lien on the property purchased in favor of the lending entity. Vehicle loans extended for the purpose of purchasing a vehicle, are also typically long-term, though generally shorter than that of housing loans.

In contrast, fixed-term or revolving credit, such as overdraft facilities in a checking account or credit cards, is usually unsecured and is granted following an underwriting process that assesses the borrower's repayment capacity. This type of credit is not short-term and may be renewed periodically. These distinctions between credit types affect, *inter alia*, the various lending conditions, including the loan amount, the interest rate and the repayment schedule.

Regulation Regarding Credit Extension

The credit sector is subject to extensive regulation grounded in legislation as well as regulatory directives, in this case - the *Proper Conduct of Banking Business* directives which establish and define standards for risk management and consumer fairness.

Credit regulation establishes, *inter alia*, obligations imposed on financial institutions regarding proper conduct toward clients, including duties of trust and care, disclosure obligations toward the client, and a

prohibition on misleading practices. It also regulates credit scoring practices, including provisions specifying the data that may or may not be considered for the purpose of determining a credit score.

Risk management of credit risks

Proper risk management is founded on three lines of defense: the business lines, an independent risk management function and internal audits. Each one of the Banking Supervision directives on credit risk management addresses issues such as corporate governance, risk appetite, the risk management framework, and the measurement, monitoring and assessment of risk.

Credit Risk

Credit risk refers to the possibility that a borrower or a lender's counterparty will fail to meet its debt obligations as agreed, resulting in a loss to the lender. Lenders try to assess the risk involved in each debt, according to the borrower's data and market conditions. Risk assessment enables lenders to make a series of decisions, including whether to approve or decline a certain loan, to set the terms applicable over the lifetime of the credit. These terms may include covenants that can trigger an acceleration or termination of the credit, as well as other conditions, such as the repayment period or the interest rate repayment structure. Credit risk assessment involves estimating both the probability of default of a counterparty (PD) and the lender's potential financial exposure in a default event (LGD). Many lenders rely on credit risk models to assess the risk of existing clients and prospective clients. Credit risk management is carried out both prior to extending credit– when evaluating its characteristics – and throughout the lifetime of the credit, by monitoring borrower behavior.

In its directives, the regulator specifies the working methods for credit risk management. Thus, the Banking Supervision Department's directives refer to the following areas: creating an appropriate risk environment (determining risk appetite and credit policy), defining the credit approval process, establishing proper and current processes for credit operations, measurement, monitoring and supervision, and ensuring appropriate controls over credit risk, as well as processes for handling distressed debts and debt restructuring.

The cost of credit, reflected in the interest rate, represents the lender's assessment of the risk associated with lending to a borrower—that is, the probability that the loan will be repaid in full and on schedule. Because the lender faces uncertainty regarding the borrower's ability to repay the loan, the interest charged (above the prevailing market rate) serves as compensation for the risk assumed. Interest rates are influenced by various factors, including the type of loan and the borrower's individual characteristics.

- Borrower characteristics: These characteristics are used to assess the risk associated with a particular borrower – namely, whether the borrower is expected to meet their obligation and repay

the loan. In credit markets, the standard method for evaluating repayment likelihood is borrower's past behavior in similar circumstances. Where there is evidence that a borrower has consistently met its obligations on time over an extended period, it is reasonable to expect that they will continue to do so in the future, which in turn influences the terms of the loan. It should be noted that the extension of credit is subject to the Prohibition of Discrimination Law (as detailed in the *Bias and Discrimination* chapter above) and, accordingly, certain characteristics may not be taken into account when deciding to grant credit or in determining its terms.

- Loan duration: Longer-term loans a are typically "more expensive," reflecting the increased risk that adverse events may occur over the extended repayment period.
- Collateral: If there is an asset serving as collateral for the loan, the price of the loan is expected to be "less expensive", since if the borrower is unable to repay, the lender can realize the collateral and use the proceeds to repay the loan.

Consumer protections

Legislation and Banking Supervision directives require supervised entities to provide clients with clear and transparent information regarding credit terms. They also mandate that the credit offered aligns with the client's purpose for the loan and his financial capacity to meet repayment obligations. Furthermore, to promote fair conduct in marketing consumer credit, a dedicated directive on consumer credit marketing came into effect in March 2024. This directive highlights any behavior inconsistent with the directive's provisions also poses a risk to the financial institution.

The directive establishes five main consumer principles designed to ensure fair conduct on the part of banking corporations:

- Obligation to conduct marketing conversations fairly: The directive prohibits aggressive marketing and pressuring clients to undertake a credit transaction.
- Obligation to tailor credit services to the client's needs and capabilities: Supervised entities are required to provide credit services responsibly to prevent household over-indebtedness. They must also ensure suitability between the terms of the credit offered and the intended purpose of the credit and the client's needs. Accordingly, supervised entities are prohibited from initiating or marketing credit services without first clarifying the client's needs.
- Fair pricing obligation: The directive requires supervised entities to price consumer credit fairly, that is to say, consistent with the client's risk profile of the borrower and their characteristics. Pursuant to the Banking Supervision directive, each supervised entity independently determines the risk profile of its clients using internally developed models. For the purpose of assessing the client's

risk, the bank is permitted to use "soft information" about the client, such as information not yet quantified regarding an expected deterioration in the client's economic situation, in addition to the criteria established in the model. The directive requires establishing guidelines for fair pricing in consumer credit policy documents, and on this basis, supervised entities are required to formulate and implement a risk-aligned pricing methodology. It should be emphasized that the relationship between risk level and interest rate must be reflected not only across the bank's overall credit portfolio but also at the level of each individual client.

- Enhanced disclosure requirements: The directive adds disclosure requirements to those established in the relevant legislation (the Fair Credit Law, the Banking (Service to Customer) Law) and the regular proper disclosure rules applicable to supervised entities. *Inter alia*, these disclosure rules were established to ensure that relevant information is provided in the marketing conversation, including disclosure of the purpose of the discussion, details of the proposed loan and related matters.
- Consumer protections: Reducing the frequency of proactive outreach and removing names from marketing lists.

Another Banking Supervision directive specifically addresses model risk management. In August 2024, the Banking Supervision Department issued the *Proper Conduct of Banking Business* Directive No. 369 on model risk management, aimed at strengthening the management of these risks across the banking system. The requirements set in this directive also apply to models that involve the use of, or reliance on artificial intelligence.

The directive is based on regulation from leading U.S. regulators (the OCC and the Fed) and outlines the key elements of effective model risk management. These include the development, implementation and proper use of models, as well as rigorous model validation, all of which are essential components of model risk management. Model risk management also encompasses corporate governance and control mechanisms, such as oversight by the board of directors and senior management, setting policies and procedures, compliance controls, as well as establishing an appropriate organizational structure and a well-designed incentive arrangement for risk mitigation. Banks are required to manage artificial intelligence model risks in accordance with the directive, placing special emphasis on the reliability and fairness of models with respect to baked-in bias or discrimination; accountability, adapted to AI models by considering their learning processes and the extent of human involvement; control design in decision-making processes based on AI models;

and transparency, including explainability of outcomes, comprehensive documentation of the model's components, and proper disclosure to ensure the rights of affected parties to receive information and appeal the results are fully preserved.

Credit Extension Process

In this section we discuss how artificial intelligence may be integrated into each of the various stages in the credit extension process.

Identifying needs

The need for credit may present itself in several ways: a client may approach the lending entity directly to request a loan; the financial entity may proactively reach out to the client with a loan offer; or a third party assisting the client with financial management may recommend taking credit. Identifying this need is a central component of understanding the client and delivering a personalized user experience.

Open banking enables the sharing of financial information, allowing client needs to be identified on the basis of aggregated data.[464] By analyzing information obtained from open financial systems, the lending entity or a third party can more efficiently recognize client requirements and tailored products and services. For example, if a client purchases a flight ticket using a credit card, the lending entity or third party can detect the potential need for financing the trip and offer a loan specifically for this purpose.

Credit underwriting

At this stage, the lender collects information about the credit applicant in order to assess his repayment capacity. For this purpose, it will examine his credit history, receive the applicant's credit score from credit bureaus, collect relevant financial information about him, including through open banking. Based on the accumulated information, the lender assesses the applicant's credit risk. It evaluates the probability that the applicant will not meet their obligations and will have difficulty repaying the credit as agreed. The applicant's risk assessment enables the lender to make a decision whether to approve or reject the credit request. In cases where the lender approves the request, a decision is also made regarding the terms of the loan, for example, the approved amount, the interest rate and the repayment schedule.

[464] Financial Information Service Law, 2021 (Israel).

Credit offer by the lender

The lender returns to the credit applicant and makes an offer including the terms and restrictions of the credit. If the lender decides not to extend credit based on the results of their risk assessment, it is required to inform the credit applicant.

Stage of acceptance of the offer by the credit applicant

When the borrower agrees to the lender's offer, they sign an agreement and commit to meet the terms as defined therein.

Transfer of money

After the offer is accepted, the lender transfers the funds to the borrower either in a one-time payment or by providing a credit facility, according to the type of credit provided.

Monitoring and management

During the life of the loan, the borrower is responsible for repaying the credit according to the terms of the agreement, which typically involves making scheduled payments of principal and interest. Throughout this period, the lender monitors the borrower's credit status and the payment performance. This oversight helps the lender manage credit risk and take action if the borrower fails to meet his obligations.

Beyond the monitoring of individual loans, lenders must implement appropriate controls and exercise continuous oversight over their entire credit portfolio, with periodic reporting to senior management and the board of directors in accordance with the institution’s internal procedures and applicable regulatory requirements. In addition, various directives set out guidance for lenders regarding the identification, classification, and disclosure of problematic loans, defined as loans where the borrower encounters difficulty in meeting the repayment schedule stipulated in the agreement. These provisions regulate both client-facing disclosures and financial statement reporting, and prescribe the actions lenders are expected to undertake to ensure that the collection process is fully and diligently pursued.

Closing and reporting

After the credit is fully repaid, the loan account at the lender is closed, and the lender reports to the credit bureau about the borrower's positive repayment history, which may improve the borrower's credit score.

AI Use in the Credit Extension Process

The potential applications of artificial intelligence span the entire lending process, from credit origination until its final repayment and extend across both back-office operations and client-facing

interactions. These applications can enhance the lender's operational efficiency, improve risk management and improve customer service as well:

Identifying client needs – Needs identification is conducted based on broad and personal information about the client (information arriving from existing data at the financial entity or collected by the financial entity through open banking). We anticipate that progress in this area will be observed in the near future. Artificial intelligence can be incorporated into the needs identification process by enabling comprehensive analysis of financial data and a client's economic behavior, including the detection of patterns and use of advanced models to identify these needs. By automatically recognizing behavioral patterns and potential risks, artificial intelligence can generate tailored recommendations and adapt financial offerings to better meet client needs.

With respect to consumer credit, there is an inherent tension between the desire for personalization, meaning adapting the specific product to the end client, and the concern regarding exclusion from the credit market. AI tools may greatly improve the lenders' ability to identify the needs of a specific borrower in real time and offer them personalized credit. However, to the same extent, it may lead to the exclusion of certain populations that are less digitally savvy. The lack of "digital" information about them may work against them (for example, clients whose main activity is in cash or who do not share information in the framework of open banking).

Credit underwriting – The underwriting process relies on the use of extensive information. Artificial intelligence offers analytical capabilities that can help analyze large volumes of data, including personal information and credit history, facilitating the creation of a comprehensive and accurate assessment of the client's loan repayment capacity efficiently and automatically. Globally, artificial intelligence is also applied to collateral evaluation within the underwriting process. This aligns with Banking Supervision directives on credit assessment and the 2024 directive on model risk management establishing a standard according to which the banking corporation must identify the risks inherent in the model throughout the life of the credit.

Despite the potential of artificial intelligence at this stage, we estimate that banking corporations will not deploy artificial intelligence in the near term for the purpose of autonomous automated underwriting. Several factors contribute to this reticence: the risk appetite of banking corporations; regulatory requirements; banks' existing experience with statistical models, as well as their dominance in the consumer credit market; the country's high level of financial inclusion; and the provisions of the Credit Data Repository Law.

A similar conclusion is drawn in a report recently published by the Basel Committee, which notes that banks currently employ artificial intelligence cautiously. At this stage, AI models are primarily used for comparison purposes or to support human decision-making rather than to fully automate processes or make autonomous decisions. This cautious approach is driven, in part, by regulatory and legal uncertainty regarding expectations for accountability, ethical standards, data privacy, fairness, transparency, and explainability, especially in the context of consumer-facing applications.

Credit offer by the lender – Looking ahead, it appears that artificial intelligence may be integrated at this stage of the process, enhancing user experiences through applications, chatbots and virtual assistants. At present, however, these tools will probably not be used in direct interaction with clients, partly due to challenges related to liability, disclosure, data protection and ethics, as discussed in previous chapters. In the initial phase, it is more likely that AI will be deployed in the form of internal support tools designed to help bank employees navigate internal systems and respond to questions on various topics, including organizational matters. Regarding generative AI models, banks currently limit their use in client-facing services and in activities defined by them as high-risk. Nevertheless, some banks are exploring or piloting internal applications aimed at improving operational efficiency.

Monitoring borrower behavior and early detection of credit failures – Greater accuracy and consistency in processing compared to manual or rule-based monitoring can enhance the identification of anomalies and the detection of unusual transactions, thereby supporting the recognition of typologies associated with suspected money laundering and terrorist financing. In addition, advanced analytical capabilities enable technology providers to analyze clients' financial behavior before they experience financial distress and to propose solutions at an early stage, which can also facilitate the borrower's timely rehabilitation.[465]

Issues Arising from AI use

The issues discussed in this report concerning the use of artificial intelligence in the financial sector are, to a large extent, also applicable in the credit sector. For example, where generative AI is used to communicate with clients, questions arise regarding the appropriate disclosure of the use of such systems. Similarly, where "black box" systems are used for credit underwriting or monitoring, explainability issues become pertinent. In addition to the general AI-related issues discussed in previous chapters, the use of artificial intelligence across the credit lifecycle gives rise to more focused concerns. Some of these overlap with issues already discussed in earlier chapters, but they take on greater

[465] Digitalization of finance, Basel Committee on Banking Supervision, May 2024, p. 8.

importance in the credit context given the essential nature of credit services, and taking the sector's distinctive regulatory framework into consideration.[466]

Exclusion from the credit sector - As stated above, using artificial intelligence for assessing the risk of loan defaults may improve the accuracy of forecasts on this matter. This has clear advantages: accurate risk assessment will help manage the bank's risks. In turn, it can help lending institutions meet the regulatory obligation to tailor loans to the client's needs and capabilities. However, personalization may also lead to exclusion from the credit market, either through the rejection of loan applications or by making credit prohibitively expensive. In such circumstances, concerns may arise regarding price discrimination in general, and particularly with respect to populations with protected characteristics.[467]

Discrimination - Even where the regulation restricts the type of information that may be used for credit scoring or borrower underwriting, AI systems are capable of inferring sensitive attributes from indirect variables ("proxies"). This may influence decisions on whether to extend credit, as well as its pricing and the treatment of the borrower throughout the life of the loan (see further discussion in the Bias and Discrimination chapter above).

Concern regarding "pushing credit" - Artificial intelligence may improve and streamline the activity of banking entities including assessing client credit needs more quickly and accurately, saving manpower for marketing the loans and other functions. Alongside these inherent advantages, there is also concern that such developments may result in more aggressive marketing of credit to the general public, and in particular to individuals with sensitive characteristics.

Concern regarding financial system stability - Additional concerns include the risk of "systemic contagion," which may arise when multiple key entities rely on the same or similar algorithms, for example in the development of credit risk models. In the event of a malfunction or error, the impact could spread across financial institutions and undermine the stability of the entire financial system (see further discussion in the chapter on Financial Stability above).

In the public comments to the Interim Report, no specific comments were submitted regarding this chapter, other than a comment that concerns about pushing credit are not unique to artificial intelligence. Within this context, the Taskforce aimed to emphasize risks likely to be exacerbated by the use of artificial intelligence, and, on that basis, no revision was deemed necessary.

---

[466] In this context, see EU AI Act and the reference to credit as a high-risk area justifying broader rules.

[467] Tel Aviv University Report, p. 19.

Credit in the Banking System – Recommendations:

- **General**: As stated above, the general issues discussed in this report relate to the credit sector as well, and hence, the general recommendations articulated in previous chapters are also relevant to the credit sector, including with respect to discrimination. In particular, attention should be paid to recommendations regarding the adaptation of regulatory tools to address risks specific to each application of artificial intelligence (see *Comprehensive and Risk-Based View* chapter above).
- **Use of existing regulation for risk management and consumer protections**: As noted in the chapter, Guiding *Principles for the Treatment of AI in the Financial Sector*, we believe that, to the extent possible, existing regulatory frameworks should be relied upon to address the challenges posed by artificial intelligence. This approach helps avoid the creation of dedicated AI regulation that may be overly restrictive, and reduces the risk of slowing the development and deployment of artificial intelligence by financial institutions. At the same time, we recognize that the issues use of this emerging technology raises unique issues and underscores the need to clarify how existing laws apply to the AI-related activities of financial entities, in order to provide greater legal certainty to the market on the matters outlined above (see the above discussion on regulatory uncertainty as a potential barrier to integrating artificial intelligence in the operations of financial entities).

As stated above, the credit sector relies heavily on risk management conducted by the financial institution, subject to the regulator's directives on this matter.

First, we believe it is essential to maintain the risk-based management approach even when artificial intelligence is involved at all stages of credit extension and oversight. **Financial institutions adopting AI applications, are required to implement risk management and client protection principles, tailored to both the scope of AI use and to the risk level that AI applications impose**.

Second, with respect to e specific arrangements for risk management, an analysis of existing regulations indicates that they are generally suitable for AI applications currently in use, as well as for those expected in the near future. Current provisions on consumer protection, risk management and model governance, do not create obstacles to deploying artificial intelligence in the sector, while providing an adequate regulatory framework for its integration in the credit market. For example, the obligation for fair marketing applies whether the interaction is conducted by a human or through generative AI. Similarly, requirements for fair and appropriate

pricing are relevant also in cases in which artificial intelligence is employed to navigate the risk assessment process. Nonetheless, regulators may need to provide clarifications regarding factors that increase or reduce risk and the mechanisms to address them, as detailed in the chapter, *Comprehensive and Risk-Based View* above. Second, with regard to specific arrangements for risk management, an analysis of existing regulations indicates that they are generally suitable for AI applications currently in use, as well as for those anticipated for the near future.

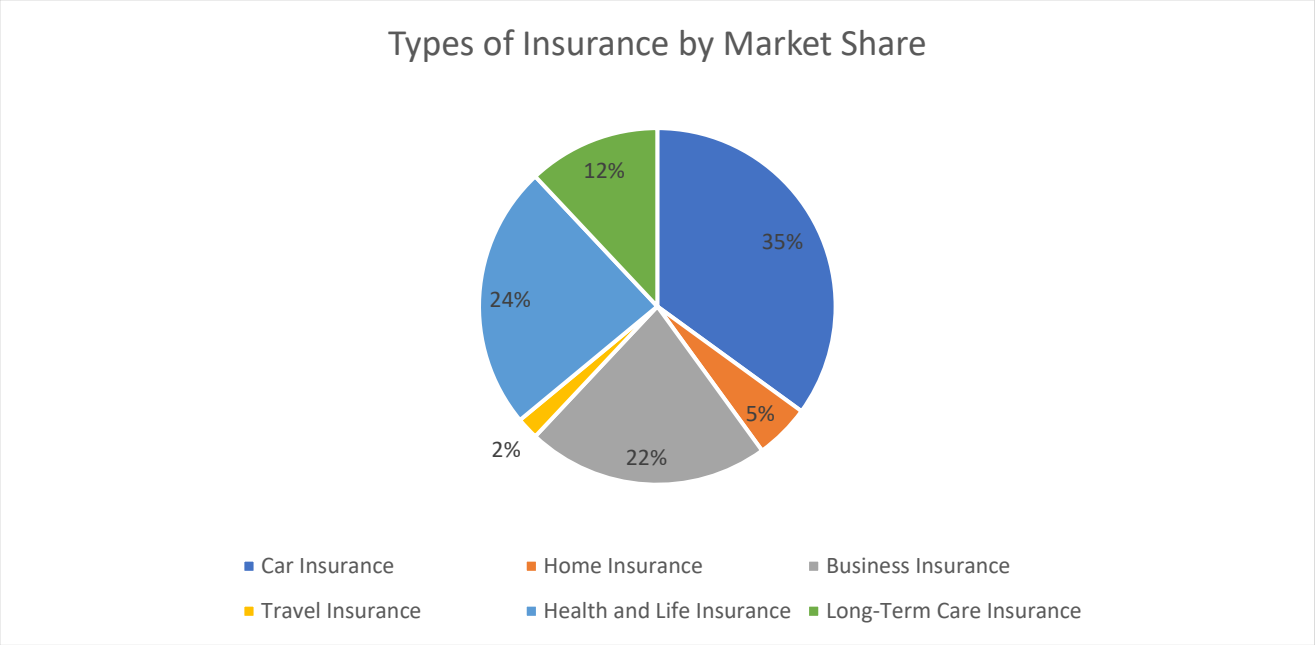
- **AI Governance**: As stated, AI governance rules already exist, alongside broader corporate governance requirements that address, *inter alia*, the roles of the board of directors and senior management in the adoption of new technologies and the development of systems and models. These rules are largely applicable to an environment in which artificial intelligence is employed. At the same time, it remains possible that dedicated AI governance tools will require further development, in line with the analysis set out in the AI Governance chapter.
- **Data Governance**: As discussed in the sub-chapter on data governance, the widespread adoption of artificial intelligence increases the reliance on data, and therefore may give rise to additional questions regarding data protection, database review, and the use of innovative or alternative data for credit risk assessment, among other matters. In this context, it appears necessary to complement existing regulation in order to safeguard data privacy and ensure that only relevant data are used, as well as to prevent the use of data that may lead to exclusion of certain populations or discrimination. Similar directives have been issued by financial regulators, such as the BIS in recent years.
- **Discrimination and exclusion risks – clarifications to existing law**: The challenges relating to discrimination were examined in detail above in the *Bias and Discrimination* chapter. As explained there, the prohibition placed on discrimination applies to financial services under general law alongside specific provisions, such as rules governing which data may or may not be considered for the credit-scoring purposes. Here too, there appears to be room for clarification that the use of artificial intelligence does not alter existing law and that financial institutions must ensure, prior to deploying AI tools, that they are able to comply with the applicable regulation– such as, avoiding the use of proxy variables for data, which under the Credit Data Law, may not be considered. In this context, mechanisms ensuring human involvement and model explainability are also important, as they can help financial institutions verify that their credit decisions are not discriminatory.

- **Concern regarding "Credit Pushing"**: Controls and disclosure relating to the use of or interaction with artificial intelligence were discussed in the *Notification and Disclosure* chapter. However, these tools also appear to be particularly relevant to concerns regarding overly aggressive credit marketing (pushing credit). Given that this is relatively novel technology, its effects on consumers and patterns of interaction are not fully understood, there may be grounds to require disclosure where artificial intelligence is used in interactions intended to market credit to clients. Such disclosure would also enable the regulator to assess whether the use of artificial intelligence has given rise to practices that raise legal concerns on this matter.

Insurance Underwriting

In the financial sector, insurance underwriting is among the areas where the potential for artificial intelligence deployment is significant. Underwriting refers to the process in which the insurance company assesses the risk of a certain insurance coverage. The key factors in the underwriting process are examining client characteristics, the characteristics of the insurance coverage they are purchasing, pricing the premium and reinsurance characteristics for that insurance coverage. The underwriting process itself includes examining the need to take steps so that the balance between all the above factors is maintained, such that the premium matches the insurance risk based on actuarial models, statistical models and algorithms.

Insurance products or policies are divided into various types and each has its own characteristics and its own risk factors. For example, health insurance covers risks of serious illnesses and expensive surgeries, home insurance covers damage to the building and property within it. As a result, each product has its own underwriting model. Below is a breakdown of market share by insurance products in 2023:



Underwriting can be divided into two separate processes: building the general underwriting model and underwriting the individual client transaction. Artificial intelligence may be integrated into each of these processes. Below is a brief description of the underwriting processes and where artificial intelligence may be deployed It should be emphasized that this is neither a set of recommendations nor a review of the existing situation, but rather an exploration of potential—some purely theoretical—avenues for AI system development.

Building a general underwriting model and determining underwriting conditions for the insurance product, including the identification and estimation of risks associated with the insurance coverage offered in the product. This process can include identifying risk groups according to various risk factors such as vehicle type or vehicle production year in vehicle insurance, or medical history in the case of critical illness insurance.

Underwriting is a complex process due to the multiplicity of insurance product types. Each insurance risk requires individualized consideration, taking into account both the availability of information and the uniqueness of the underlying risk factors. Beyond risk identification, the process of building a model and setting underwriting conditions also involves assessing the likelihood of risk occurrence and estimating the potential associated costs. This assessment can draw on the insurer's internal databases as well as external data sources.

Artificial intelligence tools, such as machine learning (ML), can be integrated into the analysis of actuarial data and pricing by leveraging the ability to process large-scale datasets from diverse sources. These tools can uncover statistical relationships and trends, supporting more accurate risk identification and estimation of both the probability and magnitude of potential losses. Another AI tool that may prove useful at this stage is Natural Language Processing (NLP), which can assist in constructing and enhancing diverse information databases.

Client underwriting or transaction underwriting is the process in which characteristics of a client or potential insurance transaction are examined using the underwriting model developed in the first stage. The process of client underwriting entails evaluating each client's characteristics, determining their suitability, categorizing them into appropriate risk groups, and making decisions regarding the insurance transaction and its conditions. This stage requires the potential client to provide relevant and reliable information regarding the type of insurance coverage, for example vehicle protection characteristics in vehicle insurance and building characteristics in home insurance.

At this stage, artificial intelligence can be integrated through intelligent interfaces, such as interactive chatbots, to manage initial client interactions and gather preliminary information. Natural Language Processing (NLP) tools can further support the processing of data collected from clients. The use of AI in client underwriting enables centralized data collection from multiple sources, including alternative sources such as online behavior, while streamlining the underwriting process. Moreover, AI can serve as a decision-support tool, offering advantages in adaptability to data, accuracy, and the consistent application of metrics across all cases, thereby reducing human bias in discretionary judgments.

In 2022, the Capital Market, Insurance and Savings Authority surveyed artificial intelligence adoption by institutional entities. Twenty institutional entities participated in the survey representing a variety of financial services in the insurance and long-term savings field as well as well as varying levels of assets under management. It was found that out of the twenty entities participating in the survey, ten already make some use of artificial intelligence models in their various business processes. In addition, it was found that additional artificial intelligence tools are under development in those entities already using artificial intelligence models as well as in several entities that have not yet deployed AI models. The survey's findings indicate integration of various artificial intelligence tools in all stages of insurance underwriting. The findings reveal that various insurance companies integrated artificial intelligence models in different underwriting stages, and that the systems serve for control and decision-making support as well. Use of artificial intelligence can be found in customer service "bots", systems for analyzing insured behavior and in systems for pricing, analyzing actuarial data and conducting risk assessment for individual transactions. The findings of this usage survey indicate the significant impact that artificial intelligence may have on the underwriting field with the advance of technological developments.

Issues and Risks

Incorporating artificial intelligence into the insurance underwriting process can offer several benefits, including improved alignment of products with client needs in terms of coverage and pricing, enhanced risk identification and assessment capabilities (including fraud detection) and, at the market level, greater accessibility to insurance products.

In addition to the advantages, however, the deployment of artificial intelligence in insurance underwriting raises several issues, including risks that while evident in the underwriting process today, may be exacerbated as a result of introducing AI tools to the process. Artificial intelligence is characterized by its ability to process large quantities of information at high processing speeds, its delivery of rapid results, its low explainability and heavy reliance on information systems. These characteristics affect the intensification of certain risks due to using artificial intelligence-based systems. The risks and issues involved in deploying AI systems have been discussed extensively in this report. We will briefly present their essentials here in the context of the insurance underwriting process:

Privacy – The first stage in the underwriting process is, as stated, information collection for the purpose of creating a risk picture of the specific client seeking to purchase insurance. Artificial intelligence integration in this process enables the collection of more detailed information, including information not necessarily relevant to then risk assessment, and from more diverse sources than do processes conducted

through other means. That is, the quantity of private information collected may grow significantly, in scope as well as in variety, which increases the risk of privacy violations, if the information is not properly protected.

Discrimination – The underwriting process in essence is a process whose purpose is to distinguish between different potential clients based on characteristics unique to them, such as the value of their property or their health condition, and based on these characteristics to adapt the insurance product to them. AI systems may assist in enhancing the accuracy of the underwriting process, *inter alia,* by finding statistical relationships between characteristics where the direct connection between them and the insurance risk is unknown. The ability to identify statistical relationships for which the explanation is not always known, may lead to creating statistical relationships between characteristics that cannot or should not be used, due to concerns of bias, creating discrimination based on this statistical relationship.

Black box and lack of explainability – One of the key advantages of artificial intelligence is the ability to locate statistical relationships which appear random to the human eye and are not based on a direct relationship or causal relationship clear even to professionals in the field. The lack of ability to explain the relationship between the data entered into the system and the results it presents may raise several challenges in the context of the insurance underwriting process. As discussed earlier in this report, the lack of explainability may make it difficult for the insurer to examine the reasonableness of AI model results or validate them. It may also make it difficult to reason a decision regarding a certain transaction. Difficulties may arise, for example, if the transaction is material to the insurer or if the client raises a question regarding it. Furthermore, a regulator, the public, or a court may request a general overview of the model's results or a detailed explanation of a specific outcome.

Model risk – Decision-support models are widely employed in the financial sector. In insurance underwriting, a quantitative-statistical model is used to estimate the likelihood of risk occurrence and the potential financial impact if such risks materialize. As noted, the use of an insurance underwriting model involves a two-stage process. The first stage focuses on model development, reliability testing, and validation, typically using information drawn from historical data on risk occurrences and insurance events. For certain insurance products, the number of past cases may be too limited to support a reliable statistical conclusion. In such instances, the model can also draw on scientific research to estimate the probability of a given risk materializing and its expected impact, as is often the case with earthquake risk. The second stage in using the model involves evaluating a specific insurance transaction. At this stage, the model incorporates transaction-specific information and assesses it against the risks covered under the insurance policy. The model's results at this stage serve the insurer in making a decision

whether to enter into the insurance transaction and in determining the transaction's terms. Among the model risks amplified by artificial intelligence are errors arising from low explainability and risks associated with the quality of data incorporated into the model, due to the diversity and multiplicity of sources.

Information system and cybersecurity and supply chain risks are additional material risks characterizing AI use. AI systems are, at their core, computer systems interconnected with other systems, relying on computing power, data transfer, and related processes, while simultaneously being exposed to the risks typical of information and cyber systems. Moreover, the development of AI systems often involves collaboration with external providers—such as domain experts offering development services or suppliers specializing in data collection and database expansion for the AI model—thereby introducing additional exposure to supply chain risks.

Artificial intelligence technology is still in the early stages of development and adoption within financial markets, and its use is expected to grow further. Integrating AI into existing operations means that, in addition to the general characteristics of AI systems, each system will possess unique attributes tied to the specific business process in which it is embedded. Consequently, risk management and the mitigation of AI-related issues will generally be conducted within the existing frameworks for managing risks in financial entities, while incorporating adaptations and updates to address the heightened risks and challenges arising from AI deployment.

Tools for Addressing the Issues and Risks

This report provides a comprehensive review of the tools needed to address the risks and challenges associated with AI use. Here, we offer a focused discussion within the context of the insurance underwriting process.

Robust and effective risk management processes are a fundamental component of sound insurance company governance, supporting careful and informed management of financial and operational activities. Such processes help ensure, with a high degree of confidence, the achievement of the company’s objectives and its capacity to meet obligations over time. The risk management framework encompasses all activities and processes within the insurance company; consequently, the integration of artificial intelligence tools into processes—such as insurance underwriting—must be implemented in a structured manner within this framework, which is designed to address the full range of risks and challenges arising from both the adoption of AI technology and its integration into the company’s

operations. Risk management tools provide the means to develop appropriate responses, prepare mitigation plans where necessary, and implement monitoring and control mechanisms to manage these risks effectively.

Corporate governance, determining accountability and human involvement serve as mechanisms to regulate the conduct of business units and various position holders within an insurance company. These mechanisms can help address a range of issues and risks. For instance, concerns such as the "black box" problem, lack of explainability, and potential discrimination can be mitigated through human oversight and the establishment of decision-making hierarchies within corporate governance policies. This ensures that the use of AI systems, and the outputs they generate, are subject to human review and discretionary judgment. Human discretion allows for the identification of results that may be unreasonable in a given context, particularly when such nuances are not fully captured in the data or models underlying the AI system. For example, in the case of a vehicle insurance renewal for individuals living in a war-affected area, discretion is required to balance the risk arising from the conflict with the insured's socio-economic circumstances and the broader social implications of renewing—or declining—the insurance. Human involvement, however, carries its own challenges: it may introduce biases and can slow down computerized processes, potentially reducing the efficiency of system development and integration. Therefore, updates to corporate governance rules, as well as the determination of accountability and the scope of human involvement, should carefully weigh both the benefits and the costs.

Explainability is a term dedicated to AI systems, as already presented in this report, and it refers to the ability to explain the outputs of an AI system in a manner that will be understandable to humans. Explainability is a technological capability developed in parallel with or during the development of the AI system, and it can assist in addressing several issues and risks. For instance, explainability enables the examination of whether an AI model considered characteristics that are typically excluded due to concerns about discrimination, while also helping to mitigate certain model risks. Explainability can further support corporate governance and disclosure mechanisms by providing insight into the factors that influenced the AI system's output, allowing for a well-informed evaluation of the results. As noted earlier in the report, explainability is not the sole means of AI system control. Other tools include running parallel statistical models on the same dataset and having experts review selected decision cases on a sampling basis.

Disclosure and notification are tools used to provide information about a financial entity's activities when a disclosure obligation applies. Transparency and information sharing facilitate external oversight

of the entity, whether by the regulator or, in some cases, the general public. In the insurance context, the results of the underwriting process are communicated to the prospective client, who receives either an insurance offer specifying the policy terms and price or, in exceptional cases, a refusal of the transaction. Insurance market regulations, which we discuss in more detail below, include directives on the information that insurers must provide to clients regarding the underwriting process, as well as additional disclosure requirements. These requirements primarily enable the regulator, and in some cases the public, to supervise and oversee both the underwriting process and the insurer's broader activities.

Existing Regulation

The Capital Market, Insurance and Savings Authority oversees and regulates insurance companies through a comprehensive framework covering all aspects of their operations. This framework includes both primary and secondary legislation, as well as directives issued by the Commissioner of the Authority (circulars). The integration of AI systems into insurance activities does not relieve insurers of their existing obligations; any new technology must be developed and operated in full compliance with the legal framework. A review of current regulations shows that most of the tools needed to address AI-related risks and challenges already exist, albeit in a general, high-level form. It is therefore necessary to assess whether AI adoption introduces specific issues that require targeted attention. We will first summarize the areas addressed by existing regulation and then outline recommendations for updating the regulatory framework.

An insurance company is required to maintain a comprehensive risk management framework and identify, assess and address all types of risks to which it is exposed, including adaptation and updating of the risk management framework so that it will also include new or developing risks. In addition to establishing a general risk framework, there are focused directives on the manner of managing various risk types including model risk, cyber and information technology risks, and outsourcing risks.[468]

Corporate governance directives establish accountability for the insurance company's governing bodies and position holders across various areas of activity, as well as decision-making hierarchies that require human deliberation. These directives are embedded throughout the regulatory framework in accordance with different risk categories, alongside a central directive assigning overarching responsibility to the board of directors, such that accountability is clearly defined across all areas of activity. In addition, the

[468] Main regulation: Risk Management – Chapter 10 of Part 1 of Gate 5 of the Consolidated Circular; Measurement of Liabilities – Chapter 1 of Part 2 of Gate 5 of the Consolidated Circular; Management of Information Technologies in Institutional Bodies – Circular 2010-9-4 dated 22.8.2010; Management of Cyber Risks in Institutional Bodies – Circular 2016-9-14 dated 31.8.2016; Outsourcing in Institutional Bodies – Chapter 6 of Part 3 of Gate 5 of the Consolidated Circular.

framework comprises disclosure and reporting obligations toward the regulator and the general public, as well as requirements to provide information to clients, including both periodic reports on ongoing activities and *ad hoc* reports relating to specific events.[469]

Discrimination and privacy topics are addressed in provisions applicable horizontally across the entire economy, and this in addition to more focused regulation applicable specifically to insurance companies. Privacy protection regulation is based primarily on the Protection of Privacy Law, 5741-1981 and regulations under it, but the Capital Market Authority also created complementary regulation for financial entities under its supervision. The regulation includes establishing confidentiality principles in information management, risk management procedures, and guidelines for encryption and information storage. This is in addition to information security and cyber directives applicable to information databases. The discrimination prevention regulation is composed of provisions of the Equal Rights for Persons with Disabilities Law, 5758-1998 and the Prohibition of Discrimination in Products, Services and Entry to Entertainment Venues and Public Places Law, 5761-2000, as well as directives issued by the Commissioner of the Capital Market, Insurance and Savings Authority. Existing regulation establishes a framework for distinguishing between clients through the underwriting process and includes restrictions on data collection and in certain cases, restricts the use of specific variables in the underwriting process. Additionally, uniform conditions and standard policies have been established for various types of insurance, such as vehicle, health and retirement savings.

[469] Main regulation: Institutional Body Board of Directors – Chapter 2 of Part 1 of Gate 5 of the Consolidated Circular; Disclosure and Reporting – Part 4 of Gate 5 of the Consolidated Circular. examine the oversight tools available to the regulator and their suitability for examining business decisions and work processes conducted through artificial intelligence.

Insurance Underwriting – Recommendations:

Guiding principles in the work of financial regulators address the need to ensure a stable and reliable financial environment and together with this, an environment in which new and innovative technologies and working methods can be developed, all for the benefit of the consumers of financial services. Creating regulation to address risks and issues arising from artificial intelligence should conform to these principles. Existing regulation on insurance companies is risk-based and is technologically neutral. As a result, an insurance company choosing to adopt artificial intelligence should develop and implement artificial intelligence technology in accordance with the regulation. However, existing regulation is not dedicated to the new technology, and the need to update it should be examined, and ensure it provides an effective response to key issues and risks in a manner consistent with the unique characteristics of artificial intelligence. Model risk management, explainability and disclosure, for example, disclosure about the use of artificial intelligence as a means of contact with the client (chatbot) are areas requiring special attention. In addition, there is a need to examine the supervision and control over insurance companies and the tools available to the regulator to implement oversight. For example, the insurance underwriting model is subject to review by the regulator upon the introduction of a new insurance product, and information regarding the model's results is also provided in connection with the handling of public inquiries. The intensification of AI use raises a need to examine the oversight tools available to the regulator and their suitability for examining business decisions and work processes conducted through artificial intelligence.

Part E

Additional Measures to Promote Financial Regulation in the Field of Artificial Intelligence

The recommendations in this section do not pertain to a specific issue related to AI activity (explainability, discrimination, and so on) or to a specific type of activity (investment advice, credit, and the like), but rather to infrastructural actions required on the regulatory side to render financial-sector AI regulation advanced and effective. These recommendations address three issues: fostering innovation; enhancing regulatory certainty and legal adaptability; and strengthening supervisory oversight.

Fostering Innovation

a. Regulatory Sandboxes and Innovation Hubs

In recent years, new technologies have fundamentally transformed financial activity across multiple domains. Routine operations—such as making payments or depositing checks—have moved from manual, in-person processes to remote execution via computers or mobile devices, while securities investment has become more readily accessible through trading platforms offered to the general public. The potential inherent in new technologies has prompted financial regulators worldwide to encourage their adoption across a wide range of activities, with the objectives of expanding financial activity, enhancing efficiency and cost-effectiveness, and facilitating the entry of new competitors into the financial system. Regulatory involvement in technology adoption can also assist regulators understand these technologies and address the risks inherent in them. In this context, financial regulation aspires to promote market development through new technologies without compromising the objectives of protecting the public, while maintaining the integrity and stability of the financial system.[470]

Given the limited adoption of AI technology in Israel's financial sector at this stage, the implementation of two regulatory techniques could be critically important to AI integration in the financial sector — establishing regulatory sandboxes and creating innovation hubs.

The term "**regulatory sandbox**" is used to describe a regulatory strategy that grants developers of financial technology a flexible, time-limited testing environment for product development, under relatively accommodating regulatory supervision that includes a reduction of burdens and supervisory requirements. Simultaneously, this environment allows the regulator to monitor the risks posed by the new technology, assess the nature and extent of required regulation in parallel with the product

[470] See more on the confrontation of financial regulation with innovation challenges: Emilios Avgouleas "Regulating Financial Innovation" The Oxford Handbook on Financial Regulation (Eds. Niamh Moloney, Eilis Ferran and Jennifer Payne)(Oxford University Press, 2015), https://ssrn.com/abstract=2458335; OECD/KDI (2021), Case Studies on the Regulatory Challenges Raised by Innovation and the Regulatory Responses, OECD Publishing, Paris, https://doi.org/10.1787/8fa190b5-en.

development stages, and adapt it to the knowledge obtained about the new activity, based on the recognition that regulators are often characterized by information and expertise gaps in the technological world relative to the market.

Since the regulatory sandbox model was first implemented in 2016 in the UK by the FCA, its use has become widespread, adopted by many regulators globally to promote technological innovation, encourage the entry of new players into the market and increase competition, while maintaining a supervised operating environment.[471]

Steps have also been taken in Israel to advance this regulatory instrument, reflecting recognition of its importance in promoting innovation in the financial sector. The primary step was drafting the Law for the Encouragement of Technology Development in the Financial Sector in Israel, 5781-2021 bill, published in January 2021 (hereinafter in this chapter: "**the Draft Law**").[472] The Draft Law sought to facilitate and establish provisions adapted to innovative activity (whether subject to licensing requirements under law or not), in order to test the business model and enable examination of how the service or product affects consumers. It was proposed that these limitations be set for a period of up to two years, with the possibility of extending them by an additional two years subject to certain circumstances. Since companies participating in the program may have an advantage over non-participating companies, and in order to increase transparency for the benefit of companies considering applying to the program as well, it was proposed to establish a publication requirement regarding various decisions made by regulators concerning the provisions to be set for participating companies. The explanatory notes to the proposal stated that the publication of decisions is also important to protect the participant's customers or investors, and is designed to enable them to manage their affairs knowing that this is a company operating within a framework of an innovative and short-term track.

---

[471] Ruth Plato-Shinar, "The Role of the Law in Developing and Designing the Payments Ecosystem in Israel", 5 **Law Society & Culture** 361-402 (2022), Ruth Plato-Shinar, "Enhancing AI in Finance Through Regulatory Sandboxes", in 8 **Law Society and Culture – Special Issues on Artificial Intelligence and the Law** 289-318 (Niva Elkin-Koren & Maayan Perel eds., 2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4666710; Inter-Ministerial Team Report on Smart Regulation, pp. 30–129, https://www.gov.il/BlobFolder/reports/corona-recommendation-report/he/smart-regulation-team-recommendations-report.pdf?fireglass_rsn=true;
Buckley, Ross P. and Arner, Douglas W. and Veidt, Robin and Zetzsche, Dirk Andreas, Building FinTech Ecosystems: Regulatory Sandboxes, Innovation Hubs and Beyond (November 1, 2019). University of Luxembourg Law Working Paper No. 2019-010, European Banking Institute Working Paper Series 2019 – no. 53, UNSW Law Research Paper No. 19-72, University of Hong Kong Faculty of Law Research Paper No. 2019/100, Washington University Journal of Law and Policy, Vol. 61, 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3455872.

[472] Government Bill for Encouragement of Development of Technology in the Financial Sector in Israel, 5781-2021 (Gov. Bill 1390). The Bill was presented following Government Resolution No. 3411, regarding the establishment of a trial environment for financial technology companies, dated January 2018, and the report of an inter-ministerial team which published a report in January 2019 regarding an outline for the establishment of a trial environment as stated.

It was further proposed to establish a special statutory committee that would serve as a one-stop-shop for companies seeking to participate in the program, and would include the financial regulators and additional government entities. This framework is designed to enable the various regulators to coordinate their treatment of companies participating in the program, prevent inconsistencies in supervision and regulatory arbitrage, and facilitate the systematic accumulation of expertise by the regulators.

A regulatory sandbox model can assist in the safe introduction of AI systems to the financial sector, as part of a dynamic learning, temporary regulatory environment. The dynamic and rapidly evolving nature of AI technologies challenges the capacity of Israel's traditional financial legislation to adequately accommodate such activity. Therefore, an efficient and flexible approach to regulation is required that will enable the potential development of this field to be maximized and even encouraged, while shaping it in a way that does not impair the objectives of financial legislation.[473] For example, entities seeking to use AI technologies to assist in making credit decisions or to build and present investment portfolios to customers will be able to do so under an adapted regulatory environment. These entities will benefit from a regulatory environment that provides the flexibility necessary for developing and operating the new activity, while simultaneously providing close supervision that will enable protection of the public and fulfillment of other regulatory objectives, as well as enhance the regulator's familiarity with the characteristics of the new activity.

This recommendation is consistent with the global trend, which indicates that the use of regulatory sandboxes may become particularly common in the context of AI regulation.[474] The European Union has emphasized the role that regulatory sandboxes may play in developing safe AI technologies, and countries such as the UK and Singapore have addressed this specifically in the context of financial services.[475] One caveat, however, is that there are question marks regarding how effective this regulatory technique actually is. A significant criticism in this context is that to date, only a small number of cases have been handled within sandboxes, and it is unclear to what extent these have contributed cumulatively to innovation.[476]

In the Israeli context, there is concern that the Draft Law, if advanced, is insufficient for promoting AI activity. It establishes a mechanism that may be cumbersome, as it requires action through a statutory

---

[473] See Policy Document, p. 100.

[474] Lawrence G. Baxter, Adaptive Financial Regulation and RegTech: A Concept Article on Realistic Protection for Victims of Bank Failures, 66 Duke Law Journal 567-604 (2016), https://scholarship.law.duke.edu/dlj/vol66/iss3/5

[475] Attrey, A., M. Lesher and C. Lomax (2020), "The role of sandboxes in promoting flexibility and innovation in the digital age", OECD Going Digital Toolkit Note, No. 2, https://goingdigital.oecd.org/data/notes/No2_ToolkitNote_Sandboxes.pdf, pp. 30-32; Plato-Shinar "Enhancing AI in Finance Through Regulatory Sandboxes ", *supra* note 471, pp. 13-16.

[476] See Building FinTech Ecosystems: Regulatory Sandboxes, Innovation Hubs and Beyond.

committee. The rationale for this mechanism was to provide a single address when an entity seeks responses from multiple financial regulators. However, AI technology may be used in core activities under the supervision of a specific regulator, so it is doubtful whether it is necessary to handle applications through a special statutory committee that includes all regulators and other entities. In addition, the Draft Law presented significant time limitations for activity within the sandbox (two years + two years). The rationale for these limitations was that the deviation from existing legal provisions was permitted only on a temporary basis, with the expectation that permanent regulation could be developed during this period. However, given that this is dynamic and evolving environment, and the time required to establish regulation (particularly through legislation) is very long, these limitations are expected to impair the effectiveness of the regulatory sandbox. These difficulties may create a negative incentive for entities, particularly large financial institutions interested in investing substantial resources in AI activity, to do so.

It is proposed to examine changes to the Draft Law, and to formulate legislative amendments to the relevant financial laws that will enable the establishment of independent sandboxes by each financial regulator within its area of responsibility, where this is not possible under existing law.

Granting such specific authority is consistent with the general approach proposed in this report, whereby increased regulatory flexibility that enables rapid adaptation, among other things, of testing environments to a wide range of AI products is required, given that this is an innovative, learning, and evolving technological field. It also stems from the recognition that AI may be used in core financial activities supervised by a single regulator, in which case, the statutory committee mechanism would be superfluous. An additional advantage of an independent sandbox may be in the dimension of regulator shepherding during the development and launch stages of the product or service, even if regulatory adaptations are ultimately not required. The Taskforce's recommendation regarding regulatory sandboxes naturally deals with AI technology, but it should also be considered with respect to other technologies as well.

Another regulatory technique, falling within the category of initiatives aimed at existing and prospective regulated entities (outreach initiatives), which the Taskforce recommends promoting, is the establishment of **innovation hubs**. Whereas a regulatory sandbox is designed to support specific initiatives seeking temporary regulatory relief, innovation hubs enable a range of broader measures intended to foster engagement, dialogue, and collaboration between the regulator and regulated entities. Both techniques can serve together, and in fact in the UK and Australia, sandboxes emerged from innovation hubs established by financial regulators. The objectives of both techniques are quite similar, and include broadcasting a message of openness and flexibility toward the market from the regulator;

promoting innovation and competition among financial institutions; and acquiring regulatory knowledge.[477]

Possible advantages of an innovation hub lie in reaching out to a broader group of entities, and in the ability to raise awareness and generate interest around a specific topic, such as AI applications (and even in a particular field, such as credit or insurance underwriting). Common methods for outreach are technological-regulatory hackathons, competitions, conferences, pilot programs, and dialogue with specific entities. In Singapore, for example, one can see a multitude of activities initiated or supported by the state to encourage the introduction of technology in general, and AI technology in particular, to the financial services industry.[478]

In Israel, an example is the Fintech Regulatory Innovation Hub established by the Israel Securities Authority in 2018. The hub's purpose is to assist and provide informal guidance to fintech companies interested in operating in Israel. The innovation hub assists entities seeking to launch operations in the local market to understand the legal framework in Israel and determine whether it is appropriate for their field of activity. Another objective of the hub is to learn and become better acquainted with the various players, their needs, the barriers facing technology companies seeking to operate in Israel, and to deepen the Authority's expertise and knowledge about technological and other issues pertaining to the companies' activities. Companies approaching the innovation hub are typically in the initial activity stages of a Business to Consumer (B2C) product and prefer, in the initial stage, to become well acquainted with the existing legal framework so that they may adapt their operations to it, sometimes even prior to obtaining a legal opinion.[479]

Another program operated by the Israel Securities Authority and the Israel Innovation Authority since 2020 is the Pilot Program, which enables Israeli technology companies in the financial sector to conduct a pilot with their end customers (the Israel Securities Authority, a financial institution, or the general public). Companies participating in the program are offered access to the Israel Securities Authority's databases and stock exchange trading data, funding at a level of 30%-50% of pilot expenses, and regulatory accompaniment by the Authority's staff on all issues related to it.

---

[477] Ibid, p. 60, 78-79.

[478] See on the Singaporean Authority website, https://www.mas.gov.sg/development/ai-in-finance.

[479] The Fintech team at the Israel Securities Authority, which operates the Innovation Hub, consists of representatives from all departments in order to provide a quick and integrative response to companies. In this framework, the Authority does not provide financial assistance or guidance on legislative matters that are not within its area of responsibility, and activity in the Hub is not a substitute for required advice, legal or otherwise. During its period of operation, the team met with many companies, accompanied them in becoming familiar with the legal framework, and referred them at times to programs relevant to their activity or to other government bodies relevant to their activity.

During 2022, the Bank of Israel joined this program as an observer, assisted in evaluating applications submitted in areas under its purview, and provided professional support to selected companies. The professional support provided by the Israel Securities Authority and the Bank of Israel enables companies a higher level of certainty regarding questions arising from the implementation of new technologies and their compliance with legal provisions. However, the Authority is not authorized to grant any exemptions or relief from legal provisions within the framework of the program. Since its establishment, 15 companies have operated within the framework of the Pilot Program, benefiting from a total funding scope of approximately NIS 14.5 million.

As part of the establishment or expansion of innovation hubs, it is proposed to examine possible connections not only between regulators and existing and potential financial entities, but also with academia and relevant entities from civil society. The advancement of AI technology demonstrates a close link between academic research and business entrepreneurship, and an innovation hub should serve to bring these parties together.[480]

The Taskforce's recommendation, therefore, is to build on existing activity and expand the use of regulatory sandboxes and innovation hubs, while emphasizing AI activity as a primary field for advancement.

b. Examining Solutions to the Information Accessibility Problem

The adoption of AI technology requires access to significant databases. New or small financial institutions may find it difficult to obtain the volume and variety of data required for developing or using an AI model. The information problem already exists today, for example, regarding the development of a statistical model for loan or insurance underwriting, and this problem may intensify with the entry of AI technology. One possible solution is opening existing databases or establishing and opening new databases for use by the relevant parties. Examples from around the world of developing databases for AI purposes span a variety of uses—from identifying health problems to predicting forest fires.[481]

---

[480] The Bank of Israel, the Israel Securities Authority, the Innovation Authority, the National Digital Directorate, the Ministry of Economy and the Ministry of Finance joined the Fintech IL – Innovation Community, operated by the Israeli Fintech Center Association. The community's vision is to promote innovation in the financial sector in Israel in order to make a wide range of advanced, high-quality and competitive services accessible to the general public and the Israeli business world, and to realize the potential of the fintech industry as a source of economic growth and employment of high-quality workers and as an export branch. The community's goals are, among others, to assist financial institutions to adopt, promote and manage innovation processes and to promote efficient work interfaces between the financial regulators and the industry, and to initiate collaborations and knowledge sharing between the various players. For more information see: https://fintech-il.org.il/.

[481] Niklas Berglind, Ankit Fadia and Tom Isherwood "The potential value of AI – and how governments could look to capture it" McKinsey (July 25, 2022), https://www.mckinsey.com/industries/public-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it.

The establishment and opening of databases correspond with an open government (open data) policy that preceded AI technology, with the aim of promoting values such as governmental transparency, realizing citizens' right to information, and encouraging economic growth.[482] In the financial context, opening databases may also be important from a competitive standpoint, so that information does not confer an exclusive advantage on large financial institutions. See in this regard the analysis above in the *Competition* chapter, which explains why the lack of access to information inputs constitutes one of the most significant competitive concerns from the entry of AI activity, and proposes action on this matter.

The challenges involved in establishing and opening databases for AI purposes are numerous. These include privacy and information security challenges, data quality, integration of information from different sources, and controls over access to information, its storage, use, and deletion. In the financial context, databases may contain personal information, and therefore their complexity is significantly greater compared to databases that do not contain such information. The possibility of taking steps to open databases that will be accessible to market players is subject to confidentiality obligations and to the principles and rules of personal data protection in databases as established by law.

The Taskforce's recommendation is to examine solutions to the information accessibility problem, particularly for small or new entities, including examining the possibility of establishing or opening databases as a basis for AI activity.

Promoting Regulatory Certainty and Legal Adaptations

c. Tools for Enhancing Regulatory Certainty

Throughout the report, various questions are presented that will arise with the entry of artificial intelligence into financial activity. At the same time, the assumption is that at this stage, dedicated regulation to govern this activity is unwarranted, and reliance should primarily be on existing law. What, then, are the ways to address these questions? There are various working tools that financial regulators already use today, which may assist in dealing with regulatory uncertainty, and even for the purpose of promoting certain financial activity. It is proposed to also use these tools for regulating technological activity in general, and artificial intelligence in particular.

One way to enhance legal certainty is through the issuance of **policy documents**. This report effectively does so, laying out guiding principles, regulatory issues, and a proposed approach for addressing them.

[482] Tehilla Shwartz Altshuler "Open government policy in Israel in the digital age" The Israel Democracy Institute, Policy Study 91 (2012), https://www.idi.org.il/media/3512/pp_91.pdf ; Anneke Zuiderwijk and Marijn Janssen "Open data policies, their implementation and impact: A framework for comparison" Government information quarterly 31, no. 1 (2014) 17-29.

In the future, it may be possible to consider formulating policy for more specific areas of activity (including those not discussed in this report) and publishing it, with the aim of further clarifying to financial institutions what is expected of them when developing or using AI technology.

A regulatory tool already available to financial regulators today is the **pre-ruling**, which is a way to `receive the regulator's position on a forward-looking issue. A pre-ruling is typically a complex question with innovative or cross-cutting aspects for which there is no structured answer or precise guidance from the regulator, and it sometimes serves regulated entities seeking to obtain the regulator's opinion regarding a transaction or certain activity they are about to embark on or that has already been conducted.[483]

Another regulatory tool is publication of **interpretive bulletins**, within which the regulator presents a professional position (for example, legal or accounting) reflecting decisions and positions on issues concerning the application of the laws within its purview. These position statements are designed to answer questions or issues that have arisen in cases brought before the regulator, which due to their nature, may also arise with other regulated entities under its authority, such as the interpretation of certain statutory provisions. The position statements further detail the considerations that led the regulator to adopt the professional position. A similar tool is the publication of **Questions and Answers** (Q&A) documents by the regulator.[484] Within this framework, answers to concrete questions that were asked and that may be relevant to additional entities are published.

Another tool for enhancing regulatory certainty, which is usually not directed at regulated entities but at the public at large and non-regulated entities, is the **publication of guides and warnings** concerning financial activity. Recently publications of this type have been issued concerning AI activity. For example, warnings concerning AI-associated fraud risks, advice regarding steps to take before investing in companies claiming to be involved in AI activity, and the like.[485]

---

[483] See for example Procedure for Preliminary Application to the Israel Securities Authority (June, 2008). The courts have recognized the importance of this regulatory tool. See the remarks of the Court regarding the fact that it is appropriate that public companies in Israel be able to apply for the preliminary position of the Israel Securities Authority and act in reliance upon it, in a manner that allows them to act dynamically and efficiently. Class Action (Economic Tel Aviv) 2484-09-12 Hatzlacha - The Consumers' Movement for the Promotion of a Fair Economic Society v. Cohen, p. 27 (Published in Nevo, 22.04.2014).

[484] A communication channel of publishing questions and answers already exists today at various financial regulators, such as the Bank of Israel, the Capital Market, Insurance and Savings Authority and the Israel Securities Authority: see for example, the Q&A page on the Bank of Israel website, the Q&A page on the subject of insurance and another page on the subject of financial information service relevant to information sources that are institutional bodies on the Capital Market Authority website, as well as the Q&A and Staff Positions page on the Israel Securities Authority website.

[485] See *supra* note 373.

A large part of the supervisory work of carried out by financial regulators employs these tools. Accordingly, they may also be helpful in dealing with AI-based financial activity, which creates uncertainty regarding the applicability of laws with respect to one activity or another. Use of these tools, already available now, provides an opportunity to interpret laws and apply them flexibly, efficiently, and variably according to need and technological development, especially where the field is not regulated from the outset in legislation. Beyond this, these tools often serve as part of the response given to risks involved in a given financial activity.

d. Authorization to Establish Special Provisions Concerning AI Activity

A central challenge in regulating technological activity in general, and artificial intelligence in particular, is the rapid pace and dynamic nature of technological development. These characteristics require the regulator to be on continual high alert in order to keep pace with developments, while addressing the changing risks posed by the technology. Sometimes there is even a need for regulation at a stage where there is still uncertainty regarding the uses and implications of the new technology, especially when dealing with generative AI technology that learns and develops rapidly, in a manner that makes hermetic regulation problematic.

The underlying assumption is that regulation aimed at overseeing the development and use of AI in the financial sector must be introduced progressively and adapted flexibly to keep pace with technological advancements. Rigid legislative and regulatory procedures that hinder timely updates to existing rules are unlikely to address these challenges effectively and may instead pose a significant obstacle to innovation and market competition. Consequently, relying solely on formal legislative processes is impractical.[486] Similarly, the regulatory tools described above, which largely depend on interpreting existing laws, also have inherent limitations in their effectiveness.

Financial regulation includes many provisions established by the regulators themselves, in a manner that enables flexibility and adaptations to technological changes and other needs arising in the market. At the same time, it is also based on primary and secondary legislation, for which making changes is much more complex and protracted as mentioned. The Israel Securities Authority in particular believes that the legal system in its supervisory domains is overly legislation-oriented and rigid, and that its rule-making powers should be expanded to facilitate timely regulatory adjustments.[487]

---

[486] See Report of the Inter-Ministerial Team for Smart Regulation, *supra* note 471, pp. 99-100.

[487] In areas where supervision was transferred to the Israel Securities Authority in recent years – financial information services and payment services – a general legislative framework was indeed established while the rest of the regulation was done by way of directives. The Israel Securities Authority proposed extending this model, which is accepted in financial supervision in Israel and abroad, to additional areas under its supervision. See in this context the proposal for expanding the Memorandum

Legislative amendments should be considered where required with the aim that regulation established by law and regulations will reflect the principle of technological neutrality and permit flexibility to operate through various technological means. Additionally, and without derogating from existing authorities of financial regulators, it is proposed to consider **granting regulatory authority to issue required instructions against the background of technological changes** where necessary and subject to legal provisions. This will achieve the dual value of providing a response to risks involved in using AI, alongside preserving its potential advantages, encouraging innovation, effective supervision, market development, and removal of excessive requirements and entry barriers.

This approach supports the principle of technological neutrality accepted in financial regulation and presented in the chapter on guiding principles for treatment of AI in the financial sector above. According to this principle, the regulatory framework should be established in a manner that is neither contingent nor conditional on a particular technology. Once the conduct of an activity must comply with financial laws given its nature and the risks associated with it, the fact that the activity is performed through AI does not derogate from the applicability of the law, even if certain adaptations are necessary. Since legislative or interpretive frameworks are often too cumbersome to accommodate frequent updates, it is recommended that financial regulators implement such adaptations flexibly through regulatory provisions, based on legal authorization and formulated in accordance with the principle of technological neutrality. Much of the existing financial regulation is already established through such regulatory provisions, so this recommendation specifically applies to laws where this approach is required.

Additional tools that can assist in achieving the objective of flexibility and response capability are **exemption powers** from regulatory provisions (which also enable granting conditional exemptions), **temporary provisions, and sunset provisions**. These authorities enable dealing with unforeseen market and regulatory needs, as well as establishing regulation for a predetermined limited period, in order to reduce risks or define a trial period.

Advancing Supervisory Activity

e. Use of Artificial Intelligence for Regulatory Purposes (SupTech)

This report primarily deals with AI applications in financial activity. There is room, however, to also draw attention to the great potential that AI has in regulatory oversight.

---

of Law for Efficiency Regulation in Securities (Legislative Amendments), 5780-2020, https://www.new.isa.gov.il/images/Fittings/isa/asset_library_pic/al_lobby/al_lobby-637f62a916518/RMP6122.pdf.

Financial regulators have long been using technological means to perform their tasks. The term Supervisory Technology, or SupTech for short, is used to describe this activity. This term includes the use of technologies for various purposes including digitization of regulatory processes such as filing reports by regulated entities, the collection of supervisory information and its analysis for regulatory purposes, and the development of monitoring and enforcement tools for regulated activity.

As part of SupTech activity, AI technology can greatly assist financial regulators in their work. As regulation and supervised activity becomes based on broader information, the supervisory challenge grows, and the financial regulator itself is required to rely on information technologies. AI may therefore assist in analyzing information about the market and regulated entities, in identifying defects in supervised activity (for example, missing reports or market manipulation), in identifying anomalies or risks in regulated entities (for example, data reflecting growth in various risks of regulated entities), and in analyzing trends and problems in the supervised market as a whole. Similar to the variety of possible uses that financial institutions have for AI in their business activity, financial regulators may also find a variety of different possible uses for AI in their activity, and generate analyses, forecasts, or other content within the framework of supervisory activity through it.[488] SupTech activity using AI also has added educational value for the regulator, which can thus become acquainted with the technology and its applications up close, in a manner that will also assist it in its oversight activity.

Around the world, AI is already being applied for supervisory purposes. A report by the International Monetary Fund (IMF) cites several examples: the Bank of Italy (Banca d'Italia) uses AI to predict credit failures; the Bank of Spain (Banco de España) employs AI to analyze ESG-related disclosures; and the European Central Bank applies AI to review fit-and-proper questionnaires and detect issues.[489] Other instances include the UK Financial Conduct Authority (FCA) using AI to manage reporting and compliance checks,[490] the Bank of Thailand reviewing board minutes of financial institutions,[491] and

[488] OECD (2021), OECD Business and Finance Outlook 2021: AI in Business and Finance, OECD Publishing, Paris, https://doi.org/10.1787/ba682899-en, in Chapter 5 (p. 121 et seq.).

[489] IMF "Powering the Digital Economy – Opportunities and Risks of Artificial Intelligence in Finance" (October 2021), https://www.imf.org/-/media/Files/Publications/DP/2021/English/PDEORAIFEA.ashx?force_isolation=true, p. 11; European Central Bank "Suptech: Thriving in the digital age" (15 November 2023), https://rb.gy/ldey1i

[490] The FCA is in the midst of a codification project (digital regulatory reporting) of the laws it is entrusted with so that they will be machine-readable and allow supervised entities to submit reports and information to the regulator in a fast, comparative, coordinated manner and also allow compliance checks within minutes. The second and third (current) stages of the project dealt, inter alia, with natural language processing. FCA "Digital regulatory reporting", https://www.fca.org.uk/innovation/regtech/digital-regulatory-reporting#section-project-overview.

[491] Financial Stability Institute "FSI Insights on Policy Implementation No 37 – Suptech tools for prudential supervision and their use during the pandemic" (December, 2021) 18, https://www.bis.org/fsi/publ/insights37.pdf?force_isolation=true.

Brazil's competition authority (CADE) using AI to identify cartels in public tenders.[492] These examples illustrate only a fraction of the current applications, with many further potential uses yet to be explored. Beyond the potential benefits for improving the capabilities and efficiency of supervisory bodies, Suptech solutions also have the potential to lower the regulatory burden imposed on supervised entities. For example, the transition to digital and structured reporting processes may ease the burden on both the supervisor and the supervised entity, and save resources on both sides.

Alongside the potential of AI for supervisory purposes, its use also entails certain risks and challenges. Some of these are not unique to AI but are common to the broader adoption of technology in public-sector applications; for example, limitations related to budgetary and human resources. However, the use of AI for SupTech purposes also presents risks and challenges specific to this context. One key challenge relates to information, which forms the essential foundation for AI operations. The system's ability to function efficiently and accurately depends on a large, high-quality, and up-to-date dataset—resources that may not always be readily accessible to the regulator. Another major challenge involves the legal aspects of deploying AI systems in regulatory work. These include questions about using information for its intended purpose, data storage and management—particularly when handled by external parties—and the admissibility of AI-generated outputs as evidence in legal proceedings. A third significant challenge concerns the "black box" problem, explainability, and the role of human oversight, issues also relevant for regulated entities. For regulators, and public authorities more broadly, the lack of full explainability can hinder transparency and complicate oversight of their own AI-driven actions.[493]

In Israel, financial regulators make extensive use of technologies within the framework of their supervisory and enforcement functions, but to date, advanced AI technologies have not been used, and this possibility is currently under examination. The Israel Securities Authority, for example, began a project in 2024 to integrate machine learning and natural language processing financial models to enable analysis and insights from data accumulated in its databases while saving supervisory and enforcement resources.

To promote the development of AI-based supervisory activities, the Taskforce considers that steps should be taken along several dimensions:

First, the feasibility of deploying AI-driven supervisory tools should be examined as part of the regulator's broader strategy for information systems and for leveraging technology in supervisory

[492] OECD Business and Finance Outlook 2021, *supra* note 463, at p. 129.

[493] For elaboration regarding these challenges and additional risks and challenges in SupTech activity of financial regulators, see Ibid. pp. 135–140.

and enforcement activities. In this context, it is proposed that regulators examine the possible needs for such activity, the possible contribution of the activity relative to existing systems, the feasibility of the activity, etc. It should be noted that, notwithstanding the natural inclination to rapidly adopt AI-based approaches, regulators may face constraints or policy considerations that counsel against doing so, and each authority should therefore undertake its own independent assessment. For example, some regulators may choose to continue relying on existing systems and defer adoption until the technology further matures, thereby benefiting from a "late-mover advantage." In other cases, budgetary or staffing constraints may limit the feasibility of such initiatives.[494]

Second, it is proposed to update, where necessary, regulators' powers to obtain information for learning, analytical, supervisory, and enforcement purposes. Financial regulators already possess a range of information-gathering authorities, some of which are general in scope and others more narrowly tailored—for example, provisions requiring the submission of details relating to a specific action by a regulated entity in order to assess compliance with legal requirements. The Israel Securities Authority has taken the view that addressing information-related challenges may require broader powers enabling the systematic collection of information from all relevant market participants, such as data concerning the activities of large traders operating algorithmic trading systems on the stock exchange. Expanded information-request powers can support supervision of a wide range of financial activities, including those involving the use of AI. At the same time, information requirements may impose regulatory burdens. Accordingly, it is essential to ensure that both the scope and implementation of such powers are confined to reasonable use and do not place disproportionate demands on regulated entities. As noted above, improving the format and methods of information submission can also help mitigate these burdens.

Third, it is proposed to allocate dedicated government resources, outside the regular budgetary frameworks, to promote the development of AI-based supervisory activities. A scenario in which the private sector rapidly adopts new technologies while the public sector lags behind is problematic, both given the critical role of public authorities—here, the protection of investors and financial consumers—and because insufficient progress in the public sector may ultimately also impede the private market's advancement.

---

[494] See also pilot for implementing AI in government ministries: https://www.gov.il/he/pages/most_ai_government_agencies_open_call_winners?fireglass_rsn=true#fireglass_params&tabid=afc880ab5bd32cd6&start_with_session_counter=5&application_server_address=isagovil-1-me-west1.prod.fire.glass.

Finally, it is proposed that cross-cutting issues arising from the use of AI in the public sector be addressed in a centralized manner, with a view to developing a uniform framework for their treatment. The deployment of AI by public authorities raises a range of challenges that may be difficult for each authority to manage independently. In this context, what unites financial regulators is not the specific subject matter of their supervision, but rather their status as public authorities that may employ AI for supervisory and enforcement purposes, much like other governmental bodies.

The Taskforce is not necessarily the appropriate forum to address these issues comprehensively and therefore considers it appropriate to recommend that relevant government entities develop consistent guidelines and practical tools for handling such uses, which would also support the regulators represented on the Taskforce. By way of comparison, in the United States, the Biden Presidential Executive Order addressed AI use in the public sector in a cross-cutting manner.[495] The advantages of this approach include savings in time and resources, including shared expertise that may not be available within each authority, as well as the establishment of a uniform and well-understood framework for addressing similar challenges.

f. Harnessing Resources of Financial Regulators and Continued Work of a Joint Task Force on the Subject

The Taskforce found many advantages in joint work on the subject of artificial intelligence in the financial sector. This is a new and complex subject that raises many shared questions, and therefore the establishment of a broad-based task force incorporating all relevant parties was essential for thorough engagement with the issue.

The establishment of the inter-ministerial Taskforce yielded several advantages, including the sharing of regulatory knowledge through a structured process of mutual learning, whereby each participating entity contributed its existing expertise on AI. The Taskforce was also able to hold numerous substantive meetings with external stakeholders, such as academic experts and foreign regulators, and to provide a platform for relevant market participants to present their perspectives on artificial intelligence. In addition, the Taskforce facilitated clarification of cross-cutting issues associated with the use of AI technology—such as privacy, corporate governance, and explainability—allowing each entity to present its position and the regulatory techniques it applies. These activities contributed to the creation of a centralized knowledge hub. In this regard, the Taskforce deliberately involved additional staff members

[495] The Executive Order.

from the participating organizations in its work, both to enrich the discussions and to broaden institutional familiarity and expertise in the field of artificial intelligence.

Considering that the Taskforce's work was done at an early stage of AI entry into the financial sector in Israel, and in light of technological and regulatory developments in the world, a logical recommendation of the Taskforce is to continue harnessing the joint resources of the Ministries of Justice and Finance and the regulators to this subject. Continued work of a joint Taskforce is expected to greatly assist in dealing with the subject and to create professional continuity in this period when AI technology is only beginning to be implemented in practice among financial institutions.

Within the framework of public comments on the Interim Report, the need arose for investing regulatory resources in the subject and advancing the regulation of regulatory sandboxes. The Taskforce agrees with these comments and believes they have been reflected in the report as well.

Summary of additional actions to promote financial regulation in the field of artificial intelligence:

Encouraging innovation

1. Updating the government bill concerning regulatory sandboxes in the financial sector, which establishes dedicated and unified regulation for all financial regulators to create a testing environment, and expanding it to increase its potential effectiveness.
2. Alongside updating the said bill, it is proposed to establish independent sandboxes by each financial regulator within its area of responsibility, and to initiate legislative amendments to the relevant financial laws to enable this where precluded by existing law.
3. Establishing innovation hubs by financial regulators and dedicating resources and focusing activity in the near term on AI applications, in addition to existing ones.
4. Examining solutions to the problem of limited information accessibility, particularly for small or new entities, including examining the possibility of establishing or opening databases as a basis for AI activity.

Promoting regulatory certainty and legal adaptations

5. Enhancing regulatory certainty in the field of AI through tools such as policy documents, responses to pre-ruling requests, interpretive positions, questions and answers, publication of guides and warnings.
6. The regulation in law and regulations should reflect the principle of technological neutrality and permit flexibility to operate through various technological means. Additionally, and without derogating from existing authorities held by financial regulators, it is proposed to consider granting regulatory powers

to issue required instructions against the background of technological changes, to grant exemptions, and to establish temporary provisions and sunset provisions, for adapting the law to AI activity and enhancing legal certainty, where required.

Advancing supervisory activity

7. Advancing supervisory activity based on AI technology (Suptech), by examining it as part of the regulator's overall strategy regarding information systems and harnessing these systems for supervisory and enforcement purposes.
8. Allocating government resources, outside regular budgetary frameworks, to advance AI-based supervisory activity.
9. Examining the need for updating powers for obtaining information for analytical, supervisory, and enforcement purposes.
10. Addressing in an orderly and centralized manner cross-cutting issues related to the use of AI in the public sector and designing a uniform treatment format.
11. Harnessing resources of the financial regulators, encouraging coordination between them, and continuing joint work on the subject.

Appendix C - List of Entities That Appeared Before the Taskforce

Bloomberg
European Securities and Markets Authority (ESMA)
European Commission
Financial Conduct Authority (FCA)
Google
Dr. Marie Briere
Prof. Mark Fagan
Microsoft
Organization for Economic Co-operation and Development (OECD)
Uri Eliavyev
Dr. Uri Hacohen
Prof. Orly Lobel.
Amplifier
Ariel Ben Porat
One Zero Bank
Israel Discount Bank
Bank Hapoalim
Mizrahi-Tefahot Bank
Dana Yafe
First International Bank of Israel
Ziv Cohen
Dr. Hofit Wasserman Rozen
Haim Ravia
Talia Gillis, Dr.
Yoav Avnstein
Yaron Hazan
Microsoft Israel Research and Development
Prof. Michal Gal
Nir Gerson
Dr. Omri Rachum-Twaig
Prof. Ofra Amir
Prof. Katharina Ligett
Shoko Insurance

Appendix D – Public Comments Received to the Interim Report