Public Consultation

Artificial Intelligence in the Financial Sector

Interim Report of the Inter-Agency Task Force on the Use of Artificial Intelligence in Israel's Financial Sector

Open for public comments until December 15, 2024.
Responses can be submitted to tomg@isa.gov.il

* This is a free translation of the Report's Executive Summary. The full report (in Hebrew) can be found here.

Artificial Intelligence in the Financial Sector –

Interim Report of the Inter-Agency Task Force on the Use of Artificial Intelligence and Machine Learning in Israel's Financial Sector

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, which recently passed comprehensive legislation 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.

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 explain how and why an AI system reaches the output it generates in a manner that can be comprehended by human beings. 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 size and complexity, 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 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.

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 (medium or 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 medium and 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)**.**
- **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.**
- **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 under the condition that the technology plays a substantial role in the decision-making process.** 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 contrast to imposing a human involvement requirement for each individual decision, there is a need to mandate human controls on the AI system's activity overall.** Such a requirement is designed to ensure the system's proper functioning from the standpoint of compliance to laws and regulations as well as minimizing errors and bias.
- **The obligation to deploy systems that allow control and human involvement is imposed on the regulated financial entity.**
- **The human involvement requirement shall be considered while taking into account other measures applying to the system, such as explainability and the ability to appeal on its decisions.**

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**. 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 private data anonymization methods with advanced methods used abroad, particularly in the financial sector.

- **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. For example, the adoption of formal requirements for providing consent, or a request to receive renewed consent from customers who had not consented in the past to process data about them for AI purposes. Another recommendation is that the use of "incremental consent" mechanisms, such that the known scope for the purpose of the consent will be aligned with the purpose of the data processing as it is known at that time, and the request for consent shall be updated when there is an increase or a revision of the specific purposes of use. **Alongside these recommendations, it is proposed to consider 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.

- **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. The report includes a call for public comment as whether to apply special obligations and rights to inferred data, such as determining that in relationships characterized by a heightened duty of good faith and fairness (such as the financial services industry), a person shall have the right to reasonable inference with respect to him. Reasonable inference means an inference that meets 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).

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**.**

- **A discussion of the discriminatory effects doctrine, is ongoing in the United Sates, *inter alia,* in the financial context in order to address prohibited discrimination risks. In this context, the task force requests public comment regarding the effectiveness and feasibility of this doctrine with respect to the uses of AI in Israel's financial sector.**

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, or from transferring it to a third party, including the customer itself. 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.

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.

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 **three-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**.
- **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 low-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.** In the questions referred to the public for comment, the task force requests, among other things, comments on whether this type of risk classification is beneficial, and whether further concretization is required at this time.

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.

- **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.**

Select 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.

This interim report is released for public comment, and consequently, the recommendations may change pursuant to the comments received. The task force encourages the submission of comments on all relevant aspects, and has included several possible questions for consideration at the end of the report.