The EU's New Law on Artificial Intelligence: *Global Implications*

The goal to create a responsible regulatory framework for artificial intelligence ("**AI**") remains an important pillar of Europe's strategy in terms of shaping its digital future. A significant milestone in this regard was recently achieved by way of a high-level political, albeit provisional, agreement ("**Provisional Agreement**") on a proposed law with respect to regulating AI (such law, the "**AI Act**").

The Provisional Agreement, arrived at in Brussels a couple of months ago, relates to a final legislative text for the AI Act. The agreement was reached pursuant to 'trilogue' negotiations among three decision-making institutions of the European Union ("**EU**") – its Parliament ("**Parliament**"), the Council of the EU ("**Council**") and the European Commission ("**Commission**"), respectively.

CURRENT STATUS AND NEXT STEPS

The Provisional Agreement underwent revision pursuant to technical meetings through December and January, further to which it was submitted to representatives from EU Member States for endorsement. According to reports from February 2024, after country-level approval was achieved, most members across key parliamentary committees ratified the agreement a few days ago. However, this ratified draft will require confirmation from both legislative bodies of the EU (*i.e.*, the Parliament and the Council, respectively). In this regard, a vote is scheduled for April. Thereafter, the confirmed text may undergo additional legal-linguistic revision before it can be formally adopted in order to become law.

Once such text has been published in the EU's Official Journal, the AI Act may enter into force 20 days later – possibly in the second or third quarter of 2024 – followed by a grace period. Since the current draft suggests that the AI Act will apply two years after its entry into force (although certain provisions could come into effect at a later date), the new law is likely to apply from 2026.

GLOBAL RELEVANCE OF THE AI ACT

The AI Act is poised to be the world's first comprehensive law on AI. Like the effect produced by the EU's General Data Protection Regulation ("**GDPR**") in terms of influencing data protection regimes and privacy trends worldwide, the AI Act may also provide a universal template for regulating AI. Further, its final provisions could have important effects on global markets, including on economies and companies in non-EU countries.

Further, the AI Act may significantly affect the use of AI systems in regulated products worldwide, subject to international and national standard-setting bodies, global markets and local requirements. The impact on high-risk AI systems for human services could be considerable, especially if such systems are built into online (or otherwise internationally interconnected) platforms. In addition, transparency requirements with respect to those AI models that interact with humans, such as chatbots and emotion detection systems, may require global disclosures on websites and apps.

THE PROVISIONAL AGREEMENT

Given that it is subject to further revision, a consolidated draft of the current text related to the AI Act has *not* yet been officially released – although certain unofficial versions were recently leaked online.

Further, various reports circulated across European print and digital media, along with press releases issued by the Parliament and the Council, respectively, provide indicative information about the current draft. As such, certain provisions in the Provisional Agreement remain identical or similar to those from previous versions of the proposed law. However, new items of discussion were introduced in the final stages of the trilogue negotiations, such as one concerning a tiered approach to regulate foundation models (discussed below).

DEFINITION OF AI

In order to ensure that the relevant definition provides clear criteria for distinguishing AI from simpler software systems, the Provisional Agreement aligns such definition with the approach proposed by the Organization for Economic Co-operation and Development ("**OECD**"). As of now, a revised version of the OECD's definition of AI has been adopted to include generative AI systems (discussed below).

APPLICABILITY AND EXTRATERRITORIALITY

In general, the AI Act applies to: (1) providers of AI systems (*i.e.*, individuals and entities, including third-parties, that (i) develop an AI system; place a developed AI

system on the market; or (iii) put such system into service under their own name or trademark), as well as authorized representatives, importers, distributors and product manufacturers; and (2) deployers of such systems (*i.e.*, any natural or legal person, public authority or other body using an AI system under its authority).

Accordingly, both developers of AI systems (*e.g.*, a system intended to screen EU residents for the purpose of ascertaining eligibility related to jobs or loans), as well as deployers of such systems (*e.g.*, a European company or banking institution) would be subject to provisions of the AI Act. Since such systems are distributed through complex value chains, the Provisional Agreement clarifies the allocation of responsibilities among different actors in such value chains (including AI providers and deployers).

Further, irrespective of their place of establishment, providers and deployers of those AI systems that are used, or which produce an effect, in the EU will be covered under the AI Act. Accordingly, applicable legal obligations will extend to AI developers and users in countries such as India or the US. Moreover, non-EU entities may be required to appoint an EU-based authorized representative. For a discussion on the AI Act's extraterritorial application and global effects, see here.

RISK-BASED APPROACH

The AI Act adopts a risk-based approach that involves evaluating AI systems based on their capacity to cause harm to society, including by viewing them through the prism of certain risk categories, such that applicable legal requirements will vary depending on the level of risk posed by the AI system concerned.

Unacceptable risk

AI systems which involve unacceptable risk are prohibited (*e.g.*, social scoring systems based on aggregate behavior or personal characteristics; emotion recognition at workplaces and educational institutions; biometric categorization to infer sensitive data such as sexual orientation or religious belief; untargeted scraping of facial images from the internet or via CCTV footage to create facial recognition databases; cognitive behavioral manipulation; and predictive policing).

High risk

AI systems may be classified as 'high-risk' when they could potentially create an adverse impact on people's safety or their fundamental rights, including those systems used in safety-critical applications, critical infrastructure and decision-making (*e.g.*, AI used across insurance, banking, education, HR or other sensitive/significant sectors). However, the use of such systems is likely to be permitted in principle – subject to

prescribed compliance requirements such as: carrying out a mandatory fundamental rights impact assessment, implementing human oversight, undertaking conformity assessments, integrating quality and risk management systems, as well as registration and post-market monitoring by surveillance authorities. Further, individuals will have a right to receive explanations about decisions based on the use of high-risk AI systems when the operation of such systems affect their rights.

On the other hand, AI systems which are *not* likely to cause serious fundamental rights violations or other significant risks may be exempt from certain legal requirements. In this regard, the Provisional Agreement seeks to introduce a new filtering system to capture genuinely high-risk applications. For example, a high-risk AI system may lose its classification as such if it is:

- only intended to perform a "narrow procedural task" (*e.g.*, transforming unstructured data into structured equivalents);
- meant to review or improve the result of a previously completed human activity (*i.e.,* merely providing an additional layer to human activity);
- purely meant to detect (i) decision-making patterns, or (ii) deviations from prior decision-making patterns (*e.g.*, for the purpose of flagging potential inconsistencies or anomalies); or
- only used to perform preparatory tasks in respect of critical use cases (*e.g.*, file handling).

Limited or medium risk

Certain AI systems which present limited risk (*e.g.*, chatbots and deepfakes) will be subject to light-touch transparency obligations, such as informing users that the content which they are engaging with is AI-generated, so that users can make informed choices.

Low or no risk

AI systems which do not fall within the scope of the risk categories mentioned above may be developed and/or used subject to existing legislation *without incurring additional legal obligations* (*e.g.*, AI-enabled spam filters or recommender systems). However, the providers and deployers of such systems may voluntarily (1) adhere to requirements under the AI Act, and/or (2) implement codes of conduct.

GENERAL-PURPOSE AI/FOUNDATION MODELS

Since systems which are deemed high-risk could be built on top of, and/or through the use of, general-purpose AI systems ("**GPAI**" or "**Foundation Models**"), GPAI providers will need to comply with a list of product safety obligations related to risk

management, data quality, security and others. An amended definition of GPAI in the current draft of the AI Act now intends to cover generative AI models (*e.g.*, ChatGPT, Dall-E, Bard). The Provisional Agreement reflects a consensus on minimum requirements with respect to GPAI deployment, such as: (i) adopting measures to ensure compliance with European copyright law, (ii) publishing detailed summaries about the datasets used for training AI models, along with information to downstream providers; (iii) preparing technical documentation in respect of GPAI use; and (iv) transparency measures (such as watermarking).

In addition, the Provisional Agreement introduces new, increased requirements for providers of GPAI when it involves 'systemic risk' (*e.g.*, powerful large language models ("**LLMs**") which exceed a prescribed threshold in terms of computational power used for training, such as GPT-4). While the present threshold may be later revised, additional criteria – *i.e.*, any criterion *other than* computational power – may be included in the future, such as user volume or an LLM's degree of autonomy.

Specifically, providers of such 'systemic risk GPAI' will need to: (i) perform model evaluation according to standardized protocols; (ii) conduct risk assessments and risk mitigation, (iii) report serious incidents or malfunctions directly to the Commission, (iv) conduct and document adversarial testing (*i.e.*, systematically evaluate their machine learning ("**ML**") model to learn how it behaves when provided with malicious or inadvertently harmful input); (v) maintain an adequate level of cybersecurity and physical infrastructure; as well as (v) track, document and report on energy consumption. Accordingly, the users of these models may need to examine their deployment on a case-by-case basis.

EXEMPTIONS

It appears that most open-source AI systems, including smaller open-source Foundation Models, will remain exempt from the AI Act's requirements. However, open-source models that are classified as 'systemic risk GPAI' will be covered by the law. In addition, AI systems which are used exclusively for (i) military or defence purposes; (ii) research and innovation; or (iii) non-professional reasons, will be exempt from the requirements of the AI Act.

MEASURES IN SUPPORT OF INNOVATION

The Provisional Agreement also encourages innovations in AI, with dedicated provisions to facilitate (i) regulatory sandboxes (*i.e.*, controlled environments for AI testing, development and validation), as well as (ii) real-world testing – subject to certain conditions and safeguards.

PENALTIES

Similar to the way fines are calculated under the GDPR, penalties for violating the AI Act will be calculated as: (i) a percentage of the liable party's global annual turnover in the previous financial year, or (ii) a fixed sum – whichever is higher. For instance, it could amount to: (1) EUR 35 million or 7% for violations which involve the use of prohibited AI applications; (2) EUR 15 million or 3% for violations of obligations under the AI Act; or (3) EUR 7.5 million or 1.5% for the supply of incorrect information. However, like with data protection laws, proportionate caps may be placed while issuing fines upon small enterprises and start-ups.

PRODUCT AND OTHER LIABILITY

In September 2022, the Commission published two draft directives intended to adapt and extend the EU's existing liability rules (adopted in 1985) to new digital technologies (including AI): *viz.*, the "**New Product Liability Directive**" and the "**AI Liability Directive**," respectively. Together, these directives aim to complement the AI Act by addressing civil liability for AI systems – including by providing remedies and redressal for affected individuals. Such changes are likely to affect the responsibilities of key stakeholders involved in the deployment of AI systems and will thus need to be carefully assessed by such stakeholders for the purpose of minimizing risks. As discussed above, under the AI Act itself, individuals and entities can lodge complaints about non-compliance and exercise their right to receive explanations about decisions based on high-risk AI systems.

PREPARING FOR THE AI ACT

Several companies have started preparing for the AI Act. Several such compliance obligations require prior planning. For instance, relevant entities should first acquire a proper understanding of new legal requirements – some of which can vary depending on (i) a company's role in the AI value chain, as well as (ii) the risk profile of an AI system. The main preparatory steps include: (i) establishing a cross-functional team; (ii) engaging expert resources; (iii) performing a comprehensive assessment of the scope and risk profile of both current and planned AI systems; (v) developing a compliance plan and related policy; (vi) establishing governance mechanisms and program management; (vii) establishing reporting and oversight procedures; (ix) maintaining comprehensive records and documentation; (ix) performing ongoing monitoring; and (x) reviewing and updating cybersecurity measures.

Companies may be able to leverage certain aspects of their data protection compliance programs to comply with requirements under the AI Act, including in respect of data governance. For instance, policies and processes for managing

personal data under the GDPR or India's Digital Personal Data Protection Act, 2023 ("**DPDP Act**") could provide a foundation for responsible data use and oversight while developing or deploying AI systems. Further, built-in checks and controls, including those related to cybersecurity and data breaches, could minimize those risks which the AI Act aims to address.

In addition, companies could adapt existing approaches and methodologies on privacy impact assessments for the purpose of evaluating AI-related risks, including in respect of use cases, training data, as well as impacts on fundamental and other rights. Regular reviews of information systems under the GDPR or the DPDP Act, such as data and process mapping, may provide clarity about how information moves into and across an organization – thus highlighting both risks and unknown factors.

Further, maintaining records of data processing activities could provide valuable documentation and audit trails to demonstrate key measures taken by an entity in respect of AI governance. Such records may also prove useful to ensure that rights related to intellectual property are appropriately addressed.

In the short term, potential impacts might include increased compliance costs along with delays in the development and deployment of AI systems, which may, in turn, reduce the pace of innovation in AI technology. However, in the medium and long term, the AI Act may produce several benefits, such as by (1) changing the way in which AI systems are developed, tested, documented and deployed, with an increased focus on trustworthy and responsible AI, as well as (2) drawing more investment on AI safety. Further, harmonized AI safety standards could stimulate the emergence of a new global framework.

*This insight has been authored by **Dr. Deborshi Barat** (Counsel) and **Reshma (Vaidya) Gupte** (Counsel). They can be reached at dbarat@snrlaw.in and rgupte@snrlaw.in, respectively, for any questions. This insight is intended only as a general discussion of issues and is not intended for any solicitation of work. It should not be regarded as legal advice and no legal or business decision should be based on its content.*