The Future of AI Regulation (Part 2): Draft Legislation from the European Commission Shows the Coming AI Legal Landscape

April 26, 2021

In this Part 2 of our series on the future of artificial intelligence ("AI") regulation, we examine the draft EU legislation. Part 1 of the series (on U.S. banking regulators' RFI) can be found here. Upcoming parts of this series will cover the recent FTC pronouncement on AI, as well as steps companies can take now to prepare their AI programs for the coming regulatory landscape.

On April 21, 2021, the European Commission published its highly anticipated draft legislation governing the use of AI, which is being referred to as the "GDPR of AI" because, if enacted, it would place potentially onerous compliance obligations on a wide spectrum of companies using AI systems. The commission proposes to regulate AI based on the potential risk posed by its intended use: AI systems that pose an "unacceptable risk" would be banned outright; AI classified as "high risk" would be subject to stringent regulatory and disclosure requirements; and certain interactive, deepfake, and emotion recognition systems would be subject to heightened transparency obligations.

Notably, the increased focus on transparency in the use of AI, coupled with specific reporting obligations for AI providers and users, will almost certainly result in more scrutiny of AI by consumers, regulators, and stakeholders. Indeed, in the same way that GDPR caused companies to significantly expand their privacy compliance, the draft AI legislation is designed to encourage companies to treat AI as an enterprise-wide risk that requires attention from their leadership in the development, deployment, and oversight of their AI systems. That encouragement is to be reinforced in the draft legislation with the prospect of severe legal and reputational consequences for companies that fail to implement robust compliance policies around their AI systems that pose risks to EU residents. In addition, the Commission proposes a labeling regime (the CE marking of conformity), whereby certain AI systems would need to be assessed and certified for conformity by a qualifying "notified body" prior to entering the market.

Although the draft legislation will probably not take effect for several years, its broad scope, and the specificity of its obligations, situate the EU as the epicenter of AI regulation and, as GDPR was for subsequent privacy laws, it will serve as the standard

against which all future AI regulations will be measured. Below we have provided a quick overview of the key features of this landmark draft AI legislation.

Key Features of the Commission's Draft AI Legislative Framework

How the Regulation Will Apply to U.S. Companies

The Commission intends the legislation to have broad extraterritorial reach, covering AI providers or users "***irrespective of whether they are established within the Union***," so long as any AI systems affect users within the EU. In particular:

- **Providers** – persons or entities that ***develop or place an AI system on the market under their own name or trademark, even if provided free of charge*** would be covered if (i) they place AI systems on the market or into service within the EU, or (ii) the output produced by the AI system is used in the EU.

- **Users** – persons or entities that ***use an AI system under their authority, other than in a personal capacity,*** would be covered if (i) they are located within the EU, or (ii) the output produced by the AI system is used in the EU.

In many instances, multiple entities are involved in the development, training, marketing, and branding of AI systems, which could result in having several "providers" for a particular AI system.

Types of AI Systems That Will Be Banned

The draft law bans the use of certain AI, including:

- **Manipulative or exploitative systems.** The legislation would prohibit AI systems that are designed to manipulate human behavior or decisions through "subliminal techniques," or to exploit vulnerabilities of groups of persons due to age, physical, or mental disability, in a manner that would materially distort their behavior and cause them or others physical or psychological harm. These are sometimes referred to collectively as "**Dark Patterns**." This prohibition will likely need further clarification because many common AI systems have been alleged to manipulate human behavior and exploit vulnerabilities (*e.g.*, AI used for gaming, advertising, social media, dating apps, etc.)

- **"Real-time" remote biometric identification systems, such as facial or gait recognition systems**. The use of these systems in public places for law enforcement purposes would be prohibited, subject to several enumerated exceptions.

- **General-purpose social scoring**. The legislation would also prohibit social scoring based on a person's social behavior or predicted personality characteristics, by or on behalf of a public authority that would lead to detrimental treatment of a person or group under certain circumstances.

AI Systems That Will Be Regulated as "High Risk"

The draft AI legislation expressly sets out in **Annex III** the applications considered to be "high risk," including:

- AI systems that evaluate consumer creditworthiness or establish their credit score, with the exception of systems provided by small entities for their own use;
- AI systems for recruiting and workplace management, including evaluating candidates through interviews, making decisions concerning promotions or termination, or monitoring and evaluating employee performance or behavior;
- AI systems for education and vocational training;
- Systems for biometric identification of natural persons, including both "real-time" and post hoc remote identification tools (other than the law enforcement uses described above that are banned);
- AI systems for management and operation of critical infrastructure;
- AI systems concerning access to public assistance benefits or to dispatch emergency first response services;
- AI systems used by law enforcement, including risk assessments, polygraphs, deepfake detection, and crime analytics.

The commission would be empowered to add AI systems to this list if they pose a risk of harm to health and safety, or adverse impact on fundamental rights. Factors that the commission will consider in determining whether to classify additional AI applications as "high risk" include: the intended purpose of the AI system, the potential impact of future harm, and the vulnerability of intended users due to an imbalance of power, knowledge, age, or economic or social circumstances. Additionally, AI systems that produce decisions that are not easily reversible, or where "for practical or legal reasons it is not reasonably possible to opt-out from [the] outcome," are also more likely to be considered high risk. Notably, the commission also states that it will consider "reports or documented allegations" of prior incidents of harm in classifying a system as "high

risk," signaling to companies that it will be carefully considering claims of AI bias or other AI incidents that may cause injury.

New Obligations for Companies Using AI

Under the draft AI legislation, the regulatory obligations imposed on AI systems will largely fall on those that are considered is "high risk."

Requirements for High-Risk AI Systems

Any company deemed to be a provider of a "high-risk" AI system will be subject to significant regulatory requirements both prior to and after placing the AI system on the market, including:

- **Mandatory AI incident reporting to regulators (Article 62)**. Perhaps the most notable feature of the commission's draft is an incident reporting requirement akin to the data breach reporting requirements in the GDPR, CCPA, and other data protection laws. This requirement, which we have not seen in other draft AI regulations, would require providers to report any "serious incidents or malfunctioning" of their high-risk AI systems to the market surveillance authorities. Providers would be required to submit this report immediately when they become aware of a "causal link" between the AI system and the incident, and in any event, no later than 15 days after becoming aware of the incident.

- **Quality and risk management procedures (Articles 9 and 17)**. Providers of high-risk AI systems must establish, implement, document, and maintain a quality management procedure, including appropriate risk management measures. In particular, any identified risks should either be eliminated or mitigated by (i) implementing adequate control measures, or (ii) providing appropriate warnings or trainings to users. Specific consideration should be given to AI systems that are likely to be accessed by (or impact) children.

- **High-quality data and data governance practices (Article 10)**. All high-risk AI systems must be developed and trained using quality data, which, according to the commission, means appropriate (i) design choices, (ii) data collection and preparation, (iii) examination of data for potential biases, and (iv) identification of data gaps or shortcomings. The training and testing data sets must also be "relevant, representative, free of errors and complete," and have the "appropriate statistical properties."

- **Robust documentation, record keeping, and provision of information upon request (Articles 11, 12, 18, 20, 23, and 50)**. The draft legislation would create significant and detailed record-keeping obligations, including to ensure that a high-

risk AI system's outputs can be verified and traced throughout the system's lifecycle. Among other requirements, the draft legislation would require the creation and maintenance of automated logs and other technical documentation concerning (i) the characteristics, capabilities, and limitations of the AI system; (ii) its algorithms and data; (iii) the development, testing and validation processes; and (iv) any anticipated risk and corresponding risk management measures. Providers would also be obligated to maintain logs generated by high-risk AI systems (to the extent those logs are under their legal control), and to provide relevant authorities with access to those logs upon request.

- **Transparency obligations (Article 13)**. High-risk AI systems will require documentation and usage instructions, including (i) the identity and contact details of the provider of the AI system, (ii) the capabilities and limitations of the AI system, (iii) specifications for the AI system's input data (when appropriate), (iv) predetermined changes to the high-risk AI systems and its performance, (v) the human oversight measures and (vi) the expected lifetime of the AI system and any necessary maintenance measures.

- **Human oversight and manual override capabilities (Article 14)**. High-risk AI systems must be developed such that they can be effectively overseen by natural persons, including by monitoring for anomalies and intervening to interrupt the system if necessary.

- **Cybersecurity and protection from malicious third parties (Article 15)**. High-risk AI systems must be protected from hacking, abuse, and exploitation by malicious actors. Given these emerging risks, the draft legislation requires companies to implement measures designed to prevent data poisoning, adversarial examples, and model flaws.

- **Consistent and accurate performance over the AI system's lifecycle (Article 15)**. The draft AI legislation requires that high-risk systems must meet a high level of accuracy that is appropriate for their intended purpose and be resilient to errors, faults, or inconsistencies, regardless of whether those are inherent in the model or introduced by human users. The technical resilience of an AI system can also be achieved through the implementation of appropriate fail-safe and backup plans.

- **Pre-market conformity assessments (Articles 19 and 43)**. The draft legislation would require providers to either undergo a self-assessment or to obtain a conformity assessment from a "notified body" (depending on the type of system) before placing their high-risk AI systems on the market, or any time the system is "substantially modified."

- **Registration on a publicly accessible EU database (Articles 51 and 60)**. Providers of high-risk AI systems (or their authorized representatives) would be required to register any high-risk AI system in an EU database managed by the commission.

- **Post-market monitoring and corrective actions (Articles 21 and 61)**. The draft AI legislation would require providers of AI systems to establish a post-marketing program to proactively analyze data from AI systems they have placed on the market and continually review their compliance. It would also require taking any necessary corrective actions with respect to issues arising with the high-risk systems, including withdrawing or recalling the system from the market.

The draft AI legislation also imposes requirements on companies acting as "importers," "distributors," "product manufacturers," and "users." In particular, users of any high-risk AI systems would be required to (i) operate any high-risk systems in accordance with their instructions, monitor their operations, and keep copies of any logs generated by those systems; (ii) ensure that any input data they control are relevant in view of the "intended purpose" of the high-risk AI system; and (iii) notify the provider if a malfunction or serious incident occurs, or if they cannot reach the provider, notify the regulator directly.

Requirements for Other AI Systems

Providers of AI systems that fall outside of the "high risk" category remain obligated to comply with certain transparency requirements. In particular, the commission notes that AI systems that interact with individuals or create content—such as chatbots, automated advisors, or deepfake generators—may "pose specific risks of impersonation or deception irrespective as to whether they qualify as high-risk." Accordingly, the draft AI legislation imposes heightened transparency requirements on such systems:

- **Providers of AI systems designed to interact with natural persons** would be required to inform individuals that they are interacting with an AI system, if not already obvious from the circumstances.

- **Users of emotion recognition or biometric characterization systems** that identify or infer an individual's emotions or intentions would be required to notify individuals of the operation of this system.

- **Users of systems that generate or manipulate images, audio, or video content to resemble existing persons, objects, places, entities, or events that would falsely appear to be authentic**, other than for the purpose of satire or free expression, would be required to disclose that the content has been artificially generated or manipulated.

Voluntary codes of conduct for non-"high-risk" AI systems

The commission also encourages both companies and industry organizations to create voluntary "codes of conduct" that would involve adopting the requirements that are applicable to "high-risk" AI systems to their other AI systems, as well as additional requirements related to environmental sustainability, disability justice, and diversity. Companies are encouraged to consult users and various stakeholders in creating such codes of conduct.

In addition, because the commission is proposing a certification regime for the pre-market conformity of certain high-risk systems, we anticipate the emergence of a third-party standards/rating industry for AI (similar to what we have seen for cybersecurity) as well as benchmark standards and best practices for internal control. Companies will want to stay apprised of these developments.

Enforcement and Potential Fines

Under the current draft legislation, market surveillance authorities would be given primary responsibility for enforcement with respect to any "high-risk" AI system. However, a specific exception is made for any AI systems put into service or used by financial institutions, where enforcement authority is granted to the appropriate financial supervisory agencies, which creates the possibility of variability in the application of the rules and enforcement practices.

Notably, market surveillance authorities would be granted "full access" to any training, validation, and testing data sets used by the provider, as well as the source code of the AI systems upon request. Additionally, other authorities and bodies that supervise or enforce EU law relating to fundamental rights would be granted the power to request any documentation maintained pursuant to the legislation, and may coordinate with a market surveillance authority to conduct further testing of the AI system.

While EU member states will ultimately set the penalties for noncompliance, the current draft includes significant fines for companies that develop or sell prohibited AI systems or provide false or misleading information to regulators. The most significant offenses would be subject to a penalty of 6 % of a company's total worldwide annual revenue or €30,000,000, whichever is greater.

Planning for Compliance: What Can Companies Do Now?

Although the draft EU AI law will not be finalized and in force for several years, companies that are developing AI models today should pay close attention to its

obligations in order to avoid having to make significant adjustments to their models (or fully decommission them) when new AI regulations come into force.

In our upcoming installments in this series on the Future of AI Regulation, we will provide a list of steps that companies can take now to limit the risks of developing AI tools that will be viewed as noncompliant with the AI regulatory landscape that is likely to take shape over the next few years. Those steps will cover overall governance, as well as:

- accountability
- appeal rights
- bias testing
- board reporting
- business continuity
- cybersecurity
- documentation
- guardrails
- human oversight
- AI inventories
- ongoing monitoring
- privacy protection
- regulatory disclosures
- risk assessments
- training
- transparency
- explainability
- vendor management

Please join Avi Gesser and Anna Gressel for a special edition of our **DSS Webcast** on **Monday, May 3, 2021 at 10:00am ET** on the **Future of AI Regulation**, as well as steps that companies can adopt now to prepare for the rapidly evolving AI regulatory landscape. You can register for the live webcast here, and for an on-demand recording here.

* * *

To subscribe to the Data Blog, please click here.

The authors would like to thank Stephen McDougall, Chief Counsel, Data & Privacy Law at Prudential Financial, for his contribution to this article, as well as Debevoise law clerks Michael Pizzi and Andres Gutierrez for their contributions.

Please do not hesitate to contact us with any questions.

NEW YORK

Avi Gesser
agesser@debevoise.com

Anna R. Gressel
argressel@debevoise.com

Steven Tegrar
sgtegrar@debevoise.com