Professional Perspective

AI Regulation Part 2: Keeping Your Company Ahead of the Curve

Peter J. Schildkraut, Arnold & Porter Kaye Scholer LLP

Bloomberg
Law

Read Professional Perspectives | Become a Contributor

AI Regulation Part 2: Keeping Your Company Ahead of the Curve

Contributed by Peter J. Schildkraut, Arnold & Porter Kaye Scholer LLP

Familiar regulatory regimes already cover artificial intelligence. Broader regulation is coming, and we are beginning to see what it will look like. While many uncertainties remain, your company can act now to get ahead of the curve. Indeed, waiting for the adoption of new regulation may make compliance harder to achieve.

The first of this pair of articles explained how the leading type of AI, machine learning, works; why people want to regulate AI; how AI already is regulated; and the approaches the U.S., the EU, and the UK are taking on more extensive regulation. This article will discuss how to prepare a company for these regulations.

Attend to the Legal and Reputational Risks of AI

Now is the time to prepare your company for AI regulations. Retrofitting compliance may be difficult, if not impossible. For example, some types of AI incorporate their training data into the model itself. However, various privacy laws permit people to withdraw consent for continuing storage of their data. If your company is subject such a law and its AI model incorporates training data, your company should make excising data from the model as easy as possible. Building ethical AI principles like this one into development, procurement, and operations can lower the chance of having to scrap your company's efforts and start over when future regulations come into force.

Despite these incentives for anticipating regulation, the trade association CompTIA found that only 14% of IT and business professionals associate "ethical problem" with AI. If your personnel are representative, they simply aren't considering the legal and reputational risks from AI.

These risks require attention, however. Working from probabilities, not certainty, your company's AI will make mistakes. A big one will attract unwelcome scrutiny from regulators, the media, and the plaintiffs' bar.

So where do you begin?

Risk Assessment and Mitigation

To start, you should audit your company's AI projects for compliance with applicable privacy laws if you aren't doing so already. Machine learning relies upon copious amounts of data. And the collection, storage, and processing of personal data often implicates these statutes. For instance, the GDPR requires a data protection impact assessment if AI systems process personal data for decisions significantly affecting individuals. GDPR, art. 35(3)(a), 2016 O.J. (L 119) at 53.

Next, you ought to make AI risk assessment an ongoing part of your company's development, procurement, and use of AI. From a 30,000-foot view, the U.S. National Institute of Standards and Technology (NIST) suggests reviewing "accuracy, reliability, resiliency, objectivity, security, explainability, safety, and accountability." In another variant, the Institute of Electrical and Electronics Engineers (IEEE) offers eight general principles to consider: human rights, wellbeing, "data agency," or control of one's own data, effectiveness, transparency, accountability, awareness of misuse, and competence.

Getting Started

As a practical matter, you'll want a checklist like The Assessment List for Trustworthy AI (ALTAI) to explore the relevant issues. Organizations like the IEEE and the Consumer Technology Association are developing AI standards that will inform future assessment lists. Of course, any list must be adapted for your company and each of its AI systems.

Before proceeding deeply into an assessment, though, you should consider what harms can result from the use, or misuse, of the AI. If the worst harm is trivial, your company may want to conduct a thorough review to improve its product. But, from a compliance perspective, an extensive assessment would waste resources. Conversely, where the AI might harm human health, safety, or welfare, careful risk assessment and mitigation become a prudent investment.

Mitigation Through Explanation

Mitigating AI risk involves many dimensions. Explainability is a good place to start. Explaining an adverse decision enables an effective appeal if an AI prediction doesn't make sense or acceptance of the outcome if it does. Either way, the affected party will be less inclined to complain to a regulator. In addition, regulators are likely to mandate explainability in certain situations.

Not all AI predictions can be explained so humans can connect the dots. If an algorithm can predict cancer more accurately than a radiologist through data patterns people can't perceive, patients probably would choose the more accurate prediction over the explainable one. Still, even if AI output eludes human understanding, your company should be able to provide some kind of explanation. The UK Information Commissioner's Office (ICO) and The Alan Turing Institute identify six main varieties:

- Rationale explanation: “an accessible and non-technical” version of the reasoning.
- Responsibility explanation: who developed, managed, and operated the AI system and how to obtain human review of decisions.
- Data explanation: what data went into the model and how they were used.
- Fairness explanation: what processes ensure the AI's decisions “are generally unbiased and fair” and a specific “individual has been treated equitably.”
- Safety and performance explanation: how designers and operators “maximise[d] the accuracy, reliability, security[,] and robustness” of the AI's decisions and operations.
- Impact explanation: the consideration and monitoring of the effects “an AI system and its decisions has or may have on an individual, and on wider society.”

Each type of explanation will not be achievable for every AI system.

Mitigating Bias

Bias should be another focus for risk mitigation. Training data should be free from bias, in both the neutral statistical sense and the damaging human sense, but that can be hard to achieve. As a Brookings Institution paper observes, “there is no simple metric to measure fairness that a software engineer can apply.” The authors suggest a bias impact statement to flag issues. A model's development and training teams should include a broad diversity of perspectives, both in lived experience and professional training, to guard against cultural blind spots. And self-testing both can root out problems and provide enforcers “a strong sign of good-faith efforts at legal compliance,” in the FTC Acting Chairwoman's words.

Moreover, changes in society can degrade a model's performance over time, so-called “concept/model drift.” Your company should monitor potential drift in the AI systems it develops and uses, retraining its models on fresh data when necessary.

Other Considerations: Tradeoffs and Outsourcing

Risk assessment and mitigation involve tradeoffs. Accuracy and explainability may clash, as in the black-box radiology example. Accuracy and fairness may be in tension: hypothetically, for instance, predictions that reflect a person's race or gender might be more accurate, at least in some sense, but also unfair. Different aspects of fairness may conflict. These tradeoffs need careful consideration. When no acceptable tradeoff can be found, should your company still release or employ the application?

If your company outsources the design or development of an AI system, that may not relieve it from liability for assessing and mitigating risk. Procurement contracts should allocate these responsibilities clearly. And be sure your vendor's risk assessment and mitigation processes are as rigorous as your own.

Oversight

Process and Structure

The novelty, complexity, and often opacity of AI systems may require changing how your company oversees risk assessment and mitigation. Regulators are signaling they want close supervision from senior management of potentially risky AI. For example, the ICO recommends:

> Your senior management should be responsible for signing-off [on] the chosen approach to manage discrimination risk and be accountable for [AI's] compliance with data protection law. While they are able to leverage expertise from … subject matter experts, to be accountable[,] your senior leaders still need to have a sufficient understanding of the limitations and advantages of the different approaches. This is also true for [data protection officers] and senior staff in oversight functions, as they will be expected to provide ongoing advice and guidance on the appropriateness of any measures and safeguards put in place to mitigate discrimination risk.

Does your company's current compliance structure assure senior management has this understanding?

You and your board need to consider whether the board has duties to monitor AI regulatory compliance and, if so, whether your board has sufficient expertise to navigate any major risks. See In re Caremark Int'l Inc. Derivative Litig., 698 A.2d 959, 970 (Del. Ch. 1996); Marchand v. Barnhill, 212 A.3d 805, 824 (Del. 2019).

Explainability for Oversight

Having the broadest possible set of explanations for each system will facilitate oversight. If the developers, purchasers, and users of an AI system can explain its operations or why they are comfortable with its predictions, you and others performing oversight will have greater confidence in the model's accuracy and its compliance with legal requirements.

When asking for explanations, inquire into their bases. The operators may not have trained or developed the model. An AI system may rely on a mix of open-source and proprietary components and code. The proprietary elements may blend customized and commercial off-the-shelf modules. The customized portions may have been produced in house or by vendors. In short, you may need to keep "peeling the onion," to arrive at well-substantiated explanations.

Documentation

Regulation of AI should spur reassessment of your company's document-retention policies too. For one thing, regulators will demand documentation about the development and use of AI systems. For another, records likely will be needed to defend against liability for harms allegedly caused by AI.

Requirements of Regulations and Standards

The ICO has detailed the GDPR's documentation requirements for users of AI involving personal data. The ICO advises recording risk assessment and mitigation decisions, especially regarding tradeoffs among risks, lines of accountability for decisions about tradeoffs, and outcomes of these decisions "to an auditable standard." The ICO goes on to recommend these records include:

- Reviews of the risks to the individuals whose personal data is processed.
- How tradeoffs among risks and values were identified and weighed.
- The rationale for choosing among technical approaches (if applicable).
- Which factors were prioritized and the reasons for the final decision.
- "[H]ow the final decision fits within your overall risk appetite."

In addition, the ICO explains that the GDPR requires users of AI involving personal data "to keep a record of all decisions made by an AI system …[,] includ[ing] whether an individual requested human intervention, expressed any views, contested the decision, and whether you changed the decision as a result." The ICO also urges AI users to analyze this information for potential problems.

Similar documentation mandates should be anticipated in other jurisdictions and from standards-setting organizations. For instance, EU country data protection agencies are likely to interpret the GDPR as the ICO has. Moreover, for high-risk AI applications, the European Commission has suggested users retain training data, or at least records about the data. plus records of the application's programming and training to enable auditing of "potentially problematic," outcomes or outputs. Beyond written policies for the operation of "autonomous and intelligent systems," the IEEE proposes:

> Engineers should be required to thoroughly document the end product and related data flows, performance, limitations, and risks of A/IS. Behaviors and practices that have been prominent in the engineering processes should also be explicitly presented, as well as empirical evidence of compliance and methodology used, such as training data used in predictive systems, algorithms and components used, and results of behavior monitoring. Criteria for such documentation could be: auditability, accessibility, meaningfulness, and readability.

Defensive Document Retention

Even if a regulator or standard does not compel documentation, your company may want to record how its AI systems were developed, trained, and used. Properly designed, trained, and functioning AI systems still make mistakes. Their outputs come with a specified probability of being correct—not a certainty—which means they also have a specified probability of being incorrect.

If the stakes of an error are high enough, sooner or later, your company should expect a regulatory investigation or a lawsuit. The novelty, complexity, and opacity of AI systems also raise the risk the factfinder will infer your company's liability from the obvious harm itself. See Restatement (Third) of Torts: Liability for Physical and Emotional Harm § 17 (Am. Law Inst. 2010 & Oct. 2020 update); Byrne v. Boadle (1863) 159 Eng. Rep. 299, 300; 2 H. & C. 722, 725. Even more than for familiar, less complicated, more explainable technologies, your company will need to produce evidence of its due care in developing or procuring, training, and using the algorithm. Similarly, if your company's AI system leads to prohibited disparate-impact discrimination, your company will want to demonstrate its good-faith efforts to prevent this outcome. Document-retention policies must balance these risks against the problems of increased preservation.

Conclusion

The arrival of AI and its regulation means your work is cut out for you. The pace of legal change demands constant vigilance. Yet, the direction in which we are heading is clear enough to define the tasks at hand. Careful risk assessments, mitigation, attention to oversight, and documentation will help your company stay ahead of the curve.

With assistance from Darrel Pae, Katerina Kostaridi, and Elliot S. Rosenwald, Arnold & Porter Kaye Scholer LLP