AI REGULATION: WHAT YOU NEED TO KNOW TO STAY AHEAD OF THE CURVE

BY PETER J. SCHILDKRAUT

Arnold & Porter

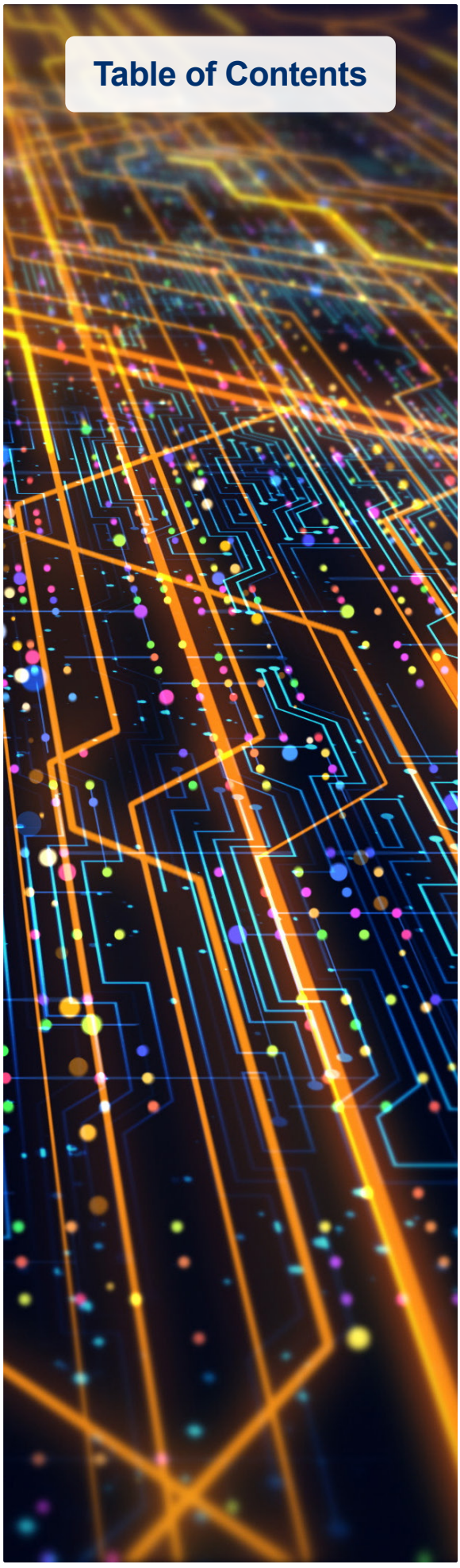


Table of Contents

An abridged version of this report was previously published in *Bloomberg Law*.

AI Regulation: What You Need to Know to Stay Ahead of the Curve

By Peter J. Schildkraut[1]

Artificial intelligence (AI) is all around us. AI powers Alexa, Google Assistant, Siri, and other digital assistants. AI makes sense of our natural language searches to deliver (we hope) the optimal results. When we chat with a company representative on a website, we often are chatting with an AI system (at least at first). AI has defeated the (human) world champions of chess and Go.[2] AI is advancing diagnostic medicine, driving cars and making all types of risk assessments. AI even enables the predictive coding that has made document review more efficient. Yet, if you're like one chief legal officer I know, AI remains on your list of things you need to learn about.

Now is the time! Right or wrong, there are growing calls for more government oversight of technology. As AI becomes more common, more powerful, and more influential in our societies and our economies, it is catching the attention of legislators and regulators. When a prominent tech CEO like Google's Sundar Pichai publicly proclaims "there is no question in my mind that artificial intelligence needs to be regulated," the questions are when and how—not whether—AI will be regulated.[3]

Indeed, certain aspects of AI already are regulated, and the pace of regulatory developments is accelerating. What do you need to know—and what steps can your company take—to stay ahead of this curve?

What Is AI?

Before plunging into the present and future of AI regulation, let's review what AI is and how the leading type works. There are many different definitions of AI, but experts broadly conceive of two versions, narrow (or weak) and general (or strong). All existing AI is narrow, meaning that it can perform one particular function. General AI (also termed "artificial general intelligence" (AGI)) can perform any task and adapt to any situation. AGI would be as flexible as human intelligence and, theoretically, could improve itself until it far surpasses human capabilities. For now, AGI remains in the realm of science fiction, and authorities disagree on whether AGI is even possible. While serious people and organizations do ponder how to regulate AGI[4]—in case someone creates it—current regulatory initiatives focus on narrow AI.

Machine Learning

One type of AI, machine learning, has enabled the recent explosion of AI applications. "Machine learning systems learn from past data by identifying patterns and correlations within it."[5] Whereas traditional software, and some other types of AI, run particular inputs through a preprogrammed model or a set of rules and reach a defined result (akin to 2+2=4), a machine learning system builds its own model (the patterns and correlations) from the data it is trained upon. The system then can apply the model to make predictions about new data. Algorithms are "now probabilistic. We are not asking computers to produce a defined result every time, but to produce an undefined result based on general rules. In other words, we are asking computers to make a guess."[6]

To take an example from legal practice, in a technology-assisted document review, lawyers will code a small sample of the document collection as responsive or not responsive. The machine learning system will identify patterns and correlations distinguishing the sample documents that were coded "responsive" from those coded "not responsive." It then can predict whether any new document is responsive and measure the model's confidence in its prediction. For validation, the lawyers will review the predictions for another sample of documents, and the system will refine its model with the lawyers' corrections. The process will iterate until the lawyers are satisfied with the model's accuracy. At that point, the lawyers can use the system to code the entire document collection for responsiveness with whatever human quality control they desire.

The quality of the training data set matters greatly. The machine learning system assumes the accuracy of what it is told about the training data. In the document review example, if, when the lawyers train the system, they incorrectly code every email written by a salesperson as responsive, they will *bias* the model towards *predicting* that every sales team email in the collection is responsive. Note that I did not say they will train the

> **THE QUESTIONS ARE WHEN AND HOW—NOT WHETHER—AI WILL BE REGULATED.**

model to identify every sales team email as responsive. The miscoded training data will increase the *probability* that the model will predict any given sales team email is responsive, but they will not make this a *certainty*. Other things about an email might overcome the bias. For instance, the lawyers may have coded every email about medical appointments as nonresponsive. As a result, the model might nevertheless predict that an email about a medical appointment is nonresponsive even if it comes from a salesperson.

Why Regulate AI?

Several characteristics of AI drive the calls for regulation.

Accuracy and Bias

AI predictions are sometimes inaccurate, which can injure both individuals and society. Poorly performing AI might underestimate a person's fitness for a job or creditworthiness. AI could crash a car by misperceiving environmental conditions or misjudging what another vehicle will do. In short, AI could harm individuals in all the ways that humans and their creations already do (and probably some novel ways too).

Policymakers may leave redress to the courts,[7] but there may be gaps in existing law that legislators decide to fill with new causes of action. Governments also may turn to regulation as a prophylactic to supplement the tort system, as many countries have done in areas such as food and consumer product safety.

The pressure may be even greater to regulate AI applications that might cause societal harm. For instance, AI can discriminate against members of historically disadvantaged groups. Imagine a human resources AI application trained to identify the best job candidates by finding those most similar to previous hires. Free from the implicit biases we all carry, it should be completely objective in selecting the best candidates for that company, right? But imagine further that the applicant pool whose resumes comprised the training set was predominantly male. Actually, you don't have to imagine. Using this method, Amazon's Edinburgh office developed a machine learning system for hiring decisions that "effectively taught itself that male candidates were preferable."[8]

Facial recognition technology also raises social-justice concerns. A study of 189 facial recognition systems found minorities were falsely named much more frequently than whites and women more often than men.[9] Privacy questions aside, using facial recognition to identify criminal suspects makes these racial differences particularly troubling because falsely identified individuals may be surveilled, searched or even arrested.[10]

GOVERNMENTS ALSO MAY TURN TO REGULATION AS A PROPHYLACTIC TO SUPPLEMENT THE TORT SYSTEM, AS MANY COUNTRIES HAVE DONE IN AREAS SUCH AS FOOD AND CONSUMER PRODUCT SAFETY.

Technology's promise is to help us escape the biases all people have. Instead, however, these examples show how AI can wind up reinforcing our collective biases. What is going wrong? First, like all of us, algorithm creators have cultural blind spots, which can cause them to miss opportunities to correct their algorithms' disparate impacts. Second, AI forms its predictions from data sets programmers or users provide. As a result, AI predictions are only as good as the data the AI is trained upon. Training data, in turn, reflect the society from which they are collected, biases included. To prevent societal biases from infecting AI predictions, developers and operators—and their attorneys and other advisors—must recognize them and determine how to adjust the training data.

Even when AI makes accurate and unbiased predictions, however, the results can be troubling. Several years ago, researchers found that Facebook was less likely to display ads for science, technology, engineering, and math jobs to women. Women were interested in the jobs, and the employers were interested in hiring women. But "the workings of the ad market discriminated. Because younger women are valuable as a demographic on Facebook, showing ads to them is more expensive. So, when you place an ad on Facebook, the algorithms naturally place ads where their return per placement is highest. If men and women are equally likely to click on STEM job ads, then it is better to place ads where they are cheap: with men."[11]

Power

In a 2018 MIT-Harvard class on *The Ethics and Governance of Artificial Intelligence*, Joi Ito relates being told that "machines will be able to win at *any* game against humans pretty soon." Ito then observes, "A lot of things are games. Markets are like games. Voting can be like games. War can be like games. So, if you could imagine a tool that could win at any game, who controlled it and how it is controlled has a lot of bearing on where the world goes."[12] It is easy to see why the public might demand regulation of this power.

Market Failures

For all its power, though, AI cannot transcend market forces and market failures (even if it might be able to win market "games"). There will be cases when AI performs accurately in a socially desirable arena, yet a market outcome may not be desirable.

Consider self-driving cars. Aside from relieving drivers of the drudgery of daily commutes, freeing time for more pleasant or productive activity, a major selling point for vehicle autonomy is safety. The underlying AI won't get tired or distracted or suffer from other human frailties that cause accidents. But how should an autonomous vehicle be programmed to pass cyclists in the face of oncoming traffic? The vehicle's occupants will be safer if the vehicle travels closer to the cyclist and further from oncoming traffic. The cyclist will be safer if the car moves closer to the oncoming traffic and further from the cyclist. Nobody wants to buy an autonomous vehicle programmed to protect others at its occupants' peril, but everyone wants other people's autonomous vehicles to be programmed to minimize total traffic casualties.[13] This is a classic collective action problem in which regulation can improve the market outcome.

How Is AI Regulated?—Application of Familiar Regulatory Regimes

Widespread regulation is coming because AI sometimes produces inaccurate or biased predictions, can have great power when accurate and remains vulnerable to market failures. (I use "regulation" expansively to include statutory constraints enforced through litigation, not just agency-centered processes.) What will this regulation look like?

Some regulation of AI will look very familiar, as AI already is regulated in certain economic sectors and activities.

Generally Applicable Law

"AI did it" is, by and large, not an affirmative defense. If something is unlawful for a human or non-AI technology, it probably is illegal for AI. For instance:

- Title VII of the US Civil Rights Act of 1964 (as amended) prohibits employment practices with "a disparate impact on the basis of race, color, religion, sex, or national origin" unless "the challenged practice is job related for the position in question and consistent with business necessity."[14] There is no carve-out for AI.
- Likewise, the US Equal Credit Opportunity Act (ECOA),[15] which also does not mention AI, "prohibits credit discrimination on the basis of race, color, religion, national origin, sex, marital status, age, or because a person receives public assistance. If, for example, a company made credit decisions [using AI] based on consumers' Zip Codes, resulting in a 'disparate impact' on particular ethnic groups, … that practice [could be challenged] under ECOA."[16]
- Antidiscrimination regimes under US state law[17] and in other countries[18] similarly apply to AI.
- The US Fair Credit Reporting Act (FCRA) requires certain disclosures to potential employees, tenants, borrowers, and others regarding credit or background checks and further disclosures if the report will lead to an adverse action.[19] Credit and background checks that rely on AI are just as regulated as those that don't. And, to comply with FCRA (or to defend against an ECOA disparate-impact challenge), a company using AI in covered decisions may need to explain what data the AI model considered and how the AI model used the data to arrive at its output.[20]
- The US Securities and Exchange Commission has enforced the Investment Advisers Act of 1940[21] against so-called "robo-advisers," which offer automated portfolio management services. In twin 2018 proceedings, the SEC found that robo-advisers had made false statements about investment products and published misleading advertising.[22] The agency also has warned robo-advisers to consider their compliance with the Investment Company Act of 1940[23] and Rule 3a-4[24] under that statute.[25]
- There should be little doubt that the US Food & Drug Administration will enforce its good manufacturing practices regulations[26] on AI-controlled pharmaceutical production processes.
- And, of course, claims about AI applications must not deceive, lest they run afoul of Section 5 of the US Federal Trade Commission Act,[27] state consumer protection statutes and similar laws in other countries. Indeed, one of the FTC's commissioners wants to crack down on "marketers of algorithm-based products or services [that] represent that they can use the technology in unsubstantiated ways" under the agency's "deception authority."[28]

> **EVEN WHEN AI MAKES ACCURATE AND UNBIASED PREDICTIONS, HOWEVER, THE RESULTS CAN BE TROUBLING.**

WIDESPREAD REGULATION IS COMING BECAUSE AI SOMETIMES PRODUCES INACCURATE OR BIASED PREDICTIONS, CAN HAVE GREAT POWER WHEN ACCURATE AND REMAINS VULNERABLE TO MARKET FAILURES.

The longer broad regulation takes to arrive, the more we should expect government enforcers to apply their existing powers in new ways. Just as the FTC has used its Section 5 authority in the absence of a federal privacy statute,[29] that agency is turning its attention to AI abuses.[30]

Express Regulation of AI

In the EU and the United Kingdom, the General Data Protection Regulation, which reaches far beyond AI, restricts the development and use of AI in connection with individuals and their personal data.[31] For instance, Article 6 specifies when and how "personal data" may be "processed,"[32] which encompasses pretty much any way one might use data with AI.[33] Articles 13-14 require "controllers" to provide clear and simple explanations about using personal data in "profiling" or other automated decision-making, including discussions of the AI's logic and the significance and anticipated consequences of the AI output for the individuals.[34] Article 22 establishes an individual's right not to be subject to fully automated decisions with significant effects on that person. To waive that right, the individual must be able to have a human hear one's appeal of the automated decision.[35] And Article 16's right to rectification of incorrect data and Article 17's right to be forgotten may require retraining or even deleting an AI model that incorporates personal data by design.[36]

US states, too, are adopting privacy and other laws expressly regulating AI.

- As amended by the California Privacy Rights Act of 2020, the California Consumer Privacy Act of 2018 contains requirements like the GDPR's (although the restrictions on automatic decision-making are left to be fleshed out by regulations).[37]
- The new Virginia privacy statute does as well.[38]
- Another California law targets intentionally deceitful use of chatbots masquerading as real people in certain commercial transactions or to influence voting.[39]
- Depending on the technology involved, the Illinois Biometric Information Privacy Act may regulate private entities' use of facial recognition technology.[40]
- Illinois also has the Artificial Intelligence Video Interview Act, under which employers must notify job applicants when they use AI to vet video interviews, explain how the AI evaluates applicants, and obtain applicants' consent to AI screening.[41]

How Will AI Be Regulated?—Development of New Regulatory Regimes

Because existing rules don't address all concerns about AI, policymakers worldwide are considering the creation of new regulatory regimes. There appears to be a loose consensus among at least the advanced democracies that the degree of regulation should be tied to an AI application's risk. For instance, a music-recommendation algorithm has much lower stakes than AI used for screening job candidates or loan applicants, diagnosing disease, or operating an autonomous vehicle.[42] The Organization for Economic Cooperation and Development (OECD) and the G20 have adopted principles embodying this high-level consensus.[43] (See box.)

OECD/G20 Principles for Responsible Stewardship of Trustworthy AI

- AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.
- AI systems should be designed in a way that respects the rule of law, human rights, democratic values, and diversity, and they should include appropriate safeguards—for example, enabling human intervention where necessary—to ensure a fair and just society.
- There should be transparency and responsible disclosure around AI systems to ensure that people understand AI-based outcomes and can challenge them.
- AI systems must function in a robust, secure and safe way throughout their lifecycles, and potential risks should be continually assessed and managed.
- Organizations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.

THERE APPEARS TO BE A LOOSE CONSENSUS AMONG AT LEAST THE ADVANCED DEMOCRACIES THAT THE DEGREE OF REGULATION SHOULD BE TIED TO AN AI APPLICATION'S RISK.

But the international consensus seems likely to fray—at least somewhat—in deciding the degree of regulation warranted by particular risks. The United States has a culture of permissionless innovation, waiting for harms to emerge and be well-understood before regulating. European governments tend to be quicker to invoke the precautionary principle: innovations should be regulated until they are proven safe. Yet, even in the United States, some warn against repeating what they see as the mistake of not regulating the internet until it was too late to prevent various harms.[44]

US Steps Toward Regulating AI

The Trump Administration

The Trump Administration sought to "promote a light-touch regulatory approach" to AI.[45] Last year, OMB published guidance on regulating AI consistent with this light-touch approach.[46] According to OMB:

- "The appropriate regulatory or non-regulatory response to privacy and other risks must necessarily depend on the nature of the risk presented and the tools available to mitigate those risks."[47] In particular, "[i]t is not necessary to mitigate every foreseeable risk. … [A] risk-based approach should be used to determine which risks are acceptable and which risks present the possibility of unacceptable harm, or harm that has expected costs greater than expected benefits."[48] This comparison should consider "whether implementing AI will change the type of errors created … and … the degree of risk tolerated in other existing systems."[49]
- Instead of "[r]igid, design-based regulations … prescrib[ing] the technical specifications of AI applications," agencies should consider sector-specific policy guidance or frameworks, pilot programs and experiments to foster creative prophylactic approaches and to learn what works, and the use of voluntary consensus standards and frameworks.[50]
- Agencies should evaluate training data quality, assess protection of privacy and cybersecurity, promote nondiscrimination and general fairness in AI application outcomes (both absolutely and compared to existing processes), and consider what constitutes adequate disclosure and transparency about using AI.[51]

The US Department of Transportation illustrated the Trump Administration's approach. As US DOT grappled with the challenges of autonomous vehicles, it prioritized voluntary, consensus-based technical standards, which can evolve more easily in tandem with technology than regulations.[52] However, when standards will not suffice and regulation is necessary, "US DOT will seek rules that are as nonprescriptive and performance-based as possible," focusing on outcomes instead of specifying features, designs or other inputs.[53]

What Will Happen Under the Biden Administration?

It remains to be seen whether the Biden Administration will have an equally light touch. In recent decades, Democrats (like Republicans) largely have supported permissionless innovation to encourage new technologies. For example, when home satellite television services were launching, both parties generally agreed to exempt them from many regulatory burdens borne by their established cable television competitors. Similarly, Democrats and Republicans alike mostly took a *laissez-faire* approach to the internet for the first quarter century of the commercial World Wide Web.

However, members of both parties have begun pressing for greater regulation of at least some internet companies and applications, arguing they have amassed too much economic, social and political power. Shaped in part by this experience, and spurred by concerns that AI and other algorithms are—under a cloak of technological objectivity—reinforcing structural biases in US society, various Democrats have called for regulation of AI even at this early stage. For example, US Senators Cory Booker and Ron Wyden and US Representative Yvette Clarke have introduced the Algorithmic Accountability Act.[54] According to its sponsors, this legislation would:

- Authorize FTC regulations requiring companies under its jurisdiction to conduct impact assessments of highly sensitive automated decision systems. This requirement would apply both to new and existing systems.
- Require companies to assess their use of automated decision systems, including training data, for impacts on accuracy, fairness, bias, discrimination, privacy[,] and security.
- Compel companies to evaluate how their information systems protect the privacy and security of consumers' personal information.

- Mandate correction of problems companies discover during the impact assessments.[55]

The Consumer Online Privacy Rights Act introduced by Senators Cantwell, Schatz, Klobuchar, and Markey similarly would require algorithmic decision-making impact assessments.[56] With an administration that is less hostile to regulation, Democratic control of the Senate, and greater public awareness of and sensitivity to systemic discrimination, such legislation may gain traction and even become law.

European Steps Toward Regulating AI

The European Commission

While the United States has yet to embrace broad regulation of AI, Europe is plowing ahead. To supplement the GDPR, the EU is moving towards additional legislation to regulate high-risk AI applications—and is likely to be more prescriptive in regulating than the United States will be.

In April 2021, the European Commission released proposed legislation regulating AI.[57] (The Commission simultaneously proposed updated legislation regulating machinery, which, among other changes, would address the integration of AI into machinery, consistent with the AI proposal.)[58] The proposed AI legislation classifies AI systems as high-risk or not based upon intended use. If adopted, the legislation would require all high-risk AI systems to meet standards for compliance programs and risk management; human oversight; documentation, disclosure and explainability; robustness, accuracy and cybersecurity; and record retention. High-risk systems would have to demonstrate compliance through conformity assessments before introduction into the European market. Some AI systems—high-risk or not—would have to meet transparency standards. Certain uses of AI would be prohibited altogether. Most current uses of AI would remain unregulated. (See box summarizing selected proposals.)[59] In the transportation sector, high-risk AI safety components, products or systems covered by eight existing legislative acts would be exempt from the proposal although those acts would be amended to take the proposal's requirements for high-risk systems into account.[60] Interested parties may submit comments during the consultation period.[61] The European Parliament and Council of the European Union then will consider the Commission's proposal in light of those comments.

TO SUPPLEMENT THE GDPR, THE EU IS MOVING TOWARDS ADDITIONAL LEGISLATION TO REGULATE HIGH-RISK AI APPLICATIONS—AND IS LIKELY TO BE MORE PRESCRIPTIVE IN REGULATING THAN THE UNITED STATES WILL BE.

Selected European Commission Proposals

Prohibited Uses

- Harmful distortion of human behavior through subliminal techniques or exploitation of age or disability.
- Real-time remote biometric identification systems in public places for law enforcement (limited exceptions).
- Governmental social scoring of individuals leading to discriminatory treatment across contexts or disproportionate to behavior.

High-Risk Uses

- Many types of safety components and products.
- Public assistance determinations.
- Law enforcement use for individual risk assessments, credibility determinations, emotion detection, identification of "deep fakes," evidentiary reliability evaluations, predictive policing, profiling of individuals, and crime analytics.
- Remote biometric identification and categorization of people.
- Evaluation of creditworthiness and credit scoring (limited exception).
- Immigration determinations.
- Admission, assignment and assessment of students.
- Emergency services dispatch.
- Judicial decision-making.
- Recruitment and other employment decisions.
- Other uses the Commission later designates as high-risk.

Selected European Commission Proposals Cont.

Requirements for High-Risk AI Systems

Compliance (Providers)

- Quality management system (compliance program), including regularly updated prescribed risk management system reflecting state of the art.
- Pre-market conformity assessment (certifications valid for up to five years absent substantial modifications) and post-market monitoring with immediate correction and reporting requirements upon reason to consider system noncompliant or risky to health, safety or fundamental rights.
- Registration in EU database.

Compliance (Others)

- Third party that sells or installs under own brand, alters intended purpose, substantially modifies system, or incorporates system into product treated as provider.
- Importers and distributors, among other obligations, must verify upstream compliance, not degrade compliance and report if system risky to health, safety or fundamental rights; distributors have correction duty upon reason to consider system noncompliant.
- Professional users must operate consistent with provider instructions, monitor operations, input only relevant data, and assess data protection impact.

Human Oversight—Humans with necessary competence, training and authority must oversee operation and be able to:

- Stop operation.
- Disregard, override or reverse the output.

Documentation, Disclosure and Explainability—To enable users to understand and control operation and to facilitate governmental oversight, providers must supply:

- Concise, complete, correct, clear, relevant, accessible, and comprehensible instructions describing:
  - Characteristics, capabilities and limitations of performance, including foreseeable unintended outcomes and other risks.
  - Human oversight and related technical safety measures.
  - Expected lifetime and necessary maintenance and care.
- Detailed prescription of continuously updated technical documentation covering (in part):
  - Descriptions of system and development process, including compliance trade-offs.
  - Monitoring, functioning and control, including system's risks and mitigation.
  - Provider's risk management system.

Robustness, Accuracy and Cybersecurity—High-risk AI systems must:

- Perform consistently at appropriate levels throughout their lifecycles, notwithstanding attempts at manipulation of training or operation or unauthorized alteration.
- Meet training, validation and testing data quality requirements.

Penalties for Violations—up to greater of:

- €30 million.
- Six percent of global annual revenue.

Retention of Records and Data

- Automatic logging of operations, ensuring traceability.
- Ten years:
  - Technical documentation.
  - Documentation of quality management system.
  - Certain conformity records.

Requirements for Certain AI Systems

Transparency—For high-risk and low-risk systems, if applicable:

- System must inform people they are interacting with an AI system unless obvious.
- Individuals exposed to emotion recognition or (unless for law enforcement) biometric categorization system must be notified.
- "Deep fakes" must be identified (qualified law enforcement and expressive, artistic and scientific freedom exceptions).

The European Parliament

Previously, the European Parliament had adopted its own recommendations for proposed legislation.[62] Parliament would not ban specific AI practices as the Commission proposed.[63] On the other hand, Parliament defined "high risk" somewhat more expansively,[64] and it suggested requiring high-risk AI not to "interfere in elections or contribute to the dissemination of disinformation" or cause certain other social harms.[65] Parliament also proposed that all conformity assessments be performed by approved third parties while the Commission would allow providers of certain types of high-risk AI to assess themselves.[66] Moreover, Parliament included an individual right to redress for violations and whistleblower protections,[67] which the Commission did not. Any of these proposals could find their way into the final legislation that ultimately is adopted.

Parliament also recommended separate legislation amending the civil liability regime for AI systems.[68] An earlier Commission report had indicated such changes might be necessary,[69] so its proposed legislation may be forthcoming.

For enforcement and other AI governance tasks, both the Commission and Parliament would look to a mix of agencies. Sectoral authorities (e.g., medical device regulators) at both the EU and member state levels would continue to implement their mandates.[70] New national authorities would fill gaps, coordinated by a new European Artificial Intelligence Board.[71] Parliament also suggests new national supervisory authorities to be coordinated by the Commission or other Union-level bodies.[72] How this mix of authorities gels (or not) will significantly influence how burdensome the contemplated regulatory regime becomes.

Whatever legislation ultimately emerges from the European Union's deliberations, exporters of AI-enabled products and services into the EU should expect to be covered.[73] The Commission's proposed legislation would even reach providers and professional users of AI systems outside the EU if the system's output is used inside the EU.[74] US and other non-European companies need to monitor the progress of this legislation and plan for compliance.

The United Kingdom

Increasing the complexity of the task, the post-Brexit UK may be forging its own path on AI regulation. The Information Commissioner's Office (ICO, the UK's data protection agency) advises developers and users of AI on their obligations under the GDPR and other legislation. When warranted, the ICO can impose significant monetary penalties for violations.[75] Moreover, the UK Department for Digital, Culture, Media & Sport (DCMS) formed the Centre for Data Ethics and Innovation (CDEI) to "investigate and advise on how we govern the use of data and data-enabled technologies, including Artificial Intelligence."[76] CDEI's mandate continues to be somewhat fluid, and the UK government promised—but has yet to propose—legislation to establish the agency more formally.[77] Nevertheless, it seems CDEI will not itself be an AI regulator. That task will remain with the ICO and sector-specific agencies. In its advisory capacity, CDEI has stated "[a]t this stage," it does not perceive "a need for a new specialized regulator or primary legislation," at least not "to address algorithmic bias."[78] However, CDEI is calling for clarification of how various statutes and regulations apply to algorithmic bias.[79]

BUILDING ETHICAL AI PRINCIPLES INTO THE DEVELOPMENT, PROCUREMENT AND USE OF AI *NOW* CAN REDUCE THE CHANCE OF HAVING TO SCRAP YOUR COMPANY'S EFFORTS AND START OVER WHEN FUTURE REGULATIONS COME INTO FORCE.

How Can You Keep Your Company Ahead of the Curve?

Generally applicable regulations cover use of AI in already-regulated activities. The United States, the United Kingdom, the European Union, and other jurisdictions[80] are—or may be—moving toward broader regulation of AI. We are beginning to see, albeit less clearly in some places, what will emerge from their deliberations. While many uncertainties remain, it is possible—indeed, imperative—for your company to get ahead of the curve.

Your company's employees probably haven't contemplated most of the issues surrounding AI regulation. The trade association CompTIA found that only 14 percent of IT and business professionals associate "ethical problem" with AI.[81] If your personnel are similar, 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 enough error will attract scrutiny from regulators, the media and the plaintiffs' bar. By being proactive, you can help your company avoid unnecessary risk.

MITIGATING AI RISK INVOLVES MANY DIMENSIONS. EXPLAINABILITY IS A GOOD PLACE TO START.

Moreover, retrofitting regulatory compliance may be difficult, if not impossible. For example, some types of AI incorporate their training data into the model itself. However, the GDPR permits people to withdraw their consent to continuing storage of their data.[82] If your company is subject to the GDPR (or another law with a similar provision) and training data are incorporated into your AI model, your company's designers should make excising data from the model upon request as easy as possible.[83] Building ethical AI principles into the development, procurement and use of AI *now* can reduce the chance of having to scrap your company's efforts and start over when future regulations come into force.

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 like the GDPR and CPRA, 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 your company's AI systems process personal data for decisions with significant effects on individuals.[84]

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 US National Institute of Standards and Technology (NIST) suggests considering characteristics including "accuracy, reliability, resiliency, objectivity, security, explainability, safety, and accountability."[85] In one of many other variants, the Institute of Electrical and Electronics Engineers (IEEE) offers eight general principles for ethically aligned design of what it terms "autonomous and intelligent systems": human rights, well-being, "data agency" (i.e., control of one's own data), effectiveness, transparency, accountability, awareness of misuse, and competence.[86]

Assessment

As a practical matter, you'll want a checklist to explore the relevant issues, but "[k]eep in mind that [any] assessment list will **never be exhaustive**. Ensuring trustworthy AI is not about ticking boxes, but about continuously identifying requirements, evaluating solutions and ensuring improved outcomes throughout the AI system's lifecycle, and involving stakeholders therein."[87] One leading list, the European Commission's High-Level Expert Group on Artificial Intelligence's *The Assessment List for Trustworthy AI (ALTAI)*, covers a broad range of topics.[88] Organizations like the IEEE are developing standards for AI that will inform future assessment lists.[89] Of course, any list must be adapted for your company generally and specifically for each AI system your company develops, procures or uses.

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 (recall the music-recommendation algorithm), 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.[90] In addition, regulators are likely to require explainability in various instances.[91]

But not all AI predictions can be explained so that humans can connect the dots. If a machine-learning algorithm can predict cancer more accurately than a radiologist through data patterns that are imperceptible to humans, I'll choose the more accurate prediction over the explainable one.[92] Even when the AI's output eludes human understanding, though, your company probably can provide other types of explanation instead. The ICO and The Alan Turing Institute identify six main varieties:

- **Rationale explanation**: "[A]n accessible and nontechnical" version of the reasons behind a decision.
- **Responsibility explanation**: Who developed, managed and operated the AI system and how to obtain a human review of a decision.
- **Data explanation**: What data went into the model and how they were used to make a particular decision.
- **Fairness explanation**: What design and implementation processes ensure the AI's decisions "are generally unbiased and fair" and that 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 system'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."[93]

Each type of explanation will not be achievable for every AI system, but any system should have at least one available.

Mitigating Bias

Bias should be another focus for risk mitigation.[94] 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. Data will tend to reflect society's biases. Unfortunately, in detecting and combating pernicious bias, "there is no simple metric to measure fairness that a software engineer can apply. … Fairness is a human, not a mathematical, determination, grounded in shared ethical beliefs."[95] Broad diversity of perspectives (both in lived experience and professional training) on the teams developing an algorithm and overseeing a model's training can eliminate blind spots in programming and curating the training data. A Brookings Institution paper suggests use of a "bias impact statement" to mitigate harmful bias.[96] And FTC Commissioner Slaughter has warned, "As an enforcer, I will see self-testing [for unlawful credit discrimination] as a strong sign of good-faith efforts at legal compliance, and I will see a lack of self-testing as indifference to alarming credit disparities."[97]

Moreover, even if the data are representative and an otherwise good fit for training an AI model at one point in time, society continues to evolve. This evolution can degrade the 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.[98]

Other Considerations: Trade-Offs and Outsourcing

Risk assessment and mitigation will involve trade-offs. Accuracy and explainability may clash (as in the black-box radiology/cancer prediction example). Accuracy and fairness may be in tension (hypothetically, for instance, predictions that consider a person's race or gender might be more accurate, at least in some sense, but might also be unfair). Different aspects of fairness may conflict. When no acceptable trade-off can be found, should your company still release or employ the application?[99]

Finally, if your company outsources the design or development of an AI system, that may not relieve it from liability for assessing and mitigating risk.[100] In procurement contracts, be sure to "include contractual terms specifying what categories of information will be accessible to what categories of individuals [both inside your company and your customers] who may seek information about the design, operation, and results of the A/IS."[101] 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 senior management to pay close attention to potentially risky AI applications. 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 technology leads and other internal or external 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.[102]

Yet, the novelty and complexity of AI may mean that your company's existing compliance structures are not well-suited to oversee the development, procurement and use of AI. If oversight responsibilities are spread across different functions or teams, nobody may have a broad enough picture to ensure regulatory compliance. Problems may be missed if they fall into the gaps between teams or if the responsibility is shared but imperfectly coordinated. Diffusion of responsibility may also mean diffusion of expertise in a new, complicated and rapidly evolving area. You need to ask whether the current compliance structure can assure that senior management has the information needed for the accountability that regulators expect.

You (and your board) also need to consider whether the board has duties to monitor AI regulatory compliance[103] and, if so, whether your board has sufficient expertise to navigate any major—and certainly any mission-critical—risks presented by AI.

Explainability for Oversight

Having the broadest possible set of explanations for each system will facilitate oversight. Even more than with risk assessment and mitigation, the explainability of AI models is important to addressing the oversight challenges they pose. 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.[104]

When asking for explanations, be sure to inquire into their bases. The system's operators may not have trained or developed the model. Furthermore, "few, if any, AI systems are built from the ground up, using components and code that are wholly the creation of the AI developers themselves."[105] 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 inside your company or by vendors. In short, you may need to keep "peeling the onion" to arrive at well-substantiated explanations for internal or public consumption.

Documentation

Regulation of AI also should spur reassessment of your company's document-retention policies. For one thing, regulators will demand documentation about the development and use of certain 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 been particularly detailed in describing the GDPR's documentation requirements for users of AI involving personal data. The ICO advises such users to record their risk assessment and mitigation decisions (especially regarding trade-offs among risks), lines of accountability for decisions about trade-offs, and outcomes of these decisions "to an auditable standard."[106] The ICO goes on to recommend that these records include:

- Reviews of the risks to the individuals whose personal data is processed.
- How trade-offs 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."[107]

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."[108] The ICO also urges AI users to analyze this information for potential problems.[109]

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 would mandate provider retention for ten years of extensive technical documentation of the system's development process, design, training, testing, validation, and risk management system as well as the provider's compliance program.[110] The Commission also proposed that providers and users retain logs of such systems' operations to ensure traceability.[111] Beyond written policies for the operation of autonomous and intelligent systems,[112] 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.[113]

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 will 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, a lawsuit or both. The harm from the error will be obvious. Whether the *res ipsa loquitur* doctrine[114] technically applies or not, the novelty, complexity and opacity of AI systems raise the risk the factfinder will infer your company's liability from the obvious harm itself. 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 legal landscape is changing rapidly and requires monitoring; 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.

Endnotes

1. Darrel Pae, Katerina Kostaridi and Elliot S. Rosenwald provided research assistance.
2. Relatively unfamiliar to a Western audience, Go is a territorial game that has been played for thousands of years. Although the "rules are quite simple," the "strategic and tactical possibilities of the game are endless," which makes Go an "extraordinary" "intellectual challenge." The International Go Federation, *About Go* (July 3, 2010), https://www.intergofed.org/about-go/about-go.html.
3. Sundar Pichai, *Why Google Thinks We Need to Regulate AI*, Fin. Times (Jan. 19, 2020), https://www.ft.com/content/3467659a-386d-11ea-ac3c-f68c10993b04.
4. *See, e.g.*, Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (2014); Max Tegmark, Life 3.0: Being Human in the Age of Artificial Intelligence (2017); Future of Life Inst., Benefits & Risks of Artificial Intelligence, https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/ (last visited Mar. 3, 2021).
5. The Committee on Standards in Public Life, *Artificial Intelligence and Public Standards* § 1.1, at 12 (2020), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF (Artificial Intelligence and Public Standards).
6. CompTIA, *Emerging Business Opportunities in AI* 2 (May 2019), https://www.comptia.org/content/research/emerging-business-opportunities-in-ai (Emerging Business Opportunities in AI).
7. *See, e.g.*, *Holbrook v. Prodomax Automation Ltd.*, No. 1:17-cv-219, 2020 WL 6498908 (W.D. Mich. Nov. 5, 2020) (denying summary judgment in suit for wrongful death of manufacturing plant worker killed by robot); *Batchelar v. Interactive Brokers, LLC*, 422 F. Supp. 3d 502 (D. Conn. 2019) (holding that broker-dealer owed a duty of care in design and use of algorithm for automatically liquidating customer's positions upon determination of margin deficiency in brokerage account); *Nilsson v. Gen. Motors LLC*, No. 4:18-cv-00471-JSW (N.D. Cal. dismissed June 26, 2018) (settled claim that car in self-driving mode negligently changed lanes, striking motorcyclist).
8. Maya Oppenheim, *Amazon Scraps "Sexist AI" Recruitment Tool*, Independent (Oct. 11, 2018), https://www.independent.co.uk/life-style/gadgets-and-tech/amazon-ai-sexist-recruitment-tool-algorithm-a8579161.html.
9. Nat'l Inst. of Standards & Tech., Press Release, *NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software* (Dec. 19, 2019), https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software; Patrick Grother et al., Nat'l Inst. of Standards & Tech., Interagency or Internal Report 8280, Face Recognition Vendor Test (FRVT): Part 3: Demographic Effects (2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf.
10. *See* Kashmir Hill, *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*, NY Times (updated Jan. 6, 2021), https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html (reporting that the three people known to have been falsely arrested due to incorrect facial recognition matches are all Black men).
11. Ajay Agrawal et al., *Prediction Machines: The Simple Economics of Artificial Intelligence* 196 (2018) (citing Anja Lambrecht and Catherine Tucker, *Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads* (paper presented at NBER Summer Institute, July 2017)).
12. Joi Ito, *Opening Event Part 1*, https://www.media.mit.edu/courses/the-ethics-and-governance-of-artificial-intelligence/ (0:09:13-0:09:52).
13. *See* Jean-François Bonnefon et al., *The Social Dilemma of Autonomous Vehicles*, 352 Science 1573 (2016); Joi Ito & Jonathan Zittrain, *Class 1: Autonomy, System Design, Agency, and Liability*, https://www.media.mit.edu/courses/the-ethics-and-governance-of-artificial-intelligence/ (~0:12:15).
14. 42 USC § 2000e-2(k)(1)(A)(i).
15. 15 USC §§ 1691-1691f.
16. Andrew Smith, Bureau of Consumer Prot., *FTC, Using Artificial Intelligence and Algorithms*, Business Blog (Apr. 8, 2020, 9:58 AM), https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms (Smith Business Blog Post). The FTC also might be able to use the unfairness prong of Section 5 of the FTC Act, 15 USC § 45(n), to attack algorithmic discrimination against protected classes. Rebecca Kelly Slaughter, Comm'r, FTC, Remarks at the UCLA School of Law: Algorithms and Economic Justice at 13-14 (Jan. 24, 2020) (Slaughter Speech).
17. *See, e.g.*, NY Dep't of Fin. Servs., Ins. Circular Letter No. 1 (Jan. 18, 2019), https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2019_01 ("[I]nsurers' use of external data sources in underwriting has the strong potential to mask" prohibited bias.).

18. *See, e.g.*, UK Info. Comm'r's Office, Guidance on AI and Data Protection 40, 43-44, 46 (Jul. 30, 2020), https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection-0-0.pdf (ICO Guidance on AI and Data Protection) (discussing the application of the UK Equality Act 2010, c. 15, to AI systems).
19. 15 USC § 1681b.
20. *See* Smith Business Blog Post; Org. for Econ. Cooperation & Dev., *Artificial Intelligence in Society* 55 (2019), https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm.
21. 15 USC §§ 80b-1–80b-18c.
22. Wealthfront Advisers, LLC, Investment Advisers Act Release No. 5086, 2018 WL 6722756 (Dec. 21, 2018); Hedgeable Inc., Investment Advisers Act Release No. 5087, 2018 WL 6722757 (Dec. 21, 2018).
23. 15 USC §§ 80a-1–80a-64.
24. 17 CFR § 270.3a-4.
25. Div. of Inv. Mgmt., SEC, *Robo-Advisers*, IM Guidance Update No. 2017-02, at 2, https://www.sec.gov/investment/im-guidance-2017-02.pdf.
26. 21 CFR pts. 210-212.
27. 15 USC § 45; *see, e.g.*, *Everalbum Inc.*, File No. 1923172, 2021 WL 118892 (FTC Jan. 11, 2021) (complaint and proposed consent order requiring company to delete or destroy facial recognition algorithm due to alleged misrepresentations about use of facial recognition on, and retention of, storage service's users' photos and videos).
28. Slaughter Speech at 13.
29. *See, e.g.*, *Facebook Inc.*, Docket No. C-4365, 2020 WL 2197924 (FTC Apr. 27, 2020).
30. *See* Elisa Jillson, Bureau of Consumer Prot., FTC, *Aiming for Truth, Fairness, and Equity in Your Company's Use of AI*, Business Blog (Apr. 19, 2021, 9:43 AM), https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai; Slaughter Speech; Smith Business Blog Post.
31. *See* Regulation (EU) 2016/679, 2016 O.J. (L 119) 1 (GDPR). With Brexit, the United Kingdom generally adopted the GDPR as domestic law. *See* Data Protection, Privacy and Electronic Communications (Amendments etc) (EU Exit) Regulations 2019, SI 2019/419 (UK).
32. GDPR, art. 6, 2016 O.J. (L 119) at 36-37.
33. *Id.*, art. 4(2), 2016 O.J. (L 119) at 33.
34. Article 29 Data Prot. Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679*, at 16, WP251rev.01 (Feb. 6, 2018) (WP251 Guidelines); *see* GDPR, arts. 13-14, 2016 O.J. (L119) at 40-42; *see also id.*, art. 12, 2016 O.J. (L 119) at 39-40. Upon succeeding the Article 29 Data Protection Working Party, the European Data Protection Board endorsed the WP251 Guidelines. European Data Protection Board, Endorsement 1/2018 ¶ 3 (May 25, 2018).
35. GDPR, art. 22, 2016 O.J. (L 119) at 46.
36. *See id.*, arts. 16-17, 2016 O.J. (L 119) at 43-44; ICO Guidance on AI and Data Protection at 69-70. The ICO advises readers that "we discuss what you must do to comply with data protection law as well as what you should do as good practice. This distinction is generally marked using 'must' when it relates to compliance with data protection law and using 'should' where we consider it good practice but not essential to comply with the law." *Id.* at 11. Please keep this advice in mind in connection with subsequent references to this guidance.
37. *See* Cal. Civil Code §§ 1798.100(c), 1798.105, 1798.106, 1798.130, 1798.140(y), 1798.185(a)(16).
38. *See* 2021 Va. Acts ch. 36 (adding Va. Code Ann. §§ 59.1-571, 59.1-573(A)(2)-(3), (5), 59.1-574(A), (C), 59.1-576(A)(3)).
39. Cal. Bus. & Prof. Code §§ 17940–17943.
40. *See* 740 Ill. Comp. Stat. 14/1; *In re Facebook Biometric Info. Privacy Litig.*, No. 15-cv-03747-JD, 2021 WL 757025, at *6 (N.D. Cal. Feb. 26, 2021) (explaining there are "genuine disputes of fact about … whether Facebook's facial recognition technology" falls within the statute's scope).
41. 820 Ill. Comp. Stat. 42/5.
42. *See, e.g.*, High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI* 5-6 (2019), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 (HLEGAI Ethics Guidelines).
43. Org. for Econ. Cooperation & Dev., *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (May 22, 2019), https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (OECD Recommendation); *G20 Ministerial Statement on Trade and Digital Economy* 11-14 (Jun. 9, 2019), https://www.mofa.go.jp/files/000486596.pdf.
44. *See, e.g.*, Mark MacCarthy, *AI Needs More Regulation, Not Less* (Mar. 9, 2020), https://www.brookings.edu/research/ai-needs-more-regulation-not-less/; Franklin Foer, *It's Time To Regulate the Internet,* The Atlantic (Mar. 21, 2018), https://www.theatlantic.com/technology/archive/2018/03/its-time-to-regulate-the-internet/556097/.
45. Michael Kratsios, *AI That Reflects American Values*, Bloomberg (Jan. 7, 2020), https://www.bloomberg.com/opinion/articles/2020-01-07/ai-that-reflects-american-values.
46. Office of Mgmt. & Budget, Exec. Office of the President, No. M-21-06, Guidance for Regulation of Artificial Intelligence Applications (Nov. 17, 2020), https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf (OMB Guidance).
47. *Id.* at 3.
48. *Id.* at 4.
49. *Id.* at 5.
50. *Id.* at 5, 7-8.
51. *Id.* at 3-6.

52. US Dep't of Transp., *Automated Vehicles 3.0: Preparing for the Future of Transportation* iv, 7 (2018), https://www.transportation.gov/sites/dot.gov/files/docs/policy-initiatives/automated-vehicles/320711/preparing-future-transportation-automated-vehicle-30.pdf (Automated Vehicles 3.0); *see also* US Nat'l Sci. & Tech. Council & US Dep't of Transp., *Ensuring American Leadership in Automated Vehicle Technologies: Automated Vehicles 4.0,* at 5, 29 (2020), https://www.transportation.gov/sites/dot.gov/files/docs/policy-initiatives/automated-vehicles/360956/ensuringamericanleadershipav4.pdf.
53. Automated Vehicles 3.0 at iv; *see also id*. at 7.
54. S. 1108, 116th Cong. (2019); H.R. 2231, 116th Cong. (2019) (collectively, Algorithmic Accountability Act). This legislation has not yet been reintroduced in the current Congress.
55. Press Release, US Sen. Cory Booker, Booker, Wyden, Clarke Introduce Bill Requiring Companies To Target Bias in Corporate Algorithms (Apr. 10, 2019), https://www.booker.senate.gov/news/press/booker-wyden-clarke-introduce-bill-requiring-companies-to-target-bias-in-corporate-algorithms. Except for data brokers, the Algorithmic Accountability Act would reach only companies (or individuals) with over $50 million in annual gross receipts or with personal information from over one million consumers or consumer devices. Algorithmic Accountability Act § 2(5).
56. S. 2968, 116th Cong. § 108(b) (2019). This legislation also has not yet been reintroduced in the current Congress.
57. *Commission Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final (Apr. 21, 2021) (EC Proposal).
58. *Commission Proposal for a Regulation of the European Parliament and of the Council on Machinery Products*, COM (2021) 202 final (Apr. 21, 2021).
59. EC Proposal arts. 5–6, 9–22, 24–29, 43–44, 50–52, 61–62, 71, annexes II–IV, recital 48.
60. *Id*. arts. 2(2), 75-82, recital 29. The amended acts are Regulation 300/2008, 2008 O.J. (L 97) 72 (EC); Regulation 167/2013, 2013 O.J. (L 60) 1 (EU); Regulation 168/2013, 2013 O.J. (L 60) 52; Directive 2014/90, 2014 O.J. (L 257) 146 (EU); Directive 2016/797, 2016 O.J. (L 138) 44; Regulation 2018/858, 2018 O.J. (L 151) 1 (EU); Regulation 2018/1139, 2018 O.J. (L 212) 1 (EU); and Regulation 2019/2144, 2019 O.J. (L 325) 1 (EU).
61. European Commission, *Artificial Intelligence—Ethical and Legal Requirements*, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Requirements-for-Artificial-Intelligence (last visited Apr. 26, 2021).
62. Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies, P9_TA(2020)0275 (Ethics Regulation Resolution).
63. *See* EC Proposal art. 5.
64. *Compare id*. art. 6, annexes II-III *with* Ethics Regulation Resolution annex B, arts. 4(e), 14(1), annex.
65. Ethics Regulation Resolution annex B, arts. 10–11.
66. *Compare EC Proposal* art. 43(1)(a), (2)-(3) *with* Ethics Regulation Resolution annex B., art. 15.
67. Ethics Regulation Resolution annex B, arts. 13, 22.
68. Civil Liability Regime for Artificial Intelligence, P9_TA(2020)0276, annex B.
69. *Commission Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics*, COM (2020) 64 final (Feb. 19, 2020).
70. *E.g.*, EC Proposal arts. 9(7), 19(2), 43(3), 61(4), 63(3)-(5), (7), 64(3); Ethics Regulation Resolution annex B, recitals 5, 15.
71. EC Proposal art. 59.
72. Ethics Regulation Resolution annex B, arts. 18, 20.
73. *See* EC Proposal art. 2(1)(a); Ethics Regulation Resolution annex B, arts. 2-3.
74. EC Proposal art. 2(1)(c).
75. *See* UK Info. Comm'r's Office, *How We Handle Concerns*, https://ico.org.uk/about-the-ico/what-we-do/how-we-handle-concerns/ (last visited Mar. 26, 2021).
76. UK Dep't for Digital, Culture, Media & Sport, *Centre for Data Ethics and Innovation: Government Response to Consultation* 5 (2018), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data//file/757509/Centre_for_Data_Ethics_and_Innovation_-_Government_Response_to_Consultation.pdf.
77. *Id*. at 5-6, 12, 14.
78. UK Ctr. for Data Ethics & Innovation, *Review into Bias in Algorithmic Decision-Making* 11 (2020), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957259/Review_into_bias_in_algorithmic_decision-making.pdf (CDEI Review); *see id*. at 109, 113.
79. *See id.* at 11, 40, 109, 116.
80. *See, e.g.*, Office of the Privacy Comm'r of Canada, *A Regulatory Framework for AI: Recommendations for PIPEDA Reform* (Nov. 2020), https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/reg-fw_202011/.
81. Emerging Business Opportunities in AI at 3.
82. GDPR, art. 17, 2016 O.J. (L 119) at 43-44.
83. *See* ICO Guidance on AI and Data Protection at 69-70.
84. GDPR, art. 35(3)(a), 2016 O.J. (L 119) at 53.
85. Nat'l Inst. of Standards & Tech., *US Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools* 8 (Aug. 9, 2019), https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf.

86. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems* 4 (1st ed. 2019), https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html (Ethically Aligned Design).

87. HLEGAI Ethics Guidelines at 31.

88. High-Level Expert Group on Artificial Intelligence, *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-Assessment* (July 2020), https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment.

89. *See, e.g.*, IEEE Standards Association, *The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS)*, https://standards.ieee.org/industry-connections/ecpais.html (last visited Feb. 10, 2021); The Use of Artificial Intelligence in Health Care: Trustworthiness, ANSI/CTA-2090 (Consumer Tech. Ass'n & Am. Nat'l Standards Inst. 2021), https://shop.cta.tech/collections/standards.

90. *See* ICO Guidance on AI and Data Protection at 72.

91. *See, e.g.,* GDPR, arts. 13-14, 2016 O.J. (L119) at 40-42; WP251 Guidelines at 16 (explaining GDPR arts. 13-14); OECD Recommendation § 1.3; *see also* Ethics Regulation Resolution recital 18 ("[C]onsumers should have the right to be adequately informed in an understandable, timely, standardised, accurate[,] and accessible manner about the existence, reasoning, possible outcome[,] and impacts for consumers of algorithmic systems, about how to reach a human with decision-making powers, and about how the system's decisions can be checked, meaningfully contested[,] and corrected."); Ethically Aligned Design at 33 ("Creators of A/IS should provide the parties affected by the output of A/IS with information on the role of the operator, the competencies required, and the implications of operator error.").

92. Nevertheless, a black-box model should be closer to a last resort than a first. UK Info. Comm'r's Office & The Alan Turing Inst., *Explaining Decisions Made with AI* 66 (May 20, 2020), https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence-1-0.pdf (Explaining Decisions Made with AI) (noting that more interpretable models may demand fewer resources because they "do not require supplemental tools and techniques for facilitating interpretable outcomes"). Black-box models also may require additional efforts at documentation, *see id.* at 67, as discussed below.

93. *Id*. at 20; *see also US Food & Drug Admin., Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan* 5 (2021), https://www.fda.gov/media/145022/download (noting "the unique challenges of labeling for AI/ML-based devices and the need for manufacturers to clearly describe the data that were used to train the algorithm, the relevance of its inputs, the logic it employs (when possible), the role intended to be served by its output, and the evidence of the device's performance").

94. *See, e.g.*, Ethics Regulation Resolution annex B, art. 17.

95. Nicol Turner Lee et al., *Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms* (May 22, 2019), https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/ (Brookings Paper); CDEI Review at 21.

96. Brookings Paper.

97. Slaughter Speech at 15. While Chairwoman Slaughter made this statement regarding ECOA enforcement, it likely reflects her attitude towards other types of algorithmic discrimination as well.

98. ICO Guidance on AI and Data Protection at 39.

99. *See* HLEGAI Ethics Guidelines at 20.

100. *See* Explaining Decisions Made with AI at 93, 98.

101. Ethically Aligned Design at 252.

102. ICO Guidance on AI and Data Protection at 47; *see also* HLEGAI Ethics Guidelines at 23; Ethically Aligned Design at 240.

103. *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); *cf.* CDEI Review at 10.

104. *See* Explaining Decisions Made with AI at 16.

105. Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 Harv. J. L. & Tech. 353, 371 (2016).

106. ICO Guidance on AI and Data Protection at 25.

107. *Id.*; *see also* Explaining Decisions Made with AI at 102.

108. ICO Guidance on AI and Data Protection at 73. *But see* GDPR art. 5(1), 2016 O.J. (L 119) at 35-36 ("Personal data shall be … (c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation'); … (e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed. …"); ICO Guidance on AI and Data Protection at 64 ("Where a model is established and unlikely to be re-trained or modified, the training data may no longer be needed. If the model is designed to use only the last 12 months' worth of data, a data retention policy should specify that data older than 12 months be deleted.").

109. ICO Guidance on AI and Data Protection at 73.

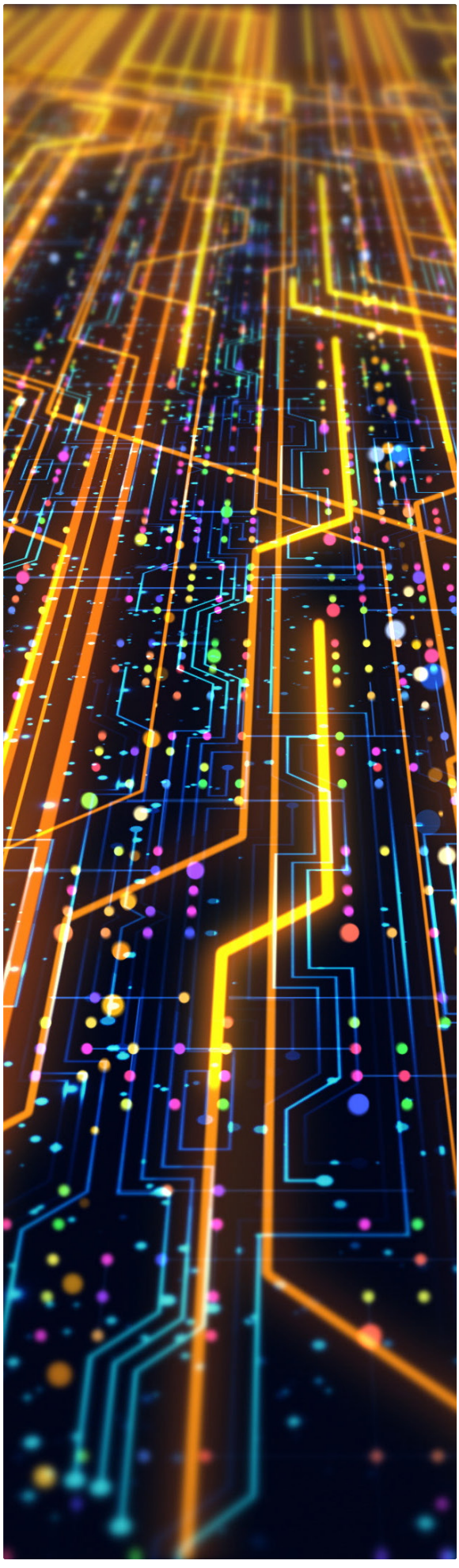
110. EC Proposal arts. 11, 17, 50, annex IV.

111. *Id*. arts. 12, 20, 29(5); *see also* Ethics Regulation Resolution annex B, art. 8(2).

112. Ethically Aligned Design at 33.

113. *Id*. at 138.

114. "The thing speaks for itself." *Res ipsa loquitur* permits inference "that the defendant has been negligent when the accident causing the plaintiff's harm is a type of accident that ordinarily happens as a result of … negligence." Restatement (Third) of Torts: Liability for Physical and Emotional Harm § 17 (Am. Law Inst. 2010 & Oct. 2020 update); *see Byrne v. Boadle* (1863) 159 Eng. Rep. 299, 300; 2 H. & C. 722, 725.



About the Author

Peter J. Schildkraut

Partner
Washington, DC
+1 202.942.5634

peter.schildkraut@arnoldporter.com

Peter Schildkraut is a co-leader of the firm's Technology, Media & Telecommunications industry team and provides strategic counsel on spectrum use, broadband and other TMT regulatory matters. Mr. Schildkraut helps clients navigate the ever-changing opportunities and challenges of technology, policy and law to achieve their business objectives at the US Federal Communications Commission (FCC) and elsewhere. He is the author of "AI Regulation: What You Need to Know to Stay Ahead of the Curve."

Mr. Schildkraut represents clients in rulemakings and administrative litigation and advises them on regulatory compliance. He also assists them in all stages of transactions, large and small, including every major US communications transaction of AT&T Inc. (formerly SBC) since 1998.

Related Articles of Interest by Peter J. Schildkraut

- How Should Technology Platforms Be Categorized?
- What New Legal Issues Are Media Companies Facing?

About Arnold & Porter's Technology, Media & Telecommunications Industry Group

Through cycles of regulation and deregulation, revolution, and maturation, Arnold & Porter's integrated transactional, litigation and regulatory insights help some of the world's most innovative and prominent technology, media and telecommunications businesses determine the character of their content, secure their infrastructures and power their platforms around the globe. We have guided our TMT clients through some of the largest mergers in history, advocated for them in their most important intellectual property and First Amendment cases, advised them in novel spectrum transactions and network-sharing deals, counseled them on how to obtain favorable regulatory treatment for their cutting-edge technologies, negotiated technology collaborations, and in countless other ways have helped them conceive, build, finance, commercialize, and protect their most important assets.

Read/hear more about our business and technical savvy, groundbreaking precedents and regulatory expertise.

Subscribe to our *TMT Time* podcast.

About Arnold & Porter

With nearly 1,000 lawyers practicing in 13 offices around the globe, Arnold & Porter serves clients across 40 distinct practice areas. The firm offers 100 years of renowned regulatory expertise, sophisticated litigation and transactional practices, and leading multidisciplinary offerings in the life sciences and financial services industries.