Why Ethical AI Initiatives Need Help from Corporate Compliance

April 4, 2022

Artificial intelligence (AI) is becoming part of the core business operations at many companies. This widespread adoption of AI has led to a proliferation of corporate "ethical AI" principles and programs, as companies seek to ensure that they are using AI fairly and responsibly, and in a manner consistent with the growing expectations of customers, employees, investors, regulators, and the public.

But ethical AI programs at many companies are struggling. Recent reports of AI ethics leaders being fired, resigning, or bringing whistleblower claims illustrate the friction that is common between ethical AI teams and executives who are trying to gain efficiencies and competitive advantages through the adoption of AI.

These struggles are not surprising to white collar lawyers who regularly work with companies on regulatory compliance, governance, and sensitive investigations. We have found that it is often difficult to achieve meaningful changes in corporate behavior solely through the adoption of ethical principles or codes of conduct. Rather, in our experience, business practices are more likely to change when companies implement concrete compliance and governance policies that are closely aligned with existing and anticipated regulatory obligations.

This dynamic can be seen across a variety of areas of corporate responsibility, including economic sanctions, money laundering, foreign bribery, campaign finance, insider trading, enterprise risk management, and workplace sexual harassment. In each case, internal codes of conduct and ethics initiatives have played an important role in shaping behavior. But the most significant changes in corporate culture are largely attributable to the compliance functions of companies developing and implementing clear policies and controls, and corporate governance practices that are based on existing and emerging legal standards and regulatory expectations. The same likely will be true for AI.

Why Ethical AI Programs Are Struggling

As AI systems have become increasingly critical to companies' strategies for growth and competitiveness, concerns over AI ethical issues have taken center stage. At universities, industry conferences, corporate boardrooms, and meetings of regulators in D.C., London, Singapore, and Brussels, there is a growing recognition that, for all its promise, AI can present serious risks to society. The concerns include invasion of privacy, increased surveillance, manipulation of human behavior, exacerbation of income inequality, and perpetuation of discrimination, as well as systemic risks to the financial markets.

To address these concerns, companies have assembled ethical AI teams, many of which have had significant positive impacts. But there are at least three reasons why ethical AI principles are often insufficient, standing alone, to meaningfully change corporate conduct:

- **Vagueness.** Ethical principles are frequently too vague to be effective in determining whether a specific type of conduct or system is or is not permissible. For example, a commitment to not perpetuate bias is an unassailable principle. As a practical matter, however, it can be very difficult to determine whether a particular AI system is producing a biased result that is unethical. For example, suppose an AI underwriting tool considers smoking habits when pricing life insurance. And suppose that women, on average, smoke more than men, and therefore women, on average, must pay higher premiums for life insurance than men. Even though it is almost certainly not unlawful, the operation of that AI tool arguably produces a biased result. But is it unethical, and on what basis does one reach that conclusion?

- **Consistency.** Relatedly, people inevitably will have different views about the proper application of ethical principles, especially since ethical principles are voluntary and not uniform across an industry, or even across different corporate functions. Business executives understandably are concerned that adhering to certain principles, as applied by their ethical AI group, will put their company at a competitive disadvantage vis-à-vis its competitors. Accordingly, it may be unrealistic to expect businesses to substantially curtail their development and deployment of a promising new AI technology, unless the rationale for restrictions is clear and consistent, and thereby widely applicable across the industry.

- **Expertise.** Companies often lack the expertise necessary to balance ethical AI principles against competing (and in some cases, mutually exclusive) corporate objectives. For example, certain techniques to reduce the risks of AI biases involve gathering and testing large volumes of personal information, which can carry

significant cybersecurity and privacy risks. Furthermore, suppose that the adoption of a particular AI program will meet a stated corporate objective of creating efficiencies and generating significant profits, but it will also result in dozens of lost jobs and have a significant carbon footprint. Without clear guidance, executives understandably may feel unable to balance these competing goals in a coherent and consistent manner. They may also question why AI programs are subject to these kinds of restraints if other business initiatives are not, especially if they see their competitors making different choices.

The Recent Emergence of a Regulatory Compliance Approach to AI

Until recently, there was insufficient AI-focused regulatory activity for companies to ground their ethical AI principles in existing or anticipated regulatory requirements. But over the last two years, there has been a flurry of AI regulatory pronouncements, draft and enacted legislation, agency guidance, enforcement actions, and court rulings, which together provide companies with a roadmap of AI rules and standards, both as they exist today and as they will likely exist in the near future.

For example, last year, the European Commission proposed a landmark draft "Artificial Intelligence Act" that would impose concrete requirements on a wide range of "high-risk" AI systems across several key sectors including credit, employment, education, workforce development, and insurance. The Artificial Intelligence Act is patterned on a product safety model, with significant requirements to ensure that any "high risk" AI is safe to place on the market. This draft law is currently proceeding through negotiations in the European Union, and may be finalized as early as 2023.

In the United States, several states and cities have passed regulations requiring companies to assess the risks of bias in certain AI models. These include a recent New York City law that bans employers from using automated employment decision tools unless they conduct and publish a yearly "bias audit" to assess the potential disparate impact of these tools. Similarly, a recent Colorado law prohibits insurers from using any external consumer data or information source, algorithm, or predictive model that unfairly discriminates on the basis of protected characteristics.

In addition, several regulators have provided guidance on the application of existing antidiscrimination laws to AI, including in the areas of fair lending, insurance, housing, hiring, and advertising. These and other recent regulatory initiatives provide strong indicators as to what is, or is likely to be, required of AI systems.

Debevoise
&Plimpton

AML Compliance as a Model

The history of anti-money laundering ("AML") regulations helps to illustrate why AI ethics programs should be more closely aligned with regulatory compliance. In the 1970s, organized crime and drug trafficking became top priorities for U.S. law enforcement. To hide their cash proceeds, the participants in these illegal activities would frequently deposit their funds at a bank and thereafter move the funds through multiple transactions. Even though the Bank Secrecy Act (the "BSA") was passed in 1970—and despite a clear ethical imperative not to allow the U.S. financial system to be used to conceal and further criminal activity—most banks did not implement a sophisticated process to detect and report these kinds of suspicious transactions throughout the 1970s and much of the 1980s.

This changed around the time Congress amended the BSA with the Money Laundering Control Act of 1986 (the "MLCA"). The MLCA clarified reporting and record-keeping obligations and required banks to establish and maintain procedures to ensure compliance with the BSA. Once the legal requirements and regulatory expectations were sufficiently clear, banks quickly adopted robust policies and procedures to ensure compliance. Once it was clear that the MLCA was going to be law, (i) the rules and regulatory expectations went from being vague to having much greater specificity, (ii) there ceased to be a concern that unilaterally restricting bank activities would result in a competitive disadvantage, since all banks faced the same obligations (at least in the United States), and (iii) banks could enlist their compliance personnel, who had the requisite expertise and internal credibility to implement the necessary policies and procedures and ensure the new rules were followed. Today, virtually all banks have highly sophisticated AML compliance frameworks in place.

Many companies are now making a similar shift with AI—moving from a reliance on ethical principles (which made sense when there was little regulatory guidance) to a regulatory compliance model with the understanding that binding AI regulations are coming.

More Effective AI Ethical Frameworks

One recurring challenge with ethical AI programs is that they advocate a broad array of principles, some of which overlap closely with existing and emerging regulatory requirements, but some of which do not. In order to successfully move to a regulatory compliance approach, it is important to tailor ethical AI programs so that they focus, at least initially, on the types of principles that undergird actual regulations, such as:

- nondiscrimination against protected classes;

- avoiding fraud, manipulation, and conflicts of interest;

- privacy and cybersecurity;

- transparency and disclosure;

- documentation and auditability; and

- human oversight and accountability.

Various other ethical principles are extremely compelling from a societal standpoint but are less likely to succeed in shaping corporate behavior as part of an AI-compliance program. These principles include:

- promotion of shared prosperity;

- maintaining human dignity;

- limiting workforce disruption; and

- ensuring broad benefit.

While many will advocate for the inclusion of these principles in the corporation's compliance program in an effort to be ethical, conscientious, and responsible, doing so can risk undermining the ethical AI program in the long term. This is because, as discussed above, these principles are often too vague to be applied consistently, are generally not mandatory under existing or anticipated regulations, and therefore put companies in the difficult position of trying to balance evolving ethics against industry competitiveness. Companies should therefore consider differentiating between ethical principles that closely align with existing or emerging legal requirements (which should be incorporated into corporate compliance and governance functions), and other principles that, however important, should be advanced through other mechanisms.

Conclusion

The AI regulatory landscape is rapidly expanding, and companies that are investing heavily in machine learning and algorithmic decision-making will need to be compliant. It is therefore important to integrate ethical AI programs and corporate compliance

programs. To be sure, despite recent developments, many AI legal standards remain somewhat vague. But lawyers and compliance professionals with experience in technology and emerging regulatory frameworks, working with ethical AI officers, can provide companies with helpful guidance as to which AI applications are likely to face significant legal scrutiny in the future, and therefore may not be worth implementing without appropriate safeguards. This is likely to be a more effective way of shaping corporate AI conduct than accusing executives of behaving unethically.

* * *

Please do not hesitate to contact us with any questions.

NEW YORK

Avi Gesser
agesser@debevoise.com

Bruce E. Yannett
beyannett@debevoise.com

Douglas S. Zolkind
dzolkind@debevoise.com

Anna R. Gressel
argressel@debevoise.com

Adele Stichel
agstichel@debevoise.com