The Value of Having AI Governance – Lessons from ChatGPT

April 5, 2023

Last month, we wrote about how many companies were implementing a pilot program for ChatGPT, as a follow up to our article about companies adopting a policy for the work-related uses of generative AI tools like ChatGPT, Bard and Claude (which we collectively refer to as "Generative AI"). We discussed how a pilot program often involves designating a small group of employees who test potential Generative AI use cases, and then make recommendations to a cross-functional AI governance committee that determines (1) which use cases are prohibited and which are permitted, and (2) for the permitted use cases, what restrictions, if any, should apply.

Lessons from ChatGPT Adoption

The process of running a Generative AI pilot program or adopting a broader Generative AI policy has resulted in companies learning several lessons about AI adoption in general, including:

The Value of AI Governance

Adoption of AI can be difficult without a proper governance structure. Even when low-risk/high-value use cases are identified, companies may still feel the need to ban Generative AI tools because they do not have the proper governance mechanisms to vet and authorize these appropriate use cases while also effectively prohibiting high-risk/low value use cases.

The Risk of Long-Term Bans

While an outright ban on generative AI tools may be a wise short-term way to reduce risk and evaluate options, for many companies, it is not a viable long-term solution. First, there is likely significant value in allowing at least some employees to experiment with Generative AI tools to find low-risk/high-value use cases that can improve productivity and deliver superior products or services. Second, because these tools are available to individual employees on their personal devices, a long-term total ban on their use risks employees using them off-platform, where the company has no ability to

monitor their use or place guardrails on the information that is input to the tools or generated from them.

The Limits of Ad Hoc Policies

Some companies have used their experience with ChatGPT as an opportunity to build out a more comprehensive AI governance program, to not only address the risks and benefits of Generative AI tools, but also to govern the adoption of AI more generally. As new AI tools become available, companies are recognizing that they will benefit from having an orderly process for quickly evaluating their risks and benefits and for adopting their use in a manner that is fair, understandable, actionable, and does not require the creation of an ad hoc or burdensome compliance structure.

A Chance to Shape Regulation

Regulators have repeatedly expressed the need for companies that are adopting AI to have proper governance and risk management. But they have not yet provided concrete guidance as to what that requires. Some regulators are likely to use their exam process or industry surveys to inform their view as to the specific AI governance requirements. Therefore, implementing an effective AI governance program is an opportunity to shape the regulatory landscape in a way that works for that particular company.

Other Benefits of Starting Early

Companies that already had AI governance programs in place were generally able to better manage their adoption of ChatGPT because they already had a committee up and running that had the mandate and the process in place to evaluate and adopt Generative AI uses cases. These companies have seen many benefits from their decision to start early with AI governance, including:

- Creating new and complicated governance structures like those required for AI deployment is resource intensive and takes time to properly set up. The process often involves some false starts and trial and error before identifying the optimal structure and processes for the company.

- Effective AI governance often requires input from several different parts of an organization, including risk, legal, compliance, IT, human resources, and various business functions. Finding the right people with sufficient bandwidth who can collaborate effectively on these issues can also take time, training, and it may require a formal internal or external search.

- Regulators and plaintiff lawyers are looking for opportunities to bring claims against companies who they believe are adopting AI in a reckless manner. Having a robust

AI governance program will reduce the risk of adopting AI that causes unexpected harm and allows the company to respond to claims that its AI tool was adopted without proper consideration and oversight.

- Similarly, adoption of AI often comes with some reputational risk. Having an AI governance program signals to customers and employees that the organization is thoughtfully considering how to responsibly deploy and monitor AI tools to ensure that their use is fair, accurate, and transparent.

Core Elements of an AI Governance Framework

For companies that are interested in adopting a broader AI governance program, here are some of the elements that should be considered:

Scope

Determine which kinds of models, algorithms, big data systems, and AI applications will be covered by the company's AI governance program, what tools are not covered, and an explanation of why. It is best to include concrete examples to assist the categorization of any new AI tools adopted.

Inventory

For each AI application that is governed by the program, document details about the application, which may include: its purpose, the problem it is intended to solve, the inputs and outputs, the training set, the anticipated benefits to the company and its customers, potential risks, who may be harmed, whether the model involves automated decision-making, any necessary safeguards, who is responsible for the application, and its risk rating.

Guiding Principles

Create a high-level set of guiding principles for design, development, and use of AI, which may include commitments to accountability, fairness, privacy, reliability, and transparency.

Code of Conduct

Draft an employee-facing code of conduct to operationalize the Guiding Principles.

Cross-Functional Governance Committee

Establish a cross-functional committee that oversees the program or other means for establishing overall accountability, including vetting new high-risk uses and identifying mitigations that will allow for their continued use; overseeing policies, procedures, and guidelines for responsible AI use; reporting to senior management or the board; and managing incidents and business continuity risks related to AI applications.

Risk Factors and Assessments

Create a list of risk factors to classify AI applications as low or high risk and determine how AI applications will be assessed for risk. This allows an organization to prioritize the highest-risk models for the cross-functional committee to review.

Risk Mitigation Measures

Establish a list of steps that the Governance Committee can recommend to reduce the risks associated with certain high-risk models, including bias assessments, stress testing, enhanced transparency, or additional human oversight, as appropriate.

Training

Provide training for individuals involved in developing, monitoring, overseeing, testing, or using high-risk AI applications on the associated legal and reputational risks.

Policy Updates

Update critical policies to address unique risks associated with AI applications, including with respect to privacy, data governance, model risk management, and cybersecurity.

Incident Response

Create a plan for responding to an allegation of bias or other deficiency in an AI application and conduct an AI incident tabletop exercise to test the plan.

Public Statements

In light of new guidance from the FTC on the risks of overselling AI applications, review the company's public statements relating to it use of AI to ensure their accuracy.

Vendor Risk Management

Review vendor policies to ensure that AI applications that are provided by third parties have been subjected to appropriate diligence and contractual provisions.

Senior Management and Board Oversight

Develop a plan for periodic reporting to senior management and the board on the program.

Documentation

Maintain documentation about the program to address concerns, respond to inquiries, and that meets regulatory expectations.

* * *

To watch an on-demand version of our webcast on how ChatGPT works, click here.

The Debevoise Artificial Intelligence Regulatory Tracker ("DART") is now available for clients to help them quickly assess and comply with their current and anticipated AI-related legal obligations, including municipal, state, federal and international requirements.

To subscribe to our Data Blog, click here.

Please do not hesitate to contact us with any questions.

NEW YORK

Avi Gesser
Partner
agesser@debevoise.com

Samuel J. Allaman
Associate
sjallaman@debevoise.com

Suchita Mandavilli Brundage
Associate
smbrunda@debevoise.com

Melissa Muse
Associate
mmuse@debevoise.com

Lex Gaillard
Law Clerk
adgailla@debevoise.com