onetrust

How to develop an AI governance program

OCTOBER 2023

About the author

Arsen Kourinian is a partner in Mayer Brown's Los Angeles office and practices data privacy and artificial intelligence law (AI). He holds the fellow of information privacy (FIP) and certified information privacy professional (CIPP) credentials through the International Association of Privacy Professionals (IAPP) for the United States (CIPP/US), Europe (CIPP/E), Canada (CIPP/C), and Asia (CIPP/A). He also holds the IAPP certified information privacy manager (CIPM) designation and Privacy Management Professional certificate through OneTrust. Arsen further helps educate the next generation of data privacy and AI professionals by authoring the CIPP/US and the IAPP's inaugural AI governance professional exams.

Arsen is a thought leader in data privacy, AdTech, and AI, having published many articles regarding nuanced issues in these fields, presenting on these topics for numerous organizations and trade association groups, and working with companies to shape the landscape in these ever-evolving areas. He was also named a member of the State Bar of California's Consulting Group on the Establishment of a Legal Specialization in Privacy Law and is a member of The Interactive Advertising Bureau's (IAB) Privacy Implementation & Accountability Task Force, which brings together leaders in the privacy community to address critical privacy issues facing the AdTech industry.

As a global counselor, Arsen advises multinational corporations regarding compliance with domestic and international data privacy and AI laws and frameworks. When advising clients, Arsen provides practical and operational guidance on how to harmonize these laws and standards into uniform policies, procedures, and practices. Arsen helps clients with organizational leadership policies, privacy and AI policies, data inventories, AdTech practices, cross-border data transfers, privacy impact assessments, mergers and acquisitions, AI governance, vendor contracts, privacy by design, and human resources data.

Arsen also advises companies' officers and directors regarding their obligations under data privacy, AI and security laws. This includes training boards and C-suite on implementing internal controls in their companies and governance plans.

Table of Contents

onetrust

In the modern era, data is the world's most valuable resource (with oil in second place). And big data is fueling artificial intelligence (AI) during mankind's fourth industrial revolution. While not new, the current iteration of AI—that is capable of machine and deep learning—is transforming the world in unprecedented ways and creating new opportunities and challenges for organizations across sectors and industries. Professionals can now ask AI to complete mundane tasks, such as inputting data and filling out paperwork, so that they can focus on matters that require human intelligence, such as providing care, education, mentorship, emotional comfort, and protecting others. Conversely, AI may also perform more complicated tasks, such as analyzing high volume of data, piecing together a complete image of the blackhole, and developing new cures for medical conditions.

For organizations to utilize the benefits of AI and maintain a competitive edge in their industry, they must ensure that their AI systems are trustworthy, which requires implementing an AI governance program. Unlike other areas of law, which may have well-established regulations and industry codes, AI is in its infancy stages in both law and technology. Until legislators pass comprehensive regulations governing AI, organizations should develop an AI governance program based on available frameworks, sector-specific laws, and even proposed regulations (e.g., the proposed EU AI Act). To help with this process, this white paper explains how to develop such a program based on global guidance materials, which require addressing the following factors:

AI Governance team: Before your organization develops and uses AI, establish a diverse AI governance team that is regularly trained on the legal, technical, and operational aspects of AI. This AI governance team will oversee the organization's AI systems and ensure that there is top-to-bottom and/or bottom-up oversight.

Data Governance: AI systems rely on data as their main input, output, and source of learning. To verify that AI systems function properly, your organization must implement data governance. This ensures that the input data is, among other things, properly collected, prepared, suitable, of good quality, updated, representative of the environment, and without bias. Through proper data governance, you can help ensure that the public trusts your AI output.

Legal compliance: Your AI governance team must identify and understand the laws applicable to your AI systems. The team should then assess whether there are gaps in compliance, and bridge these gaps to ensure that your organization's use and development of AI are lawful.

Risk management: A key responsibility of your AI governance team is to identify the risks present in the AI system, and categorize them as prohibited, high, medium, or low. Depending on the type and level of risk, your business may be precluded from developing and using AI, or have to implement further mitigation measures, such as greater level of human oversight. The AI governance team's assessment should be documented in a data and ethical impact assessment that identifies the risks and mitigation measures and concludes whether, on balance, the AI systems should be used.

Mitigation measures: Mitigating risk and maintaining trust with internal and external stakeholders requires your AI governance team to implement a number of measures in your AI systems. This means: (a) providing transparency notices regarding AI systems and explaining how AI works; (b) ensuring that AI systems are fair and without bias; (c) communicating the benefits of AI for individuals, society, and the environment; (d) confirming that AI is accurate, robust, safe, and secure; (e) incorporating privacy-by-design features; (f) adopting appropriate level of human oversight (human-in-the-loop, over-the-loop, or out-of-the-loop); (g) keeping technical documentation and logs; (h) monitoring your AI systems after they have been placed in the market and making adjustments as necessary; and (i) putting in place feedback and decision review channels.

Accountability: Your organization is ultimately accountable for your AI systems and could face regulatory scrutiny and public disrepute if they're not trustworthy. To demonstrate accountability, the AI governance team will need to maintain an auditable record addressing your AI governance program through policies and procedures.

1. Establishing and training an AI Governance team

AI governance requires a well-trained and diverse team that will oversee the organization's AI systems.

An AI governance team should be composed of individuals with different levels of seniority, skillsets, and background that has the sponsorship, support, and participation of the organization's officers and directors. The AI governance team should also have relevant expertise and representation from across the organization, including individuals responsible for data and security (e.g., Chief Data Officer, Chief Privacy Officer, Chief Information Security Officer, and/or Chief Technology Officer), research, data science, engineering, business, management, operations, procurement, human resources, and marketing. There's no one-size-fits-all for how your organization should establish its AI governance team because companies vary in their management approach (e.g., centralized versus decentralized) and the talent they employ. However, regardless of the type, size, and scale of your organization, you should strive to have a diverse AI governance team, instead of delegating AI oversight to one department (e.g., legal v. technical).

The AI governance team must also be properly trained and provided with the resources and guidance to perform their duties.

The AI governance team should receive training on, among other things:

a. The laws and regulations governing AI

b. How AI technology operates

c. AI use-cases

d. AI risks, and

e. Sector and geographic-specific issues applicable to the company

You should conduct an initial AI training for the team, followed up with additional training in regular cadence (e.g., monthly, quarterly, semi-annually, or annually). Moreover, the AI governance team should regularly monitor developments in AI technology and regulations, and provide alerts to the team and other employees in your organization if there are material developments.

Ultimately, once you establish the AI governance team, it should develop an AI leadership policy that defines the team's roles and responsibilities to demonstrate accountability.

2. Data Governance

When training, validating, and testing AI models, your teams need to implement appropriate data governance to ensure that the AI system functions properly.

AI systems rely on data for their input, output, and training. For data governance, your organization needs to adopt appropriate data collection practices. This includes the following:

a. Prepare the data (e.g., through annotation, labeling, cleaning, updating enrichment, and aggregation)

b. Formulate assumptions for the information the data is supposed to measure and represent

c. Assess the availability, quantity, and suitability of the datasets needed

d. Minimize or eliminate bias

e. Ensure that the datasets are representative of the environment

i. Free of errors

ii. Complete

iii. Good quality

iv. Respectful of the intellectual property rights of others

The AI governance team should also use different datasets for training, testing, and validation. The AI model needs to be trained using training data, the accuracy needs to be determined using test data, and validated using the validation dataset.

Next, the AI governance team needs to understand the data lineage by tracking where the data came from, how it was collected, curated, and moved within the company, and how the data's accuracy is maintained over time. This involves looking at the:

a. Data from its end-use and backdating it to its source (backward data lineage)

b. Data from its originating source and following it through to its end-use (forward data lineage)

c. Entire solution from both the data's source to its end-use and from its end-use to its source (end-to-end data lineage)

You must then maintain a data provenance record that documents the data quality (from origin to transformation), traces sources of error, updates the data, and attributes data to their sources.

Data governance is not a one-time project; the AI governance team should review the datasets periodically to ensure accuracy, quality, currency, relevance, and reliability, and update as necessary with new input data based on the advice of experts in this area, including data scientists.

3. Legal compliance

The AI governance team needs to identify which laws and regulations apply to the organization's AI systems, assess whether there are gaps in their compliance, and bridge those gaps. Your organization can't lawfully develop and use AI by only following an AI principles-based framework. Why? Because AI is regulated by AI-specific laws and laws of general applicability, which trigger different areas of law, such as data privacy, intellectual property, antitrust, and employment. The AI governance team needs to be aware of all such applicable laws and implement compliance to ensure that the AI systems are lawful, particularly considering that enforcement authorities in major jurisdictions are motivated to regulate AI under their general powers.

4. Risk management

The AI governance team needs to implement a risk management system that identifies, mitigates, and manages risk to ensure the AI systems are trustworthy. A risk management system for AI systems is a process that aims to identify, estimate, evaluate, and eliminate the potential harms that such systems can cause to people, their rights, democracy, the rule of law, or the environment, both in normal and misuse scenarios.

As a first step for risk management, the team needs to identify and rank the risks present in the AI systems, which can harm people, organizations, and the environment. Legal regimes may either outright prohibit certain AI uses (e.g., unfair and deceptive AI practices, use of subliminal techniques, exploiting vulnerable groups, certain biometric systems, social scoring, and criminal justice/law enforcement), or require a risk ranking, such as high, medium, or low. As part of its risk management, the AI governance team will need to assess the AI risks to determine whether it is prohibited, acceptable, or requires mitigation measures before the organization uses or develops AI. The AI governance team then needs to document this assessment in a data protection and ethical impact assessment.

The AI governance team also needs to establish, implement, document, and maintain the risk management system throughout the AI lifecycle. Risk management requires monitoring and evaluating emerging significant risks based on data collected after the AI system is deployed because some of the risks may not become apparent until the AI is used in the real environment. The AI governance team should ensure that any risk remaining in the AI system is acceptable and communicated to the deployer, and the risk is proportionate to the benefits and goals of the AI technology. The risk management system should also involve testing the AI systems before they are marketed or used, using appropriate metrics and thresholds. Further, the risk management system should pay special attention to the possible impacts on vulnerable groups (e.g., elderly and children) and be compatible with or integrated into any existing risk management procedures required by law.

5. Mitigation measures

The AI governance team will need to adopt measures to mitigate the risks to the extent the AI-use case is not prohibited under the law. The team should implement the mitigation measures below to ensure public trust and confidence in your organization's AI systems.

A. Transparency and explainability

The AI governance team needs to communicate your organization's use of AI systems in its products and services through appropriate plain-language public statements. The function of these public statements is two-fold.

First, the statements should be transparent about the AI systems, including how, when, and for what purpose the organization is using AI and the individual or organization responsible for the AI systems. The transparency notice should provide sufficient information so that the public understands the capabilities, limitations, and potential impacts of the AI systems.

Second, your organization needs to explain in non-technical and easy-to-understand language the AI's decision-making process. Public statements regarding AI and functionality should be tailored for its intended audience. Through transparency and explainability, your organization can build trust and confidence in its AI systems and build an open relationship with the public.

B. Fair and unbiased

The AI governance team should implement appropriate measures (including during the data governance stage) to ensure that the AI systems are fair and unbiased. To be fair, AI systems should uphold equality, equity, human rights, and democratic values, respect the rule of law, protect vulnerable people throughout the AI's lifecycle, and allow all individuals interacting with the AI to access products or services.

The AI governance team can minimize or avoid bias and discrimination against certain groups by:

a. Conducting proactive equity assessments as part of the system design

b. Using data representative of the environment

c. Ensuring that the AI systems are accessible for people with disabilities

d. Conducting disparity testing and migration

e. Implementing active oversight of the AI system

C. Human-centered and beneficial for the environment and society

The AI governance team should communicate to the public the benefits of its AI to individuals, the environment, and society. While AI offers many exciting innovations for society, the public is more concerned than excited about the proliferation of AI in their daily lives. This fear has also been fueled by technologists and AI researchers who have called for an immediate pause on training AI because "AI systems with human-competitive intelligence can pose profound risks to society and humanity." Thus, to responsibly govern AI and maintain public trust in the organization, the AI governance team should incorporate the benefits of AI to humans, society, and the environment as part of your organization's Environmental, Social, and Governance (ESG) messaging.

D. Accuracy

The AI governance team will need to test and document the accuracy of the AI systems by benchmarking how close initial observations, computations, or estimates are to true values. Your organization will then need to truthfully prepare public-facing statements regarding the AI's level of accuracy and the relevant metrics that were used to evaluate their accuracy.

E. Robustness

The AI governance team should ensure that the AI systems are robust so that they can cope with erroneous inputs or errors during execution and function correctly in non-ideal circumstances. This also includes ensuring that the AI systems perform in a manner that will not harm people or operate in unexpected settings.

The AI governance team should assess the AI's robustness by conducting ongoing testing or monitoring to confirm that the system is performing as intended, including through adversarial testing on the AI models to ensure that they are able to handle a broader range of unexpected input variables. AI systems must be stable and resilient under a variety of circumstances to remain trustworthy.

F. Safe and secure

The AI governance team should ensure that the AI systems are safe and secure throughout the AI lifecycle by continually identifying, assessing, and managing risks.

AI systems should be safe so that they don't endanger human life, health, property, or the environment. Safe AI systems require:

a. Responsible design, development, and deployment practices

b. Clear information to deployers on responsible use of the AI systems

c. Responsible decision-making by deployers and end users

d. Explanations and documentation of risks based on empirical evidence of incidents

The AI governance team should also consider and align their AI practices with applicable industry safety guidelines.

For secure AI, the AI governance team should implement appropriate cybersecurity (i.e., confidentiality, integrity, and availability) and adopt an incident response plan to address potential misuse of the AI systems and respond to third-party bad actors who may, for example, try to exploit vulnerabilities and manipulate the training dataset (data poisoning) or pre-trained components used in training (model poisoning), which may lead to harmful decision-making. Your organization should regularly test and conduct diligence on its AI systems throughout its lifecycle and update technical standards addressing safety and security. This also includes, for example, conducting internal and external red-teaming of models or systems in areas including misuse, societal risks, and national security concerns, such as bio, cyber, and other safety areas.

G. Privacy-enhanced

The AI governance team should address data privacy issues in the organization's AI systems, which includes, among other things, providing a privacy notice before using a data subject's personal information in AI systems (and, where necessary, obtaining consent or observing another lawful basis for processing), honoring data subject rights (delete, correct, access, opt-outs, etc.), entering into appropriate contracts with processors and other parties, and adopting privacy-enhancing technologies (PETs). PETs include data minimization methods, such as anonymization, de-identification, or aggregation so that personal data is not used in AI systems. Your organization should've already addressed these requirements as part of your general data privacy compliance efforts. However, the AI governance team should confirm that your organization's data privacy practices are included within their scope AI use cases.

H. Human oversight

The AI governance team needs to determine the level of human oversight necessary for the AI systems to prevent or minimize risks. To decide on the level of human oversight, the AI governance team should weigh the severity and probability of harm and consider the context of the AI use.

The AI governance team should consider three levels of human oversight for its AI systems:

a. Human-out-of-the-loop, which has no human oversight over AI decisions

b. Human-over-the-loop, which has a human involved in monitoring or supervising AI decisions with the ability to take over control when the AI encounters unexpected or undesirable events

c. Human-in-the-loop, which has active and involved human oversight, with a human retaining full control and the AI only providing recommendations or input.

If the severity and probability of harm are high (e.g., a doctor providing medical treatment to a patient), human-in-the-loop is the appropriate level of oversight. However, if the severity and probability of harm are low (e.g., product recommendations for consumers to purchase on an e-commerce website), human-out-of-the-loop is sufficient. Further, if there is a combination of low/high severity and probability of harm, organizations can consider human-over-the-loop as an option (e.g., GPS navigation recommending a route).

Lastly, the AI governance team should be aware of several other factors as part of its human oversight of AI systems, including:

a. Developing a kill switch if the AI poses a danger

b. Avoiding automation bias by understanding that AI systems are not always right

c. Understanding the capacities and limitations of the AI systems

d. Learning how to correctly interpret the AI output

e. Knowing when to disregard, override, or reverse AI decisions

I. Technical documentation and logs

The AI governance team should maintain technical documentation related to the AI systems and automatically recorded events. For technical documentation, the AI governance team should include, among other things, a description of:

a. The AI system

b. The elements of the AI system and its development process

c. The monitoring, functioning, and control of the AI system, particularly with regard to its capabilities and limitations in performance

d. The risk management system adopted

e. Any change made to the system through its lifecycle

f. The system in place to evaluate the AI systems' performance in the post-market phase

The AI governance team should also ensure that the AI systems are capable of automatically recording events (i.e., logs) while it is in operation, including the period of each AI use (start and end date and time of each use), the reference database against which the system has checked the input data, the input data for which the search has led to a match, and the persons involved in verifying the results.

J. Post-market monitoring system

The AI governance team should implement a post-market monitoring system to ensure that your organization's AI systems are in compliance throughout the data lifecycle. This involves collecting, documenting, and analyzing data about how the AI systems perform and interact with other systems or environments, and identifying and addressing any risks, defects, or non-conformities that may arise from the AI systems.

Even after developing and using an AI system, your organization's oversight of AI is not complete. The AI governance team should stay abreast of developments and feedback in the AI system to mitigate any ongoing risks.

K. Contestability

The AI governance team should establish communication channels for members of the public if they have any feedback or questions regarding the organization's AI systems. The organization should also have a decision review channel to provide individuals an opportunity to opt out and request human review of AI decisions (particularly those that have legal or similarly significant effects on them), contest decisions, and access to a person who can quickly consider and remedy any problems they may have encountered. The AI governance team should monitor the feedback and human review channels to ensure that they are accessible, equitable, effective, maintained, and do not impose an unreasonable burden on the public.

6. Accountability

Organizations are ultimately accountable for their AI systems. The AI governance team will need to ensure that measures are in place to oversee the AI systems, with accountability across the AI life cycle. The AI team needs to ensure that the above steps are properly documented through auditable policies, procedures, and practices. This may also require third-party audits to objectively test the organization's AI systems to validate its governance.

How can OneTrust help?

The road the responsible AI is one that requires all hands on deck in order to ensure fair use of AI across your organizations entire data lifecycle. With OneTrust, you can do the following with ease:

- Maintain an ML and AI inventory across your business
- Assess AI models for bias and transparency against global laws and frameworks
- Evaluate different use cases, identify risks, and govern decisions in AI development

OneTrust AI Governance helps your organization understand where AI tech is being used throughout your businesses' data lifecycle, governs its development to ensure responsible use, and mitigates risk to ensure fairness and compliance.

To learn more about how your organization can power your AI systems with trust and governance, request a demo today.

onetrust

As society redefines risk and opportunity, OneTrust empowers tomorrow's leaders to succeed through trust and impact with the Trust Intelligence Platform. The market-defining Trust Intelligence Platform from OneTrust connects privacy, GRC, ethics, and ESG teams, data, and processes, so all companies can collaborate seamlessly and put trust at the center of their operations and culture by unlocking their value and potential to thrive by doing what's good for people and the planet.