Overview

Conducting an AI Risk Assessment

Arsen Kourinian, Mayer Brown

Bloomberg Law

Read Professional Perspectives | Become a Contributor

Conducting an AI Risk Assessment

Contributed by ***Arsen Kourinian****, Mayer Brown*

Editor' Note: This document contains guidance on conducting a risk assessment that is harmonized to address requirements under domestic and international artificial intelligence laws, guidelines and frameworks, including the proposed EU Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (**EU AI Act**), which is expected to be finalized in early 2024. Until such finalization, the information provided below is subject to change.

For additional guidance on practice-specific areas of risk associated with the use of generative and other forms of AI, see our **AI Legal Issues Toolkit**. For additional information on laws, regulations, guidance, and other legal developments related to AI, visit **In Focus: Artificial Intelligence (AI)**.

Artificial intelligence (AI) can provide a great benefit to society, but it also carries risks. For this reason, AI guidance and regulations, such as the **National Institute of Standards and Technology (NIST) AI Risk Management Framework** and the proposed EU Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (**EU AI Act**) (including information about the final text of the EU AI Act from the European Commission's **Artificial Intelligence – Questions and Answers**), focus on management of risk as part of AI governance. To conduct a risk assessment, an organization's AI governance team should first identify and rank the risks as unacceptable (prohibited), high, limited, or minimal, evaluate the probability of harm, implement mitigation measures to reduce or eliminate risks, and document the risk assessment to demonstrate accountability.

While there is no one-size-fits all for conducting a risk assessment, below are steps an organization may consider as part of the risk management phase of an AI governance program. While addressing the considerations described in this overview, organizations should also review their existing infrastructure for risk assessments and consider how to leverage their current procedures to address AI. The type of risk assessment an organization may need to conduct will also depend on the laws it is subject to, such as the specific requirements of the EU AI Act when the final text is released and other future legislation.

Identify & Rank AI Risks

The first step in an AI risk assessment is to identify and rank the risks as unacceptable (prohibited), high, limited, or minimal. If an organization identifies a risk as unacceptable, it will need to stop engaging in that AI processing activity, depending on applicable laws. However, if the risk is not prohibited, the organization should rank the risk and assess its likelihood of harm, as described below in **Document the Risk Assessment**.

Unacceptable Risks

When ranking AI risks, organizations should initially assess whether the risk is unacceptable or prohibited by law. Whether an AI processing activity is prohibited will depend on the laws applicable to an organization. For example, in the US, the Federal Trade Commission, Equal Employment Opportunity Commission, Department of Justice's Civil Rights Division, and Consumer Financial Protection Bureau issued a **joint statement** on April 25, 2023, reminding the public that the use of AI could violate existing laws under certain circumstances, which are prohibited practices. In addition, the EU AI Act specifically designates certain AI systems as unacceptable. The following chart provides examples of some unacceptable AI processing activities organizations may consider as part of their risk assessment based on laws in the US and the EU AI Act.

Unacceptable or Prohibited AI Risk	**Source**
Using AI for unfair or deceptive acts or practices.	**Section 5 of the FTC Act**
Employer using AI to discriminate against an applicant or employee due to protected classifications.	**Title VII of the Civil Rights Act**
AI-based credit decisions that prevent creditors from accurately identifying the specific reasons for denying credit or taking other adverse actions.	**Equal Credit Opportunity Act**
AI-based scoring systems used to screen rental applicants based on race.	**Fair Housing Act**
Social scoring for public and private purposes.	**EU AI Act**
Exploitation of vulnerabilities of persons and use of subliminal techniques.	**EU AI Act**
Real-time remote biometric identification in publicly accessible spaces by law enforcement, subject to narrow exceptions.	**EU AI Act**
Biometric categorisation of natural persons based on biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs or sexual orientation.	**EU AI Act**
Individual predictive policing.	**EU AI Act**
Emotion recognition in the workplace and education institutions, unless for medical or safety reasons (i.e. monitoring the tiredness levels of a pilot).	**EU AI Act**
Untargeted scraping of internet or CCTV for facial images to build-up or expand databases.	**EU AI Act**

If the AI processing activity is considered unacceptable or prohibited under applicable laws, an organization should cease such practices and reconsider the proposed AI systems.

High-Risk AI

If the AI system is not prohibited, the organization should next analyze if the processing activity is high-risk, which may vary depending on applicable laws and jurisdictions. Examples of potential high-risk AI processing activities are provided below, which are based on a list of high-risk AI systems under the EU AI Act, and references to AI concerns noted in other authorities in the US and Europe, such as the UK's **A Pro-Innovation Approach to Regulating AI**, the White House's **Blueprint for an AI Bill of Rights**, the NIST AI Risk Management Framework, and various US and international privacy laws.

Activity	Privacy Laws	EU AI Act	UK, A Pro-Innovation Approach to AI Regulation	US, Blueprint for an AI Bill of Rights	NIST AI Risk Management Framework
Critical infrastructure		☒	☒		
Product safety component		☒	☒		
Biometric identification and surveillance	☒	☒		☒	
Education and vocational training		☒		☒	☒
Employment and recruitment		☒		☒	
Essential goods, services and benefits	☒	☒		☒	
Consumer rights			☒		
Law enforcement and administration of justice		☒	☒	☒	
Body scanners				☒	
Immigration and border control		☒			
Deep fakes				☒	
Sensitive personal data and domains	☒			☒	
Financial, lending, credit and economic opportunities	☒		☒	☒	☒
Housing	☒			☒	
Insurance	☒		☒	☒	
Healthcare and medical devices	☒		☒	☒	
Intrusion upon solitude, seclusion or private affairs and other privacy violations	☒		☒	☒	
Sale of personal data and data broker activities	☒			☒	

Activity (cont.)	Privacy Laws (cont.)	EU AI Act (cont.)	UK, A Pro-Innovation Approach to AI Regulation (cont.)	US, Blueprint for an AI Bill of Rights (cont.)	NIST AI Risk Management Framework (cont.)
Certain tracking activities and targeted advertising depending on applicable law	☒			☒	
Discrimination against population sub-group			☒	☒	☒
Physical or psychological harm and safety	☒		☒	☒	☒
Civil liberties or rights and democratic participation			☒	☒	☒
Harm to an organization's business operations			☒		☒
Harm to an organization from security breaches or monetary loss			☒		☒
Harm to an organization's reputation					☒
Harm to the global financial system, supply chain, or interrelated systems					☒
Harm to natural resources, the environment, and the planet					☒
Intellectual property rights			☒		

Limited & Minimal AI Risks

AI systems that are not considered high-risk may fall within a lower tier of AI risks. We understand that the final text of the EU AI Act may designate certain AI risks **as limited**, such as chatbots, which will require organizations to inform users that they are interacting with a machine so that they can make an informed decision whether to continue interacting with the AI system. Other AI systems may fall within the minimal risk category, which will not require further mitigation steps under the EU AI Act. For **minimal-risk AI systems**, organizations may voluntarily choose to apply the requirements for trustworthy AI and adhere to voluntary codes of conduct.

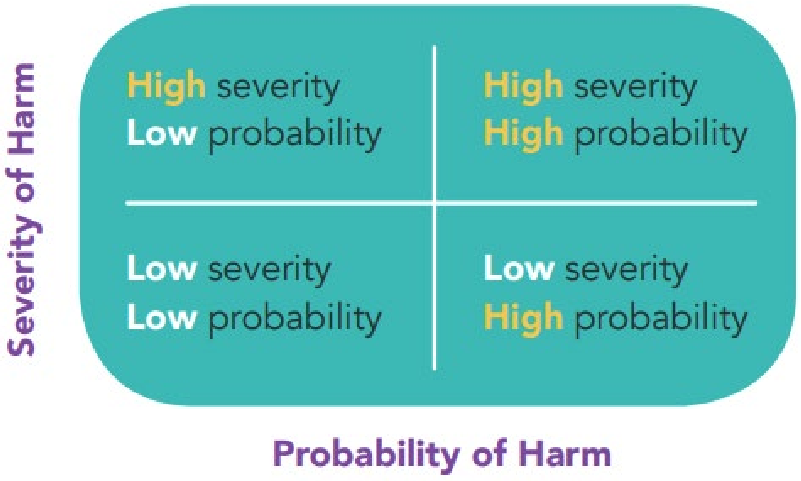
Outside of the EU AI Act, organizations should also consider the context of the AI processing activity to determine whether voluntary steps should be taken to mitigate AI risks that are not high risk. For example, depending on the context, low-risk AI systems may rise to the level of medium risk, examples of which are noted in the chart below.

AI Usage	Medium	Low
GPS navigation	Used for delivering food to homeless shelters	Used by regular consumers for daily commute
Chatbots	Used to provide career advice to students* **Note that this could also potentially be considered high-risk under the education category in the final text of the EU AI Act.*	Used to provide instructions on how to use a website
AI in video games	Used to assign scores in a video game competition involving a significant cash prize	Used to generally play the video game

Identify the Likelihood of Harm

Once an organization ranks the risks, it should then assess the likelihood that the risk will materialize, as explained under the NIST AI Risk Management Framework and the recently released International Organization for Standardization's Information Technology Artificial Intelligence Management System (ISO/IEC 42001:2023 at 6.1.2(d)(2)).

Using a risk matrix, an organization may characterize the severity and likelihood of harm as critical, moderate, or low. Depending on the risk spectrum, an organization may need to implement additional mitigation measures and safeguards. For example, **guidance** from Singapore suggests that a risk matrix may be used "to help organisations determine the level of human involvement required in AI-augmented decision-making." To help guide this assessment, Singapore offers a basic two-by-two risk matrix that categorizes risks and severity as a combination of high and low.



Another example is a three-by-three risk matrix, whereby a score of one through three is assigned for the severity of risk and probability of harm, which are then multiplied to provide a risk score.

	Probability of Occurrence (Low to High)		
Severity of Harm (Low to High)	Low Risk / Low Likelihood (1)	Low Risk / Medium Likelihood (2)	Low Risk / High Likelihood (3)
	Medium Risk / Low Likelihood (2)	Medium Risk / Medium Likelihood (4)	Medium Risk / High Likelihood (6)
	High Risk / Low Likelihood (3)	High Risk / Medium Likelihood (6)	High Risk / High Likelihood (9)

While examples of risk matrices are provided in the graphic links above, there is no single method for conducting a risk assessment. Companies may leverage their existing procedures for risk audits and assessments to assign scores to AI processing activities, depending on the severity and probability of harm. Organizations may also use a **more complex** risk matrix, which accounts for risk velocity—i.e., time to impact—and risk contagion—i.e., potential for risk in one area impacting other areas of the organization.

Document the Risk Assessment

Organizations should document AI risk assessments to demonstrate accountability. On a high level, the documentation should reflect the risks identified during the assessment, the steps taken to mitigate the risks, and whether, on balance, the mitigation measures are adequate and sufficient to address the risks for the organization to proceed with the AI processing activity. Such risks assessments are already required under existing privacy laws, such as the EU's **General Data Protection Regulation** (GDPR **Article 35**) and US comprehensive privacy laws—e.g., **Colorado Privacy Act**, **Virginia Consumer Data Protection Act**, and **Connecticut Data Privacy Act**. For an in-depth comparison of the EU GDPR and state comprehensive consumer privacy laws, including requirements pertaining to profiling and automated decision-making, see **Comparison Table - GDPR vs. State Comprehensive Consumer Privacy Laws**.

There are several sources to consult for the issues to cover in a risk assessment. The UK Information Commissioner's Office, for example, issued **guidance** on new content to include in data protection/privacy impact assessments (PIA) involving AI processing activities. In the US, the Colorado Privacy Act regulations (**4 CCR 904-3-9.06**) also identify the elements to include in a PIA for certain profiling activities that involve automated processing of personal data. In addition, the California Privacy Protection Agency (CPPA) has begun **drafting regulations** under the **California Consumer Privacy Act** (CCPA), as amended by the California Privacy Rights Act (CPRA), regarding how to document an AI risk assessment. For more information on complying with California privacy laws see **Practical Guidance: California Privacy (CCPA/CPRA)**.

Further, as **recommended** by the United Nations Educational, Scientific and Cultural Organization (UNESCO), an AI risk assessment should factor in ethical considerations as well. UNESCO states that the "impact assessments should identify impacts on human rights and fundamental freedoms, in particular but not limited to the rights of marginalized and vulnerable people or people in vulnerable situations, labour rights, the environment and ecosystems and ethical and social implications, and facilitate citizen participation in line with the values and principles set forth in [UNESCO's guidance]."

The **Ethical Accountability Framework for Hong Kong, China** similarly recommends conducting an Ethical Data Impact Assessment (EDIA), in conjunction with a PIA. Hong Kong's framework states that "[t]he EDIA is broader in scope than the typical PIA" because "all data are considered in an EDIA and not just personal data." This includes "data in the aggregate, nonidentifiable form that may be outside the scope of . . . privacy and data-protection laws." Hong Kong's framework further states that by completing both a PIA and EDIA, it will demonstrate an organization's good faith attempt to comply with privacy laws and AI accountability requirements.

Lastly, based on the recent deal reached on the EU AI Act, we understand that the final text will include **a requirement to conduct** "a mandatory fundamental rights impact assessment" for high-risk AI processing activities. The fundamental rights impact assessment needs to be completed in conjunction with a PIA, and "consist of a description of the deployer's processes in which the high-risk AI system will be used, of the period of time and frequency in which the high-risk AI system is intended to be used, of the categories of natural persons and groups likely to be affected by its use in the specific context, of the specific risks of harm likely to impact the affected categories of persons or group of persons, a description of the implementation of human oversight measures and of measures to be taken in case of the materialization of the risks." *See* European Commission, **Artificial Intelligence – Questions and Answers**. This appears consistent with recommendations by the Confederation of European Data Protection Organisations, which states in its **recent guidance** that AI governance requires the provider or user of the AI systems to complete a Fundamental Rights Impact Assessments, which analyzes the "risks to the fundamental rights and freedoms of individuals who are affected by the output."

Engage in Continuous Monitoring

Organizations should be mindful that conducting a risk assessment is not a one-and-done activity. Rather, organizations should establish, implement, document, and maintain the risk management system throughout the AI lifecycle. This includes monitoring and evaluating emerging risks once the AI system is deployed in the real environment.