

A survey of artificial intelligence risk assessment methodologies

The global state of play and leading practices identified

Executive summary

December 2021



Building a better
working world

TRILATERAL
RESEARCH



Introduction

In recognition of both the increasing importance of AI in our digital society and the wide diversity of use cases, policymakers across the globe are seeking to better understand the risks that these new AI systems might pose to society. A growing consensus is emerging in favor of risk-based approaches to regulating the use of AI. This consensus builds on the work of high-level expert groups, and studies by academics, industry associations, professional bodies and civil society.

This summary captures the main findings from EY teams survey of AI risk assessments (AIRA). The research aims to inform policymakers and regulatory stakeholders about noteworthy approaches to AIRA, including leading practices and policy trends. The findings in this report are based on a survey which has analyzed the following sources:

- ▶ Legal and regulatory approaches to date
- ▶ Current work at international bodies
- ▶ Work by standards development organizations
- ▶ Industry approaches
- ▶ Prominent approaches proposed in civil society and;
- ▶ Academic literature

Through EY teams research, this report has identified leading practices within the following areas:

- ▶ Categorization of risk
- ▶ Risk management
- ▶ Requirements for trustworthiness
- ▶ The relevant stakeholders who should be involved for identifying and mitigating AI risk

This summary represents a survey and assessment of the ecosystem of AIRA methodologies. It does not claim to be comprehensive but provides a snapshot of the AIRA landscape in 2021. Ultimately, this report can be used as a tool to understand the AIRA landscape, including emerging policy trends and leading practices.

Policy trends for the governance of AI

As of August 2021, there is no universally accepted model or common regulatory framework for AIRA, as governments are just beginning to develop policies and laws specific to AI. However, we have identified the following trends:

01

AI principles form benchmarks for AIRA

- ▶ Many countries and organizations have defined AI principles that largely align with the values-based principles published by the OECD in May 2019, and subsequently adopted by the G20.
- ▶ There is a widespread trend towards including assessment of the risks of violating the AI principles in efforts to operationalizing the AI principles.
- ▶ The broad agreement on AI principles presents an opportunity for high levels of international regulatory alignments on risk assessment.

02

AI standard development organizations are working to support risk assessments

- ▶ The technical community, including National and International Standards Development Organizations (e.g., ISO/IEC, IEEE), are making good progress on developing standards and guidance for technical implementation of AIRA. For example, draft standards such as on Artificial Intelligence and Risk Management (ISO/IEC 23894) and on Recommended Practice for Organizational Governance of Artificial Intelligence (IEEE P2863), are being developed to help support the development of risk assessments.

03

Industry is forming mixed membership associations to promulgate policy guidance

- ▶ Industry has been pursuing various approaches to AIRA and management. Of particular note are associations with a mixed membership of companies, academics, policymakers and/or civil society organizations.
- ▶ There is recognition that AIRA will need to take into account the particularities of specific sectors.

04

Academics and other researchers are playing a role in developing AIRA methodologies

- ▶ Research papers on proposed AIRA frameworks take inspiration from environmental, data protection, privacy, human rights, ethics, socio economic impact assessment methods.
- ▶ They have helped to clarify key concepts and proposed structured processes for the assessment of risks emerging from AI. For example, the European Commission appointed High Level Expert Group on Artificial Intelligence (AI HLEG).

While the AIRA ecosystem is still in flux, it is nevertheless possible to identify leading practices which could help inform AI policymaking as well as the promulgation of AIRA.

Leading practices

Based on the survey presented in this report, we have identified a number of leading practices, supported by two or more of the stakeholders that were cited in this study. These leading practices provide a good starting point for risk-based AI regulation, but any framework will need to be tailored to

fit into the legislative tapestry of the jurisdiction where it is applied. In doing so, special consideration needs to be given to the regulatory ecosystem that applies to the AI-related digital infrastructure of data, privacy legislation and information security, among others.

► Categorical dimensions of risk assessment

AI risks may be assessed along various categorical dimensions though not all of them will be applicable in all circumstances. The primary dimensions include:

- Object of analysis (e.g., overall application, role of AI subcomponent)
- Nature of risk (e.g., ethical, technical/performance, governance, communication, security)
- Metric of severity (e.g., impact on fundamental rights, frequency of use, number of affected persons, availability of (human) alternative, (ir)reversibility of harm, extent to which law provides prevention/mitigation measures)

► Risk levels

To enable a regulatory regime where the burden of compliance is proportionate to the risk posed by the AI application, leading practices suggest that risk assessments should distinguish at least three risk levels (e.g., high, medium and low). This is sometimes expanded to four levels (proposed EU AI Act suggests unacceptable, high, non-high, minimal) or five levels (Proposed German Data Ethics Committee suggests: untenable, serious, significant, some, negligible).

► Risk management

Risk assessment plays an integral part in risk management systems for AI, with leading practices suggesting that this should include:

1. identification and analysis of the known risks,
2. estimation and evaluation of new risks that are foreseeable and
3. evaluation of other, unforeseen, risks that become apparent from observations during post-market monitoring of the use and adoption of AI systems.

► Role of risk assessment is establishing trustworthiness of AI

Leading practices for establishing the trustworthiness of AI emphasize that the AI system must, in all phases of its life-cycle (design, development, and use):

1. operate in a manner that is lawful, fair, unbiased, accurate, reliable, effective, safe, secure, resilient, understandable,
2. have processes in place to regularly monitor and evaluate the AI system's performance and outcomes, and
3. be shown to afford sufficient levels of reliability, accountability, maintainability, functionality, debug-ability, evolve-ability, fragility, vulnerability, privacy, transparency, and bias.

AI risk assessment is therefore likely to play an integral part in the development and maintenance of trustworthy AI, even in the absence of a risk-based regulatory framework.

► Relevant stakeholders for identifying and mitigating AI risk

Given the scale of the challenges associated with AI, efforts to identify guidance around AI risk assessment frequently mobilize a diverse set of participants, including businesses, consumer organizations, trade unions and other representatives of civil society bodies.

Conclusion

The broad diversity of methods and applications where AI is being deployed makes it unlikely that cross-sectoral AI legislation will be able to address all the nuances that a good regulatory environment requires. It is therefore anticipated that such legislation will have to be augmented with sector-specific approaches that can address specialized AI risk factors.

Acknowledgements

This summary is an excerpt from a literature survey and report commissioned by EY teams and delivered by Trilateral Research.

Authors

Gabriella Ezeani

Research Analyst, Trilateral Research

Dr. Ansgar Koene

Associate Director, Ernst & Young LLP (UK), and EY Global AI Ethics and Regulatory Leader

Richa Kumar

Research Analyst, Trilateral Research

Nicole Santiago

Senior Research Analyst, Trilateral Research

David Wright

Director, Trilateral Research

Contributors

Douglas Bell

Director, Ernst & Young LLP (US), and EY Global Public Policy

Cathy Cobey

Partner, Ernst & Young (Canada), and EY Global Trusted AI Leader

Nigel Duffy

Partner, Ernst & Young LLP (US), and EY Global Artificial Intelligence Leader

Nina Emordi

Associate Director, Ernst & Young LLP (UK), and EY Global Public Policy

Eugène Goynes

Associate Partner, Ernst & Young LLP (Hong Kong), and EY Asia Pacific Financial Services Regulatory Lead

Josh B Konvisser

Partner, Ernst & Young LLP (US) and EY Americas General Counsel Office



Katie Kummer

Partner, Ernst & Young LLP (US), and EY Global Deputy Vice Chair, Public Policy

Ivan Mannino

Senior Associate, Ernst & Young (Belgium), and EMEIA Public Policy

Anaïs Resseguier

Senior Research Analyst, Trilateral Research

Fatima Hassan-Szlamka

Associate Director, Ernst & Young LLP (US), and EY Global Public Policy

For further information, please contact:

Dr. Ansgar Koene, EY, Global AI Ethics and Regulatory Leader
ansgar.koene@uk.ey.com



EY | Building a better working world

EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via ey.com/privacy. EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

© 2021 EYGM Limited.
All Rights Reserved.

2109-3873062
EYG no. 009900-21GbI
ED None

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.

ey.com