March 9, 2026

OECD Publishes Responsible AI Due Diligence Guidance for Multinational Enterprises

Authors:
Michael R. Littenberg , Samantha Elliott , Kelley R. Murphy

Responsible AI practices are a growing focus within many companies and in society at large. In late February, the Organisation for Economic Co-operation and Development (OECD) published Due Diligence Guidance for Responsible AI. The AI Guidance is intended for multinationals supplying inputs for AI development, actively participating in the AI system lifecycle or utilizing AI systems in their operations, products and services across all sectors. The AI Guidance is discussed in this post.

The OECD was established more than 60 years ago as a policy forum for governments to share experiences and seek solutions to common economic and social problems. Consistent with that purpose, the AI Guidance is intended to assist multinational enterprises involved in the AI system value chain in implementing both the OECD Guidelines for Multinational Enterprises on Responsible Business Conduct (MNE Guidelines) and the May 2024 revised Recommendation of the Council on Artificial Intelligence (AI Principles).

The AI Guidance can be used to inform implementation of human rights and supply chain policies and procedures that are based on the UN Guiding Principles on Business and Human Rights and the MNE Guidelines. It also may inform compliance under both current and emerging mandatory human rights due diligence legislation. Additionally, the AI Guidance is intended to complement industry-led AI risk management and governance frameworks.

For those new to this area, the MNE Guidelines are recommendations jointly addressed by governments to multinational enterprises to enhance the contribution of the business community to sustainable development and address adverse impacts associated with business activities on people, planet and society. The MNE Guidelines aim to encourage positive contributions enterprises can make to economic, environmental and social progress, and to minimize adverse impacts on matters covered by the MNE Guidelines that may be associated with an enterprise's operations, products and services. The MNE Guidelines explicitly cover key areas of business responsibility, including human rights, labor rights, environment, bribery, consumer interests, disclosure, science and technology, competition and taxation.

Chapter IX of the MNE Guidelines recommends that enterprises conduct due diligence on adverse impacts associated with science, technology and innovation. The 2023 update clarifies that the scope of this chapter, including due diligence requirements, covers development, financing, sale, licensing, trade and use of technology, including gathering and using data, as well as scientific research and innovation. The most recent update to the MNE Guidelines is discussed in this Ropes & Gray post.

The OECD has undertaken empirical and policy activities on AI since 2016. The first version of the AI Principles – which the OECD characterizes as the first intergovernmental standard on AI – was adopted in 2019. According to the OECD, they "aimed to foster innovation and trust in AI by promoting the responsible stewardship of trustworthy AI while ensuring respect for human rights and democratic values." Complementing existing OECD standards already relevant to AI, such as those on privacy and data protection, digital security risk management and responsible business conduct (RBC), the AI Principles are focused on policy issues specific to AI.

The AI Principles contain five high-level values-based principles for responsible stewardship of trustworthy AI. These relate to (1) inclusive growth, sustainable development and well-being; (2) respect for the rule of law, human rights and democratic values, including fairness and privacy; (3) transparency and explainability; (4) robustness, security and safety; and (5) accountability.

The AI Principles also include five high-level recommendations for national policies and international cooperation relating to (1) investing in AI research and development; (2) fostering an inclusive AI-enabling ecosystem; (3) shaping an enabling interoperable governance and policy environment for AI; (4) building human capacity and preparing for labor market transformation; and (5) International cooperation for trustworthy AI. In addition, the AI Principles propose a common understanding of key terms such as “AI system,” “AI system lifecycle” and “AI actors” for the purposes of the AI Principles.

The AI Principles have been revised twice since their initial adoption. They were revised in 2023 to update the definition of “AI system” for technical accuracy and to reflect technological developments, including pertaining to generative AI. The AI Principles were further revised in May 2024 to reflect technological and policy developments, including with respect to generative AI, and to further facilitate their implementation. The 2023 and 2024 updates are discussed in more detail here.

Objectives of the AI Guidance

The objectives of the AI Guidance include, among others:

- Supporting innovation, investment and growth of enterprises in the AI value chain by providing clarity on how enterprises can proactively identify and address actual and potential adverse impacts that they may cause, contribute to or be directly linked to;
- Helping enterprises navigate existing international, national, multi-stakeholder or industry-led AI risk management and governance frameworks;
- Promoting policy coherence, and where possible interoperability, between the MNE Guidelines, AI Principles and other national or international AI risk management and governance frameworks; and
- Serving as a common reference point for AI risk management frameworks across different jurisdictions.

Intended Users

The AI Guidance is specifically intended for multinationals (1) supplying inputs for AI development, (2) actively participating in the AI system lifecycle or (3) utilizing AI systems in their operations, products and services across all sectors. Each of these user groups is further described below.

“AI system” is defined in the AI Principles as a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments. The definition notes that different AI systems vary in their levels of autonomy and adaptiveness after deployment. The AI Principles describe an “AI system lifecycle” as typically involving several phases that include to: plan and design; collect and process data; build model(s) and/or adapt existing model(s) to specific tasks; test, evaluate, verify and validate; make available for use/deploy; operate and monitor; and retire/decommission. The AI Principles note that these phases often take place in an iterative manner and are not necessarily sequential.

Group 1: These enterprises provide inputs into the development of the AI system and are generally considered to be the “upstream” segment of the AI system value chain. This includes activities pertaining to the provision of inputs in

the AI ecosystem: the skills and resources, such as data, code, algorithms, models, research, know-how, training programs, governance, processes and best practices required to understand and participate in the AI system lifecycle, including managing risks. It also includes activities related to the provision of financial, logistical, administrative and hardware inputs needed to support the development of the AI system.

The AI Guidance does not cover supply chains of hardware inputs, such as mining of raw materials and manufacturing of hardware components, which are the subject of separate OECD RBC due diligence guidance.

Group 2: This group includes enterprises actively involved in the design, development, deployment and operation of AI systems.

Group 3: This group includes enterprises that use AI systems in their operations, products and services, which are generally considered to be the "downstream" segment of the AI system value chain. These include financial institutions and manufacturers and sellers of goods and services, including goods and services unrelated to AI systems or technology.

The AI Guidance indicates that enterprises in this group should consider due diligence on the AI systems that they use as part of their broader due diligence process across their operations and business relationships. According to the AI Guidance, this means prioritizing the risks of adverse impacts presented by the AI system in relation to other risks the enterprise might be causing, contributing to or linked to in its specific sectors.

Applying the AI Guidance

The AI Guidance is organized into two chapters.

Chapter 1 introduces the concept of RBC due diligence and provides an overview of the broader AI risk management policy landscape. It also describes how to use the guidance as a tool to navigate risk management frameworks.

Chapter 2, which comprises the bulk of the AI Guidance, lays out the RBC due diligence framework and practical implementation examples for enterprises involved in the development and use of AI systems.

The due diligence framework described in the MNE Guidelines and elaborated on in the OECD Due Diligence Guidance for Responsible Business Conduct is the foundation for the AI Guidance. The general concepts and framework discussed in the AI Guidance will therefore be familiar to many human rights and sustainability legal and compliance professionals. Conversely, for many in AI, tech and other corporate functions, these will be new concepts.

As noted above, the AI Guidance includes practical implementation examples. These apply the general RBC framework to AI use. The OECD notes that the practical examples have been selected to fit this context and also draw on leading AI risk management frameworks as well as desk research and consultations with experts. The practical examples are not meant to be an exhaustive or mandatory checklist.

The due diligence framework in the AI Guidance also maps to related provisions in existing frameworks. Approximately 20 are listed. The OECD notes that the RBC due diligence framework is broadly aligned with other AI risk management frameworks and there is significant overlap across many of the frameworks on key issues.

The six steps of the OECD framework are described below. This post does not include the practical examples or framework mapping from the AI Guidance, since those are detailed and context-specific. See the AI Guidance here for additional detail.

Step 1: Embed RBC into policies and management systems

1.1 Devise, adopt and disseminate a combination of policies on RBC issues that articulates the enterprise's

commitments to the principles and standards contained in the AI Principles and MNE Guidelines. Policies should include plans for implementing due diligence relevant to the enterprise's own operations and business relationships in the development and use of AI.

1.2 Seek to embed RBC issues and trustworthy AI into the enterprise's policies, oversight bodies, structures, systems, processes and teams, so that they are implemented as part of regular processes.

1.3 Incorporate RBC expectations and policies into engagement with business relationships.

Step 2: Identify and assess actual and potential adverse impacts

2.1 Carry out a scoping exercise to identify where risks may be present and where they may be most significant.

> The AI Guidance notes that multiple frameworks exist at the international, regional and national level that describe risks related to the development and use of AI systems and recommend actions companies should take to address those risks. The AI Guidance lists applications of AI have been identified by some leading AI risk management frameworks as potentially being high-risk. However, each enterprise is expected to identify its priority risk areas based on its individual circumstances.
>
> When prioritizing risks to be addressed (see Step 2.4 below), the AI Guidance indicates that enterprises should take into account that some risks are closely linked to or may enable others. The AI Guidance notes that, as with many new technologies, public and private malign actors may find ways to exploit AI systems. It also notes that the significant dual-use potential of AI systems and ability to repurpose AI systems can lead to harmful uses even when their design was intended to be innocuous.

2.2 Starting with the most significant areas of risk identified, carry out iterative and increasingly in-depth assessments of prioritized risks related to the enterprise's own activities and its business relationships (e.g., suppliers, customers and users). The AI Guidance discusses risk assessments at the data and model levels, the intersection of these and during the interaction between humans and the AI system.

2.3 Assess the enterprise's involvement with the actual or potential adverse impacts identified. Specifically, assess whether the enterprise or relevant business relationship caused (or would cause) or contributed (or would contribute) to the adverse impact or whether the adverse impact is (or would be) directly linked to its operations, products or services by a business relationship.

2.4 Drawing from the information obtained on actual and potential adverse impacts, prioritize the most significant (i.e., most salient) risks and adverse impacts for action, based on severity and likelihood. Prioritization will be relevant where it is not possible to address all potential and actual adverse impacts immediately. Once the most significant adverse impacts are identified and dealt with, the enterprise should move on to address less significant foreseeable impacts.

Step 3: Cease, prevent and mitigate adverse impacts

3.1 Cease activities that are causing or contributing to adverse impacts based on the enterprise's assessment of its involvement with the impact. Develop and implement plans to prevent and mitigate potential (future) adverse impacts.

3.2 Based on the risk prioritization, develop and implement plans to prevent or mitigate actual or potential adverse impacts directly linked to the enterprise by business relationships. The AI Guidance notes that enterprises throughout the development and use of AI might be directly linked to adverse impacts caused by other AI actors in the system lifecycle or business relationships outside of the AI system lifecycle, such as suppliers of AI inputs and users of the AI system.

Appropriate responses to risks associated with business relationships may at times include (1) continuing the relationship during risk mitigation efforts, (2) temporarily suspending the relationship while pursuing risk mitigation or (3) disengaging after failed attempts at mitigation, where mitigation is not feasible or due to the severity of the adverse impact. The AI Guidance indicates that a decision and subsequent plan to disengage should take into account potential social, environmental and economic adverse impacts and should include meaningful stakeholder engagement.

Step 4: Track implementation and results of due diligence activities

Track the implementation and effectiveness of the enterprise's due diligence activities, i.e., its measures to identify, prevent, mitigate and, where appropriate, support remediation of adverse impacts.

Step 5: Communicate actions to address impacts

Externally communicate relevant information on due diligence policies, processes and activities conducted to identify and address actual or potential adverse impacts, including the findings and outcomes of those activities. Communication can take various forms, depending on the target audience.

Step 6: Provide for or cooperate in remediation when appropriate

When an enterprise has caused or contributed to actual adverse impacts, seek to restore the affected person or persons to the situation they would be in had the adverse impact not occurred (where possible) and enable remediation that is proportionate to the significance and scale of the adverse impact.

About our Practice

Ropes & Gray has a leading ESG, CSR and business and human rights compliance practice. We offer clients a comprehensive approach in these subject areas through a global team with members in the United States, Europe and Asia. Senior members of the practice have advised on these matters for more than 30 years, enabling us to provide a long-term perspective and depth and breadth of experience that few firms can match.

Subscribe to Ropes & Gray ESG, CSR and business and human rights compliance thought leadership here.

www.ropesgray.com

This alert should not be construed as legal advice or a legal opinion on any specific facts or circumstances. This alert is not intended to create, and receipt of it does not constitute, a lawyer-client relationship. The contents are intended for general informational purposes only, and you are urged to consult your attorney concerning any particular situation and any specific legal question you may have. © 2026 Ropes & Gray LLP

ATTORNEY ADVERTISING