A Human Rights-Based Approach to Governance and Management

Guide 2 of the Responsible AI Practitioner Guides for Taking a Human Rights-Based Approach to Generative AI

BSR®

Summary

This paper provides guidance about how to take a human rights-based approach to the governance of generative AI (genAI) within organizations. It includes the following sections:

1 **Integrating Human Rights into AI Principles:** It describes the role of AI principles and how integrating human rights provides a foundation for a human rights-based approach to risk assessment and mitigation.

2 **Incorporating Human Rights Guidance into Governance and Management Structures:** Describes how existing standards and best practices for human rights governance and management apply to responsible AI governance and management.

3 **Key Resources**

Key Points

- A human rights policy commitment is the basis of embedding respect for human rights across an organization; similarly, AI principles are the foundation for embedding responsible AI across an organization.

- AI principles can ground and guide the development and deployment of AI and the remit of Responsible AI teams whose role it is to help an organization implement its principles in practice. Therefore, an important foundation for taking a human rights-based approach to genAI is including respect for human rights in organizational AI principles.

ACCOMPANYING RESOURCES

A HRA of the GenerativeAI Value Chain

Overview of the Practitioner Guide

Guide 1: Human Rights Fundamentals

Guide 2: Governance and Management

Guide 3: Impact Assessment

Guide 4: Risk Mitigation

Guide 5: Stakeholder Engagement

Guide 6: Policies and Enforcement

Guide 7: Transparency and Disclosures

Guide 8: Remedy for GenAI Related Harms

- Within companies, Responsible AI teams or cross-functional working groups are often tasked with implementing AI principles and relevant management processes. Effective collaboration between teams responsible for human rights and Responsible AI, and ensuring Responsible AI teams have human rights expertise on staff, are important for developing human rights-aligned approaches to impact assessment and mitigation.

- While both human rights and Responsible AI teams typically have a mandate to develop impact assessment and risk mitigation processes, in practice they seldom have the authority to impose any specific strategy or governing structures themselves. Therefore, it is important that governance of genAI is embedded throughout the organization. The OECD Due Diligence Guidance suggests several ways in which companies embed human rights into governance and management processes, all of which are relevant for AI.

1. Integrating Human Rights Into AI Principles and Policies

The UNGPs make clear that the basis for embedding the responsibility to respect human rights across an organization is a public "statement of policy" that describes the organization's commitment to respect human rights and how it will be implemented.[1] This policy commitment is important because it clearly communicates the expectations to both internal and external stakeholders, provides a foundation for developing the processes and systems necessary to meet that commitment, and is an important first step for embedding respect for human rights into the values of an organization.[2]

Just as a human rights policy commitment is the foundation for respecting human rights, AI principles and policies are an important foundation for ensuring Responsible AI. The majority of large technology companies involved in the development of genAI have published AI principles that are designed to guide how they develop and deploy AI products and services, including genAI. Increasingly, deployers of AI systems are also publishing AI principles. These principles are often oriented around the common set of values that have grounded the Responsible AI field: the promotion of human values and human control over technology, fairness and nondiscrimination, transparency, explainability, accountability, safety and security, privacy, and human rights. Some companies have also developed accompanying AI policies that describe how they operationalize their AI principles.

When leveraged effectively, AI principles can ground and guide the development and deployment of AI and the remit of teams charged with operationalizing the principles. Therefore, an important foundation for taking a human rights-based approach to genAI is by including respect for human rights in an organization's AI principles.

When AI principles reference human rights explicitly, they provide grounding for integrating a human rights-based approach to impact assessment and risk mitigation into Responsible AI

---

1   See Principles 16 of the UNGPs.
2   See the UNGPs Interpretive Guide guidance on Principle 16.

processes (see Guide 3: A Human Rights-Based Approach to Impact Assessment and Guide 4: A Human Rights-Based Approach to Risk Mitigation). Some examples of company AI principles that explicitly reference human rights include:

- Google's AI principles, which include a commitment to not design or deploy "technologies whose purpose contravenes widely accepted principles of international law and human rights."

- Salesforce's AI principles, which pledge to "safeguard human rights and protect the data we are entrusted with."

- NEC's AI principles, which state that their purpose is to "prevent and address human rights issues arising from AI utilization" and to "guide our employees to recognize respect for human rights as the highest priority in each and every stage of our business operations."

- Cisco's AI principles, which describe how they embed respect for human rights into their approach to Responsible AI.

If a company has a human rights policy, referencing that policy's application to AI can also provide helpful human rights grounding for Responsible AI. For example, Meta's human rights policy states, "Human rights also guide our work developing responsible innovation practices, including when building, testing, and deploying products and services enabled by Artificial Intelligence."

2. Incorporating Human Rights Guidance Into Governance and Management Structures

The UNGPs state that in order to address adverse human rights impacts, businesses should integrate findings from their impact assessments across relevant internal functions and processes. Effective integration requires that responsibility for addressing impacts is appropriately assigned, and appropriate resource allocation, reporting lines, and accountability structures are established.[3]

Within companies, management of the risks to people and society associated with genAI tends to lie with product teams and specialized Responsible AI teams or cross-functional working groups. They typically do not have human rights as an explicit part of their strategy or governing framework, unless their AI principles or commitments reference human rights. In large companies with established human rights teams, these teams may occasionally provide input into the human rights-related aspects of the Responsible AI team's policies and processes, and may conduct or commission human rights assessments. However, this is all largely done via internal collaboration and relationship-building rather than formalized governance structures. Smaller genAI companies often lack any human rights expertise, or have hired individual "leads" who have been charged with bringing a human rights perspective to broader safety or ethics efforts.

Effective collaboration between human rights and Responsible AI teams, and ensuring Responsible AI teams have human rights expertise on staff, are important for developing human rights-aligned approaches to impact assessment and risk mitigation. However, while both human rights and Responsible AI teams typically have a mandate to develop impact assessment and risk mitigation processes, in practice they lack the authority to impose any specific strategy or governing structures themselves, and must rely on collaboration and

---

3   See Principle 19 of the UNGPs.

establishing buy-in across research, product, and sales teams in order to embed human rights throughout a company's genAI approach.

Therefore, it is important that governance of genAI, including through a human rights lens, is integrated into teams other than those that oversee Responsible AI or human rights. The OECD Due Diligence Guidance suggests several ways in which companies embed human rights into governance and management processes, all of which are relevant for Responsible AI:[4]

- **Senior management responsibility**—Assign oversight and responsibility for impact assessment and implementing mitigation measures to relevant senior management.

- **Board oversight**—Assign board level oversight of an organization's AI principles, and relevant processes to exercise oversight (e.g., via a subcommittee that regularly reviews the Responsible AI strategy, policy, and management processes, challenges, and progress in meeting commitments).[5]

- **Delegation of relevant department or team responsibility**—Assign responsibility for implementing Responsible AI policies and processes across relevant departments, with particular attention given to those whose decisions increase or decrease human rights risks (e.g., product teams, research teams, trust and safety teams, legal / compliance, etc.).

- **Information collection mechanisms**—Develop or adapt existing information systems to collect information on Responsible AI processes, decision-making, and responses.

- **Communication channels**—Establish communication channels between relevant senior management and implementing departments for sharing and documenting information on AI-related risks to people and decision-making.

- **Internal alignment**—Encourage alignment across teams and business units on relevant Responsible AI policies and practices, including the role of human rights.

- **Training**—Provide training to employees to help them understand and implement Responsible AI policies and practices.

- **Incentives**—Develop incentives for employees and business units to implement Responsible AI commitments and adhere to policies / participate in processes (e.g., KPIs, performance metrics, bonuses).

- **Complaint mechanisms**—Create or utilize existing complaint mechanisms for employees to raise issues or complaints related to Responsible AI and human rights impacts.

- **Incident response and remediation**—Develop processes to respond to and where appropriate provide remedy in situations where Responsible AI policies are not followed and adverse impacts are identified (e.g., through investigations, capacity-building, or disciplinary actions).

---

4   See section 1.2 of the OECD Due Diligence Guidance.
5   Investors are increasingly asking their portfolio companies about their Responsible AI processes and AI-related risks, making board oversight of Responsible AI increasingly important.

3. Key Resources

While BSR is not aware of any companies that have effectively implemented all of the governance best practices mentioned above, below are some examples of companies that have created strong AI governance approaches that effectively integrate many of those best practices highlighted above:

- **Cisco** has developed a Responsible AI Framework to implement its AI principles, which it considers part of its broader human rights commitment. The framework is overseen by a Responsible AI Committee made up of senior executives across the company. This committee also reviews sensitive or high-risk uses of AI proposed by Cisco business units as well as incident reports for bias or discrimination. Cisco has also effectively incorporated security, privacy, and human rights principles into AI design by leveraging the existing Cisco Secure Development Lifecycle, which is required across the company. Finally, Cisco also leverages its existing security, data breach, and privacy incident response system to manage and escalate any reported AI incidents.

- **Microsoft's** Environmental, Social, and Public Policy Committee of the Board of Directors provides oversight and guidance on Responsible AI policies and programs. It also has a Responsible AI Council made up of senior leaders across the business to discuss high-level challenges and drive forward Responsible AI goals. To embed Responsible AI work across the company, Microsoft developed a network of employee Responsible AI Champions across the organization who are charged with implementing Microsoft's Responsible AI commitments. These efforts are coordinated by a horizontal Responsible AI team that develops policies and practices to uphold Microsoft's AI principles, defines roles and responsibilities, establishes governance systems, and leads "sensitive use" reviews.

BSR

BSR™ is an organization of sustainable business experts that works with its global network of the world's leading companies to build a just and sustainable world. With offices in Asia, Europe, and North America, BSR™ provides insight, advice, and collaborative initiatives to help you see a changing world more clearly, create long-term business value, and scale impact.

www.bsr.org