AI & Partners

Amsterdam - London - Singapore

EU AI Act

UNESCO Recommendations on the Ethics of Artificial Intelligence

A Mapping Exercise

May 2025

AI & Partners

Sean Musch, AI & Partners

Michael Borrelli, AI & Partners

Charles Kerrigan, CMS UK

Thomas Hoon, Nexus Alliance

Markus Krebsz, Contributor (Critical Review)

Principles

VS

AI & Partners

Amsterdam - London - Singapore

AI & Partners defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots professional services, regulatory interventions, and participating in industry groups such as AI Commons, we fight for fundamental rights in the artificial intelligence age.

This report was prepared by Sean Donald John Musch and Michael Charles Borrelli. For more information visit https://www.ai-and-partners.com/.

Contact: Michael Charles Borrelli | Director | m.borrelli@ai-and-partners.com.

This report is an AI & Partners publication.

Contents

Introduction

As artificial intelligence (AI) becomes increasingly embedded in society, ensuring its ethical and responsible development is paramount. The UNESCO AI Ethics Principles provide a comprehensive framework for governing AI in a way that respects human rights, promotes fairness, and fosters sustainability. These principles emphasize key values such as transparency, accountability, inclusivity, and environmental responsibility, ensuring that AI serves as a force for social good rather than exacerbating inequalities or causing harm.

The EU AI Act, as one of the world's first comprehensive AI regulations, incorporates many of UNESCO's ethical priorities into legal obligations, ensuring that AI systems are safe, fair, and explainable. In aligning ethical guidelines with enforceable rules, the Act strengthens AI governance, public trust, and compliance across industries. Together, UNESCO's principles and the EU AI Act shape the future of AI policy, balancing innovation with fundamental rights and democratic values.

This report explores the significance of UNESCO's AI Ethics Principles, their relationship with the EU AI Act, and the practical steps businesses and policymakers must take to ensure AI remains inclusive, transparent, and accountable. We believe that integrating ethical considerations into AI governance means stakeholders can create AI systems that are trustworthy, human-centered, and aligned with global sustainability goals.

Best regards,

Sean Musch

Founder/CEO

AI & Partners

Key questions being asked about UNESCO's Recommendation on the Ethics of Artificial Intelligence

1. What are the UNESCO AI Ethics Principles?

The UNESCO AI Ethics Principles establish a human-centered, rights-based approach to AI governance. They emphasize transparency, accountability, fairness, inclusivity, privacy, and sustainability to ensure AI benefits humanity. These principles promote ethical AI development, protect human dignity, and minimize harm by addressing risks like bias, discrimination, and environmental impact. UNESCO's framework serves as a global standard for responsible AI, guiding governments, businesses, and researchers in aligning AI innovation with fundamental rights and ethical values worldwide.

2. How do the UNESCO AI Ethics Principles relate to the EU AI Act?

The EU AI Act and UNESCO AI Principles share a commitment to ethical, transparent, and accountable AI development. The EU AI Act legally enforces many of UNESCO's ethical priorities, including risk-based AI regulation, human oversight, and bias mitigation. While UNESCO's principles serve as guidelines for responsible AI, the EU AI Act provides a legally binding framework that operationalizes these values, ensuring AI systems are safe, fair, and compliant with human rights laws in Europe and beyond.

3. Why are the UNESCO AI Ethics Principles important?

These principles ensure AI is developed and deployed responsibly, protecting human rights, social equity, and environmental sustainability. Without ethical guidelines, AI risks exacerbating inequalities, violating privacy, and reinforcing discrimination. UNESCO's principles provide a global ethical foundation for AI governance, fostering trust, innovation, and accountability. By prioritizing fairness, transparency, and human oversight, they help prevent harm, promote digital inclusion, and ensure AI serves humanity rather than controlling it.

4. How do UNESCO's AI principles address bias and discrimination?

UNESCO's framework mandates that AI systems must be fair, inclusive, and free from discrimination. Developers must use diverse datasets, conduct bias audits, and ensure AI decisions are transparent. The principles also stress that AI should not reinforce societal inequalities, calling for policies that promote accessibility and representation. Ensuring fairness in AI helps prevent discriminatory outcomes in areas like hiring, law enforcement, and healthcare, making AI more ethical and just.

5. What role does transparency play in UNESCO's AI Ethics Principles?

Transparency ensures that AI decisions, processes, and data sources are understandable and open to public scrutiny. UNESCO's principles require AI developers to provide explainability measures, allowing users to challenge AI-driven decisions. Transparency fosters trust in AI, enabling ethical oversight and preventing harmful, biased, or misleading AI applications. Without transparency, AI could undermine democracy, limit accountability, and create opaque decision-making structures that exclude public input.

6. How do the UNESCO AI Principles promote sustainability?

UNESCO's framework highlights the environmental impact of AI, calling for energy-efficient models, sustainable data practices, and AI-driven climate solutions. AI systems consume vast computational power, increasing carbon emissions and resource depletion. Ethical AI must support sustainability goals by reducing its environmental footprint and enhancing conservation efforts. UNESCO advocates for AI applications that promote sustainable development, protect ecosystems, and mitigate climate change risks, ensuring AI contributes to a greener future.

7. How do the UNESCO AI Principles ensure accountability?

UNESCO mandates that AI decisions must be traceable to human actors, ensuring clear responsibility for AI-related outcomes. Governments and businesses must establish auditing, compliance, and redress mechanisms to prevent AI misuse. Whistleblower protections, public oversight, and ethical review boards are encouraged to hold AI actors accountable. Without accountability, AI could operate unchecked, causing harm without consequences. Ensuring responsibility helps build trust, prevent abuse, and align AI with societal values.

8. What does UNESCO say about AI and human rights?

UNESCO's AI framework is rooted in human rights, emphasizing privacy, autonomy, dignity, and freedom of expression. AI should enhance, not restrict, human rights, ensuring people retain control over their digital interactions. UNESCO calls for legal safeguards to prevent AI-driven mass surveillance, data exploitation, or manipulative algorithms that infringe on freedoms. Upholding human rights in AI governance ensures that technology remains a tool for empowerment rather than oppression.

9. How does the UNESCO framework promote inclusivity?

UNESCO calls for equitable AI access, participation, and representation, ensuring that AI serves all communities, including marginalized and underrepresented groups. This includes multilingual AI systems, culturally diverse datasets, and AI education initiatives. Governments and companies must address digital divides, ensuring AI benefits rural, low-income, and developing regions. Inclusive AI development ensures that technology promotes social justice and economic equality, rather than deepening existing disparities.

10. How can businesses implement UNESCO's AI Ethics Principles?

Businesses can align with UNESCO's AI Ethics Principles by embedding transparency, fairness, and sustainability into AI systems. This includes conducting AI impact assessments, reducing biases, ensuring human oversight, and prioritizing ethical data practices. Companies should also establish AI governance frameworks, engage in multi-stakeholder collaboration, and provide public disclosures on AI usage. By committing to ethical AI, businesses enhance trust, ensure regulatory compliance, and drive responsible innovation that benefits society as a whole.

Understanding UNESCO's Recommendation on the Ethics of Artificial Intelligence

Value III.1.1

Respect, protection and promotion of human rights and fundamental freedoms and human dignity

Art.1
(Subject Matter)

What does it mean?

Respecting, protecting, and promoting human rights, fundamental freedoms, and human dignity in AI development means ensuring that AI systems do not harm or discriminate against individuals or communities. AI must recognize the intrinsic worth of every person, regardless of race, gender, age, or social status. Throughout their lifecycle, AI systems should enhance human well-being, avoiding subordination or objectification of individuals. This principle emphasizes that AI should align with international human rights laws, ensuring that technology fosters equality, autonomy, and respect for diverse cultural and personal values.

Why is it needed?

AI has the potential to reinforce biases, perpetuate inequality, or infringe on human rights if not properly regulated. Without safeguards, AI could be used for mass surveillance, discrimination in hiring, biased law enforcement, or economic exclusion. Ensuring that AI respects human dignity and freedoms protects vulnerable populations, such as children, the elderly, and individuals with disabilities. It also builds public trust in AI, fostering innovation that benefits society rather than exacerbating social divisions. A rights-based approach ensures AI remains a tool for empowerment rather than oppression.

How can it be applied?

To apply this principle, organizations and governments must integrate human rights frameworks into AI policies, design, and deployment. AI developers should conduct impact assessments to identify potential risks related to discrimination, privacy violations, or harmful biases. Ethical AI governance should include human oversight, ensuring AI-driven decisions align with fairness and dignity. Businesses and policymakers should engage civil society, legal experts, and marginalized communities to ensure AI benefits everyone.

Value III.1.2

Environment and ecosystem flourishing

Art.95
(Codes of conduct for voluntary application of specific requirements)

What does it mean?

The principle of environment and ecosystem flourishing emphasizes that AI development should not come at the cost of environmental harm. AI systems should be designed, deployed, and maintained in ways that support sustainability, protect ecosystems, and reduce environmental degradation. This includes minimizing carbon footprints, preventing excessive resource consumption, and ensuring AI applications contribute positively to ecological well-being. Since a healthy environment is essential for all life, AI should be leveraged to enhance sustainability efforts, rather than accelerate climate change or exploit natural resources irresponsibly.

Why is it needed?

AI technologies require significant computational power, energy consumption, and resource extraction, contributing to carbon emissions and environmental stress. Without sustainability-focused governance, AI could worsen climate change, increase electronic waste, and lead to biodiversity loss. The principle is necessary to ensure AI development aligns with global climate goals and sustainability initiatives. As AI becomes more embedded in industries like energy, and manufacturing, ensuring it operates in an eco-conscious manner is crucial for long-term planetary health, resource conservation, and a sustainable future for humanity.

How can it be applied?

AI developers and organizations should adopt energy-efficient models, optimize algorithms, and use renewable energy sources to power AI systems. Companies should conduct environmental impact assessments for AI projects, ensuring compliance with international sustainability laws and best practices. Governments and industries must promote AI-driven solutions for climate monitoring, conservation, and sustainable resource management. Transparency in AI's carbon footprint and environmental impact reporting should be a standard practice.

Value III.1.3

Ensuring diversity and inclusiveness

Art.95
(Codes of conduct for voluntary application of specific requirements)

What does it mean?

Ensuring diversity and inclusiveness in AI means that AI systems should be developed and deployed in a way that represents and benefits all individuals and communities, regardless of race, gender, age, socioeconomic status, or disability. AI should respect different cultural, social, and personal identities, allowing people to engage with it freely and without discrimination. Additionally, AI governance should promote participation from diverse groups, ensuring that all voices—especially those from underrepresented or marginalized communities—are considered in the design, implementation, and oversight of AI.

Why is it needed?

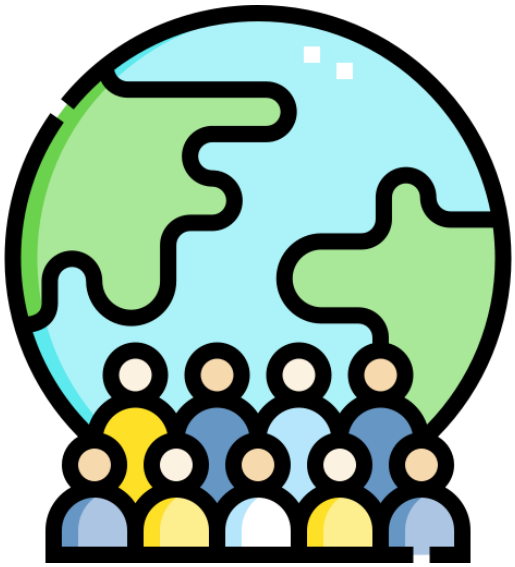
Without intentional inclusivity, AI can reinforce biases, exclude certain populations, and deepen digital divides. AI systems trained on non-representative datasets may generate discriminatory outcomes in areas such as hiring, healthcare, and law enforcement. Moreover, communities in low- and middle-income countries (LMICs) or regions with limited digital infrastructure may face barriers to AI access and participation. Ensuring diversity in AI development teams, datasets, and regulatory frameworks helps prevent bias, foster fairness, and ensure AI benefits society as a whole, rather than privileging certain groups while disadvantaging others.

How can it be applied?

To implement this principle, businesses and policymakers should prioritize diverse representation in AI design teams and require bias audits in AI systems. AI tools should be tested for fairness and inclusivity, ensuring they work equitably across different populations. Governments and international organizations should support AI capacity-building initiatives in underrepresented regions, providing access to education, digital infrastructure, and legal protections.

Value III.1.4

Living in peaceful, just and interconnected societies



Art.1
(Subject Matter)

What does it mean?

The principle of living in peaceful, just, and interconnected societies emphasizes AI's role in fostering harmony, equity, and global cooperation. AI should enhance human connections, social justice, and environmental well-being rather than creating divisions or perpetuating conflict. This principle highlights that AI systems should promote inclusivity, fairness, and mutual respect, ensuring that technological advancements support collective well-being. By recognizing the interdependence of all people and ecosystems, AI should be designed and deployed to strengthen social bonds, encourage solidarity, and contribute to sustainable development.

Why is it needed?

AI has the potential to shape societies, influencing areas such as governance, media, and social interactions. If misused, AI could amplify discrimination, misinformation, and polarization, leading to social unrest and injustice. The principle is crucial in preventing AI systems from undermining democratic values, violating human rights, or fostering division between individuals, communities, or nations. AI must be a tool for peace and cooperation, ensuring that its applications contribute to justice, inclusion, and the ethical treatment of all beings rather than creating conflict, exclusion, or environmental harm.

How can it be applied?

AI developers and policymakers should embed ethical and human rights-based safeguards into AI systems, ensuring they do not reinforce social inequalities or discriminatory practices. AI should be used to combat misinformation, enhance conflict resolution, and promote equitable access to resources. Organizations should establish diverse and inclusive AI development teams to ensure technology reflects a broad spectrum of perspectives.

Principle III.2.1

Proportionality and Do No Harm

Art.6
(Classification rules for high-risk AI systems)

What does it mean?

The principle of Proportionality and Do No Harm ensures that AI systems are used ethically, responsibly, and within justified limits. AI should not be deployed in ways that exceed what is necessary to achieve a legitimate goal, and its impact must be appropriate to the context. This principle acknowledges that AI does not automatically guarantee positive societal or environmental outcomes. Therefore, safeguards such as risk assessments, ethical evaluations, and human oversight must be in place to prevent harm, especially in high-stakes applications where AI could threaten human rights or well-being.

Why is it needed?

Unchecked AI deployment can lead to disproportionate or unjust consequences, such as mass surveillance, social scoring, discrimination, and environmental damage. AI systems, when used inappropriately, may violate privacy rights, personal freedoms, and environmental sustainability. This principle is essential to ensure AI remains a tool for benefit rather than harm, particularly in sensitive areas like healthcare, law enforcement, and governance. It also reinforces the necessity for final human oversight in irreversible or life-altering decisions, preventing AI from undermining fundamental freedoms or ethical standards.

How can it be applied?

Businesses and governments must implement strict risk assessments and ethical evaluations before deploying AI. AI systems should be proportionate to their intended purpose, ensuring they do not exceed necessity or infringe upon human rights. Organizations should adopt scientifically validated AI models that are contextually appropriate and uphold ethical, legal, and democratic values. Additionally, human oversight mechanisms should be mandated in critical decision-making processes, especially those involving life-and-death scenarios.

Principle III.2.2

Safety and security

Art.1
(Subject Matter)

What does it mean?

The principle of Safety and Security ensures that AI systems are designed, deployed, and maintained in ways that prevent harm and mitigate vulnerabilities. AI should not pose safety risks to individuals, societies, or the environment, nor should it be susceptible to cyberattacks, data breaches, or malicious exploitation. This principle emphasizes the need for strong risk management, secure infrastructure, and privacy-protective data frameworks that enable reliable AI models. By prioritizing safety and resilience, AI can function ethically and effectively, minimizing risks throughout its lifecycle.

Why is it needed?

AI systems that lack adequate security measures can be exploited for cyberattacks, misinformation, fraud, and other malicious activities. Additionally, AI models trained on poor-quality or biased data can lead to unintended safety risks, such as faulty medical diagnoses, biased legal decisions, or compromised critical infrastructure. Without a commitment to safety and security, AI could become a source of instability rather than progress. This principle ensures AI remains a trustworthy tool, fostering public confidence, regulatory compliance, and sustainable technological advancement.

s

How can it be applied?

To implement safety and security in AI, organizations should adopt robust risk assessment frameworks, conduct regular security audits, and ensure human oversight in high-risk applications. AI developers must prioritize privacy-protective data access models, using high-quality and diverse datasets to improve AI reliability. Governments and regulatory bodies should enforce AI safety standards, ensuring that AI systems meet ethical, technical, and legal requirements.

Principle III.2.3

Fairness and non-discrimination

Art.10
(Data and data governance)

What does it mean?

The principle of Fairness and Non-Discrimination ensures that AI systems promote social justice, equity, and inclusion, avoiding biases that could disadvantage marginalized communities. AI should be accessible to all, regardless of age, gender, ethnicity, language, disability, or socioeconomic status. It calls for inclusive AI development and deployment, ensuring that diverse perspectives are considered in AI design. Additionally, fairness means that AI systems should not reinforce discrimination but instead contribute to reducing digital divides and ensuring that AI benefits are shared globally, including underrepresented regions.

Why is it needed?

AI systems, if not designed carefully, can perpetuate biases, widen inequalities, and reinforce existing social injustices. Discriminatory algorithms in hiring, healthcare, education, and law enforcement have already demonstrated how AI can unintentionally disadvantage certain groups. Moreover, access to AI technologies is uneven, with rural, economically disadvantaged, and developing communities often left behind. Addressing these issues ensures that AI does not exacerbate discrimination or deepen digital divides, but instead fosters equal opportunity, representation, and fairness in its applications worldwide.

How can it be applied?

To ensure fairness, AI actors must prioritize diverse datasets, conduct bias audits, and implement transparency measures throughout AI development. Governments should promote AI accessibility, ensuring equitable access to AI education, infrastructure, and opportunities. AI systems should be designed with multilingual, culturally diverse, and disability-inclusive considerations. Regulators must enforce accountability frameworks, requiring organizations to assess and mitigate bias in AI decision-making.

Principle III.2.4

Sustainability

Art.95
(Codes of conduct for voluntary application of specific requirements)

What does it mean?

The principle of Sustainability emphasizes that AI should contribute to long-term human, social, cultural, economic, and environmental well-being rather than creating harm. AI technologies must be designed and deployed in ways that align with global sustainability objectives, such as the United Nations Sustainable Development Goals (SDGs). This means continuously assessing AI's social and environmental impact, ensuring it supports economic growth, social equity, and ecological preservation. AI should be a tool for enhancing sustainability efforts, helping societies advance in an ethically responsible and resource-efficient manner.

Why is it needed?

AI can either drive sustainability or worsen existing inequalities and environmental challenges. Unregulated AI use can increase energy consumption, environmental degradation, and social disparities, particularly in countries with weaker digital infrastructure and fewer resources. Without ongoing assessments and governance, AI may exacerbate climate change, economic divides, and cultural erosion. Ensuring AI supports sustainability goals is crucial for maintaining global stability, economic resilience, and environmental conservation, making AI a force for progress rather than exploitation.

How can it be applied?

To apply this principle, AI developers and policymakers must integrate sustainability assessments into AI life cycles. AI systems should be energy-efficient, minimize waste, and support renewable energy solutions. Governments and organizations should promote AI-driven innovations that enhance climate monitoring, smart resource management, and sustainable economic practices. AI policies must ensure equitable AI access, reducing the digital divide and fostering global cooperation in AI advancements.

Principle III.2.5

Right to privacy and data protection

Art.10
(Data and data governance)

What does it mean?

The principle of Right to Privacy and Data Protection ensures that AI systems respect individuals' privacy, autonomy, and personal data rights. AI must handle data collection, processing, sharing, and storage in compliance with international privacy laws and ethical guidelines. AI actors must ensure that data is used responsibly, securely, and with informed consent, preventing misuse or exploitation. This principle highlights the need for robust governance mechanisms to safeguard personal data, ensuring AI does not infringe upon fundamental privacy rights. Privacy protections should extend throughout the entire AI lifecycle.

Why is it needed?

AI systems rely heavily on large-scale data processing, raising concerns about unauthorized data collection, surveillance, identity theft, and privacy breaches. Without strong privacy safeguards, AI could be used to track individuals, profile users unfairly, or manipulate decision-making. Data leaks or misuse could also lead to discrimination, financial loss, or psychological harm. Establishing clear data protection regulations ensures that AI operates ethically, maintains public trust, and aligns with legal frameworks such as GDPR and international privacy standards. Protecting privacy is essential for human dignity and autonomy in the digital age.

How can it be applied?

Organizations must implement privacy-by-design principles, ensuring AI systems minimize data collection, encrypt sensitive information, and prioritize user consent. Governments should enforce strong data protection laws and require privacy impact assessments (PIAs) to evaluate AI's ethical and societal risks. AI developers must incorporate transparent data governance frameworks, allowing users to control, correct, and delete their personal data.

Principle III.2.6

Human oversight and determination

Art.14
(Human oversight)

What does it mean?

The principle of Human Oversight and Determination ensures that AI systems remain accountable to human decision-makers throughout their lifecycle. AI should assist in decision-making but must not replace human judgment or responsibility, particularly in high-stakes situations such as healthcare, law enforcement, or justice systems. This principle also extends to public oversight, ensuring that AI deployment aligns with ethical and legal standards. Governments and organizations must ensure clear attribution of responsibility, meaning AI-related actions can always be traced back to a physical person or legal entity that is accountable.

Why is it needed?

AI systems can process vast amounts of data and make autonomous recommendations, but they lack moral reasoning, ethical judgment, and social accountability. Without human oversight, AI could make biased or harmful decisions, especially in critical areas such as criminal sentencing, hiring, and healthcare. The risk of unintended AI errors or unethical use makes it essential to retain human control, particularly in life-and-death situations. Ensuring human oversight prevents AI misuse, strengthens accountability, and builds trust in AI systems by guaranteeing that humans remain the final decision-makers.

How can it be applied?

Governments and businesses should establish clear accountability structures, ensuring that AI decisions can be reviewed, contested, and corrected by humans. AI systems must include fail-safe mechanisms, allowing human intervention when necessary. Public oversight bodies should be involved in monitoring AI's societal impact, ensuring transparency and ethical compliance. Organizations should also implement AI auditing processes.

Principle III.2.7

Art.13
(Transparency and provision of information to deployers)

What does it mean?

The principle of Transparency and Explainability ensures that AI systems operate in a way that is understandable, accountable, and open to public scrutiny. Transparency means that users, regulators, and stakeholders should know when AI is making decisions, while explainability ensures that people can understand how and why AI arrived at a particular outcome. These principles are crucial for protecting human rights, ensuring fairness, and allowing individuals to challenge AI-driven decisions. AI actors must provide clear, context-appropriate explanations, especially when AI impacts human rights, safety, or fundamental freedoms.

Why is it needed?

Without transparency, AI decisions can become opaque, unchallengeable, and unaccountable, leading to risks such as biased outcomes, wrongful discrimination, and violations of fundamental rights. A lack of explainability can erode trust in AI systems, making it difficult for individuals to understand or contest decisions that affect their lives. Transparency is also essential for effective governance, reducing corruption, and ensuring ethical AI deployment. By making AI processes more understandable and interpretable, societies can strengthen democratic governance, safeguard fairness, and foster greater public confidence in AI technologies.

How can it be applied?

AI developers and organizations should disclose when AI is used in decision-making and ensure individuals have access to explanations about AI-driven outcomes. Governments should enforce transparency regulations, requiring AI actors to document how their algorithms function, what data is used, and what safeguards are in place. Companies should develop explainable AI models that provide clear reasoning for decisions, particularly in high-risk applications like hiring, and finance.

Principle III.2.8

Responsibility and accountability

Art.17
(Quality management system)

What does it mean?

The principle of Responsibility and Accountability ensures that AI actors—whether governments, businesses, or developers—are held ethically and legally responsible for the AI systems they create, deploy, and manage. AI decisions should always be attributable to humans or legal entities, ensuring that AI cannot act without oversight or consequence. This principle mandates that AI actors respect human rights, protect the environment, and comply with national and international laws. It also emphasizes the need for clear governance structures, and mechanisms to address AI-related harm throughout the entire AI lifecycle.

Why is it needed?

Without clear responsibility and accountability, AI systems can cause harm without consequences, leading to bias, discrimination, privacy violations, and environmental damage. If no human or organization is held accountable for AI-driven errors or abuses, victims may have no means of redress. Moreover, unchecked AI systems could undermine democracy, legal systems, and societal trust. This principle ensures that AI actors take responsibility for the impact of their technologies, fostering transparency, oversight, and ethical compliance while mitigating risks that could harm individuals, societies, or ecosystems.

How can it be applied?

To implement this principle, AI developers and policymakers must establish clear governance structures, ensuring that every AI-related decision can be audited and traced to a responsible entity. Governments should enforce accountability laws, requiring risk assessments, audits, and due diligence processes before AI deployment. Organizations should integrate whistleblower protections and ethical review boards to oversee AI's impact on human rights, and environmental sustainability.

Principle III.2.9

Awareness and literacy

Art.4
(AI Literacy)

What does it mean?

The principle of Awareness and Literacy ensures that individuals and communities have the knowledge and skills to understand, use, and critically evaluate AI technologies. AI should not be a mystery to the public, and people should be empowered to make informed decisions about AI's role in their lives. This requires education, digital literacy programs, and public engagement efforts that explain AI's benefits, risks, and ethical considerations. AI awareness should also account for linguistic, social, and cultural diversity, ensuring that learning materials are accessible and inclusive to all members of society.

Why is it needed?

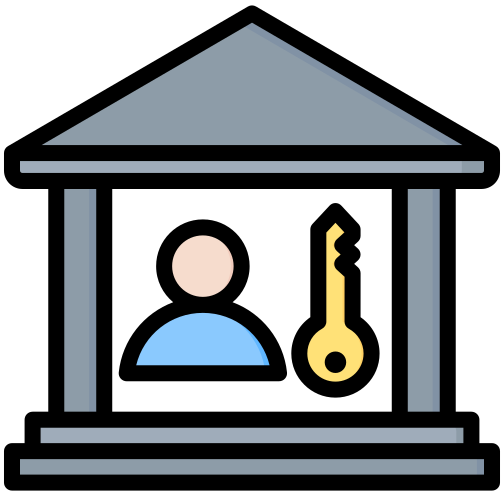
AI influences many aspects of daily life, from healthcare and employment to social media and public services. Without adequate education and awareness, individuals may be unaware of how AI systems make decisions, how their data is used, or how to challenge unfair AI-driven outcomes. Additionally, a lack of AI literacy can make people vulnerable to misinformation, algorithmic bias, and undue influence. Ensuring widespread AI education promotes ethical AI use, enhances public participation in AI policy discussions, and strengthens democracy by preventing AI-related manipulation or discrimination.

How can it be applied?

Governments, educational institutions, and technology companies should integrate AI education into school curricula, public awareness campaigns, and workforce training programs. Civic engagement initiatives should provide accessible resources on AI ethics, digital rights, and data privacy. The media should play a role in disseminating accurate AI-related information, helping the public navigate misinformation and bias. AI training should be designed to reach diverse communities.

Principle III.2.10

Multi-stakeholder and adaptive governance and collaboration



Art.95
(Codes of conduct for voluntary application of specific requirements)

What does it mean?

The principle of Multi-Stakeholder and Adaptive Governance and Collaboration emphasizes that AI governance should be inclusive, flexible, and responsive to evolving technologies and societal needs. Effective AI governance requires collaboration between governments, businesses, researchers, civil society, and marginalized communities to ensure AI benefits everyone equitably. It also recognizes the importance of respecting national sovereignty and international law when regulating data, ensuring privacy and human rights protections. This approach fosters interoperability, transparency, and ethical AI development.

Why is it needed?

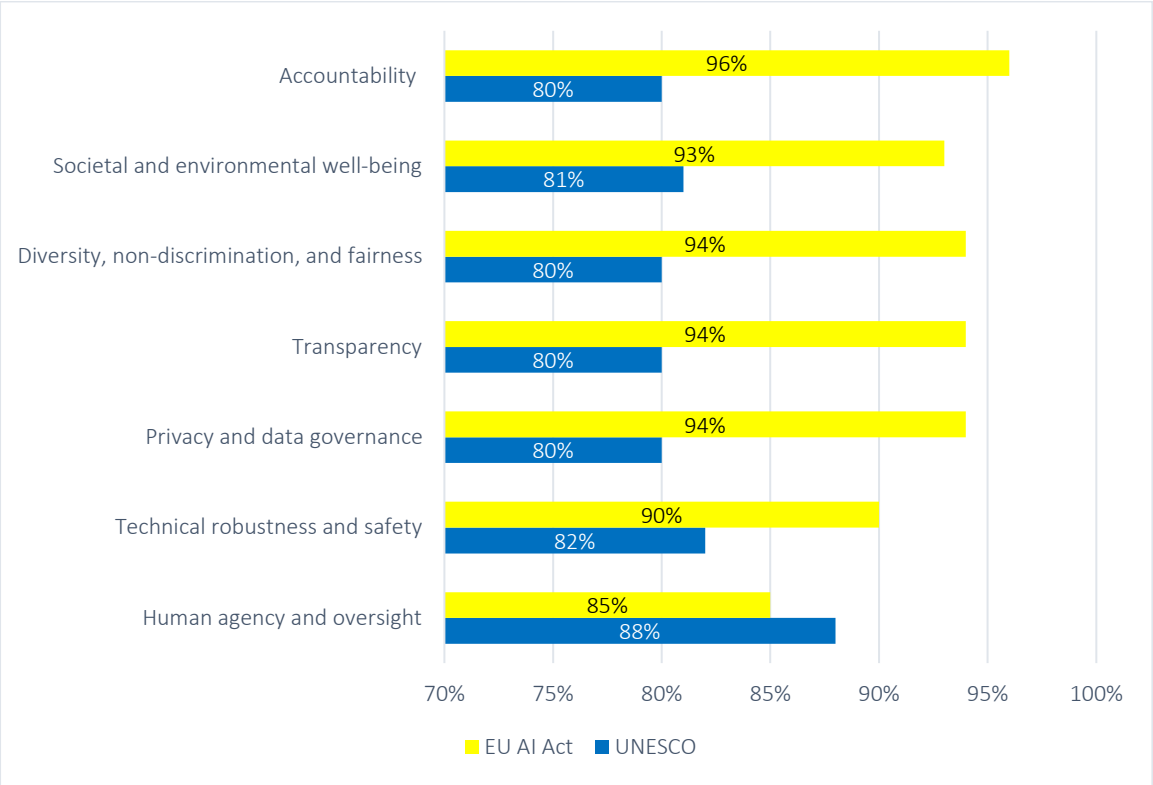
AI impacts every aspect of society, from economics and healthcare to human rights and environmental sustainability. Without broad stakeholder engagement, AI policies risk being biased, exclusionary, or ineffective. Marginalized groups and developing nations could be left behind in AI adoption and regulation, deepening global inequalities. Ensuring multi-stakeholder governance helps mitigate risks, foster innovation, and align AI with democratic values. It also allows for ongoing updates to AI policies, ensuring governance mechanisms remain adaptive to shifts in AI capabilities, societal needs, and ethical considerations.

How can it be applied?

Governments should develop AI policies through open consultations with industry leaders, researchers, civil society, and underrepresented communities. International cooperation should ensure interoperability and ethical standards across borders, while respecting national sovereignty over data governance. Open AI standards, data-sharing agreements, and ethical review boards should be implemented to facilitate collaboration and accountability. Efforts should also be made to include Indigenous communities and marginalized groups in AI policy discussions.

Measuring UNESCO's Recommendation on the Ethics of Artificial Intelligence and EU AI Act against High-Level Expert Group on AI's Ethics Guidelines for Trustworthy Artificial Intelligence

UNESCO's Recommendation on the Ethics of Artificial Intelligence and EU AI Act mapped against High-Level Expert Group on AI's ethics guidelines for trustworthy AI



Methodology

This report presents a comparative analysis of the **UNESCO Recommendation on the Ethics of Artificial Intelligence** and the **EU Artificial Intelligence Act**, measured against the **High-Level Expert Group (HLEG) on AI's Ethics Guidelines for Trustworthy AI**. The goal is to assess how each framework aligns with the HLEG's seven core principles, which are widely recognized as a benchmark for ethical AI governance in Europe and globally.

Overall Approach

The comparative matrix evaluates **two frameworks**—UNESCO's ethical recommendations and the EU AI Act's legal provisions—against the **HLEG's ethical principles**. This triangulated comparison enables a cross-sectional view of how ethics are operationalized in both voluntary (UNESCO) and binding (EU AI Act) governance mechanisms.

Each HLEG principle (e.g., transparency, human agency, technical robustness) serves as an anchor point. Both UNESCO and the EU AI Act are evaluated independently against each HLEG principle, producing two separate alignment scores per principle.

Comparison Framework

The assessment is guided by three clearly defined criteria, designed to reduce ambiguity and ensure a consistent basis for evaluation:

- **Terminological Consistency**: Does the language used in the framework directly correspond to the language and intent of the HLEG principle? This includes terms such as "human oversight," "non-discrimination," or "accountability."
  E.g., "human-in-the-loop" is explicitly mentioned in both UNESCO and the EU AI Act in the context of oversight.

- **Functional Equivalence**: Do the mechanisms or practices described in the framework fulfill the purpose of the HLEG principle, even if different terminology is used? This criterion accounts for conceptual alignment even when terms differ.
  E.g., if a framework mandates explainability of AI decisions, it is considered functionally equivalent to the HLEG's transparency requirement.
- **Regulatory Strength**: What is the normative force of the provision?
  - For the **EU AI Act**, this means whether the requirement is enforceable under EU law (e.g., mandatory for high-risk AI systems).
  - For **UNESCO**, this reflects the breadth and specificity of the recommendation and the extent to which implementation is encouraged or supported.

Each criterion was used to independently evaluate both frameworks per HLEG principle.

Scoring Mechanism

Each framework receives a **percentage alignment score (0–100%) per HLEG principle**, based on the degree to which it satisfies the three criteria:

- **100% (High alignment)**: The framework fully reflects the HLEG principle, using consistent terminology, delivering comparable functionality, and providing either legal enforceability (EU AI Act) or strong normative clarity and scope (UNESCO).
- **50–99% (Moderate alignment)**: The principle is addressed, but there are limitations—e.g., partial coverage, weak regulatory mechanisms, or general rather than specific language.
- **1–49% (Low alignment)**: The principle is weakly addressed or acknowledged only superficially without clear mechanisms or enforceable provisions.
- **0% (No alignment)**: The principle is not addressed in any meaningful way.

Descriptions in the matrix (see pages 24–26) elaborate how each criterion is satisfied or falls short for each pairing (HLEG vs UNESCO / HLEG vs EU AI Act), providing rationale behind the assigned percentage.

Calculation Method

Scoring is **qualitative and subjective**, based on a structured **suitability assessment** rather than formulaic quantification. Analysts exercised expert judgment to determine the degree of alignment, drawing on:

- the exact language and scope of each principle;
- context and intent of the provision;
- documented obligations or guidelines;
- and real-world implications for implementation or enforcement.

While not purely quantitative, this method ensures a reasoned, evidence-based comparison and allows for critical reflection on the strengths and gaps in both frameworks when benchmarked against the HLEG's ethical foundation.

HLEG		UNESCO		EU AI Act	
Principle	**Description**	**Alignment**	**Rationale**	**Alignment**	**Rationale**
Human agency and oversight	AI systems should empower human beings, allowing them to make informed decisions and fostering their fundamental rights. At the same time, proper oversight mechanisms need to be ensured, which can be achieved through human-in-the-loop, human-on-the-loop, and human-in-command approaches.	88%	The UNESCO Ethics Recommendation (88%) promotes human oversight in AI through ethical guidelines, emphasizing human-in-the-loop, transparency, and accountability to safeguard autonomy. It advocates for AI systems that empower human decision-making while ensuring ethical responsibility. However, it lacks legal enforcement and relies on voluntary adoption, making implementation dependent on governments and organizations.	85%	The EU AI Act (85%) legally mandates human oversight, transparency, and compliance in high-risk AI applications, ensuring AI remains a tool for human empowerment. While it enforces stricter regulations, it allows regulated AI autonomy in some cases. This framework balances practical compliance with human control, ensuring responsible AI governance.
Technical robustness and safety	AI systems need to be resilient and secure. They need to be safe, ensuring a fall back plan in case something goes wrong, as well as being accurate, reliable and reproducible. That is the only way to ensure that also unintentional harm can be minimized and prevented.	82%	The UNESCO Ethics Recommendation (82%) emphasizes AI resilience, security, and harm prevention through ethical principles, advocating for fail-safe mechanisms, reliability, and accuracy to minimize unintended consequences. It promotes AI systems that are trustworthy and reproducible, ensuring they do not compromise safety. However, it lacks binding regulatory measures, leaving enforcement to the discretion of stakeholders.	90%	The EU AI Act (90%) enforces strict safety, robustness, and risk mitigation requirements, particularly for high-risk AI systems. It mandates reliability testing, fallback mechanisms, and security protocols to prevent failures. By combining technical audits and legal obligations, it ensures AI systems are safe, resilient, and accountable.
Privacy and data governance	Besides ensuring full respect for privacy and data protection, adequate data governance mechanisms must also be ensured, taking into account the quality and integrity of the data, and ensuring legitimised access to data.	80%	The UNESCO Ethics Recommendation (80%) advocates for privacy protection, data integrity, and ethical data governance, emphasizing the need for secure data handling and legitimate access. It highlights the importance of transparency in data usage and ensuring AI respects fundamental privacy rights. However, it lacks legal enforcement mechanisms, relying on voluntary compliance by organizations and governments.	94%	The EU AI Act (94%) enforces strict data protection regulations, aligning with GDPR to ensure secure, transparent, and accountable AI data practices. It mandates high data quality standards, access controls, and auditability, ensuring robust governance and minimizing risks of misuse, particularly in high-risk AI applications.
Transparency	The data, system and AI business models should be transparent. Traceability mechanisms can help achieving this. Moreover, AI systems and their decisions should be explained in a manner adapted to the	80%	The UNESCO Ethics Recommendation (80%) promotes AI transparency through ethical principles, advocating for traceability, explainability, and user awareness. It emphasizes that AI	94%	The EU AI Act (94%) mandates strict transparency requirements, ensuring AI systems provide clear explanations, traceable decision-making, and disclosure of AI

	stakeholder concerned. Humans need to be aware that they are interacting with an AI system, and must be informed of the system's capabilities and limitations.		decisions should be understandable and accountable, ensuring stakeholders are informed of AI capabilities and limitations. However, it lacks binding enforcement, relying on voluntary adoption by organizations and governments.		interactions. It enforces explainability standards, business model transparency, and user notifications, particularly for high-risk AI applications, ensuring trustworthy and legally compliant AI deployment across sectors.
Diversity, non-discrimination, and fairness	Unfair bias must be avoided, as it could have multiple negative implications, from the marginalization of vulnerable groups, to the exacerbation of prejudice and discrimination. Fostering diversity, AI systems should be accessible to all, regardless of any disability, and involve relevant stakeholders throughout their entire life circle.	80%	The UNESCO Ethics Recommendation (80%) emphasizes fairness, inclusivity, and bias mitigation, advocating for accessible AI systems that prevent discrimination and marginalization of vulnerable groups. It promotes stakeholder involvement in AI development to ensure diverse perspectives are considered. However, it lacks enforcement mechanisms, relying on voluntary adoption to uphold these principles.	94%	The EU AI Act (94%) enforces strict anti-discrimination measures, requiring AI systems to undergo bias testing, fairness assessments, and transparency audits. It mandates accessibility for all users, including those with disabilities, ensuring AI systems are fair, inclusive, and free from harmful bias, particularly in high-risk applications.
Societal and environmental well-being	AI systems should benefit all human beings, including future generations. It must hence be ensured that they are sustainable and environmentally friendly. Moreover, they should take into account the environment, including other living beings, and their social and societal impact should be carefully considered.	81%	The UNESCO Ethics Recommendation (81%) promotes AI systems that benefit society and future generations, emphasizing sustainability, ethical AI development, and environmental responsibility. It encourages AI to be socially inclusive and environmentally friendly, considering its impact on ecosystems and living beings. However, it lacks binding regulations, relying on voluntary efforts to ensure AI aligns with these goals.	93%	The EU AI Act (93%) enforces legal requirements for sustainability and societal impact, mandating environmental assessments, ethical AI governance, and long-term risk evaluations. It ensures AI systems align with social responsibility standards, minimizing negative societal and ecological impacts, particularly in high-risk applications.
Accountability	Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes. Auditability, which enables the assessment of algorithms, data and design processes plays a key role therein, especially in critical applications. Moreover, adequate an accessible redress should be ensured.	80%	The UNESCO Ethics Recommendation (80%) advocates for clear responsibility in AI development and deployment, emphasizing auditability, ethical compliance, and governance. It highlights the importance of redress mechanisms and ensuring AI actors are accountable for outcomes. However, it lacks binding legal frameworks, relying on voluntary adoption to uphold these accountability principles.	96%	The EU AI Act (96%) enforces strict accountability measures, requiring AI providers to implement audit trails, risk assessments, and compliance reporting. It mandates clear liability frameworks and ensures accessible redress mechanisms for those affected by AI decisions, particularly in high-risk applications, strengthening trust and legal responsibility in AI systems.

Critical Perspectives on both UNESCO's Recommendation on the Ethics of Artificial Intelligence and EU AI Act

While both the UNESCO Recommendation on the Ethics of Artificial Intelligence and the EU AI Act represent landmark efforts in AI governance, they are not without limitations. A balanced evaluation must consider critical perspectives from industry, civil society, and academia, highlighting both structural gaps and areas of strength.

UNESCO Recommendation on the Ethics of AI

UNESCO's framework is widely appreciated for its inclusive, human-centered approach and emphasis on global equity, sustainability, and cultural diversity. It was developed through a participatory process involving over 150 countries, making it the most broadly supported global AI ethics framework to date.

However, critics argue that its non-binding nature limits its practical impact. As a set of ethical guidelines rather than enforceable rules, it relies heavily on voluntary adoption by states and organizations. This opens the door to selective interpretation, lack of accountability, and inconsistent application—particularly in jurisdictions with weak governance or limited institutional capacity. For example, while the principles of sustainability and inclusiveness are strongly articulated, there is no formal mechanism for monitoring compliance or measuring outcomes.

Academics also note the ambiguity of key terms—such as "do no harm" or "benefit all humanity"—which, while aspirational, offer limited operational guidance. As a result, organizations may struggle to translate these values into concrete governance protocols or product development standards. Constructively, UNESCO could consider issuing sector-specific implementation toolkits, model laws, or national reporting templates to support uptake and standardization. Establishing a voluntary benchmarking system (e.g., national self-assessments) could help measure progress without compromising the framework's flexibility.

EU AI Act

In contrast, the EU AI Act is a legally binding regulation that introduces enforceable requirements for AI development, deployment, and oversight, particularly for high-risk systems. Its tiered, risk-based approach has been praised for legal clarity and enforceability, positioning the EU as a global leader in responsible AI regulation. However, industry groups have raised concerns about regulatory overreach and the burden on innovation, particularly for small and medium-sized enterprises ("SMEs"). The compliance costs—estimated in some cases at hundreds of thousands of euros—may disproportionately affect early-stage AI firms and reduce experimentation.

Moreover, civil society actors have criticized the Act for lacking ambition in protecting fundamental rights. For example, several advocacy groups have called for a complete ban on remote biometric identification in public spaces—an issue the final text leaves to national discretion. There is also insufficient emphasis on intersectional fairness (e.g., systems that may discriminate based on combined factors like gender and race), and limited mechanisms for public redress when AI causes harm. A constructive improvement would involve expanding the mandate of the European AI Office to offer technical and legal support for SMEs and to monitor cross-border human rights impacts. Additionally, the Act could benefit from stronger provisions on algorithmic impact assessments and community engagement requirements for high-risk systems, ensuring that affected populations have a voice in governance.

In sum, UNESCO provides ethical breadth while the EU AI Act delivers legal depth. However, both would benefit from stronger mechanisms for implementation, stakeholder inclusion, and ongoing revision to remain responsive to AI's rapid evolution. Their alignment—and complementarity—should be seen not as static, but as part of an evolving ecosystem of global AI governance.

Sector Specific Guidance for EU AI Act Implementation

Sector	Common AI Use Cases	Risk Considerations	EU AI Act Classification	UNESCO Ethical Principles	Practical Recommendations
Healthcare	Diagnostic tools, predictive analytics, medical triage, robotic surgery	Bias in diagnosis, lack of explainability, data privacy, potential for harm	Typically **high-risk** under Annex III (e.g., medical devices using AI)	Human dignity, privacy, do no harm, inclusiveness, fairness	- Conduct algorithmic impact assessments for bias and error. - Ensure human-in-the-loop decision-making. - Obtain informed consent for AI-driven outcomes.
Finance	Credit scoring, fraud detection, automated investment advice, risk modelling	Algorithmic bias, lack of transparency, financial exclusion, data exploitation	Often **high-risk** (e.g., creditworthiness evaluation), others limited risk	Fairness, accountability, non-discrimination, human oversight	- Implement regular bias audits and fairness metrics. - Provide clear explanations for credit and loan decisions. - Strengthen data governance practices.
Education	Automated grading, adaptive learning systems, student behavioral monitoring	Reinforcement of biases, unequal access, surveillance, undermining of teacher roles	**High-risk** if impacting student performance or rights	Equity, diversity, awareness and literacy, multi-stakeholder governance	- Ensure systems accommodate diverse learning styles and needs. - Avoid over-reliance on AI for assessment. - Include educators in system design.
Public Sector	Welfare allocation, border control, surveillance, social scoring, administrative decision-making	Risk of systemic bias, lack of contestability, privacy erosion, democratic accountability	Multiple **high-risk applications**, especially in law enforcement and social services	Accountability, transparency, peaceful societies, human rights	- Introduce mandatory public consultation for AI in citizen-facing services. - Embed appeal mechanisms. - Prohibit AI-driven decisions without oversight.

Calls to action

Commit to Ethical and Transparent AI Development

Organizations must align AI development with UNESCO's AI Ethics Principles, ensuring fairness, accountability, and human rights protections. This includes clear documentation, explainability measures, and transparency in AI decision-making.

Strengthen AI Risk Management and Security

AI systems must be robust, safe, and secure throughout their lifecycle. Companies should adopt risk management frameworks that address bias mitigation, cybersecurity, and algorithmic accountability.

Promote Inclusive and Sustainable AI Adoption

AI should be developed and deployed in a way that promotes diversity, inclusion, and environmental sustainability. Organizations must ensure that AI technologies are accessible to all communities, including marginalized groups and developing nations.

Align AI Governance with Human Rights and Social Good

AI governance should be rooted in UNESCO's human rights-based approach, ensuring AI systems respect privacy, autonomy, and fundamental freedoms. Organizations must establish strong compliance mechanisms, work with international regulators.

Conclusion

The UNESCO AI Ethics Principles provide a vital framework for ensuring that AI development and deployment remain ethical, inclusive, and aligned with human rights. As AI continues to shape economies, societies, and governance structures, it is essential that stakeholders—including governments, businesses, and civil society—adhere to principles of transparency, accountability, fairness, and sustainability. Without strong ethical guidelines, AI risks perpetuating biases, deepening inequalities, and compromising fundamental freedoms.

The EU AI Act complements UNESCO's ethical framework by translating key principles into legally binding obligations, ensuring AI remains safe, fair, and explainable. The alignment of these two frameworks demonstrates a growing global commitment to responsible AI governance. However, effective implementation is crucial. AI actors must integrate ethical safeguards, conduct risk assessments, and establish clear oversight mechanisms to prevent AI-driven harm.

Public engagement, education, and regulatory adaptation will also be critical in ensuring AI systems remain aligned with evolving ethical, legal, and technological landscapes.

Moving forward, collaboration between policymakers, industry leaders, and researchers is essential to refine AI governance models that uphold human dignity, social justice, and environmental sustainability. Organizations that proactively embrace UNESCO's AI Ethics Principles will not only ensure compliance with emerging regulations but also foster public trust, innovation, and long-term societal benefits. By embedding ethics into AI from the outset, stakeholders can ensure AI remains a tool for empowerment rather than exploitation, shaping a future that is fair, inclusive, and sustainable for all.

About AI & Partners

AI & Partners

Amsterdam - London - Singapore

AI & Partners – 'AI That You Can Trust'

At AI & Partners, we're here to help you navigate the complexities of the EU AI Act, so you can focus on what matters—using AI to grow your business. We specialize in guiding companies through compliance with tailored solutions that fit your needs. Why us? Because we combine deep AI expertise with practical, actionable strategies to ensure you stay compliant and responsible, without losing sight of your goals. With our support, you get AI you can trust—safe, accountable, and aligned with the law.

To find out how we can help you, email contact@ai-and-partners.com or visit https://www.ai-and-partners.com.

Contacts

Sean Donald John Musch, CEO/Founder, s.musch@ai-and-partners.com

Michael Charles Borrelli, Director, m.borrelli@ai-and-partners.com

Authors

Sean Donald John Musch, CEO/Founder

Michael Charles Borrelli, Director

References

European Commission, (2019), 'Ethics guidelines for trustworthy AI', accessible at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (last accessed 15th March 2025)

European Parliament and The Council of the European Union, (2024), 2024/1689 Regulation (EU) 2024/1689 of the European Parliament and of The Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), accessible at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689 (last accessed 15th March 2025)

United Nations Educational, Scientific and Cultural Organization, (2022), 'Recommendation on the Ethics of Artificial Intelligence', accessible at: https://unesdoc.unesco.org/ark:/48223/pf0000381137 (last accessed 15th March 2025)

Important notice

This document has been prepared by AI & Partners B.V. for the sole purpose of enabling the parties to whom it is addressed to evaluate the capabilities of AI & Partners B.V. to supply the proposed services.

Other than as stated below, this document and its contents are confidential and prepared solely for your information, and may not be reproduced, redistributed or passed on to any other person in whole or in part. If this document contains details of an arrangement that could result in a tax or National Insurance saving, no such conditions of confidentiality apply to the details of that arrangement (for example, for the purpose of discussion with tax authorities). No other party is entitled to rely on this document for any purpose whatsoever and we accept no liability to any other party who is shown or obtains access to this document.

This document is not an offer and is not intended to be contractually binding. Should this proposal be acceptable to you, and following the conclusion of our internal acceptance procedures, we would be pleased to discuss terms and conditions with you prior to our appointment. Images used throughout the document have either been produced in-house or sourced from publicly available sources (see References for details).

AI & Partners B.V. is the Dutch headquarters of AI & Partners, a global professional services firm. Please see https://www.ai-and-partners.com/ to learn more about us.

Designed and produced by AI & Partners B.V