ETHICAL AI
MACEDONIA

PRINCIPLES FOR RESPONSIBLE ARTIFICIAL INTELLIGENCE

IN NORTH MACEDONIA

A voluntary awareness framework for responsible AI adoption

Prepared by Ethical AI Macedonia

First Edition · Version 1.0

© 2026 Ethical AI Macedonia. All rights reserved.

Contents

About Ethical AI Macedonia . . . 3

A Word from the Founders . . . 4

Introduction: Purpose and Context . . . 5

Definitions . . . 7

The Principles . . . 8

1. Human Dignity and Fundamental Rights . . . 9

2. Fairness, Equality and Inclusion . . . 10

3. Transparency, Explainability and Traceability . . . 11

4. Data Governance and Integrity . . . 12

5. Safety, Security, Robustness and Reliability . . . 13

6. Human Agency, Oversight and Competence . . . 14

7. Accountability, Governance and Remedy . . . 15

8. Participation, Local Context and Civic Legitimacy . . . 16

Acknowledgements . . . 17

About Ethical AI Macedonia

SUPPORTING AI ADOPTION THROUGH RESPONSIBLE USE

Ethical AI Macedonia brings businesses and experts together to discuss how AI can be used more widely and responsibly in North Macedonia. We focus on the decisions organizations already face: which tools to trust, what data to share, when to check an answer and how to preserve professional judgement and skills.

The Principles are our starting point. Through company research and discussion, we want to learn from real experience and improve guidance that reflects local conditions. Our particular focus is on smaller businesses and professional firms, which often have to make these decisions without dedicated technical or compliance teams.



Vladimir Boshnjakovski

Co-founder

Vladimir is a lawyer whose work focuses on personal data protection, including health data, alongside advice to software businesses and on digital services. He holds a Bachelor of Law and Master of Law from Ss. Cyril and Methodius University in Skopje, and an Executive MBA from Wirtschaftsuniversität Wien, focused on Data Science and Digital Transformation. His research examines explainability under the EU AI Act.



Angela Angjelovska

Co-founder

Angela is a lawyer with experience in online platforms, digital identity and civil society. She has held leadership roles in ELSA Macedonia and the National Youth Council of Macedonia, and holds undergraduate and master's degrees in law from Ss. Cyril and Methodius University in Skopje. She brings experience in civic engagement and organizational work to the initiative.

A Word from the Founders

AI is developing faster than the public discussion surrounding its use. We are inspired by what AI can make possible for businesses, professionals, and society. Its potential benefits include improved productivity, wider access to knowledge, and new opportunities for innovation and the development of new business models. At the same time, using AI without sufficient awareness of its limitations and risks may adversely affect fundamental rights, personal autonomy, professional skills, and an organization's credibility. Developing the capacity to understand and mitigate these risks will be a challenge on par with the technological effort required to build increasingly capable AI models. It was within this context that we created Ethical AI Macedonia, with the goal of raising awareness among businesses in North Macedonia of the importance of using AI responsibly and what responsible use means in practice.

Our initiative was not founded in opposition to technology. On the contrary, it is driven by the hope that greater AI usage responsibility will support broader and more confident AI adoption. We believe, however, that widespread and sustainable use depends on organizations understanding and addressing the risks AI may pose to privacy, autonomy, equality, reliability, security, and professional skills. Poorly informed or unmanaged use can cause avoidable harm, damage organizations, and provoke wider public concern, weakening consumer trust and even making businesses and institutions more reluctant to use or invest in AI. Risk-averse organizations, in particular, may be unwilling even to experiment without a clearer understanding of the relevant risks and how they can be managed, causing them to miss valuable opportunities and productivity gains.

These Principles are the first step on that journey. They draw on common principles reflected in established international and European frameworks, as well as national frameworks adopted in other jurisdictions. Their particular contribution is to place these widely recognized standards within the circumstances of North Macedonia. Accordingly, they have been drafted with particular attention to: i. the performance of AI in languages and dialectal varieties with more limited digital resources; ii. the levels of digitalization, data availability, access to technology, and AI literacy in North Macedonia; iii. the technical, financial, and organizational capacities of smaller businesses; and iv. the risk that systems developed for larger markets may not adequately reflect local laws, administrative practices, cultures, or social conditions.

We believe these Principles are timely. North Macedonia does not yet have a generally applicable set of ethical principles or guidelines on the responsible use of AI by businesses and other organizations, even though countries in the region operating under broadly comparable conditions have adopted similar guidance or standards. We nevertheless recognize the efforts of the Ministry of Digital Transformation, the Vezilka initiative, and the other actors active in this field. These Principles are intended to complement those efforts and contribute a practical business perspective to the national debate on how AI can be used responsibly within the circumstances of North Macedonia.

Finally, these Principles claim no regulatory authority and do not attempt to provide a final answer to every question. They are a practical starting point, not a finished product. We invite organizations, experts, institutions, civil-society organizations, and the wider community to test and criticize them so that future versions may be amended and strengthened through practical experience, constructive debate, and continuing technological and legal developments.

V. Boshnjakovski

Vladimir Boshnjakovski
Co-founder, Ethical AI Macedonia

A. Angjelovska

Angela Angjelovska
Co-founder, Ethical AI Macedonia

Introduction: Purpose and Context

Our mission. These Principles seek to support the confident, informed, and human-centred adoption of artificial intelligence by businesses and other organizations in North Macedonia. They provide a shared ethical reference for understanding what responsible AI use means in practice and for considering the benefits, limitations, and risks of AI before consequential decisions are made.

International foundations. The Principles are inspired in part by existing and well-established international and national documents. They draw on the common themes reflected in the UNESCO Recommendation on the Ethics of Artificial Intelligence, the OECD AI Principles, the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, the European Union AI Act, and practical risk-management frameworks of standard-setting organizations.

These instruments differ in their scope and legal character. Some establish binding obligations, while others provide ethical principles, policy commitments, or practical methods for managing AI-related risks. Nevertheless, they share a common understanding: AI should be developed and used in a manner that respects human dignity and fundamental rights, preserves meaningful human agency, addresses foreseeable risks, and maintains accountability for its use and consequences.

The Principles do not seek to replace or compete with these frameworks. Nor do they claim to introduce an entirely new ethical model. They build upon their widely recognized foundations and seek to express them in a form that is understandable, relevant, and usable by organizations operating in North Macedonia.

Regional context. Countries of different sizes and levels of technological development are adopting their own approaches to responsible AI. Serbia has adopted ethical guidelines for the development, application, and use of reliable and responsible AI. Albania has introduced a methodology and technical standards concerning the use of AI by public institutions. Slovenia has incorporated trustworthy and human-centred AI into its national policy.

These initiatives reflect the fact that responsible AI cannot be addressed solely through documents developed for the world's largest economies, markets, or technology companies. Each country must also consider how AI interacts with its own institutions, languages, economy, technological capacity, and social conditions.

The North Macedonian context. North Macedonia is not starting from zero. SMART/MK 2030 places AI within the country's wider digital-development agenda. The Vezilka initiative is intended to strengthen domestic AI capacity, while work concerning a national AI strategy and other public-policy initiatives is continuing. North Macedonia has also signed the Council of Europe Framework Convention on Artificial Intelligence and is expected, as part of the European integration process, to align its legal and institutional framework with the EU AI Act. Domestic actors have additionally contributed through research, education, capacity-building, and sector-specific work, including guidance concerning the responsible use of AI in journalism.

These initiatives are valuable, but they perform different functions. A national strategy, infrastructure project, international commitment, or sector-specific policy document does not necessarily provide businesses and other organizations with a generally applicable ethical reference for the practical decisions they are already making. At the time of publication, the official sources reviewed did not identify a comparable, accessible, and cross-sector set of principles focused specifically on the responsible use of AI by businesses and other private organizations in North Macedonia.

The present gap is therefore not an absence of technological activity or ambition. It is the absence of a shared, locally grounded reference that organizations can use when selecting AI tools, procuring AI services, developing or adapting systems, processing information, relying on AI-generated outputs, or considering how their use of AI may affect employees, customers, and other people.

The particular contribution of the Principles. The Principles translate established international standards into the circumstances in which organizations in North Macedonia operate. Accordingly, they give particular attention to:

- the performance of AI in Macedonian, Albanian, and other languages and dialectal varieties with more limited digital resources;
- differing levels of digitalization, connectivity, data availability, access to technology, and AI literacy;
- the technical, financial, and organizational capacities of smaller businesses and institutions; and
- the possibility that systems developed for larger foreign markets may not adequately reflect local laws, administrative practices, professional standards, cultures, or social conditions.

This local perspective does not alter the universal importance of human dignity, equality, autonomy, safety, and accountability. It seeks to make those values more meaningful by considering the practical conditions in which they must be applied.

Nature and status. These Principles are a voluntary awareness framework and a starting point for continued discussion. They are not legislation, official government policy, certification, a compliance manual, a technical standard, or an implementation roadmap. Following these Principles does not, by itself, establish legal compliance, and they are not a substitute for applicable law, professional responsibility, contractual obligations, or specialist advice.

The Principles are intended to help organizations identify relevant questions, recognize who may be affected, and determine where safeguards, human oversight, or additional expertise may be required. They should be applied proportionately, with greater care where AI may materially affect fundamental rights, safety, livelihoods, access to opportunities, or other important interests.

Continued development. This first version is deliberately open to review and improvement. Following publication, the Principles will be presented through consultations, roundtables, workshops, and other public and professional discussions. Businesses, AI developers and service providers, chambers of commerce, professional associations, public institutions, universities, researchers, civil-society organizations, workers and their representatives, and experts in relevant fields will be invited to examine and criticize them.

The purpose of this process is to test whether the Principles are understandable, relevant, and realistic. Future versions may clarify, amend, expand, or reorganize them in response to practical experience, constructive criticism, and continuing technological, legal, and social developments.

Purpose of the Principles. The Principles are intended to:

- raise awareness of the opportunities, limitations, and risks associated with AI;
- establish a shared vocabulary for discussing responsible AI use;
- help organizations ask appropriate questions before and during the use of AI;
- encourage safeguards proportionate to the possible effects on people and organizations;
- support broader and more confident adoption by reducing uncertainty and avoidable harm; and
- contribute a practical organizational perspective to the national discussion concerning AI.

Definitions

These working definitions are included only to make the Principles easier to read. They do not replace definitions in applicable law.

- **AI system** - a machine-based system that uses inputs to generate outputs, such as predictions, content, recommendations, or decisions, which may influence physical or digital environments.
- **AI-assisted decision or process** - a decision or process in which AI recommends, ranks, filters, generates, or otherwise materially influences an action, even where a human formally makes the final decision.
- **Affected person or group** - any person or group whose rights, interests, opportunities, working conditions, safety or access to services may be influenced by an AI system, directly or indirectly.
- **AI literacy** - the knowledge and practical understanding needed to use, evaluate and oversee AI appropriately, including awareness of its capabilities, limitations and risks.
- **Consequential use or decision** - a use or a decision capable of materially affecting a person's rights, safety, livelihood, employment, education, access to essential services, legal position, or other similarly important interests.
- **Meaningful human oversight** - oversight by a person with sufficient information, competence, time, authority, and practical ability to question, override, suspend, or stop the use of an AI system.
- **Personal data** - information relating to an identified or identifiable person. The binding meaning and requirements are determined by the Law on Personal Data Protection and other applicable law.
- **Provider** - an internal or external party that develops or supplies an AI system, AI model or related service.
- **Using organization** - an organization that selects, procures, integrates, operates or uses an AI system in a professional or institutional activity.

The Principles

Eight groups of principles

The eight groups of principles below are intended to be read together. They provide an ethical reference for reflection and discussion; they do not replace applicable legal or professional duties.

1 Human Dignity and Fundamental Rights

2 Fairness, Equality and Inclusion

3 Transparency, Explainability and Traceability

4 Data Governance and Integrity

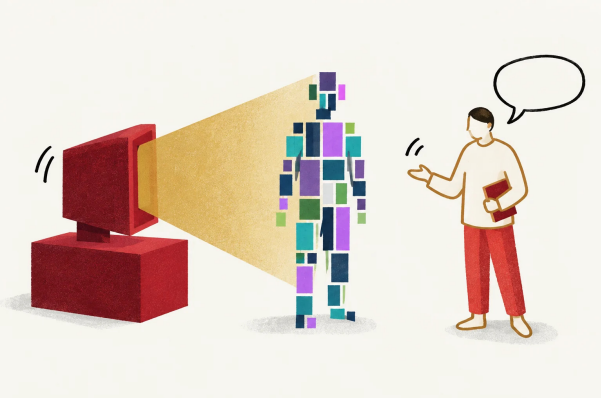
5 Safety, Security, Robustness and Reliability

6 Human Agency, Oversight and Competence

7 Accountability, Governance and Remedy

8 Participation, Local Context and Civic Legitimacy

1. Human Dignity and Fundamental Rights



1.1 Respect for human dignity

Artificial intelligence shall be developed and used in a manner that recognizes the inherent worth of every person. A person must never be treated merely as a data point or as a means for achieving an organizational objective. AI systems shall not objectify, exploit, or dehumanize individuals, nor shall they assign moral worth or social value to a person based on inferred traits, behaviour, or group membership.

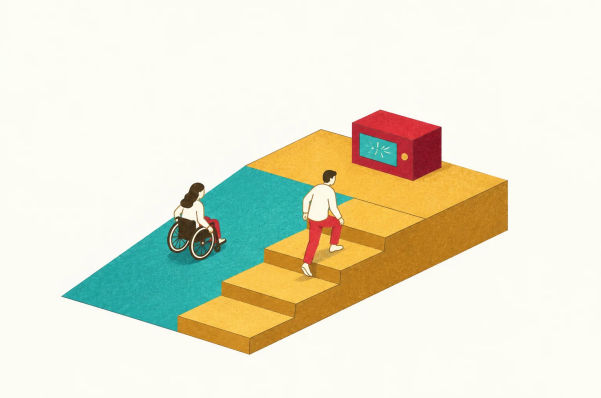
1.2 Protection of fundamental rights and freedoms

The development and use of artificial intelligence shall be compatible with fundamental rights and freedoms. Organizations shall identify the rights that may be affected, assess both direct and indirect interference, and adopt safeguards capable of preventing or reducing adverse effects. This responsibility applies throughout the lifecycle of the system and cannot be discharged solely through contractual terms, user consent, or reliance on a technology provider.

1.3 Human autonomy and self-determination

Artificial intelligence shall support, rather than displace, the ability of individuals to form, express and revise their own choices. Systems shall not use covert manipulation, exploitative persuasion or the strategic use of vulnerability to steer a person towards decisions that primarily serve the interests of the organization or another third party.

2. Fairness, Equality and Inclusion



2.1 Non-discrimination and prevention of bias

Artificial intelligence shall not create, reinforce or legitimize unjustified disadvantage based on protected characteristics or comparable personal and social circumstances. The prohibition extends to the use of variables that operate as proxies for characteristics the system is not permitted to use directly.

2.2 Substantive fairness

Artificial intelligence shall be assessed not only by whether it applies the same formal rule to everyone, but also by whether it produces justifiable outcomes in the real conditions in which people live. Different circumstances may require differentiated safeguards or support so that individuals have a genuinely fair opportunity to access a right, service, benefit, or professional assessment.

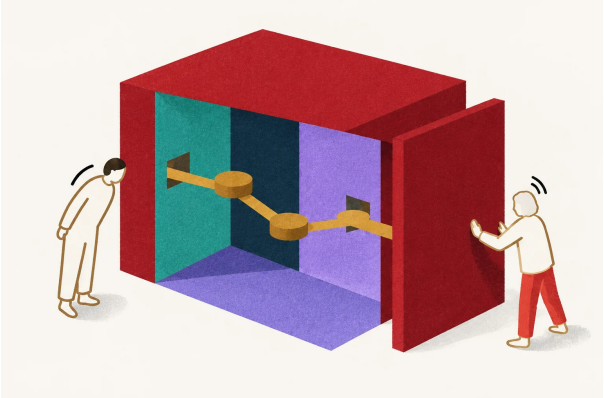
2.3 Procedural fairness

Consequential AI-assisted procedures shall provide timely notice, access to relevant information, understandable reasons, and a meaningful opportunity to contest an error or present additional facts before an impartial and authorized human decision-maker.

2.4 Special protection and accessibility for children and vulnerable persons

Artificial intelligence affecting children, persons with disabilities, elderly persons, economically dependent individuals or others in positions of heightened vulnerability shall be subject to stronger safeguards and accessible design. Organizations shall account for reduced bargaining power, limited literacy or technical ability, dependence on essential services, connectivity and equipment barriers, and the possibility that consent or formal choice does not reflect genuine freedom.

3. Transparency, Explainability and Traceability



3.1 Disclosure of AI use and manipulated content

Individuals shall be clearly informed when they are interacting with an AI system or when artificial intelligence materially influences a process that affects them. AI-generated or materially manipulated text, images, audio or video shall be clearly identified where their artificial origin is relevant to authenticity, public understanding, legal proof or accountability, and appropriate provenance information shall be preserved. Disclosure shall be timely, prominent and expressed in language appropriate to the audience.

3.2 Transparency of purpose, role and scope

An organization shall clearly define and communicate why an AI system is used, the task it performs, the decisions or activities it may influence, and the boundaries within which it is authorized to operate. System-level transparency shall describe the system's architecture, capabilities, limitations, intended users, and material dependencies sufficiently to enable responsible persons to understand and govern its role. A material expansion of the system's functions, users, or influence over decisions shall require renewed approval and shall not occur through informal use or gradual function creep.

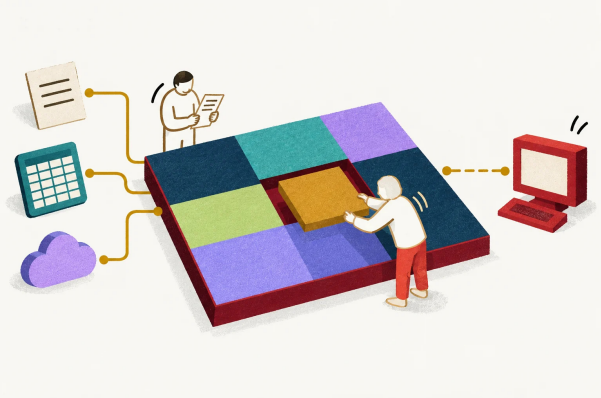
3.3 Decision explainability

Organizations shall ensure that persons who use or rely upon an AI system can understand why a particular decision, recommendation or outcome was produced and independently assess whether it should be accepted. The acceptable accuracy and speed of the explanation shall be proportionate to the importance and urgency of the decision. For consequential decisions, an adequate and timely explanation shall form an integral part of the decision-making process.

3.4 Decision traceability

Material AI-assisted decisions shall be traceable so that the process and outcome can be reconstructed. Relevant records shall identify the system and model version, material inputs and instructions, output, human intervention or override, and the final action taken.

4. Data Governance and Integrity



4.1 Data governance and lifecycle management

Organizations shall establish proportionate governance for data used to develop, validate, test and operate AI systems, covering collection, origin, preparation, labelling, access, security, use, sharing, correction, retention and deletion. Material data shall remain traceable through transformations and restrictions, with original, inferred, synthetic and human-created information appropriately distinguished.

4.2 Data quality

Data used to develop, validate, test or operate an AI system shall meet technical standards for accuracy, completeness, consistency, timeliness and integrity. Quality shall be measured throughout the data pipeline, with controls for collection, labelling, transformation, validation, correction, versioning and drift proportionate to the system's intended function and the consequences of error.

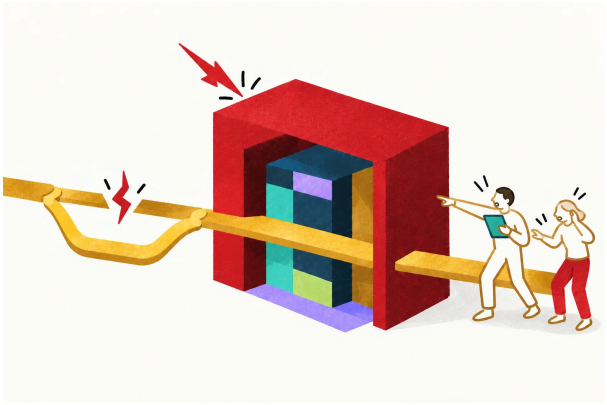
4.3 Privacy, processors and data transfers

Personal data shall be processed in accordance with applicable data-protection law. Organizations shall remain accountable when engaging AI providers or other processors that process personal data on their behalf and shall ensure appropriate instructions, contractual safeguards, security, control of subprocessors, and transparency regarding where and by whom data are processed. Transfers to other jurisdictions shall occur only where the required legal basis and safeguards are in place.

4.4 Control over personal data in models

Organizations shall recognize that deleting source data may not remove information learned, inferred or memorized by an AI model. Personal data used for training, fine-tuning or operation shall therefore be minimized and protected through measures that prevent unauthorized retention, extraction and reuse and, where feasible, enable correction, deletion or unlearning. Sensitive processing should take place within a controlled environment. Locally deployed or open-weight models may be used where they provide greater control over data and can be securely governed and maintained.

5. Safety, Security, Robustness and Reliability



5.1 Safety and prevention of unacceptable risk

Artificial intelligence shall be designed, deployed and operated so that foreseeable hazards are identified and controlled before they result in unacceptable harm. Harm includes physical, psychological, economic, legal, reputational and social injury, as well as harm to groups, communities and institutions. Safety requirements shall be defined in relation to the actual environment, affected persons, connected systems and reasonably foreseeable misuse, rather than only normal or intended operation.

5.2 Accuracy and precision

AI systems shall be evaluated against performance criteria appropriate to their intended use and the consequences of error. For generative AI, evaluation shall address, where relevant, factual accuracy, grounding in reliable sources, completeness, relevance, consistency and source fidelity, rather than relying on a single aggregate metric.

5.3 Reliability and consistency

An AI system shall provide sufficiently stable and dependable performance over time. Comparable inputs and circumstances shall not produce materially inconsistent outcomes without a justified reason, particularly where the output contributes to legal, professional or administrative decisions. Where variability is inherent in the system, the organization shall control and communicate it and shall establish procedures that prevent arbitrary treatment.

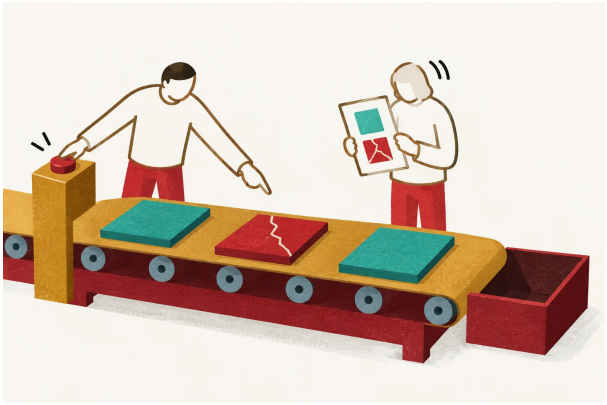
5.4 Robustness and resilience

Artificial intelligence shall tolerate foreseeable data errors, unusual inputs, environmental changes and partial failures without producing unacceptable consequences. The system shall recognize conditions that differ materially from those for which it was validated and shall not continue to produce confident output when its assumptions no longer hold. Organizations should maintain the ability to continue essential functions when an AI system or external service is unavailable, compromised, or materially changed.

5.5 Cybersecurity and access security

AI systems, models, datasets, interfaces, credentials and supply-chain components shall be protected against unauthorized access, disclosure, extraction, alteration and disruption. Security controls shall correspond to the sensitivity of the information, the importance of the function and the consequences of compromise. Security should cover both conventional cyber threats and risks specific to AI-enabled systems.

6. Human Agency, Oversight and Competence



6.1 Meaningful human command, oversight and intervention

Organizations shall determine whether, why, and under what conditions artificial intelligence is used. Oversight shall be proportionate to the system's context and risks and shall be meaningful only where the responsible person is informed of its purpose, limitations, uncertainty, and foreseeable failures; can access the relevant facts and evaluate its output independently; and has the competence, authority, and practical ability to question, override, suspend, or terminate its operation.

6.2 Preservation of human judgment and professional skill

The introduction of artificial intelligence shall not erode the knowledge and judgment needed to supervise the system, detect errors and perform essential work when technology fails. Organizations shall identify the competencies that must remain human and shall organize training and work allocation so that those competencies continue to develop. Junior professionals shall receive sufficient opportunity to acquire foundational skills before AI is allowed to replace the learning processes through which sound judgment is formed.

7. Accountability, Governance and Remedy



7.1 Clear ownership and lifecycle accountability

Every AI system shall have an identified organizational owner responsible for its purpose, approval, operation, monitoring, material changes and withdrawal. Responsibilities among management, professional users, technical teams, data owners and providers shall be explicit.

7.2 Impact and risk assessment by design

Before deployment and after any material change, the responsible organization shall assess the intended benefit, affected persons, impacts on rights and equality, data limitations, safety and security, organizational dependencies, and available safeguards. The assessment and its findings shall shape design and procurement from the outset, and serious unresolved risks shall prevent or restrict use.

7.3 Effective remedy and redress

A person harmed or materially disadvantaged by an AI-assisted process shall have access to a timely and effective remedy, including explanation, contestation, human reconsideration, correction, restoration and compensation where appropriate. Redress shall address both the individual outcome and any underlying defect that may affect others.

7.4 Institutional and individual accountability for AI-assisted decisions

An organization remains accountable for the AI systems it develops, procures, integrates or uses, including their outputs and consequences. Where an AI output informs a particular decision, the person who makes or adopts that decision remains responsible for independently evaluating the output and for the final decision, and shall have the information, competence and authority necessary to do so. Neither the organization nor the decision-maker may avoid responsibility by attributing the outcome to the AI system, a vendor, supplier, developer or other third party.

8. Participation, Local Context and Civic Legitimacy



8.1 Stakeholder and affected-community participation

Organizations shall identify and meaningfully involve the persons, institutions, and communities that develop, operate, regulate, are affected by, or bear the risks of an AI system, particularly affected, marginalized, and vulnerable groups. Participation shall begin while it can still influence the system's purpose, design, safeguards, and conditions of use, continue as new effects emerge, and remain accessible across differences in language, resources, and digital access.

8.2 Local contextualization

An AI system shall be evaluated against the laws, institutions, administrative practices, infrastructure, and organizational capacity of the environment in which it will operate. Validation in another country or organization does not by itself establish suitability for North Macedonia. Local adaptation shall address the system's actual operation, governance, and integration with existing institutions and infrastructure, rather than only translation of its interface.

8.3 Institutional capacity and maturity

An organization shall not deploy artificial intelligence beyond its capacity to govern data, supervise providers, evaluate performance, exercise human oversight, and respond to errors or incidents. Before deployment, it shall assess its people, processes, infrastructure, authority, and financial ability to sustain responsible operation.

8.4 Linguistic and cultural inclusion

AI systems used in North Macedonia shall be evaluated for the Macedonian language, its dialectal varieties and Cyrillic script, as well as for Albanian and the other languages, scripts and cultural contexts of communities living in the country. Organizations shall identify and address linguistic, cultural or data gaps that produce weaker performance, exclusion or misinterpretation and, where necessary, provide an effective human or non-AI alternative.

8.5 Data fitness for the local context

Organizations shall ensure that data used by an AI system adequately reflect the local legal, institutional and social context in which it operates. Particular attention shall be given to uneven digitalization, fragmented records, unwritten practices and relevant information that exists outside formal digital systems. The absence of digital data shall not be treated as proof that a fact or circumstance does not exist, and material gaps shall require additional human verification or restriction of the system's use.

Acknowledgements

Ethical AI Macedonia thanks the organizations that contribute to its work. Their support includes academic input, company outreach, event collaboration and the contributions of the founders' legal practices.

Embassy of Sweden in Skopje

Funding support for the November event on responsible AI in business.

Swedish-Macedonian Chamber of Commerce

Company outreach, circulation of the business questionnaire and event collaboration.

Lund University

Academic input on the initial principles and participation in the November event.

Boshnjakovski Law Office

The legal practice of co-founder Vladimir Boshnjakovski and an initiative partner.

Angela Angjelovska Law Office

The legal practice of co-founder Angela Angjelovska and an initiative partner.

Future contributions

These principles will continue to develop through comments and practical experience. We invite businesses, researchers, practitioners and civil-society organizations to contribute examples, criticism and suggestions.

aimacedonia.org | info@aimacedonia.org

Correspondence: DTC Paloma Bjanka, Dame Gruev 16/10
1000 Skopje, North Macedonia