Structured Document Review Report (Task 3.1 – Partner)

(For completion by each partner institution as part of WP3 Policy and Regulatory Framework Analysis)

Section 1: Partner & Submission Info

Purpose: To capture basic information about your institution and submission.

Instructions: Please fill in all fields below. Indicate who coordinated the review within your institution.

Partner Institution	Country	Reviewer(s)	Date of Submission
Rzeszów University of Technology	Poland	Łukasz Paśko (EU & national: E1, D1, D2), Marcin Olech (Local directives: D3-D7)	23.01.2026

Section 2: List of Documents Reviewed

Purpose: To provide metadata for all documents selected for review.

Instructions: List the 4–6 national and institutional documents your institution analysed. Include official titles and details. Use one line per document.

For each document, provide brief metadata in a table:

Doc ID	Title	Issuing Authority	Year	Type (Law / Strategy / Policy / Guideline)	Scope (National / Institutional)	Short Description
D1	Policy for the Development of Artificial Intelligence in Poland from 2020	Ministry of Digital Affairs	2020	Strategy / Policy	National	The initial national document on AI – a strategic rather than legally binding policy. It describes the actions that Poland should implement and the goals that should achieve. The document establishes goals and actions across six domains: society, business, science, education, international cooperation, and public sector. Two chapters directly address HE: Chapter 3 "AI and Science" and Chapter 4 "AI and Education." Each of them defines short-term (by 2023), medium-term (by 2027), and long-term objectives in the areas of science and education from Poland's national perspective.
D2	Project of the Policy for the Development of Artificial Intelligence in Poland up to 2030	Ministry of Digital Affairs	2025	Strategy / Policy	National	This strategic document serves as an action plan for creating optimal conditions for AI development in Poland. It highlights the importance of an AI ecosystem in which universities, research centers, and R&D institutions are the key actors. In addition to defining goals, it outlines possible methods for achieving them. However, these formulations remain highly general. It contains a set of

						indicators for monitoring the achievement of the goals; however, these indicators are specific to the Polish context.
D3	Directive no. 70/2024 according to "Principles of using artificial intelligence-based tools in teaching and research"	Cardinal Stefan Wyszyński University in Warsaw	2024	Policy	Institutional	The directive constitutes rules for the use of Gen-AI in the following areas: teaching, research, and publication in the CSWU publishing at Cardinal Stefan Wyszyński University in Warsaw. It presents a specific list of tasks that can be accomplished using Gen-AI, as well as a description of threats when using these tools.
D4	Directive no. 14/2024/2025 according to "Rules of applying AI tools in education process in WSB University"	WSB University in Dąbrowa Górnicza	2025	Policy	Institutional	The directive presents general rules according to the use of AI tools (Gen-AI included) by academic staff as well as students in HE. It presents some specific tasks where AI tools can be used, but in scope less compared to D3. It contains how the application of the AI tools should be marked in the academic work. There is also a list of some specific tasks which are forbidden to use with AI tools.
D5	Directive no. 5 according to "The introduction of rules for the use of artificial intelligence in the preparation of written assignments"	SGH Warsaw School of Economics	2024	Guideline	Institutional	The document presents detailed information according to what activities are allowed and forbidden in the specific area of application of the AI system. The areas presented in the document are: conceptual activities, literature review, writing process, text editing, generating graphics
D6	Guidelines according to responsible applying AI in teaching process in AGH University of Krakow	AGH University of Krakow	2025	Guideline	Institutional	The document presents rules and good practices for using AI tools in the teaching process in HE. The sections of the document focus on the explanation of keywords, recommendations for students, recommendations for teaching staff, and ethical and transparency issues.
D7	Levels of applying the AI in teaching process in AGH University of Krakow	AGH University of Krakow	2025	Guideline	Institutional	The document is novelty according to the other documents. Describes various levels of the application of AI tools with descriptions of goals, skills, benefits, and specific tasks examples. The proposed levels are as follows: No AI use; AI use for text editing process only; AI limited usage: Planning with AI, Collaboration with AI, Meet the AI; AI no restriction usage.
E1	ETHICS GUIDELINES FOR TRUSTWORTHY AI	High-Level Expert Group on Artificial Intelligence / European Commission	2019	Guideline	EU	The document contains soft-law ethical and operational guidelines aimed at supporting organisations in the design, deployment, and assessment of AI systems in line with the concept of Trustworthy AI. It concentrates on the ethical principles and technical–societal robustness dimensions. It articulates four core ethical principles and translates them into seven operational requirements intended to guide organisations across sectors. The guidelines are cross-sectoral and technology-neutral, making them applicable to public administration, industry, research, and education, including higher education contexts.

Section 3: Cross-Document Analysis by Category

Purpose: To summarise what the reviewed documents say under each key WP3 category.

Project: "Promoting Accountable and Transparent Human-centered AI in Higher Education" Project n. 2025-1-PL01-KA220-HED-000359600
Co-funded by the Erasmus+ Programme of the European Union

Instructions: Use this table to capture relevant extracts or themes found across your reviewed documents.

For all the reviewed documents together, partners summarise findings under each coding category:

Category	Key Extracts (Doc IDs) D1	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	The document places particular emphasis on the ethical aspects of AI, as reflected in the set of ethical principles defined therein. This set constitutes a fundamental basis for the design of all artificial intelligence systems, which must comply with the principles of Trustworthy AI, including human oversight, technical robustness, privacy and data protection, diversity, non-discrimination, societal and environmental well-being, transparency, accountability, and responsibility. The document stresses that all AI solutions must serve human beings, placing human dignity and fundamental rights at the forefront. It recognises the value of the European, human-centric approach to Trustworthy AI, which is based on the establishment of comprehensive ethical frameworks. The document further emphasises the comprehensive engagement of all stakeholders within the AI ecosystem, including citizens, entrepreneurs, research institutions, and public administration, while simultaneously safeguarding ethical principles and the paradigm of human sovereignty over AI. This document demonstrates that ethics is prioritised over purely technical considerations. The document assumes active participation in the work of international organisations involved in developing consensus on the ethical use of AI for civilian and security-related purposes. Engagement in the creation of a global AI ethics code and in deliberations on human autonomy vis-à-vis the automation of digital machines is identified as one of the policy directions outlined in the document. The document calls for intensified Polish involvement in the activities of	**Coverage**: Very broad and repeated references to AI ethics and AI integrity, present both in the diagnostic section and in the operational objectives. The document is grounded in European guideline frameworks and includes recommendations aimed at promoting Trustworthy AI at both the national and international levels. It emphasises the importance of ethics throughout the entire AI lifecycle. **Strengths**: • strong anchoring in the European Trustworthy AI paradigm, • very strong positioning of AI within a human-centric paradigm and the protection of human dignity, • references to European ethical standards and international normative frameworks, • explicit linkage between AI and human rights, • recognition of ethics as an element of public policy, • a plan to develop a Polish AI ethics code considering the Constitution and human rights, • commitment to supporting research on AI ethics,

	international organisations such as the European Union and the United Nations, with a particular focus on human dignity, fundamental rights, and practical methods for implementing core ethical values within technical and non-technical frameworks for AI assessment. The establishment of standards for the deployment of AI solutions—especially those ensuring compliance with AI ethics—should be a responsibility of public administration. One of the medium-term objectives is the development of a dynamic AI ethics code aligned with the Polish Constitution and the Charter of Fundamental Rights of the European Union. The analysis of the ethical consequences of AI implementation and its impact on the sphere of human rights, support for research on AI ethics through grants and competitive funding schemes, assessment of the societal impact of AI systems on human rights and freedoms, the development of methods for their independent auditing, as well as consultation of future national and international regulations (OECD, UN, EU, Council of Europe) in the field of AI ethics, are defined as short-term objectives. The section on AI and education emphasises the need to uphold AI ethics principles within digital education strategies and to develop standards and best practices for the use of AI solutions in education that are consistent with AI ethics and protect the rights of participants in the education system.	• clear identification of seven ethical principles within the definition of AI. **Weaknesses**: • limited concrete guidance on the practical implementation of ethical principles in AI systems, • absence of mandatory ethical procedures for AI projects, • lack of enforcement mechanisms for ethical principles and the absence of defined sanctions for violations of AI ethics, • absence of ethical metrics or audit mechanisms, • undefined institutional competences and responsibilities for monitoring AI ethics.
Data Privacy & Security	The document treats privacy protection and data governance as one of the seven key pillars of Trustworthy AI. The Polish AI ecosystem should engage various stakeholders with due attention to cybersecurity, which constitutes a significant element of data protection in the context of innovation and technological development. The document foresees support for work on principles of personal data processing resulting from the GDPR (especially data minimisation) and for the rigorous risk assessment of AI systems. The section on "AI and international cooperation" assumes that trusted data spaces	**Coverage**: Privacy and personal data are explicitly addressed, primarily from a legal perspective. The document includes plans for legislative changes to support secure data sharing. It refers to the GDPR and personal data protection, with particular emphasis on sensitive data. **Strengths**: • linkage to the GDPR and the European standard of data protection,

	will be created within European cooperation, where data from European public institutions or the private sector would be made available in a reciprocal and equitable manner. Medium-term objectives mentioned in the document include, among others, updating legislation on access to sensitive data (e.g., medical data), conditions for the operation of trusted data-sharing spaces with due regard to privacy protection, preparing legislation for cloud computing, edge computing and IoT in the context of industry (sharing industrial data), the collection of public data, as well as promoting data openness through portals such as Open Data and Digital Sandbox. The document also talks about Open Data programme, the obligation to create public APIs with respect for personal data protection, and the introduction of medical data pilots, while emphasising compliance with legal provisions on the protection of classified information, trade secrets, statistical confidentiality, and other legally protected secrets, as well as the use of appropriate protection techniques (e.g., anonymisation or pseudonymisation). Short-term objectives in the "AI and innovative enterprises" section emphasise strengthening the opening of public administration data, incentives for sharing non-public data in trusted data spaces (data trusts), reciprocal data sharing, and standards for commercial solutions that respect trade secrets. Policy directions indicate the protection of trade secrets and the standardisation of data governance. The document also defines long-term objectives such as the availability of public data via APIs with protection of trade secrets and the free flow of non-personal data, safeguarding citizens' rights to privacy protection, and supporting projects in the field of cybersecurity and combating disinformation.	• recognition of the importance of sensitive data and mechanisms for its protection, • the principle of data minimisation as an element of AI policy, • promotion of the concept of secure data trusts and open data via APIs with respect for privacy, • plans for legislative changes supporting secure data sharing, • inclusion of anonymisation and pseudonymisation as protection techniques, • mention of risk assessments for AI systems. **Weaknesses**: • lack of detailed technical standards for data protection in AI systems, • absence of clear guidance on security audits, • unexplained role of the private sector in data protection, • absence of clearly defined sanctions for data security breaches, • lack of specific measures addressing cybersecurity of AI infrastructure, • no reference to data security across the entire AI lifecycle, • lack of emphasis on the importance of training AI operators in security, • simultaneous emphasis on data openness and privacy.
Equity & Inclusion	In the document, equality and inclusion are embedded in the general definition of Trustworthy AI, where one of the key principles is "diversity, non-discrimination and fairness." In the	**Coverage**: The document covers the area of Equity & Inclusion from ethical principles, through the

	diagnostic section on challenges, it is noted that the development of AI may exacerbate existing inequalities between more and less developed regions and between individuals with varying levels of digital skills if protective and compensatory mechanisms are not implemented. The section "AI and Society" adopts a broad approach to inclusiveness, through labour market protection measures, democratisation of access to AI, the enhancement of civil servants' competencies, and public awareness campaigns. The document explicitly identifies disadvantaged groups. Equity & Inclusion are pursued through information, education and retraining programmes targeted at the most vulnerable. The document also draws attention to equal opportunities in access to an academic career in AI. Inclusion is pursued through scholarships, reimbursements and publication support for students and doctoral candidates. It lists support instruments for pupils and students regardless of their place of origin, including scholarships and participation in research projects. In the "AI and Science" section, educational inclusion is emphasised through the flexibility of the academic system (individualised study paths), interdisciplinarity and the consideration of diverse competencies of candidates. This section also links inclusion with open access to educational materials, academic mobility and internationalisation. The document highlights the need for equal access to data as a condition for AI development. This concerns inclusion in access to data resources. In the section describing the Polish AI ecosystem, it is emphasised that it should operate with respect for social solidarity and sustainable development, which also implies the obligation to consider the situation of less privileged groups in access to technology and digital services. In the part dedicated to international ethical frameworks for AI, it is emphasised that	labour market, education, science, data access, and up to international cooperation. Inclusion is primarily understood as equal opportunities in access to AI and education, as well as counteracting social exclusion. **Strengths**: • clear recognition of the risk of exacerbating social inequalities, • linkage of AI to labour market policy: plans for retraining programmes for workers threatened by automation, • consideration of regional and territorial inequalities, • strong anchoring in labour market and education policies, • emphasis on the democratisation of access to AI and data, • striving for compliance with EU and UN frameworks, • engagement in training public officials to prevent discrimination. **Weaknesses**: • lack of explanation of how discrimination will be monitored in practice, • absence of specific metrics to measure the reduction of inequalities, • lack of references to minorities, • lack of tools for assessing the impact of AI on vulnerable groups (limited attention to people with disabilities in terms of access to AI).

	organisations such as the EU and the UN are working on regulations that explicitly refer to ensuring the absence of social exclusion and avoiding accidental discrimination, and that Polish representatives should advocate for practical methods of implementing these values in technical and non-technical rules.	
Transparency & Explainability	The document identifies transparency as one of the ethical principles to which AI systems should adhere. It treats transparency as a normative value embedded within the framework of Trustworthy AI. Transparency and explicability are directly linked to social trust. The document emphasises their relationship with the human-centric AI approach and with the prevention of algorithmic errors, but it remains at the level of general principles rather than technical requirements. The document states that transparency is a prerequisite for societal acceptance of AI and requires iterative improvement and institutional support. The document states that transparency and explicability should be part of AI regulation, especially in systems affecting human rights and fundamental freedoms. It also highlights the need to develop knowledge and tools related to transparency through research funding. The document states that citizens' and consumers' understanding of how AI systems operate is a prerequisite for building social trust. Explicability in this context does not refer to algorithmic architecture, but rather to users' ability to assess the effects of AI system operation. This is associated with the need to create "model explanations" designed for the recipients of AI systems. Transparency also concerns decision-making processes in public administration. Explicability is linked here to the right to appeal decisions made using AI, particularly when they affect civil rights.	**Coverage**: The document addresses transparency and explicability, particularly in the context of the public sector, AI regulation, user education, and building social trust. Transparency & Explainability are present both as ethical principles and as objectives of public policy. **Strengths**: • recognition of the need to understand AI algorithms, • recognition of explicability as an element of protecting civil rights, • transparency necessities for AI systems in the public sector, • plan to develop model explanations for AI decisions, • the right to appeal decisions made with AI support, • support for research on model interpretability, • inclusion of user and citizen education. **Weaknesses**: • lack of specification on how transparency would be achieved, • lack of detailed standards regarding the level of explicability required for different types of AI, • unclear how the right to explanation will be implemented in practice,

		• lack of a timeline for implementing model explanations, • few specific guidelines on how citizens will access explanations, • focus on principles rather than design requirements, • lack of measurable criteria for evaluating the transparency of AI systems.
Governance Structures	The document integrates the state administration into the Polish AI ecosystem. It identifies the need for a central coordination mechanism for AI policy. Governance also refers to the monitoring and evaluation of AI systems. It includes regular assessment of the effectiveness of AI policy instruments and their compliance with the adopted strategic objectives. The document places responsibility on the Polish government to develop a dynamic ethical AI code, in cooperation with AI experts and with regard to the Constitution and the EU Charter of Fundamental Rights. The Council of Ministers is also responsible for regular reviews of AI policy. The document outlines a formal division of roles and responsibilities and describes the institutional architecture of AI governance. It envisages the establishment of multiple specialized advisory, monitoring, and legislative bodies, indicating a multi-layered governance model. The document also addresses AI policy management: planning, monitoring, reporting, financing, and coordinating actions across the entire public sector. The AI Policy Task Board is to monitor changing realities and adjust objectives, metrics, additional actions to be taken, and corrections in the implementation of AI Policy. The AI Policy Task Board will be responsible for the operational aspects of implementing AI Policy. The implementation of the actions described in this document will be coordinated and carried out by the minister responsible for digitalization. The document emphasizes cross-ministerial involvement in AI policy. It identifies a broad list of	**Coverage**: Describes governance structures at the government level, including central coordination, institutional architecture, division of roles, and inter-ministerial cooperation. **Strengths**: • clear governance structure with defined responsibilities, • formal embedding within government structures, • involvement of multiple ministries, • establishment of appropriate task boards, • identification of strategic partners for each area, • recognition of the need for an evaluation and update system for AI Policy. **Weaknesses**: • lack of an independent supervisory body, • lack of external control mechanisms, • high structural complexity and risk of bureaucratization, • lack of mechanisms for resolving competency conflicts and for communication between expert teams and the government side,

	ministries and public institutions as partners in achieving AI policy objectives in the areas of: society, innovative companies, science, education, international cooperation, and the public sector. For each of these areas, a list of responsible ministries is defined.	• lack of performance indicators for governance structures, • weak role of the academic sector in formal governance structures.
Technical Standards & Interoperability	The document emphasizes that technical robustness and security are fundamental quality requirements for AI systems, on which the trustworthiness of AI depends. The document highlights the need to regulate the principles for sharing AI algorithms and solutions (e.g., new types of licensing). It also stresses regulatory flexibility as a condition for effective implementation of technical standards. Polish AI solutions must comply with global technical standards. In the context of cooperation within international initiatives, interoperability is indicated at the level of data infrastructure and distributed systems. The document also assumes the creation of international digital innovation hubs and AI technology testing centers within the country. The document explicitly identifies technical standards and interoperability as directions of state policy. It emphasizes the need to establish norms, certification mechanisms, compliance protocols, and interoperability rules. Operational objectives include interoperability of existing e-health systems, the development of API standards for access to public data, standards for sharing non-public data, and mutual recognition of certificates and compliance protocols. It also envisages the creation of a trusted public cloud for the public sector to store and process data of Polish citizens, including the use of edge computing. The document promotes shared architectures, models, and datasets as elements that foster interoperability and reusability of AI solutions. It emphasizes the importance of education on the conscious use of methods and tools derived from computer science, which is intended to improve users' understanding of AI technologies. The document specifies particular programming	**Coverage**: The document addresses technical standards and interoperability at a general level. It mentions the need to develop appropriate technical norms, data standards, and system architectures, and emphasizes the importance of international interoperability and technological education. **Strengths**: • plans for standardizing interfaces for data access, • recognition of the importance of open interoperability standards, • commitment to harmonization with European standards, • references to specific technologies and tools in the context of AI education. **Weaknesses**: • lack of identification of formal standardization institutions, • lack of metrics for compliance with standards, • absence of references to specific technical standards, • lack of information on mechanisms for enforcing standards, • unexplained certification and accreditation process.

	languages, environments, and libraries as recommended tools for disseminating practical knowledge about AI at all educational stages (embedded systems, Python, Jupyter Notebook, Scikit-learn, PyTorch, TensorFlow). It also points to the necessity of using advanced software tools and continuously verifying their market relevance.	
Accountability & Liability	The document explicitly identifies accountability and responsibility as fundamental principles that AI systems should fulfil. At the procedural and institutional level, it emphasizes the need for systematic assessment of AI systems' impact on human rights and freedoms, as well as the development of methods for independent auditing. Accountability is understood not only as ex-post liability, but also as a process of continuous monitoring, evaluation, and support for research on accountability mechanisms. The document introduces a distinction between roles and liability regimes. It states that AI producers should be held liable based on a duty of care, whereas operators should be liable under a risk-based regime. At the same time, it clearly separates the responsibility of end-users from that of operators. Responsibility for AI systems should always be attributed to a human or a legal entity. The document introduces an obligation of ex-ante self-assessment, which should include problem identification, allocation of responsibilities, potential errors (including algorithmic bias), and remedial measures. Responsibility is thus understood as conscious, prior risk management rather than merely reacting to harm after the fact.	**Coverage**: The document includes references to accountability and responsibility, particularly in the context of the public sector. It addresses ethical, legal, procedural, and institutional dimensions. It introduces a differentiated responsibility framework across different stakeholders (producers, operators). **Strengths**: • Consistency with a human-centric approach: clear attribution of responsibility to humans and organizations, • Exclusion of "algorithmic responsibility", • Mandatory ex-ante assessment for public systems (self-assessment, audits), • Differentiation of responsibility between producers (duty of care) and operators (risk-based), • Funding for research on accountability, • Emphasis on building a culture of accountability and trust. **Weaknesses**: • Absence of legal liability frameworks, • Lack of detailed rules for attributing fault and remedying harm, • Lack of enforcement procedures for accountability,

		• No mention of AI risk insurance, • Lack of precise definitions of harm caused by AI, • Limited reference to civil and criminal liability, • No reference to specific national or international legal acts.
Sustainability	The document links sustainable development with AI ethics, identifying social and environmental well-being as one of the core principles of Trustworthy AI. Sustainability is not understood solely in ecological terms, but also encompasses the social consequences of AI development (social well-being). The document indicates that the capacity to create and implement AI solutions translates into an increase in socio-economic development. It highlights Poland's potential as a testing ground for AI solutions in areas crucial for long-term social and environmental well-being (energy, environmental protection, water retention, agriculture and the food sector, healthcare, sustainable cities, smart factories, transport). Consequently, adequate funding and support for innovative companies, development of managerial competencies, and strengthening the ability of the academic system to support long-term innovation and workforce education are required. The document provides a particularly extensive description of AI's potential in the energy transition. AI systems enable increased energy efficiency, integration of renewable energy sources, and improved stability of power systems. The document emphasises the role of AI in reducing losses, improving energy supply quality, and creating more flexible and resilient systems, which directly aligns with climate and environmental objectives. Medium-term objectives indicate the need for cooperation with other countries, including the use of AI in infrastructure and energy sectors.	**Coverage**: the document includes references to sustainable development, particularly in the context of AI applications. It highlights the potential of AI to support climate goals and environmental monitoring. It also addresses the social dimension of sustainability. **Strengths**: • Clear recognition of AI's role in achieving climate objectives and the Sustainable Development Goals (SDGs). • Commitment to environmental monitoring and protection. • Mention of smart energy grids and renewable energy management. • Commitment to sustainable growth and social wellbeing. • Sectoral approach to sustainability (energy, agriculture, health, cities). **Weaknesses**: • Lack of reflection on the environmental costs of AI (energy consumption, carbon footprint). • Absence of specific targets and metrics for sustainable AI. • No reference to the sustainability of materials and electronics used in AI.

		• Unexplained mechanism for monitoring the implementation of sustainability objectives. • Lack of clear sanctions for non-compliance with sustainability standards.

Replicate table for each document

Category	Key Extracts (Doc IDs) D2	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	The document emphasizes that technological development in AI must be combined with adherence to ethical standards. It assigns the AI Fund Council a role in actively shaping the direction of ethical and safe AI development. The document stresses that the implementation of regulation (including the AI Act) should promote ethical AI design rather than merely imposing restrictions. Regulatory sandboxes are presented as a tool that enables testing of solutions aligned with ethical principles in a controlled environment. The document suggests that, in order to ensure Trustworthy AI, codes of ethics for AI applications should be developed. It also highlights the need for training and guidance materials for specific professional groups, particularly in critical infrastructure areas. Ethics & Integrity are treated here as a set of competencies and procedures that must be implemented in everyday organisational practice for the ethical use of AI systems. The document identifies a specific ethical problem—disinformation amplified by AI systems—and indicates that the response to this risk should be the principle of ethics by design. Public administration is expected to play a special role in ethics as a model of ethical AI deployment. Consistent implementation processes based on ethics by design are necessary, suggesting that the public sector should act as a leader in integrity and ethical accountability of AI systems. The document also notes that AI users and society are important in this context. Therefore, it emphasises the	**Coverage:** Ethics and integrity are clearly present as a principle of Trustworthy AI. The document covers a wide spectrum of ethical issues and also identifies concrete actions that need to be taken. It places particular emphasis on the "ethics by design" initiative and the Council of Europe's Framework Convention. **Strengths:** • Explicit reference to human rights and grounding in international standards, • Identification of practical actions necessary for creating ethical AI, • Implementation of the "ethics by design" principle, • Preparation of sectoral codes of ethics, • Institutional anchoring of ethics (AI Fund Council), • Emphasis on education, training, and practical tools, • Integration of ethics into regulatory and innovation processes (regulatory sandboxes). **Weaknesses:** • Lack of concrete enforcement mechanisms for ethical codes,

	importance of education and communication in reducing irrational fears and promoting responsible attitudes towards AI.	• It does not specify who will enforce ethical standards nor what the consequences of their violation will be, • Lack of precise indicators for measuring ethics, • Dominance of declarations over descriptions of control procedures.
Data Privacy & Security	The document places security and trust among the fundamental prerequisites for the technological development and treats personal data protection as a core element of trustworthy AI. It refers to specific principles of personal data protection and emphasizes the need to assess the compliance of AI systems with the fundamental principles of the GDPR and the AI Act. The Office for Personal Data Protection is expected to play a significant role in this regard. AI security is also perceived in a supranational dimension. The document stresses the importance of cooperation within the EU and embeds Data Privacy & Security in international normative frameworks (OECD principles of 2019). It refers to the ISO/IEC 27090 security standard and highlights the need to adapt it in public administration. It envisages strengthening research institutions focused on system resilience to attacks and security verification, as well as cooperation with cybersecurity experts, preparing training materials for management staff regarding threats, and systematic monitoring of AI-related incidents. Security must be a design feature of AI models created for public administration. Security is not limited to infrastructure but also concerns the AI models themselves and their adaptation to the specific needs of the public sector. The document points to the national cloud, the Common State IT Infrastructure, and the Government Cloud as tools ensuring regulatory compliance, protection of sensitive data, and technological sovereignty.	**Coverage:** The document covers Data Privacy & Security very comprehensively, combining personal data protection, AI system security, cybersecurity, cloud infrastructure, regulations (GDPR, AI Act), technical standards, and sectoral use cases—particularly in healthcare and critical infrastructure. **Strengths:** • Clear linkage between AI and personal data protection and data security. • Recognition of GDPR and its relationship with the AI Act and OECD principles. • Detailed references to security infrastructure (Common State IT Infrastructure and Government Cloud). • Consideration of sensitive and medical data. • Reference to international standards (ISO/IEC). • Creation of an oversight system for supervising data processing in the context of AI.

	An important issue is the management of shared and open data, with a clear reservation regarding copyright protection. The document signals a tension between data openness and its protection, indicating that the sharing of cultural resources must be structured and legally compliant. In the case of high-risk systems, AI security is associated with audits, compliance with the AI Act, and incident management and monitoring. In the case of health data, the management model should give the patient control over their data. The document emphasizes the importance of education on the GDPR and privacy protection, indicating that data security depends not only on technology but also on administrative practices.	• Implementation of initiatives raising employee awareness. **Weaknesses:** • Lack of detailed enforcement mechanisms in the event of security violations. • Limited information on sanctions and accountability for incidents. • Identification of the conflict between data minimization (GDPR requirement) and the Big Data needs of AI, without proposing a solution.
Equity & Inclusion	The document treats inclusivity as one of the main principles of AI policy, embedding it in the international OECD standards. It lists specific groups: children, seniors, socially excluded people, and persons with disabilities. Inclusivity should be expressed through the design and implementation of AI that increases accessibility, safety, and quality of life, as well as protects against the negative impacts of technology. One dimension of equality is social justice in the context of labour market changes. Support is needed for workers affected by AI-driven transformation. The document highlights the need for research that considers both the potentials and risks of AI, as well as the situation of vulnerable groups. Public policy must be consciously designed with the weakest participants of the labour market in mind. Territorial equality of access to AI infrastructure is also emphasized. Inclusion is treated as enabling entrepreneurs and researchers to use computational power regardless of region. The document also emphasizes equal access to AI tools in public education. All schools and public universities will be able to use AI technological solutions. The document points to the need to develop AI skills across society, starting from early education, through teacher training, to supporting research teams. Incentives	**Coverage:** The document addresses equality and inclusivity across several dimensions: access to AI education for all, support for marginalized groups, regional distribution of infrastructure, and international cooperation. **Strengths:** • Linking AI with social policy, • Explicit identification of vulnerable social groups – children, seniors, and people with disabilities, • Formulation of indicators monitoring the increase in AI-based solutions for vulnerable communities, • Digital accessibility for people with disabilities as a policy element, • AI education at all levels, • Free access to AI tools in schools, • International cooperation to support linguistic and cultural diversity in Europe. **Weaknesses:**

	are also envisaged for students and researchers to engage in AI issues and access to open datasets. Inclusivity is reflected in lowering entry barriers to AI research. Support may also be necessary for entrepreneurs, especially startups and AI-developing companies.	• Lack of detailed analyses regarding inequalities in AI access across different regions, • Lack of discussion of algorithmic bias in the context of discrimination, • Limited discussion of measures to improve access for lower-income populations, • Indicator 4.6 requires a 20% increase in AI-based solutions for vulnerable communities but does not define what constitutes a "solution".
Transparency & Explainability	In the document, transparency and explainability are treated as fundamental normative principles of Poland's AI policy, deriving directly from OECD recommendations. Explainability is particularly important in public administration. AI systems must be not only technically explainable but also adapted to the organizational context of various institutions – from local to central levels. Users of AI within the administration must understand the principles of how it operates. It is planned to create a public register of AI systems used in the administration, including descriptions of their functions and basic technical parameters, which is intended to increase public trust and enable social oversight of AI usage. The document identifies lack of transparency as a regulatory issue requiring legislative intervention. It also refers to the development of the scientific and technical infrastructure for explainable AI (XAI) and highlights the need for research into explainability methods. Transparency also concerns the commercial sector, particularly in the context of personalization algorithms. The document signals the need to regulate how algorithms influence consumer decisions by increasing transparency of personalization mechanisms.	**Coverage:** The document covers transparency at multiple levels: public administration, AI systems, personalization algorithms, and automated decisions. It places particular emphasis on transparency for citizens using public services. **Strengths:** • a concrete transparency tool: a public register of AI systems, • alignment with OECD standards, • emphasis on the public sector: AI support for public officials should be transparent and explainable, • emphasis on developing guidelines for transparency of algorithms personalizing purchase offers, • implementation of "ethics by design," which should pay attention to system transparency, • recognition of the lack of transparency as a regulatory challenge.

		Weaknesses: • lack of technical details on how explainability should be implemented, • long deadline for guidelines on transparency of algorithms (2028), • no discussion of the trade-off between explainability and model accuracy, • limited discussion on accessibility of explanations for non-technical users, • lack of standards regarding the level of explainability required for different types of decisions.
Governance Structures	The document precisely defines the central institutional hub for AI in Poland. The minister responsible for informatization acts as the coordinator of the AI ecosystem, overseeing legislation (AI Act), infrastructure, cross-sector cooperation, and the establishment of key research institutions. The implementation of the AI Development Policy will be the responsibility of the minister responsible for informatization under the supervision of the Digital Committee, while noting that achieving all policy goals will require cooperation between the government and public sector entities. The mentioned minister should monitor the implementation of the AI Policy annually, while a review of the AI Policy (in cooperation with members of the Council of Ministers) will be conducted every two years. The document defines a performance indicator for the AI policy based on international rankings: the Tortoise Global AI Index, the Government AI Readiness Index, and the Stanford HAI Global AI Vibrancy Tool. These rankings serve as an external mechanism for assessing the quality of AI development governance. The establishment of a supervisory authority for the AI systems market suggests preparing structures to enforce regulations, including the AI	**Coverage:** The section covers governance very broadly: institutions, roles, procedures, monitoring, evaluation, financing, standards, and multi-level cooperation. **Strengths:** • clear role of government administration, • clear allocation of responsibilities, • emphasis on inter-ministerial cooperation, • establishment of a formal supervisory body; creation of the Artificial Intelligence Fund Council for coordination – a supra-ministerial institution, • linkage of governance with financing and international standards. **Weaknesses:** • limited formal role of civil society organizations within decision-making structures,

	Act. AI governance at the highest level also includes planning, allocating, and scaling public funds, which is also specified in the document along with an estimate of expenditures on AI development in Poland.	• lack of details regarding the independence of the supervisory authority, • little information on mechanisms for resolving competency conflicts, • lack of details regarding the powers of supervisory bodies.
Technical Standards & Interoperability	The document refers directly to the OECD 2019 principles and recommends shaping an interoperable AI environment. The implementation of coherent technological solutions is to be the result of cooperation between the AI Fund Council and representatives of key institutions financing AI projects. The document lists infrastructure elements (AI factories and gigafactories, supercomputers, data centers, high-throughput networks) and assumes their redundancy and scalability. Interoperability is supported through central computing resources and joint infrastructure initiatives (the Common State IT Infrastructure and the Government Cloud), which enable a coherent technological environment. The document also highlights the international dimension of interoperability. Poland participates in the development of a common European infrastructure, which promotes compatibility of solutions in large language models and computational models. The importance of Poland's participation in standardization processes is also emphasized. The PLLuM project is to serve as an example of a national component of linguistic interoperability, ensuring consistency of language technologies in public administration while considering the Polish context. The document also formulates measurable indicators of standardization activities and indicates that the development of technical standards in public administration will be quantitatively monitored.	**Coverage:** The document addresses technical standards at the infrastructure level (ISO/IEC), model level (PLLuM), and system level (interoperability between systems). Interoperability is emphasized in the context of OECD and European standards. **Strengths:** • Awareness of the importance of standards, • Reference to specific ISO norms (27090, 42001), • Formulation of measurable standardization indicators, • The Polish language model PLLuM as a national standard for public administration, • Strong embedding in international frameworks (OECD, EU), • Clear emphasis on infrastructure interoperability and modularity, • Developed vision of a shared infrastructure. **Weaknesses:** • Indicator 2.6 refers to the "development" of standards, but not their "implementation,"

		• Limited discussion of interoperability with systems of other countries (outside the EU), • Lack of discussion about standards for AI models, • No information on how conflicts will be resolved when standards are inconsistent (e.g., between Polish standards and EU norms), • Few references to interoperability of commercial systems.
Accountability & Liability	The document states that responsibility is one of the key elements of the OECD set of principles within the AI Policy. Responsibility is presented as a component of trust and integrity that must be supported through educational and communication activities. The requirement for accountability is to be a feature of AI systems deployed in public administration. The introduction of a complaints system implies recognition of an individual's right to challenge the functioning of an AI system. This is a clear element of the infrastructure of responsibility and potential institutional accountability. The document also refers to responsibility for the consequences of AI operation and calls for conducting international exercises in responding to AI-related crises. This implies recognition of the possibility of serious AI-related incidents and the need to prepare public institutions for response.	**Coverage:** The document covers normative responsibility (OECD), institutional responsibility (public administration), social responsibility (awareness and education), operational responsibility (crisis management), and procedural responsibility (complaints mechanism). Special emphasis is placed on responsibility in public administration. **Strengths:** • Strong anchoring of responsibility within international frameworks (OECD). • Consideration of both prevention and response to AI-related harms. • Implementation of a complaints mechanism enabling individuals to file objections regarding the functioning of AI systems. • Collection and analysis of information on AI-related incidents. **Weaknesses:**

		• Lack of clear consequences for failure to uphold responsibility. • Lack of discussion on vendor liability for AI system errors. • Lack of discussion on AI liability insurance. • Unclear how causality will be determined in the event of an AI error. • Lack of discussion on responsibility in the context of AI-generated disinformation. • No references to civil or criminal liability.
Sustainability	The document introduces sustainability as a key condition for the deployment of AI and indicates that Poland's long-term success in the field of AI depends on its ability to build an ecosystem that delivers benefits both to the economy and to social well-being. The development of AI that generates such benefits is conditional upon the creation of appropriate legal, financial, technical, and educational frameworks. With regard to the environmental dimension of sustainability, the document explicitly identifies the high energy consumption of AI—particularly LLMs—as a significant risk to the stability of energy systems and to the environment. At the same time, it points to concrete directions for action, including energy-efficient algorithms, green infrastructure, cooperation with renewable energy sources, and advanced cooling systems. These measures align with EU climate policy on green AI infrastructure. The document emphasizes that AI solutions should support the reduction of negative human impact on the environment and foster technological development aimed at mitigating both the causes and effects of climate change. The document also refers to potential applications of AI models in sustainability-related domains, such as advanced transport systems, energy, smart city solutions, and services in	**Coverage:** The document addresses sustainability primarily in the energy dimension (green AI infrastructure), but also in social and environmental terms. **Strengths:** • recognition of the potential of AI to support the Sustainable Development Goals (SDGs), • linking innovation with social well-being, • a strong emphasis on energy-related issues, • assuming development of green AI infrastructure. **Weaknesses:** • lack of concrete guidelines on energy consumption by AI infrastructure, • absence of measurable sustainability indicators, • no discussion of the carbon footprint of AI systems,

	healthcare and environmental protection. Sustainability is thus understood as the capacity of AI to support sustainable development goals.	• no analysis of the AI hardware lifecycle, from production to disposal, • no concrete commitments to renewable energy sources, • no discussion of the water footprint of AI, particularly in relation to water-based cooling systems, • lack of timelines and environmental impact reporting obligations.

Category	Key Extracts (Doc IDs) D3	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	Directive points out that it is crucial to know and understand what and why some threats may appear when applying the Gen-AI tool. Academic teachers, students, and PhD students should know exactly what allowed and forbidden activities are if applying the Gen-AI tools during their academic work. Because of the University specification, there is also pressure on the importance of relations between people, which can be violated or limited when applying Gen-AI tools or similar.	**Coverage:** Threats description (pp. 2-3) when applying the Gen-AI tools in HE connected mainly to ethics. Limitation and responsibilities of teaching staff and students connected to the teaching process (pp. 3-4) and making academic works if using Gen-AI tools (p. 6). Detailed presentation of allowed and forbidden tasks to be accomplished by using Gen-AI tools (pp. 3-4, 7-8). Description of application of Gen-AI tools in the publication process at UKSW Publishing House at various stages of the process (pp. 9-10). **Strengths:** + Detailed list of allowed/forbidden activities in presented areas and justification of possible threats if applying Gen-AI tools + Focus on the importance of relations between people, which can be strongly limited if using the AI Tools + The need for academic discussion according to the ethical aspect of AI Tools in HE + Suggestion to generate sample answer from the Gen-AI System to help the teacher detect unauthorised usage of Gen-AI tools in student's works when performing verification process of the student's works **Weakness:** - Lack of some terminology explanation – e.g. ethics

		behaviour, academic honesty, highest standard of academic ethics - No metrics/measures provided to measure quality/quantity of relation between the people - There is no detailed instruction on how the verification process should be done if answers from Gen-AI tools are generated or if the teacher wants to compare student work with the content generated by the AI tool - No consequences description if ethic behaviours are violated by inappropriate usage of Gen-AI tools - No further details according to academic discussion in the "Ethics" area of the Gen-AI tools application
Data Privacy & Security	All people should keep in mind the data privacy and security issue when applying Gen-AI tools, especially in the area of: classify, top secret or sensitive data. Moreover, generating new content can violate the copyright issue. All activities involving Gen-AI tool application should follow the GPRD regulation.	**Coverage:** Limitations according to applying Gen-AI tools if using: classify, top secret, or sensitive data (p. 3). Copyright issues when new content is generated (p. 2, 9). Data security according to the GPRD if applying Gen-AI tools (p. 4, 11). **Strengths:** + Important direct explanation why classified, top secret, and sensitive data cannot be used when applying Gen-AI tools + Copyright issue included, plagiarism involved + Focus on the importance of getting clear permissions to use different people's data when used in Gen-AI tools + The need for academic discussion according to the privacy & security aspect of AI in HE **Weakness:** - There are no direct directions according to the issue that Gen-AI tools should meet requirements provided by the GPRD directive - No information about the consequences if the Gen-AI tool was used in an improper way and causes a violation of data privacy & security - No further details according to academic discussion in the "Data Privacy & Security" area of the Gen-AI tool application
Equity & Inclusion	All Gen-AI tools should be available to all people without the exclusion of any group. This must be monitored at the institutional level. Users of Gen-AI tools should be aware that generated results can be	**Coverage:** The rules and recommendations presented according to equity and inclusion aspects of applying Gen-AI tools concern all academic works made by academic teachers, PhD students, and students (p. 3, 7, 11) **Strengths:**

	burdened with some ideological aspects – it is important to know what the reasons are.	+ Detailed information why and what kind of ideological burden can be included in the Gen-AI knowledge base model that causes the generated answers to exclude or favour a specific group of interests + Confirmation that Gen-AI tools can: enhance teaching effectivity, fulfil individual student needs, help to create personalised teaching paths + Introduction of institutional (University) responsibility for monitoring if the application of Gen-AI tools will not exclude any student's or PhD student's group + The need for academic discussion according to equity & inclusion issues if applying Gen-AI tools in HE **Weakness:** - No detailed information on how institutional monitoring of applying Gen-AI tools should be performed to cheque exclusion issue - No further details according to academic discussion in the "Equity & Inclusion" area of Gen-AI tools
Transparency & Explainability	All academic activities involving Gen-AI usage should be performed in terms of transparency rules. It should focus on reporting issues to make the reader know, what elements of the work are Gen-AI tool generated. On the other hand, Gen-AI tools can also be used by teachers to improve the evaluation process, and here the issue of transparency and explainability should also be addressed.	**Coverage:** Focus on how to provide transparency rules in teaching process, academic works creation process, research paper as well as publications published by UKSW Publishing House (p. 3, 4, 5, 6, 9, 10) **Strengths:** + Importance of reporting Gen-AI tools usage in teaching, research and publishing area is strongly highlighted + Prompts and answers used by Gen-AI tools should be provided to get full transparency & explainability + Suggestion of applying various Gen-AI tools to help a teacher cheque whether student's or PhD student's work is original and individual + Importance of personal verification if auto-evaluate Gen-AI systems are used to evaluate student's or PhD student's work results + The way to explain what content has been generated by Gen-AI tools is proposed. **Weakness:** - Specific transparency rules are not provided in detail, all issues (despite the last) presented in "Strengths" are not precisely described in the document. They are only stated in general form without explanation.

Governance Structures	The rules and recommendations proposed in the document concern all people involved in the University. They are some duties according to Gen-AI tools usage spread to academic teachers and University to specify what activities should be made according to governance issue.	**Coverage:** Entities responsible to follow the rules and recommendations are: University authorities, teaching staff, students, and PhD students (p. 1, 7, 11) **Strengths:** + Recipients of rules and recommendations are specified + The division of responsibilities is provided between various stakeholders + Different activities according to Gen-AI management have been specified and assigned to stakeholders **Weakness:** - The activities connected to management of used Gen-AI systems are only mentioned without further specification how they should be performing, e.g.: no information how data gathered should be secure, how Gen-AI-tools systematically analyse should be done, how technical and ethical problems with AI tools should be reported and to whom. - No consequences for violating rules provided in the document were specified in details, only general remarks are presented.
Technical Standards & Interoperability	Important academic works should be inspected by using specific software like Anti-plagiarism system, which should be able to detect improper usage of content generated by Gen-AI tool. It can also participate in typical academic works if the teacher suspects that the student's work is not individual.	**Coverage:** The interoperability between systems is mainly related to the final and doctoral thesis, but also less formally to the academic works of the student and the PhD student (p. 6) **Strengths:** + Recommendation to verify student's independence in various academic works using dedicated anti-plagiarism systems equipped with AI detection feature **Weakness:** - There is no detailed information on how the verification process should be made - No technical data according to different IT systems cooperation was provided to ensure that the detection of improper Gen-AI content will be possible
Accountability & Liability	The teacher who does not follow the rules according to the Gen-AI tools application can be punished disciplinarily. The student who is suspected to commit not independent	**Coverage:** Following the rules provided in the document concerns academic teachers, researchers, students, and PhD students (p. 5, 6, 8, 9, 11) **Strengths:**

	academic work can be called to provide explanations and his work is suspended until investigation is over. All authors are responsible for ethics and liability violations and can be punished disciplinarily. The rules provided do not exclude other directives, regulations, and laws in applying Gen-AI tools in teaching, scientific research, academic work, and publication of articles.	+ List of subjects who are responsible for following the regulations presented in the document **Weakness:** - No detailed description of consequences because of regulations violation - Simplified and general way of proceeding in a controversial issue
Sustainability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document

Category	Key Extracts (Doc IDs) D4	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	The goal of applying Gen-AI tools in HE should support the teaching process as well as extend the digital competences of students and teaching staff. It is important to indicate what Gen-AI tools application is allowed in the curriculum. The rules provide what sample activities connected to Gen-AI tools are allowed. Content generated by Gen-AI systems is not considered as an independent student's work if systems are used to generate texts for master's, engineering, bachelor's, doctoral theses, papers, student projects, coursework, articles, and all	**Coverage:** The directive focusses on teaching process only (p. 1), no research or publication area is included. The rules presented (p. 1), recommendations (p. 2), and limitations (p. 3) only concern students and the academic teacher (p. 1). **Strengths:** + Specific examples of allowed activities for Gen-AI tools are provided + Recommendations for limiting academic works that can be easily accomplished by using Gen-AI tools + Recommendations for the preparation of academic works related to: practical knowledge and/or critical thinking - a sample of those works is provided + Recommendations for the application of Gen-AI content detectors by the teacher with named samples services **Weakness:** - There is no research/publication area involved in the directive

	other content, unless the course guidelines allows to use AI systems in that way (e.g., classes on prompt creation or AI applications, etc.). There is also need to know what main limitations of Gen-AI tools are e.g. hallucinations. The usage of Gen-AI tools is strictly forbidden during all verification activities such as exams or tests. It is prohibited to use AI technologies to falsify identities, create or use deepfakes, synthetic voices, images and content.	- There is no justification why some activities are allowed while others are prohibited
Data Privacy & Security	It is forbidden to enter personal, classified and top-secret data, as well as other information that is restricted to copy-right.	**Coverage:** The directive is very limited according to the area of data privacy & security (p. 3) **Strengths:** + Sample data types forbidden to use with AI systems are provided + Links to other regulations, directives, or laws are provided **Weakness:** - The proposed application of Gen-AI content detectors (Ethics & Integrity) can violate data privacy & security due to the way the detectors work - There is no justification why some specified data cannot be used with Gen-AI tools - Links to other regulations are very general; there is no list of specific documents the application of AI systems according to the data privacy & security area should follow
Equity & Inclusion	The application of AI systems allows students to improve their competencies in the teaching process such as inspiration, innovation, critical thinking, and digital competencies. It is recommended to share information about the ethical, secure, and responsible use of	**Coverage:** The directive focusses on equal access to Gen-AI tools and share information about it according to advantages when applying it (pp. 1-2) and some limitations (p.2). **Strengths:** + Very limited but present information explaining why Gen-AI tools can be useful to equal opportunities between different people **Weakness:**

	the Gen-AI system among all students and academic teachers. The knowledge and abilities of all academic teachers should be continuously extended, especially in terms of the quality of the prompts Users of Gen-AI tools should be aware that the generated content can be affected by bias, non-ethical, or negative stereotypes.	- There is no justification why Gen-AI tools can generate results which can violate equity & inclusion aspect - There is no specific information on the courses for people working with AI systems
Transparency & Explainability	The application of Gen-AI tools in academic work should be transparently and clearly marked: a) in the case of a thesis - in the introduction or in the chapter where research methods are presented, b) in other academic works - in the introduction or bibliography. The marking should include the name of the AI system (e.g. Chat GPT), a link to the system, the date when the content was generated, the purpose of use and, if applicable, how much material was generated by the AI system and information on if the fragments of generated content was used in its original form or was modified by the student.	**Coverage:** The directive forces only how different content (e.g. diploma thesis and passing works) should be marked if it involves Gen-AI generated content (pp. 1-2). **Strengths:** + The way of marking according to diploma thesis and passing work is defined + Precise information about how the marking should be made **Weakness:** - No specification what type of academic work should be marked when using the AI tool and what is the proper way to do this - Transparency & explainability area is not involved if other type of academic activities are considered, e.g.: generating teaching contents, samples, exercises...
Governance Structures	Detailed rules and scope for the application of AI systems should be provided by the academic teacher. All of these should be incorporated into the syllabus	**Coverage:** The directive focusses on the academic teacher as the main governance structure (p. 1), who formulate the detailed rules, scope, and limitation of applying the AI systems during the classes and is responsible for verification (p. 2). **Strengths:**

	or presented during the first classes. The verification process of academic work is up to the academic teacher.	+ The way in which specific rules, scope, and limitations are proposed in applying AI systems in teaching is proposed **Weakness:** - No detailed recommendation or examples that help the teacher design detailed rules of applying the AI system during the teaching process - No official template what elements according to the rules, scope, and limitation of applying AI system during the classes should be presented
Technical Standards & Interoperability	There is a general recommendation to use generated content detectors to help the teacher identify if the student's work is not individual.	**Coverage:** Only general recommendation to use the detectors, which can help to identify the AI systems generated content, which result in interoperability of the different system, but in manual way (p. 2) **Strengths:** + Sample detector names are provided **Weakness:** - No specific technical standards were formulated - No integrated solution was established - No verification process is presented
Accountability & Liability	The academic teacher should inform students according to ethical and legal consequences if Gen-AI systems are applied, especially in forbidden way. The responsibility for no individual academic work or citation to inexistent source in diploma thesis or passing work is up to the student. If the diploma thesis or passing work violates the rules covered in the current directive, there are consequences applied described in study regulations. Serious violation of the regulation will result in disciplinary penalty according to the study regulations and other legal regulations.	**Coverage:** The rules according to accountability and liability concern academic teachers (p. 1,3 7) and students (p. 2, 3) **Strengths:** + Duties of academic teachers are defined, + Responsibilities of students and teachers and consequences if the directive rules are violated with relation are defined in relation to other legal acts, **Weakness:** - There is no explanation what "serious violation of regulation" is,

Sustainability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document

Category	Key Extracts (Doc IDs) D5	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	The application of AI systems will allow students, PhD students, postgraduate students, and academic staff to solve complex theoretical and practical problems and develop their intellect in a much more efficient way. However, the way the AI systems work can bring many threats such as disinformation, problems with intellectual honesty, and with copyrights. An inacceptable application of AI systems will be harmful in creating various academic works in an independent way. The directive presents the rules of applying AI according to the written assignments in SGH. It is important to know that AI systems are only an auxiliary tool, but not the substitution of people creativity.	**Coverage:** High focus on presenting detailed rules according to the following area of academic activity (HE): conceptual, literature review, writing process, text editing, generating graphics, programming, data analysis, mathematical and economic modelling where Gen-AI tools can or cannot be applied (pp. 2-5). The rules concern students, PhD students, postgraduate students, and academic staff (p. 1). **Strengths:** + Specification of nine areas where AI systems can be potentially applied + Each area contains a list of specific activities which are allowed or forbidden if the AI system is applied. + In some areas, there are also some recommendations and/or remarks when using AI systems **Weakness:** - The directive concerns mainly written academic assignments related to teaching activities - No specific rules and tasks description according to the scientific research or publication process - There is no extended justification of threats if AI systems are used in an improper way - No consequences are defined if the ethical aspect of applying the AI system is violated
Data Privacy & Security	The application of AI tools can violate copyrights.	**Coverage:** Data Privacy & Security aspect of applying AI tools is very limited in the directive: only copyrights are mentioned (p. 1). **Strengths:** Not covered in the document **Weakness:** - The aspect of data privacy and security is very limited in the directive - violation of copyrights is barely mentioned

		- There are no links to other laws and regulations that govern data privacy and security - No consequences are defined if the data privacy and security aspect of applying the AI system is violated
Equity & Inclusion	Applying AI systems can be beneficial to all academic societies: students, PhD students, postgraduate students, and academic staff in terms of effective development of intellect and solving complex theoretical and practical problems.	**Coverage:** Equity and Inclusion aspect of the application of AI tools is limited in the directive. Only group of interest and basic justification, why AI systems can be beneficial are presented (p. 1). **Strengths:** + Recommendation of applying AI tools to equalise the chances of different people according to the intelligence to develop and solve complex problems **Weakness:** - There is no information about the bias in the AI tools and the source of it - No recommendations to allow equal access to AI tools with no exclusion of any groups - No consequences are defined if the equity and inclusion aspect of applying the AI system is violated
Transparency & Explainability	Because of the need for transparency if applying the AI tools, the directive contains a dedicated section about how to report AI tools applications. Reporting needs to provide a detailed description of the way and scope of applied AI tools in the separate section of the written assignment. The appropriate information should be placed near the elements of the work, which were created with the help of AI tools using footnotes and/or descriptions of the graph/model. In addition, reporting needs to store the applied prompts and received responses from the AI tools, which can be shared with the supervisor of academic work.	**Coverage:** Transparency & Explainability aspect of applying AI tools is limited only to reporting issues if written assignments are considered (p. 2, 5). **Strengths:** + The way in which reporting should be provided in the written assignment is presented **Weakness:** - No information about what exact elements should be provided when reporting, - Transparency & Explainability aspect presented in the directive concerns only written assignments prepared by a student; no academic teacher (teaching process) nor researcher (scientific research process) role is considered for AI tools application

Governance Structures	The rules and limitations provided in the directive are imposed by the SGH Rector. Besides: supervisor, academic teacher, program council of the field of study, deans of: bachelor's degree, master's degree, doctoral studies, and directors of postgraduate studies are authorised to introduce new rules and limitations according AI tools application. Academic teachers or supervisors of academic works are authorised to set an additional way of reporting the AI Tool application.	**Coverage:** The directive lists all roles which are authorised to provide new rules and limitations according to applying AI tools during making written assignments (p. 1, 5). **Strengths:** + A quite long list of people authorized to incorporate new rules and limitations in the AI tools application area causes high level of flexibility **Weakness:** - No connections between rules and limitations proposed by different people, which potentially will lead to inconsistency and some chaos when applying AI systems at different fields of study and various subjects - No hierarchical dependency between various source of rules and limitations is provided
Technical Standards & Interoperability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document
Accountability & Liability	The full responsibility for the created content is up to the author, who needs to be careful and use a conscientious judgement if applying AI tools	**Coverage:** The Accountability & Liability aspect of applying various AI tools in written assignments is highly limited only to the author, and soft and very general criteria are formulated (p. 2). **Strengths:** + Precise point out who is fully responsible for applying the AI tools in written assignments **Weakness:** - There are no definitions of what "careful" and "conscientious judgement" mean in practice - No consequences are defined if the accountability and liability aspect of applying the AI system is violated - No links to other legal regulations and directives
Sustainability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document

		Weakness: Not covered in the document

Category	Key Extracts (Doc IDs) D6	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	AI brings new possibilities to teaching and research areas. Technology should be used in a creative, responsible, and conscious way regarding academic honesty. Conscious use of AI tools helps to build a better future where man and technology will support each other. AI should support education and development in the way that ethical standards and the relationship between people in the education process are maintained. Specific recommendations for students and teaching staff are presented. The AI generated content should be verified in a careful way to cheque if source documents exist in reality and the content can be developed further. It is connected to hallucinations phenomenon. Students should be aware that some assignments would be accomplished by using AI, but to cheque student abilities and/or learn a new way to solve the problem, the application of the AI tools is forbidden. Key student competencies cannot be achieved if AI tools are applied too often. In addition, if AI models are not	**Coverage:** The document is focus on students and academic staff to support key competences development like: independence, reliability, and critical thinking with the usage of AI tools (pp. 2-4) **Strengths:** + Sample specific AI applications and platforms are presented according to their functionality: text creation, image creation, prepare the presentation, sound/music production, programming support, research support + Specific recommendations for students and academic staff are directly explained in terms of advantages and threats + Example of switching the typical evaluation process of the student's work into evaluation regarding the AI tools application + Samples of non-ethical activities which can be found in HE area **Weakness:** - Recommendations are described in various ways, no template is used - The structure of the document is not so well: allowed and forbidden examples of applying AI tools are mixed through whole document, so it is hard to gather all valuable information together - No consequences are defined if the AI tools are applied in an unequally responsible way

	available, the student will not be able to perform even basic tasks. The teacher should use the AI responsibly. The teacher should show how the AI application is valuable in the teaching process. The teacher should point out the restrictions and consequences of the irresponsible application of AI. The teacher should adapt current evaluation methods to consider the possibilities of AI application. The conscious and responsible applying the AI tools should choose the right balance between technology support and the contribution of the work itself.	
Data Privacy & Security	The application of AI tools should be avoided if personal, classified, or copyright data are used.	**Coverage:** Limited only to basic information on the data privacy and security issue if applying the AI tools (p. 3) **Strengths:** + The data privacy & security issue is mentioned in the document **Weakness:** - No citations to other directives or regulations governing the data privacy and security area (e.g. GDPR) - No consequences are described if the data privacy and security area is violated
Equity & Inclusion	All students will develop their key competencies. The output content generated by AI tools can be affected by bias due to the characteristic of the training data. Applying the AI tools should consider various perspectives:	**Coverage:** General and brief view of equity and inclusion aspect of applying the AI tool in HE (pp. 1, 3, 4) **Strengths:** + The equity and inclusion area is addressed in the document **Weakness:** - There is no more detail on the aspect of equity & inclusion when applying the AI tools

	social sensitivity includes, e.g.: equity and inclusion.	- No consequences are defined if the equity and inclusion aspect of applying AI tools is violated
Transparency & Explainability	The application of the AI tools by the student should be transparent and reported (marked) if the application is beyond the spell cheque. The teacher should maintain transparency (see the student's recommendations) and equilibrium when interacting with the AI. Any restrictions and recommendations according to the AI application should be fully justified by the academic teacher. The AI application should be fully and transparently presented by: a description of the goal and scope which justified why and how the AI tools were applied.	**Coverage:** Focus on student and teacher role and their basic tasks if applying the AI tools (pp. 2, 3) **Strengths:** + Emphasis on justification and understanding the issue of transparency and explainability when applying AI tools **Weakness:** - No detailed information provided on how different types of AI generated content should be reported or marked, - The level of interaction between men and AI is defined in an arbitrarily way - There is no more information on the possible or sample goals and scope description in the document (it can be founded in the D7 document) - No consequences are described if the transparency and explainability area is violated
Governance Structures	Students should follow the restrictions set by the academic teacher. The academic teacher should present restrictions and recommendations according to the AI application in an obvious way.	**Coverage:** Coverage limited only to academic teacher as a subject responsible and authorised to formulate rules and limitations according to the application of the AI tools in HE (pp. 2, 3) **Strengths:** + The governance structures issue is addressed in the document **Weakness:** - No other institutions or regulation bodies are specified in the document, even on local, university level - The way of presentation of rules, restrictions, and recommendations by academic teacher is not specified in details
Technical Standards & Interoperability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document

Accountability & Liability	The author is responsible for the generated content and the author should verify it carefully.	**Coverage:** Focus on student responsibility for application of AI tools (p. 2), which must be consistent with University's study regulations. The document is very limited according to the accountability and liability aspect of applying the AI tools (pp. 2, 3) **Strengths:** + Citation to AGH study regulations that can address the accountability and liability aspect of dishonest work **Weakness:** - No other formal, legal documents are specified in the document - The way of verification of generated content is not described in detail. - No consequences are described if the accountability and liability aspect of applying AI tools is violated
Sustainability	The issue not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document

Category	Key Extracts (Doc IDs) D7	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	Depending on the chosen level of application of AI in the teaching process, various activities will be allowed or forbidden. It is important to know that every proposed level of integrating AI into the education process will develop specific abilities and bring specific benefits to students.	**Coverage:** The document mainly addresses students who can use the AI Tools at different stages of integration (pp. 1-3) **Strengths:** + Proposed levels are characterised by: description, main goal, developed skills, gained benefits + At each level, concrete examples of tasks/activities are presented + Proposed levels are equipped with icons that will be useful for marking artefacts created with or without the AI tools support **Weakness:** - No threat description connected to the various levels of applying the AI tool
Data Privacy & Security	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document

		Weakness: Not covered in the document
Equity & Inclusion	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document
Transparency & Explainability	The applying AI tools for the editorial task level does not require reporting. The same is true for no restriction level. In contrast, reporting is essential if planning with AI, cooperating with AI, or meeting AI activities are performed.	**Coverage:** The document points out what levels of application the AI tool needs reporting (pp. 2-3). **Strengths:** + Each proposed level of applying the AI tool is characterised according to the reporting needs **Weakness:** - No more details on the transparency & explainability aspect of applying the AI tools in addition to the reporting issue
Governance Structures	If planning with the AI, cooperation with AI or meeting the AI activities are performed, the students are obliged to receive agreement from academic teacher	**Coverage:** The academic teacher is mentioned as a governance structure in some activities connected to the applying the AI tool (p. 2). **Strengths:** Not covered in the document **Weakness:** - There are no specific details on how the agreement should look like - No more information on the governance process itself - No other roles or institutions are mentioned to provide governance according to the application of AI tools
Technical Standards & Interoperability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document
Accountability & Liability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:** Not covered in the document **Weakness:** Not covered in the document
Sustainability	The issue is not presented in the document	**Coverage:** Not covered in the document **Strengths:**

Project: "Promoting Accountable and Transparent Human-centered AI in Higher Education" Project n. 2025-1-PL01-KA220-HED-000359600
Co-funded by the Erasmus+ Programme of the European Union

		Not covered in the document **Weakness:** Not covered in the document

Category	Key Extracts (Doc IDs) E1	Observations (coverage, strengths, weaknesses)
Ethics & Integrity	The document treats Ethics and Integrity as one of the three pillars of Trustworthy AI (alongside lawfulness and technical robustness). Ethical challenges related to the use of AI in society concern its impact on people, human decision-making capacity, and safety. The document formulates four ethical principles, rooted in fundamental rights, that must be respected by AI systems: Respect for Human Autonomy, Prevention of Harm, Fairness, and Explicability. AI ethics is defined as a form of applied ethics addressing issues related to the development, deployment, and use of AI, with the aim of assessing AI's impact on the *"good life,"* human autonomy, and democracy. The document places strong emphasis on an ethical mindset. The development of ethical AI is presented as a continuous process of reflection, supported by education, public debate, and organizational culture. It also highlights the importance of education and stakeholder participation as prerequisites for sustainable AI ethics and an ethical mindset. Stakeholders include AI developers, users, and all individuals affected by a given AI solution. EU fundamental rights and human dignity constitute the foundation of AI ethics. Ethics and Integrity are closely linked to a human-centric approach. Human dignity is described as an absolute and non-negotiable value in the design and deployment of AI systems. This requires unconditional protection of an individual's physical, psychological, identity-related, and cultural integrity, as well as the capacity to recognize situations in which the development or deployment of AI should not proceed if it violates absolute values such as human dignity. Ethics should stimulate reflection on both the	**Coverage**: Very broad: ranging from axiological foundations and ethical principles to governance mechanisms, education, and the boundaries of acceptable technological use. **Strengths**: • Assessment List on fundamental rights, human agency and human oversight that can be applied to AI systems (p. 26), • strong grounding in fundamental rights and human dignity, • clear distinction between legality and ethical acceptability, • explicit recognition of value tensions and ethical conflicts, • clearly defined limits of ethical compromise (human dignity is non-negotiable), • emphasis on ethical culture and mindset, not solely on regulation, • linking human autonomy to specific, identifiable technological risks. **Weaknesses**: • high level of abstraction of ethical principles, • lack of hard enforcement mechanisms, • limited sector-specific guidance (no concrete examples for different industries),

	protection of individuals and social groups and the support of innovation. In cases of value conflicts between AI system requirements, a rational and well-documented ethical risk analysis should be conducted. Ethics and Integrity require transparency of decision-making and a willingness to suspend or terminate an AI project when no ethically acceptable solution exists. The document identifies automatic identification of individuals and AI-enabled surveillance as areas of particularly high ethical risk. Other significant risks are associated with human-like AI systems and robots. Ethics and Integrity also encompass the protection of the value of *"being human"* against symbolic and social erosion.	• absence of sanctions for ethical violations (beyond the recommendation not to deploy unethical systems).
Data Privacy & Security	Privacy and data governance are clearly identified in the document as one of the seven pillars of Trustworthy AI. A key principle is the prevention of harm to individuals caused by AI systems, which translates into an obligation to implement active safeguards against misuse and to anticipate potential abuses or inappropriate uses of AI systems. Data protection and the security of AI systems must be ensured throughout all stages of the AI system lifecycle. The document recommends applying a privacy by design and security by design approach as a standard design principle for AI. This means that privacy and security should be embedded in the system architecture from the very beginning. A particularly important recommendation is the implementation of mechanisms for the safe shutdown of systems and their controlled restart after incidents. Testing of AI systems should be continuous, multi-dimensional, and conducted by diverse teams, and should also include adversarial testing. The document also points to the possibility of involving external entities through bug bounty programmes, in order to increase system resilience to attacks and errors. It recommends implementing safeguards covering data, AI models, and technical infrastructure to	**Coverage**: The document covers privacy protection, data governance, technical security, organisational security, testing and validation mechanisms, crisis response, and risk assessment. **Strengths**: • Assessment List on technical robustness and safety that can be applied to AI systems (p. 27), • Assessment List on privacy and data governance that can be applied to AI systems (p. 30), • integration of legal, technical and social perspectives, • coverage of the entire lifecycle of AI systems, • technical recommendations, • strong grounding in the EU fundamental rights framework, • privacy-by-design and security-by-design are

	protect against vulnerabilities and security weaknesses. AI systems should be designed to be resilient both to technical failures and to negative societal impacts, such as loss of trust, erroneous decisions, or misuse. The document requires strict access control to personal data, limiting access exclusively to individuals with appropriate competencies and a justified need. It introduces an obligation to test and document data at every stage of the AI lifecycle, including in the case of systems procured from external suppliers. The document introduces the requirement for fallback plans as an integral component of secure AI systems, in order to minimise the risk of AI systems being used in ways that are inconsistent with their original purpose. It also recommends defining measurable quality and security indicators for AI systems.	presented as necessary foundations for AI design. **Weaknesses**: • the soft-law nature of the document, • lack of sector-specific requirements.
Equity & Inclusion	Equality and non-discrimination are defined as one of the seven key requirements of trustworthy AI. The document clearly states that AI systems require particular caution in contexts involving vulnerable groups, such as children, people with disabilities, or historically marginalised populations. It also emphasises the need to analyse power or information asymmetries, for example in employer–employee or company–consumer relationships. Inclusiveness should be ensured at all stages of the AI system lifecycle – from design to deployment. The diversity of teams developing and implementing AI is identified as a key factor for the quality and fairness of systems. This includes both demographic diversity and interdisciplinary competencies. The document explicitly recommends applying universal design principles to enable everyone to use AI, regardless of age, gender, ability, or background. The guidelines clearly identify sources of bias in data (historical biases, incompleteness, poor data governance) and emphasise the need for oversight and transparent analysis of system decisions. Particular emphasis is placed on the inclusiveness of training data and on addressing	**Coverage**: vulnerable groups and the risk of exclusion, non-discrimination and algorithmic fairness, power and information asymmetries, inclusive data and universal design, diversity of design and development teams. **Strengths**: • Assessment List on diversity, non-discrimination and fairness that can be applied to AI systems (pp. 29-30), • strong grounding in EU fundamental rights, • clear identification of vulnerable groups, • comprehensive approach to fairness (data, teams, deployment), • linking inclusiveness with the AI lifecycle, • consideration of socio-economic aspects. **Weaknesses**:

	the needs of groups at risk of exclusion, in order to prevent the reproduction of historical inequalities. The document also highlights the principle of fairness, defined both as the absence of discrimination and as the fair distribution of benefits and costs resulting from the use of AI. Appropriately designed AI systems can actively increase social equality, for example by improving access to education, services, and technology. The principle of harm prevention, which should apply to AI systems, includes the obligation to provide increased protection for people vulnerable to the negative effects of AI.	• lack of measurable criteria for assessing inclusiveness, • lack of concrete methods for testing systems for inclusiveness, • absence of sector-specific deployment guidelines (e.g., for high-risk sectors), • limited guidance on enforcement of the principles, • limited discussion of access to AI technologies for low-income populations or people living in smaller towns, • the soft-law nature of the document.
Transparency & Explainability	Explainability is recognized as one of the four fundamental ethical principles of AI. Transparency, in turn, is formally identified as one of the seven key requirements of Trustworthy AI. Transparency is defined broadly — it covers data, the AI system itself, and the business models in which it is embedded. Transparency should enable users to make informed and autonomous decisions. This means not only providing information that the user is interacting with an AI-based decision, but also providing tools that allow users to understand, evaluate, and — where possible — challenge the system's operation. AI systems should not impersonate humans, and users have the right to know that they are interacting with AI. The guidelines indicate that transparency also includes informing users about the system's accuracy level and the likelihood of errors, especially when they cannot be completely eliminated. Communicating the capabilities, limitations, and accuracy level of the system should be adapted to the specific use case. In contexts such as citizen scoring (e.g., in education), full transparency is required,	**Coverage**: transparency of systems, data, and business models; communication with users; informed decision-making; Explainable AI (XAI); high-risk systems (e.g., scoring). **Strengths**: • Assessment List on accuracy, traceability, explainability, communication that can be applied to AI systems (pp. 27-29), • Assessment List on technical robustness and safety that can be applied to AI systems (p. 27), • clear connection between transparency and trust as well as accountability, • consideration of black-box systems, • broad understanding of transparency (data, system, business),

	including information about the system's process, purpose, and methodology. Special emphasis is placed on traceability and auditability, especially in critical contexts. The document strongly emphasizes documenting data, processes, and algorithms to enable tracing AI decisions. Traceability should support both auditing and explaining errors, as well as preventing their recurrence. Explicability is considered a condition for trust and the ability to challenge AI decisions. Technical explainability means that AI decisions can be understood and traced by humans. For black-box systems, alternative explainability measures are required (e.g., auditability, documentation, communication). The document recognizes Explainable AI as a key research area necessary to understand AI systems' behavior and build trust.	• strong grounding in user rights to information and appeal, • requirement to communicate system limitations, • clear indication that AI systems must be identifiable. **Weaknesses**: • lack of precise technical standards for explicability, • limited implementation guidance for practitioners, • lack of differentiation of obligations across sectors, • lack of detailed technical guidance on XAI, • limited discussion of the trade-off between explainability and accuracy, • the assessment list (Chapter III) contains questions about explainability but they are very general, • soft-law nature of recommendations.
Governance Structures	The guidelines recommend that the assessment of Trustworthy AI should be embedded in existing governance structures of the organization or carried out through new governance processes tailored to the organization's size and resources. A key factor for effectiveness is the involvement of top management, combined with engagement at the operational level and a broad group of stakeholders, which increases acceptance and the real impact of implemented principles. The document explicitly identifies governance frameworks as a key mechanism for ensuring ethical accountability. Both internal and external structures are recommended, including appointing persons responsible for AI ethics or establishing dedicated advisory bodies (committees, panels, or ethics councils).	**Coverage**: Very limited information, only mentioning the integration of Trustworthy AI into existing organizational structures and the importance of different management levels. **Strengths**: • Emphasises the role of top management in effective implementation, • Flexible approach (adaptation to the size and resources of the organisation), • Emphasis on involving stakeholders and operational level staff,

		• Requirement for multi-level organisational management. **Weaknesses**: • Limited reference to control mechanisms and sanctions, • Lack of guidance on the composition of an ethics board — who should be involved and what their responsibilities should be, • Chapter III assessment list contains questions but they are too general, • No guidance on conflict of interest management within governance structures, • No discussion of cooperation between governance structures of different organisations, • Soft-law nature of recommendations without enforceable authority.
Technical Standards & Interoperability	The document emphasises that the technical design of AI systems should not assume a single universal user. AI systems should interact with assistive technologies and accommodate different user profiles. The document highlights the importance of established design standards, such as privacy-by-design and security-by-design. The requirement for resilience against attacks, fail-safe shutdown, and the ability to resume operation after an incident indicates the need for interoperability with security infrastructure. Proposed tools include white lists (behaviours or states that the system should always follow), black lists (restrictions on behaviours or states that the system must never transgress), or more complex formal behavioural guarantees. Compliance monitoring processes for AI systems with these guidelines are also necessary. The document introduces the idea of measurable quality-of-service indicators as a technical	**Coverage**: the document covers standards and interoperability by referring to Trustworthy AI architectures, "by-design" design standards, quality indicators and technical metrics, as well as AI standardization and certification. **Strengths**: • Promoting formal, measurable quality criteria, • Supporting international harmonization (ISO, IEEE), • Privacy-by-design and security-by-design, • Fail-safe shutdown requirement, • Certification as a mechanism to confirm trustworthiness.

	reference point for evaluating AI systems. The reference to classical software engineering metrics (performance, reliability, security, maintainability) is an attempt to integrate AI with existing ICT standards. An important issue addressed in the document is certification, which should apply standards developed for different application domains and AI techniques, appropriately aligned with industrial and social standards and diverse contexts. References to ISO and IEEE P7000 demonstrate an effort to harmonise technical and ethical AI standards and to create a common language for interoperability. The document emphasises the obligation to clearly inform users about human–AI interaction. Systems must be designed in such a way that they disclose, in a standardised and verifiable manner, that AI is being used.	**Weaknesses**: • Lack of references to compatibility between AI systems from different providers, • Lack of specific standards for interoperability between AI systems, • No discussion of API standards or data format standards, • No guidance on how to test compliance with ISO/IEEE standards, • Certification is treated as optional — there is no mandatory requirement.
Accountability & Liability	The document establishes accountability as a requirement for Trustworthy AI. Accountability should be institutionally anchored within governance structures. The appointment of roles, ethical committees, or external panels indicates the need for formal governance frameworks that enable responsibility to be assigned for AI-related decisions. The document expands the concept of accountability to include active management of the impacts of AI. Organizations using AI systems are required not only to respond to harms, but also to identify, document, and mitigate them, especially from the perspective of individuals affected by AI decisions. Redress mechanisms are recognized as a condition for trust in AI. Accountability means ensuring accessible and effective remedies, with particular attention to vulnerable groups. The document emphasizes the principle of proportionality: the less human control there is over AI decisions, the more rigorous the testing, governance procedures, and ex-ante accountability mechanisms must be. The document highlights the key connection between accountability and the ability to challenge AI	**Coverage**: the document covers issues such as ex ante and ex post accountability of AI systems, the four components of accountability, human oversight of AI, and the right to appeal decisions made by AI. **Strengths**: • Assessment List on reliability and reproducibility that can be applied to AI systems (pp. 27-28), • Assessment List on accountability that can be applied to AI systems (p. 31), • Clear definition of the four components of accountability (auditability, minimisation of negative impact, trade-offs, redress), • Requirement for both internal and external auditability — especially for systems affecting fundamental rights,

	decisions. Effective redress requires identification of the responsible entity and explainability of the decision-making process. Accountability encompasses responsibility both before the deployment of an AI system and after its launch. Reproducibility and reliability of AI systems are presented as technical prerequisites for accountability. Without repeatable results and stable operation, effective auditing, responsibility assignment, and harm prevention are not possible. Auditability is a key tool for enforcing accountability. Certification can support accountability by confirming compliance with specific standards, but it cannot replace it.	• Protection of whistleblowers reporting on AI systems, • Requirement to document trade-offs between requirements imposed on AI systems, • Requirement for redress mechanisms, • Requirement for continuous review. **Weaknesses**: • Lack of precise assignment of legal responsibility, • Limited references to liability in the civil law sense, • Lack of sanctioning mechanisms, • No discussion of third-party liability, • Lack of clarity on who is responsible for what, • No guidelines on liability insurance, • No procedures for redress, • No discussion of causality — how to establish that an AI system caused harm.
Sustainability	The document clearly places social and environmental well-being among the core requirements of trustworthy AI and emphasizes the role of AI in achieving the United Nations SDGs. AI is expected to support, among other things, gender equality, the sustainability of healthcare systems, combating climate change, rational use of natural resources, and monitoring progress in social cohesion and sustainable development. Trustworthy AI is presented as a tool that supports the well-being of individuals and entire societies through economic growth and improved quality of life. Responsible AI design requires considering the impact of systems on the natural environment	**Coverage**: the document addresses social and environmental well-being, the SDGs, the ecological responsibility of technology, the long-term social effects of AI use, potential sectors in which AI may be applied, and concerns related to energy and resource consumption. **Strengths**: • Assessment List on societal and environmental well-being that can be applied to AI systems (pp. 30-31),

	and other living beings. The guidelines mandate an assessment of environmental impact not only of the AI system's operation but of the entire process of its creation and use, including the supply chain. Special attention is given to energy and resource consumption during model training and the need to choose less harmful alternatives. Trustworthy AI is identified as a tool that supports reducing human environmental impact through the optimization of energy and resource use. Examples include smart energy grids, public transport, and systems supporting decarbonisation and energy efficiency.	• strong linkage between AI and the 2030 Agenda and SDGs, • a positive, pro-innovation vision of sustainability, • a holistic approach to the AI life cycle and supply chain, • integration of social and environmental aspects, • requirement to assess energy consumption during training, • requirement to consider future generations, • requirement for social impact assessment. **Weaknesses**: • lack of measurable environmental impact indicators (e.g., carbon footprint, water footprint), • limited references to technological trade-offs, • lack of mandatory environmental assessments for AI, • absence of concrete commitments to renewable energy, only general guidelines, • lack of discussion on e-waste, • lack of discussion on competition between different sustainability goals (e.g., energy consumption vs. model accuracy), • lack of concrete guidelines on how to measure social impact.

Section 4: Gaps, Overlaps, and Inconsistencies

Purpose: To identify where your national/institutional frameworks are strong, missing, or contradictory.

Instructions: Use bullet points or short paragraphs to describe gaps, overlaps, and inconsistencies. Refer to specific document IDs when possible.

- **Gaps:** Elements not covered in any of the reviewed documents (e.g., no reference to AI in teaching).
- **Overlaps:** Common elements repeated across multiple documents.
- **Inconsistencies:** Conflicts between national vs. institutional policies or contradictions within documents.

Doc ID:	D1
Gaps:	• Lack of discussion of the risks of AI disinformation • No discussion/procedures regarding AI plagiarism/fraud in academic work • No specific standards/procedures for AI-based curricula • Superficial treatment of higher education (lack of concrete actions) • Lack of regulatory procedures for AI in universities • Lack of specific budgets for higher education • Lack of ethical oversight procedures for research • Lack of coordination mechanisms between universities • Lack of procedures for technology transfer from universities • Lack of alignment with the AI Act (2024) • Lack of specific indicators for higher education • Lack of implementation and enforcement mechanisms (soft law) • Lack of KPIs, timelines, and operational structures • Lack of reference to generative AI (due to time constraints)
Overlaps	• Education and competence development in the context of AI (overlaps with: D2, D3, D4, D5, D6, D7, E1) • Support for research and innovation (overlaps with: D2, D3, D5, D6, D7, E1) • Human-centred approach and respect for human dignity and rights (overlaps with: D2, D3, D5, D6, D7, E1) • Emphasis on the importance of international cooperation (overlaps with: D2) • Open data and open AI models (overlaps with: D2) • AI based on six pillars: society, business, science, education, cooperation, and the public sector (overlaps with: D2) • Ethics and responsibility in the design and use of AI (overlaps with: D2, D3, D4, D5, D6, D7, E1) • Transparency, explainability, and reporting of AI use (overlaps with: D2, D3, D4, D5, D6, D7, E1) • Equality, non-discrimination, and protection of vulnerable groups (overlaps with: D2, D3, D6, E1) • The importance of computing infrastructure (overlaps with: D2) • Identification of key sectors for AI development (overlaps with: D2) • Science–business technology transfer (overlaps with: D2) • Regulation and legal frameworks for AI (overlaps with: D2, D3, D4) • Author responsibility for AI-generated content (overlaps with: D2, D3, D4, D5, D6, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D2, D3, D4, D5, D6, D7, E1) • Data protection and privacy (overlaps with: D2, D3, D4, D6, E1)
Inconsistencies	• **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **National (D1, D2):** Strategic documents promote the development of AI as a supporting tool. **VS. Local (D3, D4, D5, D6):** Default prohibition on recognizing AI-generated content as independent work.
Doc ID:	D2
Gaps:	• No procedures for interdisciplinary research • No discussion/procedures regarding AI plagiarism/fraud in academic work

	• No discussion of job displacement in the academic staff labor market (but the problem of exclusion in labor market caused by AI is mentioned) • No specific standards/procedures for AI-based curricula • Shallow focus on AI in higher education • Lack of procedures for universities regarding AI in teaching • Lack of cybersecurity procedures for AI • Lack of implementation and enforcement mechanisms (soft law)
Overlaps	• The importance of computing infrastructure (overlaps with: D1) • Education and competence development in the context of AI (overlaps with: D1, D3, D4, D5, D6, D7, E1) • Identification of key sectors for AI development (overlaps with: D1) • Regulation and legal frameworks for AI (overlaps with: D1, D3, D4) • Science–business technology transfer (overlaps with: D1) • Open data and open AI models (overlaps with: D1) • Human-centred approach and respect for human dignity and rights (overlaps with: D1, D3, D5, D6, D7, E1) • Transparency, explainability, and reporting of AI use (overlaps with: D1, D3, D4, D5, D6, D7, E1) • Equality, non-discrimination, and protection of vulnerable groups (overlaps with: D1, D3, D6, E1) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D3, D4, D5, D6, D7, E1) • Support for research and innovation (overlaps with: D1, D3, D5, D6, D7, E1) • Emphasis on the importance of international cooperation (overlaps with: D1) • AI based on six pillars: society, business, science, education, cooperation, and the public sector (overlaps with: D1) • Author responsibility for AI-generated content (overlaps with: D1, D3, D4, D5, D6, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D3, D4, D5, D6, D7, E1) • Protection of copyright and intellectual property (overlaps with: D3, D4, D5, D6) • Data protection and privacy (overlaps with: D1, D3, D4, D6, E1)
Inconsistencies	• **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **National (D1, D2):** Strategic documents promote the development of AI as a supporting tool. **VS. Local (D3, D4, D5, D6):** Default prohibition on recognizing AI-generated content as independent work.
Doc ID:	D3
Gaps:	• Deepfakes and identity fraud are not directly prohibited in D3 while D4 states it explicitly (Eth&I) • Sustainability aspect when applying AI tools is not defined in the document (Sus) • Technical standards when applying AI tools is not defined in the document (TechS&I) • No discussion/procedures on the application of AI in distance learning (Eq&I) • No specific AI detection tools is presented – only general recommendations are presented (Transp&Expl) • No procedures for joint projects with other universities/institutions (Eq&I) • No discussion/procedures for professional and postgraduate education (Eq&I)

Co-funded by the European Union
Project: "Promoting Accountable and Transparent Human-centered AI in Higher Education" Project n. 2025-1-PL01-KA220-HED-000359600
Co-funded by the Erasmus+ Programme of the European Union

	• No procedures for resolving conflicts between guidelines and regulations (Acc&Liab) • No procedures for interdisciplinary research (Eq&I) • No discussion of accessibility for people with disabilities (Eq&I)
Overlaps	• Transparency, explainability, and reporting of AI use (overlaps with: D1, D2, D4, D5, D6, D7, E1) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D2, D4, D5, D6, D7, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D2, D4, D5, D6, D7, E1) • Protection of copyright and intellectual property (overlaps with: D2, D4, D5, D6) • Data protection and privacy (overlaps with: D1, D2, D4, D6, E1) • Human-centred approach and respect for human dignity and rights (overlaps with: D1, D2, D5, D6, D7, E1) • Education and competence development in the context of AI (overlaps with: D1, D2, D4, D5, D6, D7, E1) • Support for research and innovation (overlaps with: D1, D2, D5, D6, D7, E1) • Regulation and legal frameworks for AI (overlaps with: D1, D2, D4) • Author responsibility for AI-generated content (overlaps with: D1, D2, D4, D5, D6, E1) • Equality, non-discrimination, and protection of vulnerable groups (overlaps with: D1, D2, D6, E1) • It is forbidden to generate complete works (D3, D4, D5, D6) (Eth&I) • There is the obligation to report the use of AI (D3, D4, D5, D6) (Transp&Expl) • It is forbidden to process personal data (D3, D4, D6) (DataPriv&S) • AI-generated content must be verified (D3, D4, D5, D6) (Eth&I) • Allowed auxiliary uses of AI tools for: brainstorming, literature search, language correction, bibliography formatting (D3, D5, D6) (Eth&I) • Emphasis on developing critical thinking skills (D3, D4, D6) (Eth&I) • The use of AI tools in exams is strictly prohibited (D3, D4) (Eth&I) • Some activities allowed and forbidden when applying AI tools presented in the document are similar to the activities allowed/forbidden by other documents (D3, D4, D5, D6, D7) (Eth&I)
Inconsistencies	• D3 allows teachers to automatically assess short answers with human verification, while D4 does not require human verification during automatic evaluation. • **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **Local (D5):** There is no need to report the use of AI for idea generation. **(D3, D5):** There is no need to report the use of AI for "text operations" (language correction, stylistics). **VS. Local (D4):** Requires very detailed reporting, including: the name of the system, a link to the system, the date the content was generated, and an indication of whether the text has been modified by the student. • **EU (E1):** AI systems must guarantee privacy and data protection throughout a system's entire lifecycle - information initially provided by the user, and the information generated about the user during interaction with AI (e.g. outputs that the AI system generated for specific users or how users responded to particular recommendations). **VS. Local (D3, D5):** The academic teacher or supervisor has the right to request a file with the prompts and the answers obtained before accepting a student's work.

	• **National (D1, D2):** Strategic documents promote the development of AI as a supporting tool. **VS. Local (D3, D4, D5, D6):** Default prohibition on recognizing AI-generated content as independent work. • **EU (E1):** AI does not have legal personality, so it cannot be an author; it is merely a tool. **VS. Local (D3, D4, D5, D6):** The regulations include the statement that AI can be the author of a work, which is not consistent with EU nomenclature. • **EU (E1):** Requires AI developers to ensure that systems are explainable. **VS. Local (D3, D4, D5, D6):** The obligation to provide explanations is shifted to students. Students are required to "explain and justify" every claim generated by AI. • **EU (E1):** The EU requires systematic training for all stakeholders in Trustworthy AI. **VS. Local (D3, D4):** Insufficient staff training, just general recommendations.
Doc ID:	D4
Gaps:	• No procedures/guidelines for experimental/laboratory/empirical research if applying AI tools (Eth&I) • No procedures/guidelines for scientific publications (Eth&I) • No procedures/guidelines for Ph. D. students using AI tools (Eq&I) • No procedures for interdisciplinary research (Eq&I) • No discussion/procedures on the application of AI in distance learning (Eq&I) • No procedures for joint projects with other universities/institutions (Eq&I) • No discussion of AI hallucinations beyond fictional sources (Eth&I) • No discussion/procedures for professional and postgraduate education (Eq&I) • No discussion of accessibility for people with disabilities (Eq&I) • No procedures for resolving conflicts between guidelines and regulations (Acc&Liab) • Sustainability aspect when applying AI tools is not defined in the document (Sus)
Overlaps	• Transparency, explainability, and reporting of AI use (overlaps with: D1, D2, D3, D5, D6, D7, E1) • Author responsibility for AI-generated content (overlaps with: D1, D2, D3, D5, D6, E1) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D2, D3, D5, D6, D7, E1) • Education and competence development in the context of AI (overlaps with: D1, D2, D3, D5, D6, D7, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D2, D3, D5, D6, D7, E1) • Data protection and privacy (overlaps with: D1, D2, D3, D6, E1) • Regulation and legal frameworks for AI (overlaps with: D1, D2, D3) • Protection of copyright and intellectual property (overlaps with: D2, D3, D5, D6) • It is forbidden to generate complete works (D3, D4, D5, D6) (Eth&I) • There is the obligation to report the use of AI (D3, D4, D5, D6) (Transp&Expl) • It is forbidden to process personal data (D3, D4, D6) (DataPriv&S) • AI-generated content must be verified (D3, D4, D5, D6) (Eth&I) • Emphasis on developing critical thinking skills (D3, D4, D6) (Eth&I) • The use of AI tools in exams is strictly prohibited (D3, D4) (Eth&I) • Some activities allowed and forbidden when applying AI tools presented in the document are similar to the activities allowed/forbidden by other documents (D3, D4, D5, D6, D7) (Eth&I)

Project: "Promoting Accountable and Transparent Human-centered AI in Higher Education" Project n. 2025-1-PL01-KA220-HED-000359600
Co-funded by the Erasmus+ Programme of the European Union

Inconsistencies	• Using AI content detectors presented in the D4 can violate "Data Privacy & Security" aspect if the academic work contains personal, sensitive, classified data. • D3 allows teachers to automatically assess short answers with human verification, while D4 does not require human verification during automatic evaluation. • **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **Local (D5):** There is no need to report the use of AI for idea generation. **(D3, D5):** There is no need to report the use of AI for "text operations" (language correction, stylistics). **VS. Local (D4):** Requires very detailed reporting, including: the name of the system, a link to the system, the date the content was generated, and an indication of whether the text has been modified by the student. • **National (D1, D2):** Strategic documents promote the development of AI as a supporting tool. **VS. Local (D3, D4, D5, D6):** Default prohibition on recognizing AI-generated content as independent work. • **EU (E1):** AI does not have legal personality, so it cannot be an author; it is merely a tool. **VS. Local (D3, D4, D5, D6):** The regulations include the statement that AI can be the author of a work, which is not consistent with EU nomenclature. • **EU (E1):** Requires AI developers to ensure that systems are explainable. **VS. Local (D3, D4, D5, D6):** The obligation to provide explanations is shifted to students. Students are required to "explain and justify" every claim generated by AI. • **EU (E1):** Equal access to technology and innovation. **VS. Local (D4):** The use of any AI devices (watches, glasses) during exams is prohibited and will result in the exam being invalidated. • **EU (E1):** The EU requires systematic training for all stakeholders in Trustworthy AI. **VS. Local (D3, D4):** Insufficient staff training, just general recommendations.
Doc ID:	D5
Gaps:	• Automatic evaluation of assignment by using the AI tools is not mentioned in the D5, while D3 and D4 allow it. (Transp&Expl) • Deepfakes and identity fraud are not directly prohibited in D5 while D4 states it explicitly. (Eth&I) • No procedures/guidelines for experimental/laboratory/empirical research if applying AI tools (Eth&I) • No procedures/guidelines for scientific publications (Eth&I) • No procedures/guidelines for Ph. D. students using AI tools (Eq&I) • No procedures according to personal/sensitive/data security when applying AI tools (DataPriv&S) • No procedures for interdisciplinary research (Eq&I) • No discussion/procedures on the application of AI in distance learning (Eq&I) • No specific AI detection tools is presented (Transp&Expl) • No procedures for joint projects with other universities/institutions (Eq&I) • No discussion of accessibility for people with disabilities (Eq&I) • No discussion of the consequences of violating the rules (Acc&Liab) • No discussion/procedures for professional and postgraduate education (Eq&I) • No procedures for resolving conflicts between guidelines and regulations (Acc&Liab)

	• Detection software of AI created content is not involved in the document (Transp&Expl) • Sustainability aspect when applying AI tools is not defined in the document (Sus) • Technical standards when applying AI tools is not defined in the document (TechS&I) • No discussion of AI hallucinations (Eth&I)
Overlaps	• Transparency, explainability, and reporting of AI use (overlaps with: D1, D2, D3, D4, D6, D7, E1) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D2, D3, D4, D6, D7, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D2, D3, D4, D6, D7, E1) • Protection of copyright and intellectual property (overlaps with: D2, D3, D4, D6) • Human-centred approach and respect for human dignity and rights (overlaps with: D1, D2, D3, D6, D7, E1) • Education and competence development in the context of AI (overlaps with: D1, D2, D3, D4, D6, D7, E1) • Support for research and innovation (overlaps with: D1, D2, D3, D6, D7, E1) • Author responsibility for AI-generated content (overlaps with: D1, D2, D3, D4, D6, E1) • It is forbidden to generate complete works (D3, D4, D5, D6) (Eth&I) • There is the obligation to report the use of AI (D3, D4, D5, D6) (Transp&Expl) • AI-generated content must be verified (D3, D4, D5, D6) (Eth&I) • Allowed auxiliary uses of AI tools for: brainstorming, literature search, language correction, bibliography formatting (D3, D5, D6) (Eth&I) • Some activities allowed and forbidden when applying AI tools presented in the document are similar to the activities allowed/forbidden by other documents (D3, D4, D5, D6, D7) (Eth&I)
Inconsistencies	• Some activities presented (9. Artificial intelligence as a subject of research) are forbidden in other documents. They are allowed here due to specific context of the AI application. • **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **Local (D5):** There is no need to report the use of AI for idea generation. **(D3, D5):** There is no need to report the use of AI for "text operations" (language correction, stylistics). **VS. Local (D4):** Requires very detailed reporting, including: the name of the system, a link to the system, the date the content was generated, and an indication of whether the text has been modified by the student. • **EU (E1):** AI systems must guarantee privacy and data protection throughout a system's entire lifecycle - information initially provided by the user, and the information generated about the user during interaction with AI (e.g. outputs that the AI system generated for specific users or how users responded to particular recommendations). **VS. Local (D3, D5):** The academic teacher or supervisor has the right to request a file with the prompts and the answers obtained before accepting a student's work. • **National (D1, D2):** Strategic documents promote the development of AI as a supporting tool. **VS. Local (D3, D4, D5, D6):** Default prohibition on recognizing AI-generated content as independent work. • **EU (E1):** AI does not have legal personality, so it cannot be an author; it is merely a tool. **VS. Local (D3, D4, D5, D6):** The regulations include the

	statement that AI can be the author of a work, which is not consistent with EU nomenclature. • **EU (E1):** Requires AI developers to ensure that systems are explainable. **VS. Local (D3, D4, D5, D6):** The obligation to provide explanations is shifted to students. Students are required to "explain and justify" every claim generated by AI.
Doc ID:	D6
Gaps:	• Automatic evaluation of assignment by using the AI tools is not mentioned in the D6, while D3 and D4 allow it. (Transp&Expl) • Deepfakes and identity fraud are not directly prohibited in D6 while D4 states it explicitly. (Eth&I) • No procedures/guidelines for experimental/laboratory/empirical research if applying AI tools (Eth&I) • No procedures/guidelines for scientific publications (Eth&I) • No procedures/guidelines for Ph. D. students using AI tools (Eq&I) • No procedures for interdisciplinary research (Eq&I) • No discussion of accessibility for people with disabilities (Eq&I) • No discussion of the consequences of violating the rules (Acc&Liab) • No discussion/procedures on the application of AI in distance learning (Eq&I) • No discussion/procedures for professional and postgraduate education (Eq&I) • No procedures for joint projects with other universities/institutions (Eq&I) • No procedures for resolving conflicts between guidelines and regulations (Acc&Liab) • Detection software of AI created content is not involved in the document (Transp&Expl) • Sustainability aspect when applying AI tools is not defined in the document (Sus) • Technical standards when applying AI tools is not defined in the document (TechS&I) • No discussion of AI hallucinations (Eth&I)
Overlaps	• Human-centred approach and respect for human dignity and rights (overlaps with: D1, D2, D3, D5, D7, E1) • Transparency, explainability, and reporting of AI use (overlaps with: D1, D2, D3, D4, D5, D7, E1) • Data protection and privacy (overlaps with: D1, D2, D3, D4, E1) • Author responsibility for AI-generated content (overlaps with: D1, D2, D3, D4, D5, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D2, D3, D4, D5, D7, E1) • Education and competence development in the context of AI (overlaps with: D1, D2, D3, D4, D5, D7, E1) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D2, D3, D4, D5, D7, E1) • Equality, non-discrimination, and protection of vulnerable groups (overlaps with: D1, D2, D3, E1) • Support for research and innovation (overlaps with: D1, D2, D3, D5, D7, E1) • Protection of copyright and intellectual property (overlaps with: D2, D3, D4, D5) • It is forbidden to generate complete works (D3, D4, D5, D6) (Eth&I) • There is the obligation to report the use of AI (D3, D4, D5, D6) (Transp&Expl) • It is forbidden to process personal data (D3, D4, D6) (DataPriv&S) • AI-generated content must be verified (D3, D4, D5, D6) (Eth&I)

	• Allowed auxiliary uses of AI tools for: brainstorming, literature search, language correction, bibliography formatting (D3, D5, D6) (Eth&I) • Emphasis on developing critical thinking skills (D3, D4, D6) (Eth&I) • Some activities allowed and forbidden when applying AI tools presented in the document are similar to the activities allowed/forbidden by other documents (D3, D4, D5, D6, D7) (Eth&I)
Inconsistencies	• **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **National (D1, D2):** Strategic documents promote the development of AI as a supporting tool. **VS. Local (D3, D4, D5, D6):** Default prohibition on recognizing AI-generated content as independent work. • **EU (E1):** AI does not have legal personality, so it cannot be an author; it is merely a tool. **VS. Local (D3, D4, D5, D6):** The regulations include the statement that AI can be the author of a work, which is not consistent with EU nomenclature. • **EU (E1):** Requires AI developers to ensure that systems are explainable. **VS. Local (D3, D4, D5, D6):** The obligation to provide explanations is shifted to students. Students are required to "explain and justify" every claim generated by AI.
Doc ID:	D7
Gaps:	• Specific document which is oriented to different levels of applying AI and providing examples of allowed activities. Gaps identified in other documents refers to this document as well.
Overlaps	• Transparency, explainability, and reporting of AI use (overlaps with: D1, D2, D3, D4, D5, D6, E1) • Education and competence development in the context of AI (overlaps with: D1, D2, D3, D4, D5, D6, E1) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D2, D3, D4, D5, D6, E1) • Human-centred approach and respect for human dignity and rights (overlaps with: D1, D2, D3, D5, D6, E1) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D2, D3, D4, D5, D6, E1) • Support for research and innovation (overlaps with: D1, D2, D3, D5, D6, E1)
Inconsistencies	• "Application is permitted with no restrictions" level of applying AI tools with examples of tasks is contradictory to other regulation where these tasks are generally forbidden – D3, D4, D5
Doc ID:	E1
Gaps:	• No specific guidelines for higher education and academic education • No procedures/guidelines for experimental/laboratory/empirical research if applying AI tools • No procedures/guidelines for scientific publications • No procedures according to personal/sensitive/data security when applying AI tools (only overall requirements are mentioned) • No discussion/procedures regarding AI plagiarism/fraud in academic work • No discussion of job displacement in the academic staff labor market (but the problem of exclusion in labor marked caused by AI is mentioned) • No discussion of AI-generated content in theses/research • No consideration of AI hallucinations in academic research • No procedures for verifying the originality of academic work • No specific measures for evaluating implementation • Lack of a specific national or institutional context (general, horizontal nature)

	• Lack of KPIs, timelines, and operational structures • Lack of reference to generative AI (due to time constraints)
Overlaps	• Human-centred approach and respect for human dignity and rights (overlaps with: D1, D2, D3, D5, D6, D7) • Transparency, explainability, and reporting of AI use (overlaps with: D1, D2, D3, D4, D5, D6, D7) • Equality, non-discrimination, and protection of vulnerable groups (overlaps with: D1, D2, D3, D6) • Ethics and responsibility in the design and use of AI (overlaps with: D1, D2, D3, D4, D5, D6, D7) • Education and competence development in the context of AI (overlaps with: D1, D2, D3, D4, D5, D6, D7) • Author responsibility for AI-generated content (overlaps with: D1, D2, D3, D4, D5, D6) • Intellectual autonomy, independence, and critical thinking (overlaps with: D1, D2, D3, D4, D5, D6, D7) • Data protection and privacy (overlaps with: D1, D2, D3, D4, D6) • Support for research and innovation (overlaps with: D1, D2, D3, D5, D6, D7)
Inconsistencies	• **EU & national (E1, D1, D2):** Promote the widespread use of AI in education and administration. **VS. Local (D3, D4, D5, D6):** Universities prohibit citing LLM as sources and prohibit the generation of parts of theses. • **EU (E1):** AI systems must guarantee privacy and data protection throughout a system's entire lifecycle - information initially provided by the user, and the information generated about the user during interaction with AI (e.g. outputs that the AI system generated for specific users or how users responded to particular recommendations). **VS. Local (D3, D5):** The academic teacher or supervisor has the right to request a file with the prompts and the answers obtained before accepting a student's work. • **EU (E1):** AI does not have legal personality, so it cannot be an author; it is merely a tool. **VS. Local (D3, D4, D5, D6):** The regulations include the statement that AI can be the author of a work, which is not consistent with EU nomenclature. • **EU (E1):** Requires AI developers to ensure that systems are explainable. **VS. Local (D3, D4, D5, D6):** The obligation to provide explanations is shifted to students. Students are required to "explain and justify" every claim generated by AI. • **EU (E1):** Equal access to technology and innovation. **VS. Local (D4):** The use of any AI devices (watches, glasses) during exams is prohibited and will result in the exam being invalidated. • **EU (E1):** The EU requires systematic training for all stakeholders in Trustworthy AI. **VS. Local (D3, D4):** Insufficient staff training, just general recommendations. • **EU (E1):** The EU recommends a uniform ethical framework for all sectors. **VS. Local (D3, D4, D5, D6):** Local regulations are fragmented, they have different levels, scales and scope.

Replicate the table for each document

Section 5: Identified Risks

Purpose: To identify any potential ethical, legal, pedagogical, social, or technical risks that emerge from the policies you analysed.

Instructions: List any risks revealed by the documents or by the absence of clear guidance. Indicate whether the institution has a response or mitigation measure.

Project: "Promoting Accountable and Transparent Human-centered AI in Higher Education" Project n. 2025-1-PL01-KA220-HED-000359600
Co-funded by the Erasmus+ Programme of the European Union

Risk ID	Description (of Risk)	Level (High / Medium / Low)	Potential Impact	Source Doc ID(s)	Mitigation or Policy Response
R.1.1	Use of artificial intelligence systems to generate academic content such as essays, coursework, theses, or scientific publications, which may lead to **violations of academic integrity (plagiarism, undisclosed use of AI by students or academic staff, and publication of unverified, erroneous, or misleading content)**. Such practices may result in falsely certified knowledge and competences, undermining the fundamental principles of academic integrity.	High	Reduction in the quality of education and scientific research, and undermining the credibility of assessments, diplomas, and academic degrees. In the long term, this may result in loss of trust in higher education institutions, devaluation of academic achievements, and weakening of institutional reputation. At the individual level, consequences include disciplinary sanctions for students and staff, as well as ethical and professional liability risks.	D3, D4, D5, D6	Introduction of institutional rules regulating permissible forms of AI use in teaching and research activities, including mandatory disclosure of AI use by students and staff. Application of tools supporting detection of AI-generated content and updating academic codes of ethics. Modification of learning outcome verification methods, such as increased use of oral examinations and practical projects. Definition of clear disciplinary sanctions for violations of academic integrity.
R.1.2	Excessive reliance on generative AI tools in the learning process may lead to **reduced intellectual autonomy of students and limited development of key competences** such as critical, analytical, and creative thinking.	High	Students may fail to acquire skills necessary for independent academic and professional work, reducing their preparedness for the labor market and further education. In the long term, this risk may lead to technological dependency, weakened critical evaluation of information, and reduced competitiveness of graduates.	D3, D4, D6	Promotion of responsible and conscious use of AI. Inclusion of content on AI hallucinations and limitations (bias) in curricula. Strengthening courses that develop creativity and critical thinking (e.g., case studies, team projects, discussions). Banning AI-generated content alone is difficult to monitor effectively and should be complemented by educational measures.
R.1.3	**The use of deepfake technologies and synthetic voices in the context of education and assessment** poses a threat to the credibility of academic processes. These tools may be used to **impersonate other individuals during remote examinations, oral presentations, or assessments based on audio-video recordings**, thereby undermining the authenticity of the assessed person's identity.	High	Falsification of assessment results, violations of academic integrity principles, and erosion of trust in remote forms of examination. In the longer term, this may lead to a reduction in the credibility of diplomas and conducted examinations.	D4	Introduction of explicit prohibitions on the use of deepfake technologies and synthetic voices in assessment processes. Application of proctoring tools, identity verification mechanisms, and technologies supporting the detection of audio-video manipulation. Adaptation of assessment formats (e.g. synchronous examinations, interactive elements) in order to reduce susceptibility to such abuses.
R.1.4	**Unclear and inconsistent boundaries between permitted "editing" or**	High	Unintentional violations of academic regulations, committed in good faith but	D3, D4, D5, D6	Introduction of clear, shared definitions distinguishing editing, language correction,

	language support and prohibited "content generation" by AI tools create uncertainty regarding the correct manner of their use in student work.		based on incorrect or imprecise interpretations of permissible AI use. This may lead to unfair disciplinary consequences, interpretative disputes, and a decline in trust in the assessment system.		and stylistic support from the actual generation of substantive content by AI. Development of coherent standards at the institutional or sectoral level, supplemented with practical examples of permitted and prohibited uses of AI.
R.1.5	**Unclear or non-existent procedures** defining how academic staff should **specify and communicate rules for the use of AI** within their courses.	High	Inconsistent practices across courses and instructors, student confusion regarding applicable rules, risk of unequal treatment, and difficulties in enforcing accountability in cases of violations.	D6	Development of central university-level guidelines defining minimum standards for the use of AI in teaching, while allowing instructors limited autonomy at the course level. Introduction of a requirement to clearly communicate rules on AI use in course syllabi. Provision of training support and model policy statements (templates) for teaching staff.
R.1.6	**Lack of established procedures for assessing the impact of AI tools on students' mental well-being**, including the absence of systematic monitoring activities and mechanisms for identifying and responding to negative effects of AI use.	Medium	Increased risk of technology addiction, higher stress levels, and deterioration of students' mental health. In extreme cases, this may result in mental health crises requiring specialist support.	D6	Development and implementation of procedures for assessing the impact of AI on students' mental well-being. Introduction of a monitoring system. Definition of concrete support measures for students (e.g. access to psychological counselling, preventive programmes, and support pathways).
R.1.7	**Insufficient preparedness of academic teaching staff for AI –** inadequate competencies of university educators in the field of AI.	High	Ineffective teaching and inadequate supervision of students (e.g., inability to detect student misconduct). Errors in AI-related educational content. Delays in the implementation of modern teaching methods. In the long term, this may lead to a deterioration in the quality of education.	D1, D2, D3, D4, D5, D6	Organising training sessions and courses for academic staff (upskilling programmes in AI). Exchange of best practices between universities. Development of institutional guidelines and manuals for educators.
R.1.8	**Violation of intellectual property protection** – the requirement to share prompts and AI-generated responses with the supervisor may raise concerns regarding data privacy and the protection of	Medium	Breaches of students' confidentiality and intellectual property may discourage creative and original approaches to academic work. It may also lead to conflicts regarding copyright and result in the loss of innovative ideas that	D1, D2	Establish clear guidelines on sharing prompts and AI outputs with supervisors, considering the protection of students' intellectual property. Specify which information is mandatory to share and which may be anonymised or redacted. Implement procedures to safeguard the confidentiality of ideas and

	solutions/ideas/concepts developed by the student.		could have commercial or scientific potential.		concepts and provide legal support regarding copyright and intellectual property rights.
R.2.1	**Substantive errors and misinformation in academic work** – large language models may generate false or outdated information (so-called "hallucinations"), which makes unverified scientific content unreliable.	High	The production of articles and reports with low credibility, containing erroneous research conclusions, fabricated study results, or citation mistakes. As a consequence, the quality of the university's research may be undermined.	D3, D5, D6	Introduce a requirement to verify sources and information generated by AI. Provide training for researchers on methods of data validation. Recommend using AI only as a supporting tool, not as the primary research instrument. Develop specific procedures and tools for verifying AI usage.
R.3.1	**Unequal access to AI tools and educational resources** – disparities between institutions (e.g., public vs. private) or regions, as well as specific student circumstances (international students, students with disabilities), may result in some students not acquiring the digital competencies required in the labour market.	Medium	Risk of exacerbating educational divides (smaller universities/students with limited resources may face systemic disadvantage). Additionally, collaboration between academic institutions may be hindered.	D1, D2	Monitor the use of AI tools across different types of institutions and student groups to identify potential access inequalities. Provide financial and organisational support to equalise access to AI tools and educational resources. Address the accessibility of AI technologies for students with disabilities by implementing appropriate digital accessibility standards and procedures that support equal opportunities.
R.4.1	**Loss of trust and transparency in AI** – insufficient disclosure regarding the functioning of AI systems (black box) may undermine public trust and lead to unethical decisions.	Medium	Declining public acceptance of AI, risk of discriminatory decisions by unaudited systems, and legal and reputational conflicts arising from the inability to explain decisions made with the involvement of AI.	E1, D6	Introduce transparency requirements encompassing auditability of AI systems, explainability of decisions, and the adoption of certification mechanisms aligned with recognised standards. Promote user and public education on how AI systems operate. Establish ethical codes and effective appeal mechanisms.
R.5.1	**Changes in the labour market caused by AI** – AI advancements may displace certain occupations (including within higher education).	High	An increase in technological unemployment in some sectors, which may lead to heightened social resistance to AI adoption.	D1, D2	Implement retraining and educational programmes for individuals at risk of automation. Equalise educational opportunities in the digital domain (courses coupled with grants to fund participation). Support SMEs in AI adoption (e.g., tax incentives, venture funding).
R.6.1	**Lack of a clear procedure for resolving conflicts between the seven requirements of Trustworthy AI** (e.g.,	Medium	Confusion in the AI implementation process and the adoption of inconsistent decisions. Inconsistent application of Trustworthy AI	E1, D1, D2	Introduce a clear mechanism for resolving conflicts between Trustworthy AI requirements. Develop a procedure that defines priorities and decision-

	between transparency, privacy, security, and fairness). As a result, there is no unambiguous method for prioritising requirements or making decisions when these requirements conflict.		principles and a reduction in the quality and credibility of deployed solutions.		making processes in cases of conflicting requirements. Identify responsible individuals or teams for decision-making in such situations. Develop tools to support the analysis of trade-offs.
R.6.2	**Extreme fragmentation of AI-related procedures across different universities.** Each institution applies its own rules and approaches.	High	Lack of consistency in regulations and practices concerning AI use in the academic environment. Confusion for students transferring between institutions. Unequal quality of education and the absence of a unified standard for AI implementation and oversight, which hinders the comparability of learning outcomes and inter-university collaboration.	D3, D4, D5, D6	Introduce mechanisms for coordinating and harmonising AI procedures across universities. Establish a central coordinator or an inter-university team responsible for developing common guidelines, standards, and best practices.
R.6.3	Risk of **inconsistencies between strategic documents (national, international) and local regulations** in force at individual universities.	Medium	Differences in the interpretation and application of provisions, leading to confusion among staff. Difficulties in implementing policies and procedures. As a result, inconsistent application of principles may occur.	E1, D1, D2, D3, D4, D5, D6, D7	Conduct a compliance analysis between strategic documents and local regulations. Develop procedures for ongoing compliance monitoring. Create common interpretative guidelines. Define responsibilities for updating documents. Establish a procedure for resolving discrepancies.
R.6.4	**Lack of clear provisions regarding the consequences of violating AI-related rules.**	Medium	Students will not perceive tangible consequences for breaching regulations. As a result, this may lead to an increase in violations and a reduction in the effectiveness of policies governing AI use.	D5, D6	Clarify procedures and consequences for violations in existing documents or develop new ones. Introduce clear and consistent provisions outlining actions to be taken in the event of violations, and specify concrete mechanisms for enforcing the rules.
R.7.1	**Violation of personal data protection** – inputting confidential or personal data into AI systems (e.g., student information from grade records) may lead to data leakages, unauthorised use, or irreversible loss of anonymity.	High	Legal sanctions related to GDPR violations, loss of public trust in the university, reputational damage to the institution, and harm to individuals whose data has been disclosed.	E1, D1, D2, D6	Prohibit the inclusion of personal, confidential, classified, or copyright-protected data in AI tools (ensure content anonymisation). Provide GDPR training for AI users. Conduct audits of AI usage.
R.7.2	**Cyber threats and AI manipulation** – AI systems may be targets of	High	Disruption of system operations (e.g., critical infrastructure), data	E1, D1, D2	Invest in AI cybersecurity (security testing, training for AI developers). Foster

	cyberattacks (e.g., manipulation of training data) or may themselves generate security risks (e.g., against network security).		leakages, and loss of technological sovereignty.		international cooperation on security measures. Maintain an AI incident registry. Implement sandboxes for secure AI testing.
R.8.1	**Unclear procedures for technology transfer from universities to industry involving AI.**	Medium	Difficulty translating research outcomes into industry applications and challenges in collaborating with external partners. Limited effectiveness of innovation commercialization. Consequently, there may be a loss of the economic and social potential derived from scientific research, as well as a weakening of the competitiveness of universities and industry.	D1, D2	Develop and implement clear technology transfer procedures that consider the specific characteristics of AI projects. Introduce principles regarding copyright and intellectual property. Define mechanisms for collaboration with industry, including licensing terms, profit-sharing arrangements, and protection of know-how. Provide legal and administrative support to researchers during the commercialization process.
R.1.1	**Use of artificial intelligence systems to generate academic content** such as essays, coursework, theses, or scientific publications, which may lead to violations of academic integrity (plagiarism, undisclosed use of AI by students or academic staff, and publication of unverified, erroneous, or misleading content). Such practices may result in falsely certified knowledge and competences, undermining the fundamental principles of academic integrity.	High	Reduction in the quality of education and scientific research, and undermining the credibility of assessments, diplomas, and academic degrees. In the long term, this may result in loss of trust in higher education institutions, devaluation of academic achievements, and weakening of institutional reputation. At the individual level, consequences include disciplinary sanctions for students and staff, as well as ethical and professional liability risks.	D3, D4, D5, D6	Introduction of institutional rules regulating permissible forms of AI use in teaching and research activities, including mandatory disclosure of AI use by students and staff. Application of tools supporting detection of AI-generated content and updating academic codes of ethics. Modification of learning outcome verification methods, such as increased use of oral examinations and practical projects. Definition of clear disciplinary sanctions for violations of academic integrity.

Add table rows for each risk identified

Section 6: Additional Checks – Policy Implementation and Adoption

Purpose: To trace how EU-level AI and digital education frameworks are implemented at the national and institutional levels.

Instructions: Provide short written answers (2–4 sentences per field) explaining how implementation occurs and to what extent.

Check Area	Guiding Question	Partner Assessment / Notes (brief)
Tracking Implementation	How do EU-level policies (AI Act, GDPR,	Policies formulated at the EU- level are generally mentioned in documents at the national level. However, their further

(EU → National → Institutional)	DEAP) translate into your national and institutional documents?	implementation at the local (university) level is not complete. Of course, possible explanation may be that international and EU directives are dedicated to a very broad category of applications, and people involved in formulating AI policy at the university level may have considered these areas irrelevant to higher education. To sum up, it seems reasonable to develop a single coherent reference policy on the application of AI tools in higher education.
Extent of Implementation	Was implementation partial or full? Provide brief reasoning and examples.	The implementation of EU regulations at the national level can be considered complete (but the regulations in the national level are soft law in nature). The national documents often refer directly to EU regulations. However, as discussed, further implementation at the local (university) level is only partial. For example, certain areas like "Ethics and Integrity," "Privacy and Security," and "Transparency and Explainability," are generally well characterized, but their description varies in scope and level of details or references to other applicable regulations and policies. On the other hand, areas such as "Equality and Inclusion," "Accountability and Liability," and "Governance Structures" are not specified in all documents or their description is very limited. Finally, areas such as "Technical standards and interoperability" and "Sustainable development" are essentially omitted entirely at the local (university) level.
Institutional Adoption of AI Regulations	How has your institution implemented or enforced AI-related guidelines (e.g., academic policies, staff guidance, committees)?	According to the local (university) level the implementation of regulations governing the use of AI tools in local institutions is not consistent. The D3 document presents the rules in the form of a university rector's order. The D4 and D5 documents present the rules in the form of attachments to the rector's order. The D6 and related D7 documents present the rules as guidelines announced by the university's vice-rector for education. The bodies that created the rules are not specified in D3, D4, and D5. In the case of D6 and D7, the guidelines were developed by a 10-person team responsible for preparing AI guidelines for AGH University.

Section 7: Partner Reflections

Purpose: To gather qualitative reflections and recommendations from each partner.

Instructions: In 1–2 short paragraphs, reflect on the overall strengths and weaknesses of your policy environment regarding AI in higher education and any recommendations for UM to include in the consolidated Policy Analysis Report.

- General comments on the policy landscape in your country/institution.
- Initial recommendations for what should be considered in the Policy Analysis Report.

Document E1 provides the normative and ethical foundations of Trustworthy AI, D1 sets out the general vision and development direction, while D2 focuses on the operationalisation and institutional implementation of solutions. However, from the perspective of higher education (HE), these documents have an indirect character. Their primary focus is on AI systems in general. None of these documents

explicitly provides procedures or guidelines on how AI should be applied in HE, nor do they directly address the specificity of academic processes such as teaching and learning, student assessment, academic autonomy, or the teacher–student relationship. Nevertheless, some of the guidelines and good practices contained in these documents can be translated into AI systems used in HE. References to education appear implicitly in examples related to citizen scoring, skills development, and stakeholder participation, which can be readily mapped to higher education contexts.

One of the most important components of E1 is the *Assessment List for Trustworthy AI*, a practical checklist designed to support organisations in operationalising the guidelines. This checklist can be used by HE institutions as a self-assessment, governance, and risk identification tool when designing, procuring, or deploying AI systems. Document D2 is explicitly built upon D1, expanding and operationalising selected issues. However, when using D1 and D2 as part of a Policy Analysis Report, it is important to note that both documents are grounded in the Polish national context and primarily address the Polish AI ecosystem.

The final policy on the use of artificial intelligence tools in higher education should take into account the following elements:

- It is important to define the levels of AI tool application so that, depending on the specific nature of the task to be performed, it is possible to determine what use is ethical and what is unethical – examples of levels and assignments of activities can be found in the appendix: “D7 - Appendix.docx”.
- It is important to indicate which specific activities can and cannot be supported/replaced by AI – examples of activities can be found in the following appendices: “D3 – Appendix.docx”, “D5 - Appendix.docx”, “D6 – Appedix.docx”, “D7 - Appendix.docx,” “D4 - Appendix.docx.”
- It is very important to justify the risks of using AI; understanding these issues by the user should significantly reduce the unethical use of such a tool – an explanation can be found in the appendix: “D3 - Appendix.docx.”
- It is also important to know the benefits and risks of using Gen-AI tools in an ethical or unethical manner, as explained in the attachment “D6 - Appendix.docx.”

It is also important to note the prohibited use of AI when personal, sensitive, secret, confidential, or copyrighted content is involved. The use of AI tools in such cases may result in the commission of a crime and a lack of control over the further fate of the data fed into the AI model.

Submission Instructions

Please complete all sections in this document and submit as a Word or PDF file named: T3.1_PolicyReview_[PartnerName]_[Country].docx by 30 November 2025 via the Teams folder (WP3 → Task 3.1 → Partner Reports).