Created by: SMK College of Applied Sciences · Fachhochschule des Mittelstands · Turība University · Akademia WSEI · Aspira University of Applied Sciences · Universidade da Maia · Batman Üniversitesi

AI-ERIT:

AI Integration Framework for Higher Education:
Ensuring HEIs' Readiness for Inescapable Transformation

AI AUDIT PROCESS

guidelines for HEIs

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

CONTENT

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Document Structure

1. Introduction

The introduction sets the context for AI auditing in higher education, highlighting both the transformative potential and inherent risks of AI applications. It establishes why a systematic audit process is necessary to ensure fairness, transparency, and compliance with legal and ethical standards.

- AI's role in transforming teaching, research, and administration.
- Risks: bias, data privacy breaches, opaque decision-making, inequitable access.
- Rationale for a structured AI audit process.
- Recommendation: Draft an **AI Audit Policy Statement** outlining scope and governance.

2. Foundational Concepts

This section defines the core concepts, ethical principles, and legal frameworks necessary to understand and conduct AI audits in HEIs. It bridges theory and practice, guiding institutions in aligning with global best practices.

- **AI Definition in HEIs**: Educational, research, and administrative domains.
- **Governance Challenges**: Rapid tech change vs. slow policy adaptation.
- **Ethical Principles**: Accountability, fairness, autonomy, privacy, safety, inclusivity, transparency.
- **Legal Frameworks**: GDPR, FERPA, AI Act, ESG standards.
- **CRAFT Framework**: Rules, Access, Familiarity, Trust, Culture.
- Translating ethics into measurable audit indicators.

3. AI Audit Process Framework

This framework outlines a seven-phase audit methodology, ensuring comprehensive evaluation from planning to continuous monitoring. It is cyclical, supporting ongoing improvement.

- **Phase 1**: Planning and Scoping.
- **Phase 2**: Stakeholder Engagement.
- **Phase 3**: Data and Model Evaluation.
- **Phase 4**: Risk Assessment.
- **Phase 5**: Compliance Review.
- **Phase 6**: Reporting and Recommendations.
- **Phase 7**: Implementation and Continuous Monitoring.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

4. Detailed Step-by-Step Guidelines

Each phase is explained with actionable steps and roles, offering a clear operational blueprint for audit teams.

- Define scope, objectives, and success criteria.
- Assemble a multidisciplinary audit team.
- Conduct stakeholder workshops and surveys.
- Evaluate data quality, bias, and algorithm performance.
- Map compliance with legal frameworks.
- Prioritize risks by severity and urgency.
- Prepare phase-specific checklists.

5. Auditing Steps

Domain-specific auditing procedures ensure that each area of AI application is thoroughly evaluated. Checklists and risk tables are provided for practical use.

- **Research Processes**: Evaluate bias, reproducibility, and ethical approval compliance.
- **Educational Processes**: Assess fairness, accessibility, and learning outcomes.
- **Administrative Processes**: Review transparency, efficiency, and service equity.
- **AI Systems/Tools**:
  - **Technical Evaluation**: Performance, robustness, scalability.
  - **Ethical Evaluation**: Bias detection, transparency, explainability.

6. Evaluating the Overall Audit Process

This stage consolidates findings from all domains into an institutional risk profile. A multidisciplinary approach ensures balanced decision-making.

- Aggregate results from all audit areas.
- Classify risks as high, medium, or low.
- Use consensus-based review panels.
- Set timelines for corrective actions.
- Document justifications for decisions.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

7. Overall Audit Report

The report template organizes audit results into a strategic document for institutional leadership. It identifies achievements, risks, and priority actions.

- Summarize findings for each domain.
- Highlight strengths and best practices.
- List areas requiring urgent attention.
- Provide actionable recommendations.
- Integrate into the Annual Quality Report.

Terms and Definitions

The basic terms and definitions used in this audit guideline are presented below.

AI Auditing: Systematic evaluation of AI systems to ensure alignment with ethical principles, legal regulations, and institutional objectives, covering both technical performance and governance aspects.

AI Systems: Technological solutions that use artificial intelligence methods to perform tasks requiring human-like intelligence. In higher education, they support teaching, research, and administration through tools such as admissions algorithms, automated grading, learning analytics, and predictive modeling.

Accountability and Responsibility: An ethical principle requiring clearly assigned human oversight and decision-making responsibility for AI systems, with documented mechanisms for addressing AI-related risks.

Bias and Fairness: The requirement that AI systems be designed, tested, and monitored to prevent discrimination and ensure equitable outcomes across diverse user groups.

Human Autonomy and Agency: The principle that AI should support human decision-making rather than replace or undermine it, ensuring that users retain control over critical outcomes.

Privacy and Data Protection: The safeguarding of personal data in AI systems, ensuring collection, processing, and storage in compliance with applicable privacy regulations such as GDPR and FERPA.

Safety and Security: Ensuring AI systems operate reliably while minimizing risks to physical, psychological, and digital safety, including robust protection against malicious attacks.

Inclusivity: The design and deployment of AI systems to be accessible and usable by diverse populations, including marginalized or vulnerable groups.

Transparency and Explainability: The principle that AI decision-making processes should be understandable to users and stakeholders, with clear documentation and disclosure of AI's role in outcomes.

Model Drift: The decline in an AI system's performance caused by changes in data patterns, user behavior, or environmental conditions that differ from those used during the model's training, resulting in reduced accuracy, fairness, or relevance over time.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

1. Introduction

Artificial Intelligence (AI) is increasingly reshaping the core functions of higher education institutions, influencing how they deliver instruction, manage administration, conduct research, and engage with students. These advancements create valuable opportunities but also generate challenges, including data privacy risks, algorithmic bias, ethical concerns, and regulatory compliance (Bates et al., 2020). To ensure that AI technologies are implemented responsibly and remain aligned with institutional values, universities must conduct regular audits that extend beyond technical performance alone.

This guideline is intended as a clear and practical resource for higher education leaders, academic staff, and administrative personnel, including those without a background in AI. It provides a structured, step-by-step approach to auditing AI systems, supported by checklists, sample tools, and plain explanations. By following this framework, institutions can promote the ethical, fair, and transparent use of AI, thereby fostering more inclusive, trustworthy, and effective learning and working environments.

AI is already embedded in diverse academic functions, from admissions assessments to the personalisation of instructional content. However, the inner workings of these systems—particularly the data they rely upon and the ways they shape outcomes—are often opaque, raising significant concerns about transparency and accountability (Cheong, 2024). These concerns highlight the necessity of systematic AI auditing to ensure that institutional reliance on AI remains fair, ethical, and consistent with core educational values.

1.1. Purpose and Goals

AI technologies are transforming higher education by enhancing student services, optimising administrative processes, and creating more engaging learning environments. While these developments offer considerable benefits, they also introduce challenges such as bias, privacy breaches, and accountability concerns (Crompton & Burke, 2023). This guide supports higher education institutions by presenting a structured framework for conducting AI audits, offering tools and strategies to ensure that AI systems align with ethical, legal, and institutional goals, and providing mechanisms to promote trust, equity, and transparency.

The auditing process evaluates not only the technical performance of AI systems but also their alignment with principles such as fairness, inclusivity, and accountability. For instance, an automated grading system that disproportionately emphasises spelling and grammar may disadvantage students with strong cognitive skills but weaker writing mechanics (Johnson et al., 2022). An audit can lead to a more balanced framework that gives appropriate weight to the quality of content. Similarly, a scholarship allocation algorithm that relies solely on grade point averages, without considering socio-economic circumstances, may exclude students in genuine financial need (Turahman, 2024). Through targeted audits, such shortcomings can be identified and addressed, enabling AI systems to operate more equitably, transparently, and responsibly.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Accordingly, this guide has been developed to ensure that AI systems in higher education function both effectively and in line with ethical, legal, and institutional principles. Its scope includes admissions, grading, scholarship allocation, research support, distance learning, and administrative automation. The auditing process encompasses ethical evaluation, risk analysis, data protection, and user experience, going beyond technical performance alone. All assessments are aligned with the Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG, 2015).

The primary groups responsible for implementing this guide include IT and software development teams, quality assurance offices, ethics committees, data protection officers, and relevant academic and administrative leaders. By equipping these units with actionable tools, checklists, and sample templates, the guide aims to strengthen the safe, fair, and transparent governance of AI systems in higher education institutions.

1.2. Importance of AI Audits in HEIs

Systematically integrating AI into teaching, research, and operations has proven difficult for many higher education institutions. A significant number lack staff with the necessary expertise to deploy and oversee AI effectively (AI in Education: A Microsoft Special Report, n.d.). Obstacles include data privacy concerns, algorithmic bias, intellectual property exploitation, academic integrity risks, and the ethical use of AI by both students and teachers (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2023a). The digital divide is further exacerbated by regional regulatory disparities and unequal access to AI tools, particularly in low- and middle-income countries. Researchers and educators are also concerned that AI may replace or diminish some of their responsibilities, adding pressure to already heavy workloads.

Although students value AI support, they continue to place greater importance on human elements of the teacher–student relationship (UNESCO, 2023a). Similar to the experimentation-driven approach found in industry, institutions are adopting generative AI cautiously and inconsistently. A shift towards a comprehensive and well-supported adoption paradigm is, however, required. Despite the growing recognition of values such as ethics and integrity, there remains a significant gap in equipping leaders, educators, and students with the resources needed to integrate AI successfully into academic and operational processes.

AI audits play xa crucial role in ensuring compliance with regulations such as the GDPR, FERPA, and institutional policies, while mitigating risks such as bias, discrimination, and data breaches. They foster ethical AI use by upholding fairness, accountability, and inclusivity, and they enhance decision-making by enabling data-driven insights grounded in transparency and accuracy. In this context, AI auditing refers to the systematic evaluation of AI systems within higher education institutions to ensure alignment with ethical, legal, and institutional principles. The process identifies and addresses risks such as data privacy violations, algorithmic bias, lack of transparency, and potential harm to users. It also assesses how AI applications align with institutional goals, affect students and staff, and contribute to decision-making. Through such audits, AI systems can be implemented efficiently, fairly, and in ways that build trust and accountability.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Well-structured AI audits protect the rights and privacy of students, faculty, and staff by ensuring compliance with ethical and legal standards. They foster trust through transparency, accountability, and equitable practices (Crompton & Burke, 2023). By enhancing decision-making with accurate and unbiased insights, audits help institutions align AI systems with their missions and values, mitigate risks such as bias and data breaches, and reduce reputational and legal liabilities. Furthermore, audits help institutions remain competitive by responsibly leveraging AI to improve educational outcomes and operational efficiency, creating more inclusive and effective learning environments (Fernsel et al., 2025).

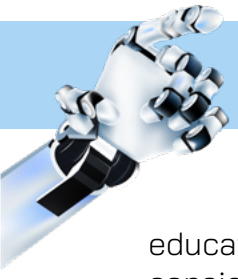
As a result of structured AI audits, institutions can achieve more secure and transparent systems, reduce risks of algorithmic discrimination, and build greater stakeholder confidence. Audits also enable institutions to detect weaknesses early, refine internal policies, and ensure compliance with evolving regulatory frameworks (Adeoye et al., 2025). Ultimately, they support the development of an ethical, inclusive, and sustainable AI ecosystem within higher education, reinforcing institutional commitments to responsible innovation and long-term academic excellence.

1.3. Overview of Key Benefits

In the context of AI integration, several goals should be prioritised to ensure responsible and effective implementation. Improved transparency provides a clear understanding of how AI systems function, enabling stakeholders to trust their processes and outcomes. Enhanced equity requires proactive identification and mitigation of bias, ensuring fair and inclusive results. Building stakeholder trust is essential, strengthening confidence among students, faculty, and administrators through ethical and accountable practices. Finally, adopting sustainable AI practices establishes a foundation for continuous improvement, ensuring that systems remain adaptable, reliable, and consistent with institutional values over time.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union



2. Foundational Concepts

This section establishes the theoretical and conceptual basis for AI auditing in higher education institutions (HEIs). It clarifies key definitions, guiding principles, and governance considerations essential for the responsible integration, monitoring, and evaluation of AI systems in education, research, and administration.

2.1. Definition of AI in Higher Education

Artificial Intelligence in Higher Education is the systematic implementation of computational systems that simulate human cognitive processes to enhance, automate, or augment institutional functions across research, education, and administrative domains while maintaining ethical, legal, and institutional compliance standards (Crompton & Song, 2021). In HEIs AI applications include, (a) in the educational process: adaptive learning systems, automated assessments, learning analytics, personalised tutoring; (b) in the research process: data analysis, predictive modelling, simulation, literature mining; (c) in the administrative process: enrolment management, resource allocation, human resources, strategic planning.

2.2. Challenges in AI Governance and Auditing Considerations for HEIs

The integration of AI in HEIs presents specific governance challenges including multi-stakeholder complexity with diverse interests of faculty, students, administrators, and regulators, rapid technological change versus slow policy adaptation, regulatory compliance requiring alignment with GDPR, AI Act, and national regulations, decentralised implementation through independent adoption by faculties or units without central oversight, and cross-border issues involving data flows and jurisdictional conflicts in international collaborations. Addressing these governance challenges requires a coordinated and proactive approach that combines robust policy frameworks, cross-departmental collaboration, and adaptive regulatory mechanisms. Without such alignment, the ethical, legal, and operational risks associated with AI in higher education may undermine institutional integrity and stakeholder trust (Al-Omari, et al., 2025).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

2.3. Ethical Principles for AI Implementation in Higher Education

The ethical integration of AI in HEIs should follow a structured set of principles that safeguard academic integrity, human rights, and societal benefit. These principles, aligned with recommendations on the ethical implementation of AI, are as follows:

Accountability and Responsibility

AI systems should have clearly assigned human oversight and decision-making responsibility. Audit criteria should verify the presence of documented accountability mechanisms, including reporting channels for addressing AI-related risks or harms (Lazcoz & Hert, 2023).

Bias and Fairness

AI needs to be designed and monitored to prevent discrimination and ensure equitable outcomes across all user groups. Audits should assess dataset diversity, bias detection processes, and corrective measures (Hasanzadeh, et al., 2025).

Human Autonomy and Agency

AI should support, not undermine, human decision-making, ensuring that users maintain ultimate control over critical outcomes. Audits should confirm that manual override options and user consent mechanisms are in place (Westphal, et al., 2023).

Privacy and Data Protection

Personal data should be collected, processed, and stored in compliance with applicable privacy regulations. Audit checks should verify encryption use, data minimization practices, and transparent consent procedures (Murdoch, 2021).

Safety and Security

AI systems are expected to operate reliably, minimizing risks to physical, psychological, and digital safety. Audits assess incident response plans, system robustness, and protection against malicious attacks (Salhab, et al., 2024).

Inclusivity

AI should be accessible and usable by diverse populations, including marginalized or vulnerable groups. Audit criteria should examine accessibility features, multilingual support, and equitable resource distribution (Acosta-Vargas, et al., 2024).

Transparency and Explainability

Ensuring that AI decision-making processes are understandable to users and stakeholders is essential. Audits verify the availability of clear documentation, explanation tools, and disclosure of AI's role in outcomes (Grimmelikhuijsen, 2022).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

2.4. Legal Frameworks for AI Implementation in Higher Education

HEIs operate within a complex regulatory environment governing the ethical and lawful use of AI systems. The General Data Protection Regulation (GDPR) sets strict requirements for data privacy and security, while the Family Educational Rights and Privacy Act (FERPA) safeguards student records. Local legislation adds jurisdiction-specific obligations, creating multiple layers of compliance. The AI Act classifies systems by risk level—from unacceptable to minimal—and imposes obligations on high-risk systems, including transparency, record-keeping, and risk assessment (European Parliament & Council, 2024; DataGuard, 2024).

Regulatory alignment should also extend to the Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG), which provide a framework for both internal and external quality assurance (ENQA, 2015). ESG requires HEIs to maintain a publicly available, systematic quality assurance policy integrated into strategic management. Key elements include programme design and approval, student-centred teaching and assessment, transparent admission and certification, fair evaluation and development of faculty qualifications, adequate learning resources and student support, effective information management, transparent institutional communication, and regular programme monitoring and review.

Integrating AI into higher education necessitates adapting these ESG principles to address AI-specific operational, ethical, and regulatory challenges while preserving academic quality. This adaptation translates broad quality assurance principles into actionable requirements for AI design, deployment, monitoring, and evaluation. Adapted standards focus on algorithmic fairness, data governance, transparency, stakeholder engagement, and continuous quality improvement.

Table 1. ESG-adapted AI integration standards for HEIs

ESG Standards	ESG-Adapted AI Integration Standards for HEIs
1.1 Institutions should have a policy for quality assurance that is made public and forms part of their strategic management. Internal stakeholders should develop and implement this policy through appropriate structures and processes, while involving external stakeholders.	**AI Quality Assurance Policy** Institutions should establish clear policies for the quality, security, and ethical use of AI systems and integrate these policies as part of their strategic management.
1.2 Institutions should have processes for the design and approval of their programs. The programs should be designed so that they meet the objectives set for them, including the intended learning outcomes. The qualification resulting from a program should be clearly specified and communicated, and refer to the correct level of the national qualifications framework for higher education and, consequently, to the Framework for Qualifications of the European Higher Education Area.	**AI-Supported Program Design** Institutions should ensure that AI tools supporting program design processes are transparent, objective, and aligned with educational objectives.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

1.3 Institutions should ensure that the programs are delivered in a way that encourages students to take an active role in creating the learning process, and that the assessment of students reflects this approach.	**AI-Enhanced Personalized Learning** Institutions should ensure that AI-based learning systems are student-centered and provide personalization while protecting student privacy.
1.4 Institutions should consistently apply pre-defined and published regulations covering all phases of the student "life cycle", e.g. student admission, progression, recognition and certification.	**AI in Student Lifecycle Management** Institutions should ensure that AI systems in student admission, progression, and evaluation processes are fair, transparent, and non-discriminatory.
1.5 Institutions should assure themselves of the competence of their teachers. They should apply fair and transparent processes for the recruitment and development of the staff.	**AI Competency Development** Institutions should support faculty in acquiring necessary competencies for ethical and effective use of AI tools.
1.6 Institutions should have appropriate funding for learning and teaching activities and ensure that adequate and readily accessible learning resources and student support are provided.	**AI Resource Management and Support** Institutions should ensure that AI-based student support services are accessible, reliable, and supportive of student welfare.
1.7 Institutions should ensure that they collect, analyse and use relevant information for the effective management of their programs and other activities.	**AI Data Management and Analytics** Institutions should ensure that data collected by AI systems is used securely, ethically, and in alignment with institutional objectives.
1.8 Institutions should publish information about their activities, including programs, which is clear, accurate, objective, up-to date and readily accessible.	**AI Transparency and Accountability** Institutions should transparently share AI system operations, decision-making processes, and usage policies with stakeholders.
1.9 Institutions should monitor and periodically review their programs to ensure that they achieve the objectives set for them and respond to the needs of students and society. These reviews should lead to continuous improvement of the program. Any action planned or taken as a result should be communicated to all those concerned.	**AI System Continuous Monitoring** Institutions should continuously monitor AI system performance, ethical compliance, and alignment with societal needs and implement improvements accordingly.
1.10 Institutions should undergo external quality assurance in line with the ESG on a cyclical basis.	**AI External Audit and Evaluation** Institutions should regularly undergo independent external audits of AI systems and integrate them into quality assurance processes.

The ESG-adapted AI integration standards bridge the gap between traditional quality assurance frameworks and the complex requirements of AI governance, offering HEIs a pathway to adopt technology while maintaining compliance with institutional quality standards and emerging regulations.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

2.5. Principles for Responsible AI Use

Responsible AI integration in HEIs requires not only adherence to ethical principles but also the use of practical frameworks that guide daily operations. Beyond compliance with transparency, fairness, accountability, privacy, and human oversight, HEIs benefit from structured approaches that translate values into practice. transparency, fairness, accountability, privacy, and human oversight, HEIs benefit from structured approaches that translate values into practice.

One such approach is the CRAFT framework (Liu & Bates, 2025), which identifies five areas for effective AI integration:

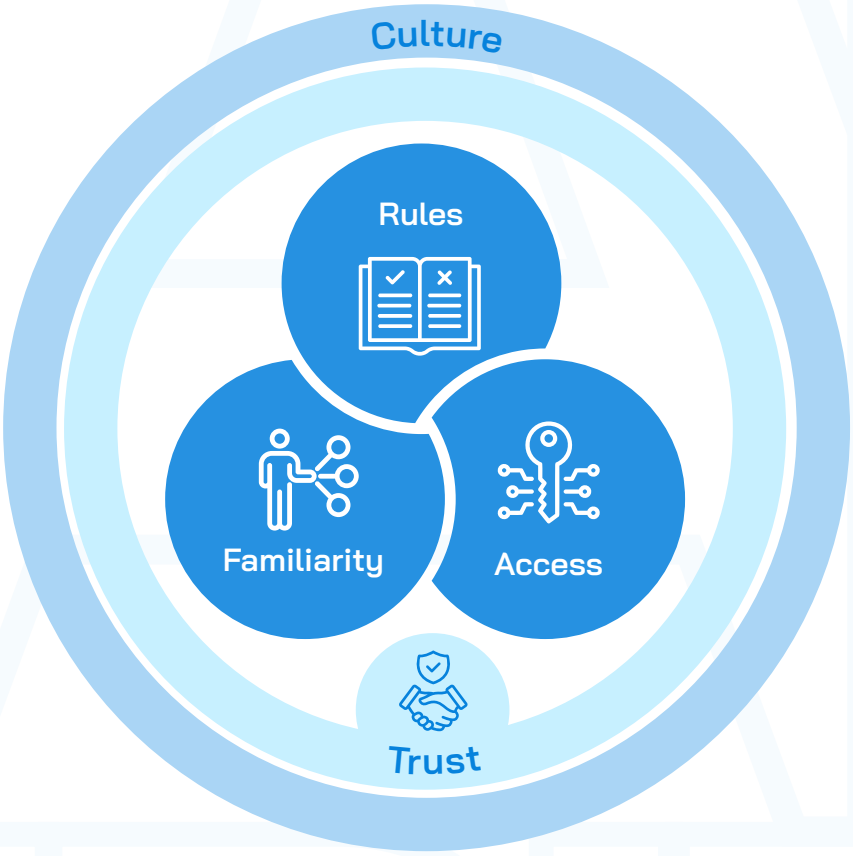


Figure 1 CRAFT Framework - Five core areas needed to safely address Gen-AI in HEI's - adapted from (Liu & Bates, 2025)

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

This framework, identifies five key areas for effective AI integration in education, research, and administration:

1. Rules – Establish clear institutional policies, principles, and guidelines to govern both individual and organizational use of AI, preventing misuse and ensuring ethical consistency.

2. Access – Ensure equitable access to AI tools and resources for all stakeholders, removing technical, financial, and procedural barriers.

3. Familiarity – Provide targeted training and awareness programs to enhance AI literacy among faculty, staff, and students, enabling informed and effective use.

4. Trust – Build institutional trust in AI systems through transparency in their design, deployment, and outcomes, supported by ongoing evaluation and stakeholder engagement.

5. Culture – Foster a positive institutional culture that embraces innovation while maintaining academic integrity, human-centered values, and inclusivity.

By adopting such structured implementation frameworks alongside ethical guidelines, HEIs can ensure that AI technologies are integrated in a manner that is both operationally effective and socially responsible.

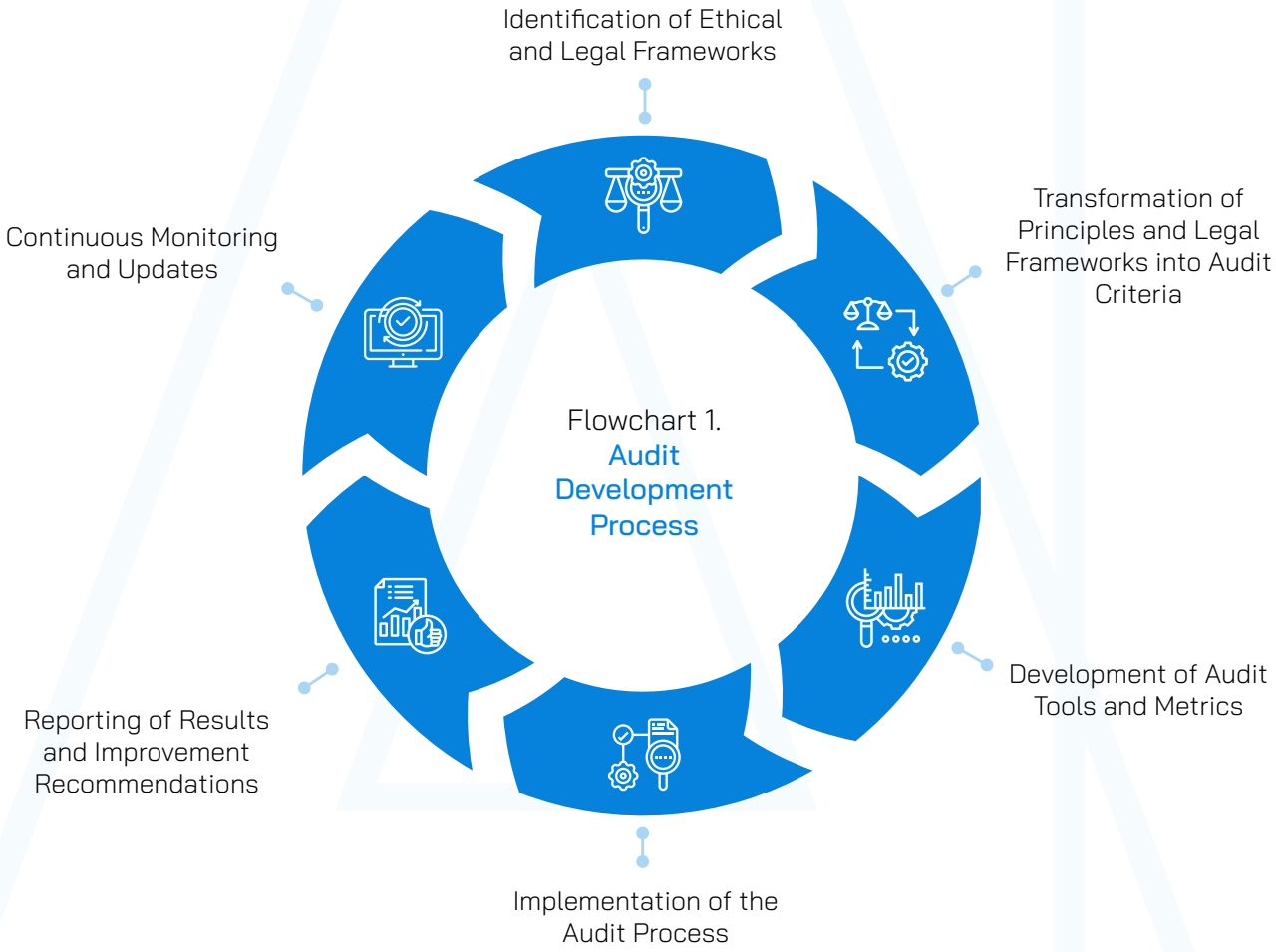
2.6. Transforming the Ethical Principles and Legal Frames into Audit Processes

AI auditing in higher education must be comprehensive, addressing technical performance metrics, ethical compliance, legal requirements, and institutional policy alignment. This multidimensional approach ensures that AI enhances rather than undermines the values of higher education, while maintaining trust and integrity (Luo et al., 2025).

The translation of ethical principles and legal requirements into practical audit procedures is a critical component of AI governance in HEIs. The cyclical steps of audit development are presented in Flowchart 1.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union



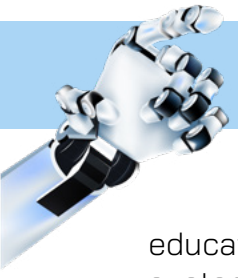
Flowchart 1. **Audit Development Process**

As shown in Flowchart 1, the transformation process begins with identifying ethical principles such as fairness, transparency, accountability, and human dignity. These are considered alongside legal mandates from the GDPR, FERPA, and the AI Act. The principles are then systematically converted into measurable audit criteria. The development of audit tools and metrics enables institutions to operationalise these principles, creating standardised procedures to evaluate AI across contexts.

Implementation produces empirical data that reveal compliance gaps and ethical risks. Reporting findings and recommendations generates a feedback loop, addressing immediate issues while refining principles and criteria for future evaluations (Brown et al., 2021). Continuous monitoring and updating ensure that the audit framework remains responsive to evolving ethical standards, legal obligations, and technological capabilities.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union



3. AI Audit Process Framework

As artificial intelligence (AI) systems become increasingly embedded in higher education, the need for structured evaluation and oversight has grown significantly. These systems—spanning admissions algorithms, learning analytics platforms, chatbots, and predictive modelling tools—are reshaping institutional operations, pedagogy, and student support. While such innovations provide substantial benefits, they also present notable risks, including the amplification of bias, infringements on privacy, opacity in decision-making, and potential misalignment with institutional values and regulatory frameworks (Yan & Tang, 2025).

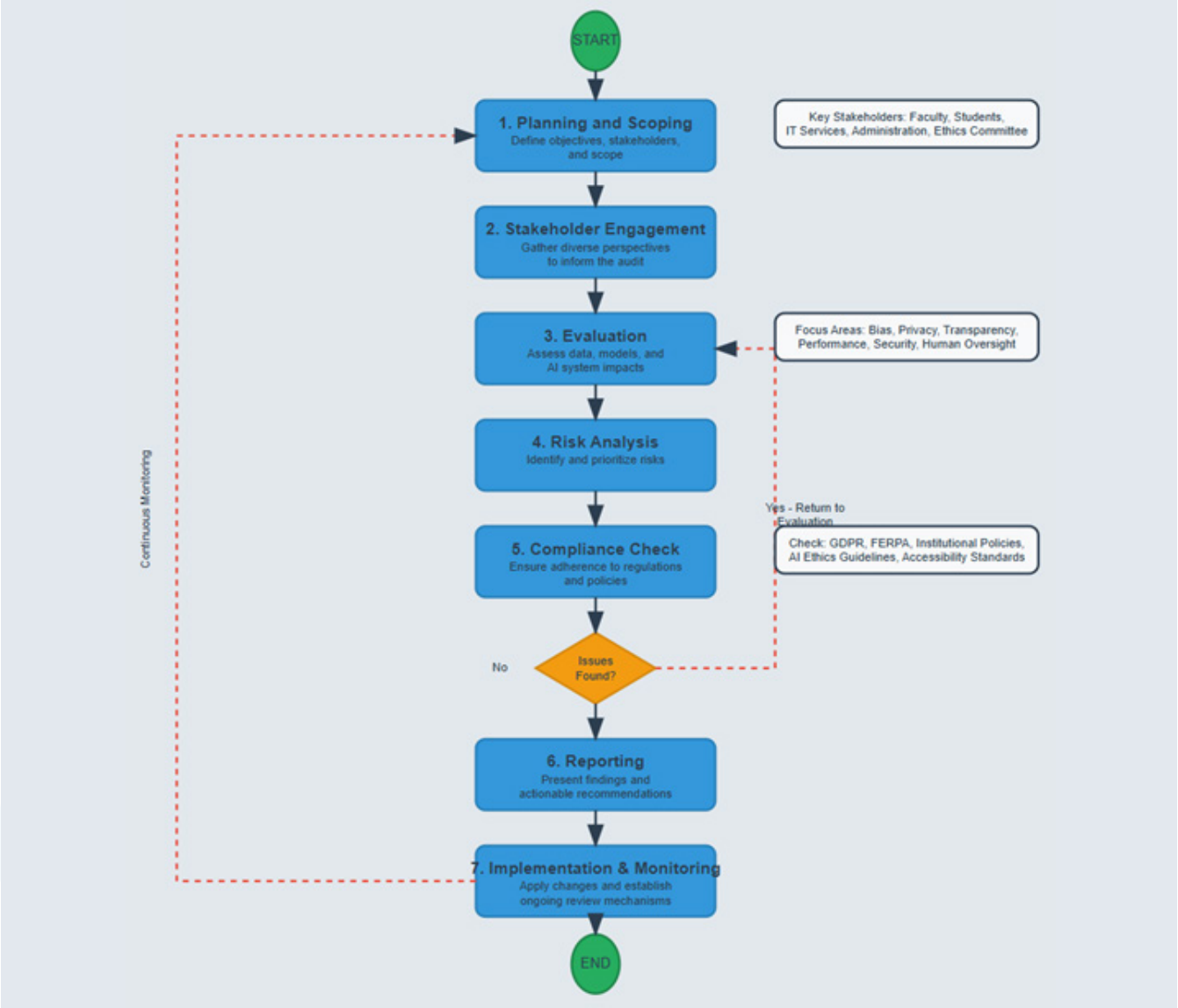
An AI audit constitutes a structured, multidisciplinary evaluation that examines not only technical performance but also ethical integrity, legal compliance, and alignment with institutional objectives. Unlike conventional IT audits, AI audits address broader concerns such as human rights, fairness, transparency, data protection, academic integrity, and adherence to regulatory requirements, including the General Data Protection Regulation (GDPR), the Family Educational Rights and Privacy Act (FERPA), and accessibility standards (Farley & Lansang, 2025; Fernsel et al., 2025).

Given the complexity of AI auditing in higher education, a coherent visual framework is essential to guide institutions through the process. Research on auditability frameworks underscores the importance of structuring evaluations around verifiable claims, supporting evidence, and accessible validation mechanisms (Fernsel et al., 2025). Broader institutional AI governance often builds on principles such as explainability, accountability, fairness, and data privacy, echoing models proposed in consulting practices for responsible AI adoption in higher education (Huron, 2025).

Flowchart 1 presents a seven-phase methodology tailored to the sector. This structured model supports audit teams in navigating sequential and interrelated stages, identifying critical decision points, and maintaining methodological consistency across diverse AI system evaluations. By embedding feedback loops, the framework enables institutions to address deficiencies and implement mechanisms for continuous improvement, ensuring adaptability to different institutional contexts and AI applications.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union



Flowchart 1. Phases of the Audit Process

The AI audit process unfolds across seven interlinked phases. Each phase is designed to ensure a comprehensive evaluation and the continuous improvement of AI systems in higher education. It begins with Planning and Scoping, where institutions define audit objectives, determine the systems under review, and identify relevant stakeholders. Stakeholder Engagement follows, integrating perspectives from students, faculty, administrators, IT services, and ethics committees to capture diverse concerns and expectations. The Evaluation phase examines data quality, algorithmic performance, decision-making transparency, and the broader impact of AI systems on academic integrity, equity, and institutional effectiveness. Risk Analysis identifies and prioritizes potential harms or vulnerabilities based on their likelihood and severity, providing a clear framework for mitigation. Compliance Check verifies adherence to regulatory requirements such as GDPR, FERPA, and the AI Act, and ensures consistency with institutional policies, accessibility obligations, and ethical guidelines. Reporting synthesizes findings into actionable, well-structured recommendations tailored for leadership and relevant operational units. Finally, Implementation and Monitoring establishes mechanisms for applying corrective actions, tracking performance, and creating feedback loops that allow governance structures to remain adaptive to evolving technologies and institutional priorities. This cyclical process ensures that AI audits are not one-off assessments but an ongoing, responsive governance practice.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

4. Detailed Step-by-Step Guidelines

This section outlines the specific tasks associated with each phase of the audit process, as illustrated in Flowchart 1.

Table 2. Step-by-Step Guidelines

Phase	Purpose	Detailed Steps
Phase 1 – Planning and Scoping	Establish the scope, objectives, and governance structure of the AI audit to ensure clarity, independence, and institutional alignment.	**1. Define the Scope** • Identify AI systems to be audited (e.g., admissions algorithms, grading tools). • Describe each system's purpose, operational context, and institutional impact. **2. Set Objectives** • Align goals with institutional priorities. • Consider technical, ethical, legal, and value-based standards. **3. Assemble the Audit Team** • Include IT, legal, ethics, academic reps. • Ensure independence: no prior involvement in audited system. • Engage external experts when possible. • Rotate auditors in cyclical audits. **4. Develop Project Plan** • Timelines, milestones, deliverables. • Coordinate via relevant units: – Research: R&D, Ethics Committee, Tech Transfer. – Teaching: EdTech, Academic Quality Assurance. – Administration: HR, Financial, Student Affairs. • Ensure relevant expert participation. **5. Secure Senior Leadership Support** • Obtain leadership commitment. • Establish governance for cross-department coordination.
Phase 2 – Stakeholder Engagement	Ensure audit reflects diverse perspectives, addresses concerns, and integrates user experiences.	**1. Identify Stakeholders** • Map all affected groups: students, faculty, administrative staff, and governance bodies. • Include underrepresented/vulnerable populations. **2. Develop Engagement Plan** • Use surveys, workshops, focus groups, interviews. • Collect feedback on: – Fairness & transparency. – Data privacy & security. – Accessibility & user experience. – Understanding of AI decision-making. – Human oversight expectations. – Institutional impacts (academic freedom, employment, culture). – Specific use cases & pain points. **3. Translate Feedback into Action** • Refine audit scope & criteria. • Revise goals based on priorities & risks. • Plan ongoing communication & updates. **4. Document & Store Feedback** • Secure structured records. • Link feedback to stakeholder categories.
Phase 3 – Data and Model Evaluation	Assess inputs (data) and processing mechanisms (models) for fairness, accuracy, transparency, and compliance.	**1. Define Scope** • Data: student records, performance data, demographics, LMS analytics, research datasets, admin records, training & operational data. • Models: ML models, statistical, rule-based, NLP, recommendation algorithms. **2. Assess Data Quality** • Verify accuracy, completeness, diversity. • Ensure demographic representation. • Correct missing/outdated/inconsistent data. **3. Analyse Model Performance** • Detect bias across demographic groups. • Validate accuracy, error rates, and fairness metrics. • Robustness testing under varied conditions. • Review transparency/explainability.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

		4. Review Governance & Compliance • Check GDPR, FERPA compliance. • Ensure consent management, retention limits, anonymization. • Align with ethics & accessibility standards.
Phase 4 – Risk Assessment	Transform findings into prioritized, actionable risks.	**1. Expert Consensus Methodology** • Multidisciplinary review (technical, ethical, legal, operational). • Collaborative classification. • Ensure independence from system creators. **2. Classification Framework** • High: threats to reputation, legal, safety, integrity (⩽30 days fix). • Medium: inefficiencies, moderate compliance issues, dissatisfaction (60–90 days). • Low: usability/docs gaps (⩾90 days). **3. Risk Register** • Record source, failed checklist, rationale, stakeholders, impacts, expert assessments. **4. Mitigation Planning** • Define actions, assign responsibili**ty, set timelines, list resources.** **5. Monitoring & Escalation** • Progress checkpoints, success metrics, escalation rules.
Phase 5 – Compliance Review	Verify compliance with regulations, policies, and ethics.	**1. Legal & Regulatory Compliance** • GDPR, FERPA, AI Act (risk categories). • Identify high-risk systems & safeguards. **2. Institutional Policy Alignment** • Check governance, ethics, and accessibility. • Integrate AI Act transparency/accountability duties. **3. Gap Identification** • Document deficiencies, propose updates. • Ensure alignment with legal & strategic priorities. **4. Documentation** • Keep detailed compliance records securely (privacy-by-design).
Phase 6 – Reporting and Recommendations	Communicate results and guide improvement.	**1. Compile Findings** • Risks, compliance gaps, performance outcomes. **2. Actionable Recommendations** • Technical (data, models) + governance (policies, oversight). • Prioritize & set timelines. **3. Executive Summary** • Strategic focus for leadership. • Highlight high-priority risks & resource needs. **4. Data Protection** • GDPR-aligned retention, encryption, access control, audit trails.
Phase 7 – Implementation and Continuous Mon	Implement and sustain improvements over time.	**1. Corrective Actions** • Assign roles using a RACI matrix to define who is Responsible, Accountable, Consulted, and Informed for each task, and track milestones. For example: - Project Manager: Coordinates activities and reports progress. - Technical Expert: Implements AI updates, including retraining and code adjustments. - Legal/Compliance Expert: Ensures alignment with legal and regulatory requirements. - Ethics Expert: Reviews adherence to ethical standards. - Stakeholder Representatives: Faculty and student representatives provide feedback on changes. • Weekly status reports with milestone checkpoints, early risk indicators, and timeline monitoring to prevent delays. **2. Monitoring & Metrics** • Define KPIs for fairness, transparency, accuracy, compliance. • Use risk indicators. **3. Review Frequency** • High: 6 months. • Medium: 1 year. • Low: 2 years. • Extraordinary audits when triggered. **4. Model Drift Detection** • Monitor concept/data drift. • Retrain or adjust when needed. • Example triggers: hybrid learning, economic shifts, AI-assisted cheating, job market changes. **5. Escalation** • Low: resolve within 30 days. • Medium: fix in 90 days. • High: urgent intervention. **6. Continuous Improvement** • Use results to refine policies, strengthen oversight, and improve safeguards.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Model Drift Scenarios in Higher Education Institutions

Model drift occurs when an AI system's performance declines because the conditions it was trained on no longer match the current reality. In higher education, this can manifest in several ways:

Scenario 1 – Student Success Prediction System

Initial Situation: Model trained on 2019 face-to-face education data.
Change: Transition to hybrid education during the COVID-19 pandemic.
Effect: Model cannot interpret online participation, digital assignment submissions, and virtual exams due to lack of prior exposure.
Result: Error rates in predictions increase by 30–40%.

Scenario 2 – Scholarship Allocation Algorithm

Initial Situation: Model trained during a period of economic stability.
Change: Economic crisis leading to higher unemployment and altered family income patterns.
Effect: Model fails to recognize new indicators of financial need.
Result: Unfair scholarship distribution and inefficient resource allocation.

Scenario 3 – Academic Integrity Detection System

Initial Situation: Model trained to detect traditional cheating methods.
Change: Widespread use of AI tools such as ChatGPT.
Effect: Model fails to identify AI-assisted plagiarism or ghost-writing.
Result: Significant gaps in upholding academic integrity.

Scenario 4 – Course Recommendation System

Initial Situation: Model trained on pre-pandemic student preferences.
Change: Shift in demand toward new technology courses aligned with evolving job market needs.
Effect: Model continues recommending outdated course patterns.
Result: Recommendations misaligned with current skill requirements.

Recommendation:

As part of continuous monitoring, institutions should establish a model drift detection mechanism to identify and address performance declines promptly. This ensures AI systems maintain accuracy, fairness, and institutional relevance. A three-tier monitoring approach is detailed in Section 5.3.1 (Technical Evaluation of AI Systems/Tools).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

5. Auditing Steps

The following framework provides a systematic approach to conducting AI audits in higher education institutions. It outlines the preparatory phases that form the foundation for effective evaluation. The auditing process begins with clearly defining institutional objectives and expectations for the audit, followed by comprehensive preparation activities including scope determination, stakeholder identification, and audit team assembly.

5.1. Preparing for the Audit

Preparing for the audit primarily involves defining audit goals and assembling teams with clearly assigned tasks:

Defining the audit goals

• What specific objectives does your institution have for this audit?
• What is the nature of the audit:
- The audit focuses on checking whether the institution's use of AI follows ethical principles and legal requirements;
- The audit focuses on evaluating whether the institution has the structures, policies, resources, and competencies in place to use AI ethically and effectively;
- The audit evaluates how well students, staff, researchers, and administrators understand the use of AI within the institution.
- All of the above
Other:.....

• Align the audit's objectives with:
– Institutional code of ethics
– Internal quality assurance systems
– ESG standards, particularly:
o ESG 1.1 (Policy for Quality Assurance)
o ESG 1.3 (Student-Centred Learning and Assessment)
o ESG 1.5 (Teaching Staff – AI literacy and agency)
o ESG 1.7 (Information Management and AI-generated data use)
Other:.....

Assemble Audit Teams

Effective AI auditing requires a multidisciplinary team with diverse expertise to ensure comprehensive evaluation of technical, legal, and ethical aspects of your AI systems. This team may include quality assurance professionals, technical specialists, legal experts, ethics officers, and domain-specific AI professionals.

Team Composition:

- ☐ Quality Assurance Leadership
- ☐ IT Specialist
- ☐ Legal Advisors
- ☐ Ethics Committee Officers
- ☐ AI systems/research/education/administration experts for specific auditing goals

Assign roles for your team:
1. Name: _____ Role: _____
2. Name: _____ Role: _____

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

5.2. Auditing Ethical AI Use

This ethical auditing framework comprises three specialized audit checklists designed to assess the ethical integration of artificial intelligence systems across core higher education functions: AI-Assisted Research Audit Checklist, AI-Assisted Education Audit Checklist, and AI-Assisted Administrative Processes Audit Checklist. Each checklist addresses the critical need for systematic oversight of AI implementation, ensuring technological advancement aligns with educational values and institutional accountability.

The checklists are structured around seven fundamental ethical principles: Accountability and Responsibility, Bias and Fairness, Human Autonomy and Agency, Privacy and Data Protection, Safety and Security, Inclusivity, and Transparency and Explainability. Each evaluation item is systematically mapped to ESG standards specifically adapted for AI integration in higher education institutions, providing a robust framework for risk assessment, policy development, compliance monitoring, and continuous improvement of AI deployment practices.

Common Checklist Instructions

For each item, select one option only: "Yes," "No," or "N/A" (Not Applicable).

- **"Yes"**: The practice/policy exists and is documented with verifiable evidence
- **"No"**: The practice/policy does not exist or lacks proper documentation
- **"N/A"**: The item is not applicable to your institution's current AI usage

Checklist Requirements

- Base all responses on documented facts and current practices, rather than assumptions or planned activities
- For "No" responses, include internal notes indicating planned remedial actions
- Answer all applicable questions; no item should be left blank
- Upon completion, conduct a full review for accuracy and consistency

Verification and Notes

For "Yes" responses: Supporting documentation should include policies, procedures, training records, audit reports, or other official institutional documents that demonstrate compliance. Briefly mention these in the "Checklist's Notes" column.

For "No" responses: Use the Notes column to indicate relevant details such as:

- Planned remedial actions and timelines
- Reasons for current gaps (e.g., "Policy under development," "Budget constraints," "Technical limitations")
- Priority level for addressing the deficiency
- Responsible department/person for follow-up
- Any interim measures in place

For "N/A" responses: Brief explanatory notes may be added when clarification would be helpful for future reference.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

AUDIT INFORMATION

Institution/Organization: _____

Audit Conducted By:

- Internal Institution/Department: _____
- External Institution/Department:

Audit Team Leader: _____

Position/Title: _____

Audit Team Members:

Audit Period: From _____ to _____

Date of Audit Completion: _____

Target Groups/Faculties Covered:

Before completing the checklist, institutions should first consider the following precondition question.

Precondition question:

1. Does your institution currently use integrated AI systems in research processes with established procedures or policies?

☐ Yes ☐ No

If "No," this checklist is not applicable. If "Yes," proceed with the full assessment.

2. Does your institution currently use integrated AI systems in educational processes with established procedures or policies?

☐ Yes ☐ No

If "No," this checklist is not applicable. If "Yes," proceed with the full assessment.

3. Does your institution currently have integrated AI systems in administrative processes with established policies or procedures?

☐ Yes ☐ No

If "No," this checklist is not applicable. If "Yes," proceed with the full assessment.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

5.2.1. Auditing AI-assisted research process

This section evaluates the integration and governance of AI technologies within institutional research activities. The audit examines whether AI tools used in research processes comply with academic integrity standards, ethical research guidelines, and data protection regulations. Key focus areas include the availability of institutional policies for AI research usage, accessibility and inclusivity of AI tools for all research communities, transparency in AI-driven research methodologies, and the establishment of proper training and feedback mechanisms. This ensures that AI contributes to research productivity while upholding scientific rigor and ethical standards. The following recommendations can serve as a reference for determining audit frequency and team composition for AI-assisted research auditing.

Recommended Audit Periods

Research-critical AI systems: Annually
General research support tools: Every 2 years
Project-specific audits: At project initiation and completion
Triggered audits: Upon ethical concerns, data breaches, or significant research misconduct allegations

CHECKLIST FOR AI-ASSISTED RESEARCH AUDITING

This comprehensive audit checklist framework consists of three sequential sections that must be completed following the provided guidelines:

1. **Audit Information:** Document institutional details, audit team composition, scope, and target groups.
2. **AI-Assisted Research Audit Checklist:** Systematic evaluation of AI integration across six key domains (Accountability & Responsibility, Bias & Fairness, Human Autonomy & Agency, Privacy & Data Protection, Safety & Security, Inclusivity, and Transparency & Explainability).
3. **AI-Assisted Research Assessment Results and Risk Analysis Prioritization:** Assessment of identified risks and development of corrective action strategies.

Instruction: Complete each section in the specified order, ensuring all relevant fields are filled and appropriate checkboxes are marked before proceeding to the next section.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Table 3. AI-Assisted Research Audit Checklist

	Items	Related ESG Standards	Yes	No	N/A	Notes
Accountability and Responsibility	Do institutional strategies and policies exist for integrating AI tools into research processes?	1.1, 1.2				
	Are potential risks (e.g., data protection, bias, intellectual property issues) associated with AI usage being evaluated?	1.4				
	Are intellectual property rights policies clearly defined with respect to the use of AI tools in the research process?	1.1, 1.8				
Bias and Fairness	Is there a documented institutional procedure to review datasets for diversity and representation before their use in AI-supported research?	1.1, 1.7				
	Are researchers required to record and report any observed biases or fairness concerns during AI-supported research?	1.1, 1.8				
	Is there a follow-up process to address and document actions taken when bias or fairness issues are identified in AI-assisted research? (new)	1.1, 1.9				
Human Autonomy and Agency	Is there a documented procedure ensuring that human approval or intervention is mandatory before AI-generated recommendations are adopted in research?	1.8				
	Are researchers required to document how final decisions in AI-assisted research remain under human authority?	1.8				
	Does the institution provide systematic training to enable researchers to critically evaluate and, where necessary, override AI outputs?	1.5				
Privacy and Data Protection	Have data security policies been established concerning the use of AI tools in research?	1.1				
	Are personal or sensitive data provided to AI anonymised or otherwise protected with appropriate documentation?	1.7				
	Are data retention and deletion policies for AI-processed research data clearly defined?	1.1, 1.7				

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Safety and Security	Have AI models been tested for accuracy, reliability, and the risk of unexpected outcomes before deployment in research?	1.9				
	Are incidents affecting the validity or reliability of AI-assisted research outputs documented, together with corrective actions taken?	1.7, 1.8				
	Has the institution established predefined steps to address potential AI-related failures in research?	1.7				
	Are procedures in place to respond to potential security breaches in AI-assisted research?	1.7				
	Is there a protocol to prevent or mitigate potential data loss in AI-assisted research?	1.7				
	Are AI tools regularly updated to maintain security standards and address emerging threats in research processes?	1.9				
Inclusivity	Are the AI tools used in research accessible to disadvantaged groups (e.g. students with disabilities, socio-economically disadvantaged individuals, linguistic minorities, refugee or migrant students, and those from rural or remote areas) in accordance with the principle of inclusivity?	1.3				
	Has the inclusivity of AI tools been evaluated in terms of language support?	1.3, 1.6				
	Are AI tools assessed for accessibility, particularly for users with disabilities or other disadvantages?	1.3, 1.6				
	Is the user-friendliness of AI tools systematically evaluated to ensure inclusivity?	1.3, 1.6				
	Has the inclusivity of AI-generated outputs been assessed regarding whether they reinforce or reduce existing inequalities in research?	1.7				
Transparency and Explainability	Has clear information been provided about the data sources used to develop the AI tools and their overall performance?	1.7, 1.8				
	Are AI-generated outputs clearly documented in a comprehensible way for researchers?	1.8				
	Are the limitations and known weaknesses of AI models documented?	1.2, 1.5, 1.6				
	Are these documented limitations and weaknesses shared with relevant research stakeholders?	1.2, 1.5, 1.6				

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Risk Analysis and Mitigation Plan Required: Yes No
If Yes, Risk Analysis and Mitigation Plan Details:
- Plan Development Responsible: _____
- Target Completion Date for Mitigation Actions: _____
- General Priority Level: High Medium Low

Follow-up Review Date: _____

AI-ASSISTED RESEARCH ASSESSMENT RESULTS AND RISK ANALYSIS PRIORITIZATION

Following completion of the AI-assisted research audit checklist, the next critical step is to analyse and prioritise any identified gaps or deficiencies. The risk prioritisation table below translates checklist findings into actionable priorities, categorising issues by their potential impact and urgency for research integrity and compliance.

This systematic approach ensures that the most critical vulnerabilities are addressed first, while providing a clear timeline for the comprehensive enhancement of the AI-assisted research governance framework.

Table 4. Risk Prioritization - Research Processes

Identified Issue	Area	Risk Level	Rationale	Urgency
No ethics guidelines for AI use	Research Ethics	High	Threat to research integrity and academic reputation	Urgent
No AI training provided for researchers	Capacity Development	Medium	Risk of misuse and decline in research quality	60 days
Inadequate AI research data security	Data Security	High	Intellectual property loss and potential security breach	Urgent
No researcher feedback collected	Continuous Improvement	Low	Weakness in continuous development processes	90 days
Unclear intellectual property policies	Legal Framework	Medium	Legal uncertainty and increased risk of disputes	45 days

AI-Assisted Research Process Risk Analysis and Mitigation

The most critical risks identified in the research domain are the absence of ethics guidelines and deficiencies in data security. These weaknesses directly undermine academic integrity and institutional reputation.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Table 5. Mitigation Plan - Research Processes

Issue	Target Level	Intervention Strategy	Responsible	Deadline	Resource Requirements
AI ethics guidelines deficiency	Comprehensive ethics guideline document	Establishment of a multidisciplinary committee and drafting of guidelines	Ethics Committee and Research Office	30 days	External ethics expertise, consultancy fees
Research data security	ISO 27001 compliant security framework	Strengthening of security infrastructure	IT Security and Research Office	25 days	Security software licences, infrastructure costs
Researcher AI training	⩾ 90% participation rate	Mandatory AI literacy programme	Researcher Development and Education	60 days	Training materials, programme development costs
Intellectual property uncertainty	Clear and updated IP policy	Legal consultancy and policy revision	Legal Counsel and Technology Transfer Office (TTO)	45 days	Legal consultation services, advisory fees
Feedback mechanism	Systematic feedback system Framework	Online survey system, analysis protocol	Quality Assurance and IT	90 days	Survey platform subscription, software costs

5.2.2. Auditing AI-assisted education process

This section evaluates the implementation and management of AI systems within teaching and learning environments. The audit places particular emphasis on institutional strategies for AI integration in education, equitable access to AI-enhanced learning opportunities, and the safeguarding of student data security and privacy. Key areas of examination include the establishment of clear guidelines for the use of AI in academic activities, the provision of ethics training for both staff and students, and the effectiveness of feedback mechanisms for continuous improvement. The purpose of this evaluation is to ensure that AI technologies support educational excellence while protecting student welfare and academic integrity. The following recommendations may serve as a reference for determining audit frequency and the appropriate composition of audit teams for AI-assisted education processes.

Recommended Audit Periods

Student assessment AI systems: Annually

Learning management and analytics systems/Educational support tools: Every two years

Triggered audits: In response to student complaints, academic integrity concerns, or accessibility issues

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

CHECKLIST FOR AI-ASSISTED EDUCATION AUDITING

This comprehensive audit checklist framework consists of three sequential sections that must be completed following the provided guidelines:

1. **Audit Information:** Document institutional details, audit team composition, scope, and target groups.
2. **AI-Assisted Education Audit Checklist**: Systematic evaluation of AI integration across six key domains (Accountability & Responsibility, Bias & Fairness, Human Autonomy & Agency, Privacy & Data Protection, Safety & Security, Inclusivity, and Transparency & Explainability).
3. **AI-Assisted Education Assessment Results And Risk Prioritization**: Assessment of identified risks and development of corrective action strategies focusing on academic integrity, educational equity, and teaching quality.

Instruction: Complete each section in the specified order, ensuring all relevant fields are filled and appropriate checkboxes are marked before proceeding to the next section.

AI-ASSISTED EDUCATION AUDIT CHECKLIST

Table 6. AI-Assisted Education Audit Checklist

Items		Related ESG Standards	Yes	No	N/A	Notes
Accountability and Responsibility	Do institutional strategies and policies exist for integrating AI tools into teaching and learning processes?	1.1, 1.2, 1.4				
	Have clear and explicit rules regarding the use of AI in assignments and classroom activities been established?	1.6				
	Are training programs provided to enhance the ethical awareness of faculty staff and students regarding AI usage?	1.5				
	Do students have a reporting mechanism for issues encountered while using AI tools?	1.3				
	Do educators have a reporting mechanism for issues encountered while using AI tools?	1.3				
Bias and Fairness	Are AI-generated educational content and recommendations regularly reviewed to ensure they do not perpetuate stereotypes or cultural biases?	1.2, 1.8				
	Do AI tools used in education provide equal opportunities for students from diverse socio-economic and cultural backgrounds?	1.3, 1.4				
Human Autonomy and Agency	Are students clearly informed that they have the right to accept or reject AI-generated suggestions in their learning activities?	1.8				
	Are students provided with explicit opportunities to critically evaluate and question AI-generated content in their learning activities?	1.5				

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Privacy and Data Protection	Have data security and confidentiality policies been established regarding the use of AI tools in education?	1.6				
	Are students clearly informed about how their personal data will be collected, used, and stored when engaging with AI-assisted educational tools?	1.7, 1.8				
	Are relevant staff (e.g. academic and administrative) clearly informed about how their personal data will be collected, used, and stored when engaging with AI-assisted educational tools?	1.7, 1.8				
	Is explicit consent obtained from students before their personal data are used in AI-assisted learning activities?	1.7				
Safety and Security	Are there predefined procedures to ensure continuity of teaching and learning in case of AI tool malfunctions or service interruptions?	1.7				
	Have measures been implemented to protect educational processes from potential misuse of AI tools by students or staff?	1.9				
Inclusivity	Are AI tools accessible to disadvantaged groups (e.g. students with disabilities, socio-economically disadvantaged individuals, linguistic minorities, refugee or migrant students, and those from rural or remote areas) in accordance with the principle of inclusivity?	1.3				
	Are training programmes provided to enhance the ethical awareness of students regarding AI usage?	1.5, 1.6				
	Are training programmes provided to enhance the ethical awareness of academic staff regarding AI usage?	1.5, 1.6				
	Are AI-assisted educational materials designed to accommodate diverse learning needs and styles?	1.3				
	Are examples, case studies, and content in AI-assisted learning resources representative of diverse cultural and social backgrounds?	1.2				

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Transparency and Explainability	Is feedback collected from students on the effectiveness and usefulness of AI tools in the learning process?	1.7				
	Is feedback collected from academic staff on the pedagogical aspects of using AI tools in education?	1.7				
	Is feedback collected from academic staff on the ethical aspects of using AI tools in education?	1.7				
	Are pedagogical measures being implemented to improve AI-supported educational environments?	1.8				
	Are ethical measures being implemented to improve AI-supported educational environments?	1.8				
	Are the roles and contributions of AI tools in the learning process clearly communicated to learners at the beginning of the course or activity?	1.8, 1.2				

Risk Analysis and Mitigation Plan Required: Yes No

If Yes, Risk Analysis and Mitigation Plan Details:

- Plan Development Responsible: _____
- Target Completion Date for Mitigation Actions: _____
- General Priority Level: High Medium Low

Follow-up Review Date: _____

AI-ASSISTED EDUCATION ASSESSMENT RESULTS AND RISK PRIORITIZATION

After completing the AI-Assisted Education Audit Checklist, it is essential to evaluate and prioritise identified deficiencies or gaps in institutional AI governance for education. The following risk prioritisation table translates audit findings into a strategic action plan, ranking issues according to their potential impact on academic integrity, student equity, and educational quality.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Table 7. Risk Prioritization - Educational Processes

Identified Issue	Area	Risk Level	Rationale	Urgency
Lack of student awareness of AI usage rules	Academic Integrity	High	Risk of academic misconduct, unfair grading	Urgent
Inequitable access to AI tools among students	Educational Equity	High	Risk of opportunity inequality and discrimination	Urgent
Lack of faculty training in AI ethics	Teaching Quality	Medium	Insufficient guidance, inconsistent application	60 days
Deficiencies in student data security	Data Protection	High	FERPA violation, privacy breaches	Urgent
Insufficient quality control of AI-assisted course content	Academic Quality	Medium	Decline in teaching standarts	45 days

AI-Assisted Education Process Risk Analysis and Mitigation

The most critical risks in education concern academic integrity, educational equity, and student data security. These areas directly affect student rights and the quality of education.

Table 8. Mitigation Plan - Educational Processes

Issue	Target Level	Intervention Strategy	Responsible	Deadline	Resource Requirements
Student AI usage rules	Clear guidelines + ⩾ 90% awareness	Development of AI Academic Integrity Policy, mandatory student training	Academic Affairs + Student Affairs	30 days	External ethics expertise, consultancy fees
Inequality in access to AI tools	100% equal access	Institutional AI tool licences, student support programme	IT + Financial Affairs + Student Affairs	25 days	Security software licences, infrastructure costs
Student data security	Full FERPA compliance	Strengthening of data protection infrastructure, mandatory training programme	Data Protection Office + IT + Education	60 days	Training materials, programme development costs
Faculty AI ethics training	⩾ 90% participation and certification	Comprehensive faculty development programme	Faculty Development	45 days	Legal consultation services, advisory fees
AI course content quality control	Standardised quality protocol	Peer review system, structured quality checklist	Education Quality Office + Curriculum Committee	90 days	Survey platform subscription, software costs

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

5.2.3. Auditing AI-assisted administrative processes

This section examines the deployment and oversight of AI systems in institutional administrative operations. The audit evaluates governance frameworks for administrative AI usage, compliance with data protection and privacy regulations, and the effectiveness of staff training programmes on ethical AI implementation. Key areas of assessment include institutional AI policies for administrative functions, risk management procedures, and accessibility standards. In addition, mechanisms for continuous improvement, informed by staff feedback, are reviewed. This evaluation ensures that AI enhances administrative efficiency while maintaining transparency, accountability, and service quality for all institutional stakeholders. The following recommendations may serve as a reference point for determining audit frequency and team composition in the auditing of AI-assisted administrative processes.

Recommended Audit Periods

Human Resources and Payroll AI systems: Every 6 months
Critical Student Decision Systems (admissions, financial aid, academic evaluation): Every 6 months
General Student Services (information systems, scheduling, support): Annually
General administrative automation: Every 2 years
Triggered audits: Upon staff complaints, operational failures, or compliance violations
External Compliance Auditor (when required)

CHECKLIST FOR AI-ASSISTED RESEARCH AUDITING

This comprehensive audit checklist framework consists of three sequential sections that must be completed following the provided guidelines:

1. Audit Information: Document institutional details, audit team composition, scope, and target groups.

2. AI-Assisted Research Audit Checklist: Systematic evaluation of AI integration across six key domains (Accountability & Responsibility, Bias & Fairness, Human Autonomy & Agency, Privacy & Data Protection, Safety & Security, Inclusivity, and Transparency & Explainability)..

3. 3. AI-Assisted Administration Assessment Results and Risk Analysis Prioritization: Assessment of identified risks and development of corrective action strategies focusing on human resources compliance, financial system security, and operational efficiency.

Instruction: Complete each section in the specified order, ensuring all relevant fields are filled and appropriate checkboxes are marked before proceeding to the next section.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Table 3. AI-Assisted Research Audit Checklist

Items		Related ESG Standards	Yes	No	N/A	Notes
Accountability and Responsibility	Do institutional strategies and policies exist for integrating AI tools into administrative processes?	1.1				
	Are accountability measures implemented in AI-assisted administrative processes, incorporating user feedback?	1.9				
	Are roles and responsibilities clearly defined for monitoring and managing AI-assisted administrative tasks?	1.1				
	Are training programmes provided to support the ethical use of AI in administrative processes?	1.5				
Bias and Fairness	Are AI-assisted recruitment and hiring processes regularly audited for potential bias related to demographic characteristics (e.g. gender, age, ethnicity, disability status)?	1.4, 1.9				
	Is there a process to review AI-assisted administrative decisions to ensure they do not systematically disadvantage any group?	1.4, 1.9				
Human Autonomy and Agency	Do administrative staff have a mechanism to report issues encountered when using AI tools?	1.7, 1.9				
	Is feedback systematically collected from administrative staff regarding the use of AI tools?	1.7, 1.9				
	Are staff empowered and authorised to override AI-generated suggestions or decisions in administrative processes when necessary?	1.8				
Privacy and Data Protection	Have data security and privacy policies been established for the use of AI tools in administrative processes?	1.1				
	Are procedures in place for securely deleting or anonymising administrative data processed by AI tools after its intended use?	1.1, 1.7				

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Safety and Security	Are administrative staff regularly informed of their responsibilities for protecting sensitive data when using AI tools?	1.5				
	Are potential risks associated with AI usage regularly evaluated?	1.9				
	Are contingency measures defined to maintain critical administrative functions in the event of AI system failures or security breaches?	1.1, 1.9				
	Is there a regular review process to ensure AI tools remain protected against emerging security threats?	1.9				
Inclusivity	Are there institutional policies to ensure that AI tools used in administrative processes comply with the principles of accessibility and inclusivity?	1.3				
	Have administrative AI tools been tested to confirm their usability for individuals with varying levels of digital literacy?	1.3, 1.6				
Transparency and Explainability	Are administrative staff informed about which AI tools are used in their work processes and how these tools affect their daily tasks?	1.8				
	Are administrative AI tools accompanied by clear, non-technical explanations of how they function and how outputs should be interpreted?	1.8				
	Are the limitations and potential errors of AI tools in administrative processes clearly documented and communicated to relevant staff?	1.5, 1.8				

Risk Analysis and Mitigation Plan Required: Yes No

If Yes, Risk Analysis and Mitigation Plan Details:

- Plan Development Responsible: _____
- Target Completion Date for Mitigation Actions: _____
- General Priority Level: High Medium Low

Follow-up Review Date: _____

AI-ASSISTED ADMINISTRATIVE PROCESS ASSESSMENT RESULTS AND RISK PRIORITIZATION

The most critical risks in administrative processes relate to bias in human resources systems and deficiencies in financial system security. These issues require immediate intervention as they pose significant risks of legal sanctions and institutional liability.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Table 10. Risk Prioritization - Administrative Processes

Identified Issue	Area	Risk Level	Rationale	Urgency
Bias detected in HR decision processes	Human Resources	High	Potential breach of employment law; risk of discrimination	Urgent
Data security deficiencies in financial systems	Financial Affairs	High	Risk of financial loss and regulatory non-compliance	Urgent
Insufficient staff AI training	Capacity Development	Medium	Decline in operational efficiency	60 days
No AI system risk assessment conducted	Risk Management	Medium	Late detection of potential risks	45 days
Lack of staff complaint mechanism	Accountability	Low	Staff dissatisfaction; reduced process improvement	90 days

AI-Assisted Administrative Process Risk Analysis and Mitigation

The most critical risks in administrative processes relate to bias in human resources systems and deficiencies in financial system security. These issues require immediate intervention as they pose significant risks of legal sanctions and institutional liability.

Table 11. Mitigation Plan - Administrative Processes

Issue	Target Level	Intervention Strategy	Responsible	Deadline	Resource Requirements
HR system bias	⩾ 95% fair outcome rate	Deployment of bias detection applications; revision of algorithms	HR + IT + Legal	20 days	Bias analysis tools, testing costs
Financial system security	Banking-grade security	Advanced encryption; multi-factor authentication	Financial Affairs + IT Security	25 days	Security infrastructure costs, system fees
Staff AI training	⩾ 90% staff training rate	Department-based AI literacy programme	HR Training + IT	60 days	Training materials, programme costs
AI risk assessment	Systematic risk protocol	Establishment of risk assessment framework	Risk Management + IT	45 days	Risk management expert services, consulting fees
Staff complaint system	Effective feedback channel	Implementation of a digital feedback platform	HR + IT Support	90 days	Feedback platform subscription, software costs

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

5.2.4. Rationale for Benchmark Values in The Guideline

In this guideline, both percentage-based and non-percentage figures are adopted as institution-set benchmark values—policy-defined, auditable targets that translate governance principles into measurable objectives for higher education institutions. The referenced frameworks emphasise objective-setting, awareness, risk management, monitoring, and continual improvement, yet they generally do not prescribe numerical thresholds. Accordingly, the figures in this guideline serve as institutional benchmarks that operationalise those requirements into transparent and testable practice (International Organization for Standardization and International Electrotechnical Commission [ISO/IEC], 2022; ISO/IEC, 2023; National Institute of Standards and Technology [NIST], 2023; UNESCO, 2023b; UNESCO, 2023c).

Accessibility and inclusion are treated as matters of binary legal compliance. Public-sector web and mobile services are required to be accessible under Directive 2016/2102 of the European Parliament and of the Council, implemented via the harmonised standard EN 301 549 of the European Telecommunications Standards Institute (ETSI). Because legal conformity cannot be interpreted along a continuum, targets relating to digital access are set at 100% equal access, reflecting that partial compliance cannot constitute an acceptable steady state (European Parliament & Council, 2016; ETSI, 2021).

For training, awareness, and certification, governance frameworks require institutions to establish objectives, ensure stakeholder awareness of roles and responsibilities, and implement mechanisms for continual improvement, but they do not prescribe participation rates. A benchmark of at least 90% coverage for students, staff, faculty, and researchers is therefore adopted. This figure approximates "near-universal" reach whilst remaining operationally feasible in academic contexts characterised by turnover, sabbaticals, and leave. The percentage represents an institutional choice that implements the awareness requirements of ISO/IEC 27001, the objective-setting and continual-improvement expectations of ISO/IEC 42001, and aligns with UNESCO's emphasis on capacity building for the safe, ethical, and responsible use of artificial intelligence (ISO/IEC, 2022; ISO/IEC, 2023; UNESCO, 2023b; UNESCO, 2023c).

For fair outcomes in automated decisions, authoritative instruments require risk management and bias mitigation but do not prescribe a single fairness threshold. This guideline therefore adopts a stricter institutional benchmark of at least 95% fair-outcome rate to minimise disparate impacts and reduce measurement variability in contexts such as staff recruitment or student admissions. The "four-fifths rule" (80%) established by the United States Equal Employment Opportunity Commission (EEOC) Uniform Guidelines is acknowledged as a heuristic for identifying potential adverse impact but is not treated as a binding threshold in this context (European Parliament & Council, 2024; EEOC, 1978).

For technical performance in decision-support models, lifecycle frameworks require the definition and governance of fit-for-purpose metrics, yet do not establish universal thresholds. To ensure comparability and escalation clarity across heterogeneous systems, a benchmark of at least 80% is adopted for accuracy, precision, and recall. Stricter thresholds may be established for high-risk applications. This value is explicitly identified as an institutional benchmark designed to promote governance clarity and consistent monitoring (NIST, 2023; ISO/IEC, 2023).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Data quality and drift monitoring are presented as diagnostic controls rather than externally mandated thresholds. Indicators such as 15% data missingness or a 25% increase in outliers are interpreted as observed values rather than prescriptive limits. Since the acceptability of missing data depends on the underlying mechanism—such as missing completely at random, missing at random, or missing not at random—and on the analytic remedy, no universal "acceptable percentage" is identified in statistical literature. The commonly cited aspiration of below 5% missing data is recognised as a quality-assurance benchmark rather than a strict rule. Distribution shift is monitored using the Population Stability Index (PSI), where practice-based heuristics typically interpret values below 0.10 as small change and 0.20–0.25 or above as substantial drift warranting investigation or model refresh. These cut-offs are regarded as practice-based heuristics rather than legal requirements (Little & Rubin, 2019; du Pisanie, Allison, & Visagie, 2023; Yurdakul & Naranjo, 2020).

Where observed performance falls below a benchmark (e.g. a 65% accuracy baseline against the ⩾ 80% deployment threshold), time-bound improvement plans (e.g. 12–24 months) are implemented. This ensures that thresholds function as progressive benchmarks rather than static pass/fail gates. The approach aligns with management-system expectations to set objectives, demonstrate continual improvement, and document governance decisions and progress (ISO/IEC, 2023).

6. Evaluating the Overall Audit Process

This section presents a structured framework for evaluating AI audits in higher education, determining institutional risk profiles, and formulating improvement strategies. Findings from four audit areas—AI systems/tools, research, education, and administration—are consolidated to assess AI governance maturity. The methodology applies a multidimensional analysis of risk impact and likelihood, supported by explicit prioritisation criteria. Each area is evaluated separately before results are integrated into an institutional profile that incorporates action plans and monitoring mechanisms for continuous improvement. The table below summarises the core components of the risk assessment methodology, detailing its purpose, approach, process steps, risk classification levels, and integration into institutional improvement plans. It serves as a practical reference for audit teams to ensure actionable outcomes across all AI audit areas.

Table 12. Risk Assessment Components

Component	Details
Purpose	To ensure consistent, expert-driven risk evaluation across all AI audit areas, supporting institutional decision-making.
Approach	Standard methodology using expert consensus rather than purely mathematical scoring.
Process	**1. Identification** – Failed items from checklists (Sections 5.2.1, 5.2.2, 5.3, 5.4, 5.5) are listed. **2. Multidisciplinary Review** – Technical, ethical, legal, and operational experts assess items from their perspectives. **3. Consensus Meeting** – The team discusses differences and agrees on risk classification. **4. Decision Recording** – Majority vote taken; rationale documented.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Risk Levels & Required Action Time	**High priority** – Immediate threats to reputation, compliance, safety, academic integrity, or finances → **Action within 30 days**. **Medium priority** – Operational inefficiencies, performance decline, minor compliance issues → **Action within 60–90 days**. **Low priority** – Minimal operational impact, usability issues, documentation gaps → **Action within 90+ days**.
Decision Roles	- **Team Leader** – Moderates discussion, sets agenda. - **Technical Expert** – Evaluates technical aspects. - **Ethics Expert** – Assesses ethical/social impact. - **Legal Expert** – Reviews legal implications. - **All Members** – Participate in classification and vote.
Documentation	All classifications recorded with justifications, responsible parties, realistic timelines, and resource needs.
Integration into Mitigation Plan	Risk classifications directly inform corrective actions, resource allocation, and monitoring schedules, ensuring that high-priority risks are addressed first whilst maintaining transparency and accountability.

This risk assessment framework provides a structured, expert-driven process for evaluating failed audit items across all AI governance areas. By combining multidisciplinary review, consensus-based prioritisation, and clear role assignments, it ensures consistent classification of risks into high, medium, and low priority levels. Its transparent and accountable approach enables higher education institutions to address immediate risks effectively while strengthening long-term adaptability and continuously improving their AI governance maturity.

7. Overall Audit Report

The overall audit report template is compiled from the contributions of all audit teams. A joint summary report is thus produced, integrating each team's feedback on its respective audit area. This consolidated structure facilitates monitoring for decision-makers and provides an objective basis for informed institutional decisions. The audit report template below highlights institutional strengths, areas for improvement, and critical issues, while also incorporating final evaluations and recommendations. Completed audit checklists, risk prioritisation tables, and mitigation plans are appended to this overall audit report, which is submitted to the institution's highest governing body (e.g. rector, vice-rector, or quality commission chair).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Audit Report Template

Table 13. Findings Based on Audit Results

Audit Area	Strengths	Areas for Improvement	Critical Issues
Ethics/Legal Audit of AI			
AI-Assisted Research Processes Audit			
AI-Assisted Educational Processes Audit			
AI-Assisted Administrative Processes Audit			
Technical Audit of AI Systems and Tools			
Ethics/Legal Audit of AI Systems and Tools			

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Recommendations and Action Planning

Building upon the comprehensive findings outlined above, the following table translates identified areas for improvement and critical issues into specific, actionable strategies. These recommendations are prioritised based on risk assessment outcomes and are designed to strengthen institutional AI governance while addressing immediate compliance and operational concerns.

Table 14. Recommendations for Institutional Improvement and Critical Issues

Audit Area	Recommendations for Areas for Improvement	Recommendations for Critical Issues
Ethics/Legal Audit of AI		
AI-Assisted Research Processes Audit		
AI-Assisted Educational Processes Audit		
AI-Assisted Administrative Processes Audit		
Technical Audit of AI Systems and Tools		
Ethics/Legal Audit of AI Systems and Tools		

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

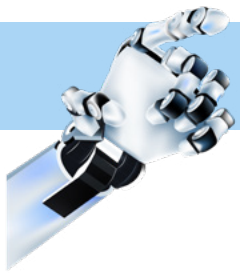
Co-funded by the European Union

Final Evaluation:

Additional Notes or Considerations

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union



Appendix

Technical Auditing of the AI systems/tools

This section presents two complementary approaches for the comprehensive evaluation of artificial intelligence (AI) systems and tools used in higher education institutions. First, systems are assessed using measurable criteria such as technical performance, accuracy rates, speed metrics, and operational efficiency. Second, ethical evaluation is undertaken within the framework of algorithmic transparency, fairness, inclusivity, and accountability principles. This two-dimensional approach is designed to ensure that AI systems operate both technically reliably and ethically responsibly. The following recommendations may serve as a reference for determining audit frequency and team composition for AI systems and tools.

Recommended Audit Periods

- **High-risk systems:** Annualy (admission algorithms, grading systems, financial aid decisions)
- **Medium-risk systems:** Every 2 years (learning analytics, chatbots, recommendation systems)
- **Low-risk systems:** Every 3 years with total auditing (general information tools, basic automation)
- **Triggered audits:** Upon system updates, performance issues, security incidents, or regulatory changes

Recommended Audit Team Members

IT Systems Administrator (Lead)
Data Security Specialist
AI/Machine Learning Technical Expert
Quality Assurance Officer
Legal Compliance Advisor
Ethics Committee Representative
End-user Representative (faculty/staff/student)

Technical evaluation of the AI systems/tools

This checklist evaluates the core technical performance indicators of your AI systems to ensure they meet institutional standards for accuracy, efficiency, and operational requirements.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Scalability and Maintenance:

Metric	Description	Measure (M)	Benchmark values	Evaluation (Success/Fail)
Accuracy	Accuracy is the ratio of correctly predicted examples to the total number of examples (If you have 100 examples and your model correctly predicts 85 of them, then the accuracy value is 85%.).	(1 to 100)	If (M>80%) success else fail	☐ Success ☐ Fail
Reliability (Precision)	Precision is the ratio of truly positive examples among the examples predicted as positive by the model. It answers the question 'how accurate are the model's positive predictions? (If your model made 'positive' predictions for 100 examples and 80 of these are actually positive, then the precision value is 80%)	(1 to 100)	If (M>80%) success else fail	☐ Success ☐ Fail
Recall	Recall is the ratio of correctly predicted positive examples among all the actually positive examples. It answers the question 'what proportion of the actual positives were found by the model?(If your dataset contains 100 examples that are actually positive, and your model correctly identifies 75 of them as positive, then the recall value is 75%.)	(1 to 100)	If (M>80%) success else fail	☐ Success ☐ Fail
Efficiency				
• **Inference Time**	Inference Time is the duration an already trained model takes to produce a prediction for a new input (example). In other words, it is the time required by the model to complete the prediction process (If a natural language processing model takes 0.2 seconds to analyze a sentence, the inference time is 200 ms).	**Successful if below X ms**	If (M<200ms) success else fail	☐ Success ☐ Fail
• **Training Time**	The time required for the model to learn meaningful patterns, language structures, and semantic relationships from text data	**1 to 10** **1:** 14+ days *(Very slow, unacceptable)* **2:** 10-14 days *(Quite slow)* **3:** 7-10 days *(Slow)* **4:** 5-7 days *(Below average)* **5:** 3-5 days *(Average)* **6:** 2-3 days *(Above average)* **7:** 1-2 days *(Good)* **8:** 12-24 hours *(Very good)* **9:** 6-12 hours *(Near perfect)* **10:** <6 hours *(Excellent)*	If (M>8) success else fail	☐ Success ☐ Fail

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

• **Training Time**	The time required for the model to learn meaningful patterns, language structures, and semantic relationships from text data	**1 to 10** **1-2:** Very weak understanding ability **3-4:** Weak understanding ability **5-6:** Medium level understanding ability **7-8:** Good understanding ability **9-10:** Excellent understanding	If (M>8) success else fail	☐ Success ☐ Fail

Can the AI systems/tools adapt to growing needs?
☐ Yes ☐ No

Are the AI systems improved with the feedbacks?
☐ Yes ☐ No

Does the AI follow data minimization principles?
☐ Yes ☐ No

Are secure storage practices implemented?
☐ Yes ☐ No

Is regular data breach testing conducted?
☐ Yes ☐ No

AI Systems/Tools Technical Performance Risk Analysis

Following the completion of the technical performance metrics assessment, identified deficiencies are systematically prioritized based on their potential impact on institutional operations, student outcomes, and system reliability. This risk analysis translates performance gaps into actionable priorities for technical remediation.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Table 15. Risk Prioritization - Technical Performance

Identified Issue	System/Area	Risk Level	Rationale	Urgency
Student admission algorithm 65% accuracy (benchmark 80%)	Student Affairs	High	Incorrect admission/rejection decisions, legal liability	Urgent
Chatbot NLU performance 4/10 score	General Information System	Medium	User dissatisfaction, time loss	60 days
Assignment evaluation system precision ⩾ 80%	Academic Assessment	High	Unfair grading, academic complaints	Urgent
Library recommendation system inference time 250ms	Library Services	Low	Minor user experience issue	90 days
Financial system NLU understanding capacity 6/10	Financial Affairs	Medium	Transaction errors, operational delays	45 days

AI Systems/Tools Risk Analysis and Mitigation

As seen from these examples, technical failures in critical decision-making processes such as student admission and academic evaluation are classified as high risk because they directly impact individual rights and institutional responsibilities. Issues in support systems like general information chatbots can be evaluated in the medium risk category.

Table 16. Mitigation Plan - Technical Performance

Issue	Target Level	Intervention Strategy	Responsible	Deadline	Resource Requirements
Student admission algorithm low accuracy	⩾ 80% accuracy	Model retraining, data quality improvement	IT Dept. + Data Science Team	30 days	Data scientist support, development costs
Assignment evaluation system bias	⩾ 80% precision, fair scoring	Algorithm review, test set expansion	Academic IT + Assessment Unit	25 days	External consultant services, consulting fees
Chatbot understanding problem	8/10+ NLU score	Knowledge base update, NLP model training	IT Support + Content Team	60 days	NLP expert support, training costs
Financial system NLU capacity	9/10+ understanding score	Domain-specific lexicon, specialised model	Financial Affairs + IT	45 days	Software licensing fees, licensing costs
Library recommendation speed	<100ms inference	Server optimization, cache mechanism	IT Infrastructure	90 days	Hardware upgrade costs, equipment fees

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

AI Model Drift Detection Mechanism

The model drift detection mechanism operates through a structured three-layer approach that balances thoroughness with practicality. Each layer serves a specific purpose in maintaining AI system reliability while requiring minimal resources from institutional staff.

Layer	Frequency	Duration	Key Activities	Primary Focus
Layer 1	Weekly	5 minutes	Monitor accuracy, complaints, processing time	Performance indicators
Layer 2	Monthly	30 minutes	Data quality checks, PSI calculation	Data integrity
Layer 3	Quarterly	2 hours	Full evaluation, stakeholder feedback	Strategic assessment

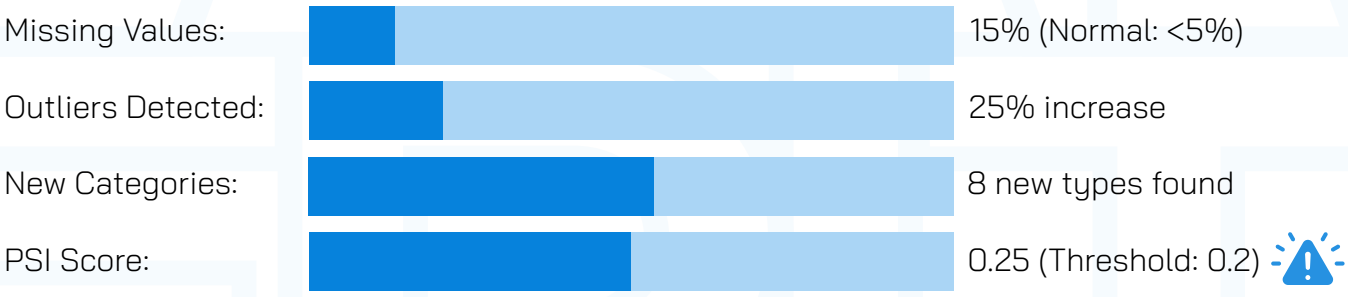
First Layer: Weekly Performance Assessment

The weekly performance check serves as the frontline defense against model degradation. During this assessment, institutions monitor three critical indicators: system accuracy rate, user complaint count, and processing time average. When system accuracy drops below 80 percent from an established baseline of 90 percent or higher, this indicates significant performance degradation. A 50 percent increase in user complaints suggests technical issues or changing user needs, while doubled processing time often signals that the model is struggling with new types of data.

Second Layer: Monthly Data Integrity Evaluation

The monthly data health check focuses on ensuring data quality and consistency. This evaluation compares current input data with original training data to identify significant changes. The cornerstone of this layer is the Population Stability Index - when PSI exceeds 0.2, this indicates substantial data drift requiring model adjustment.

Data Health Monitoring:



Third Layer: Comprehensive Quarterly Assessment

The quarterly review represents the most thorough evaluation, including full model performance analysis, systematic stakeholder feedback collection, and strategic decision-making regarding model updates based on accumulated evidence.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Risk-Based Alert Classification System

The detection mechanism categorizes system status into three distinct levels based on severity and scope of identified issues.

Status Level	Criteria	Required Actions	Response Time
Stable Operations	All metrics within range	Continue normal monitoring	N/A
Elevated Monitoring	1 metric crosses threshold	Increase checks to weekly, investigate	24-48 hours
Critical Intervention	Multiple metric failures	Immediate investigation, manual override	0-24 hours

During stable operations, institutions continue normal operations while maintaining regular monitoring schedules. Elevated monitoring occurs when one performance metric crosses its warning threshold, requiring increased surveillance and preliminary investigation. Critical intervention emerges when multiple metrics simultaneously fail, demanding immediate root cause investigation and urgent corrective measures.

Implementation Strategy

Implementation begins with installing a basic monitoring dashboard and establishing baseline metrics from current system performance data. The institution designates one IT staff member as the primary monitor responsible for weekly checks.

Implementation Timeline:

Week 1-2: Setup Phase

- Install monitoring dashboard
- Establish baseline metrics
- Assign responsible staff

Week 3+: Operations

- Weekly: 5-min dashboard check
- Monthly: 30-min stakeholder meeting
- Quarterly: 2-hour comprehensive review

Operational Phase

Each week (e.g., monday mornings), the designated staff member spends five minutes checking the dashboard for alerts and logging issues. Monthly meetings bring together stakeholders for performance trend review and action planning. Quarterly reviews provide opportunities for comprehensive system evaluation and strategic decisions about model updates.

Resource Requirements and Response Protocol

Human resource requirements include one IT staff member contributing five minutes weekly and one supervisor providing 30 minutes monthly for oversight. Technical infrastructure needs remain modest, requiring only basic dashboard software, automated data export capabilities, and simple email alert functionality.

When the monitoring system detects concerning changes, institutions follow a structured 15-day response timeline:

Response Protocol (PSI > 0.2 OR Accuracy < 80%): Investigate root causes Collect current data samples Retrain model (old + new data) Deploy updated model Verify improvement

This approach ensures rapid response while maintaining quality control throughout the remediation process. The streamlined methodology delivers significant drift detection benefits while requiring minimal complexity, making it highly practical for higher education institutions operating under typical resource constraints.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

References

Acosta-Vargas, P., Salvador-Acosta, B., Novillo-Villegas, S., Sarantis, D., & Salvador-Ullauri, L. (2024). Generative Artificial Intelligence and Web Accessibility: Towards an Inclusive and Sustainable Future. *Emerging Science Journal*. https://doi.org/10.28991/esj-2024-08-04-021.

Adeoye, O., Alimi, A., Agboola, O., Akindele, A., Arulogun, O., & Adigun, G. (2025). Advancing Higher Education through Artificial Intelligence (AI): A Framework for Teaching, Assessment, and Research Integration. *East African Journal of Education Studies*. https://doi.org/10.37284/eajes.8.2.2946.

AI in Education: A Microsoft Special Report, n.d.).

Al-Omari, O., Alyousef, A., Fati, S., Shannaq, F., & Omari, A. (2025). Governance and Ethical Frameworks for AI Integration in Higher Education: Enhancing Personalized Learning and Legal Compliance. *Journal of Ecohumanism*. https://doi.org/10.62754/joe.v4i2.5781.

Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, 17. https://doi.org/10.1186/s41239-020-00218-x.

Brown, S., Davidović, J., & Hasan, A. (2021). The algorithm audit: Scoring the algorithms that score us. *Big Data & Society*, 8. https://doi.org/10.1177/2053951720983865.

Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, 6, 1421273. https://doi.org/10.3389/fhumd.2024.1421273

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 20, 1-22. https://doi.org/10.1186/s41239-023-00392-8

Crompton, H., & Song, D. (2021). The Potential of Artificial Intelligence in Higher Education. *Revista Virtual Universidad Católica del Norte*. https://doi.org/10.35575/RVUCN.N62A1.

DataGuard. (2024, August 20). The EU AI Act and obligations for providers. DataGuard.
du Pisanie, J., Allison, J. S., & Visagie, J. (2023). A proposed simulation technique for population stability testing in credit risk scorecards. *Mathematics*, 11(2), 492. https://doi.org/10.3390/math11020492

ENQA. (2015). Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG). European Association for Quality Assurance in Higher Education, Brussels, Belgium.

European Parliament & Council. (2016, December 2). *Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies. Official Journal of the European Union*, L 327, 1–15. https://eur-lex.europa.eu/eli/dir/2016/2102/oj/eng

European Parliament & Council. (2024, July 12). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union*. http://data.europa.eu/eli/reg/2024/1689/oj

European Telecommunications Standards Institute. (2021, March 10). *EN 301 549 V3.2.1: Accessibility requirements for ICT products and services*. https://www.etsi.org/deliver/etsi_en/301500_301599/301549/03.02.01_60/en_301549v030201p.pdf (see also overview page: https://www.etsi.org/human-factors-accessibility/en-301-549-v3-the-harmonized-european-standard-for-ict-accessibility)

Fernsel, L., Kalff, Y., & Simbeck, K. (2025). Audits for Trust: An Auditability Framework for AI-Based Learning Analytics Systems. , 51-62. https://doi.org/10.5220/0013254300003932.

Grimmelikhuijsen, S. (2022). Explaining why the computer says no: algorithmic transparency affects the perceived trustworthiness of automated decision–making. *Public Administration Review*. https://doi.org/10.1111/puar.13483.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

Hasanzadeh, F., Josephson, C., Waters, G., Adedinsewo, D., Azizi, Z., & White, J. (2025). Bias recognition and mitigation strategies in artificial intelligence healthcare applications. *NPJ Digital Medicine*, 8. https://doi.org/10.1038/s41746-025-01503-7.

International Organization for Standardization & International Electrotechnical Commission. (2022). *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection—Information security management systems—Requirements*. https://www.iso.org/standard/27001

International Organization for Standardization & International Electrotechnical Commission. (2023). *ISO/IEC 42001:2023 Artificial intelligence management system—Requirements*. https://www.iso.org/standard/4200

Johnson, M., Liu, X., & McCaffrey, D. (2022). Psychometric Methods to Evaluate Measurement and Algorithmic Bias in Automated Scoring. *Journal of Educational Measurement*. https://doi.org/10.1111/jedm.12335.

Lazcoz, G., & Hert, P. (2023). Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities. *Comput. Law Secur. Rev., 50*, 105833. https://doi.org/10.1016/j.clsr.2023.105833.

Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley. https://doi.org/10.1002/9781119482260

Liu, D., & Bates, S. (2025). *Generative AI in higher education:Current practices and ways forward*.

Luo, X., Wang, X., & Jiang, T. (2025). Application of AI technology in audit risk assessment and control: Taking internal audit of higher education institutions as an example. *Journal of Infrastructure, Policy and Development*. https://doi.org/10.24294/jipd10125.

Murdoch, B. (2021). Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Medical Ethics*, 22. https://doi.org/10.1186/s12910-021-00687-3.

National Institute of Standards and Technology [NIST]. (2023). *Artificial Intelligence Risk Management Framework* (AI RMF 1.0) (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1 (PDF: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf)

Salhab, W., Ameyed, D., Jaafar, F., & Mcheick, H. (2024). A Systematic Literature Review on AI Safety: Identifying Trends, Challenges, and Future Directions. *IEEE Access*, 12, 131762-131784. https://doi.org/10.1109/ACCESS.2024.3440647.

Turahman, I. (2024). Decision Support System for Achievement Scholarship Recipients at Vocational High Schools with the Analytical Hierarchy Process Method. *Engineering: Journal of Mechatronics and Education*. https://doi.org/10.59923/mechatronics.v1i1.14.

UNESCO (2023a). *Guidance for generative artificial intelligence in education and research* (F. Miao & W. Holmes, Eds.). https://unesdoc.unesco.org/ark:/48223/pf0000386693 (landing page with last update: April 14, 2025)

UNESCO Ethical Impact Assessment. (2023b). Tool for assessing AI systems' ethical impact.

UNESCO Readiness Assessment. (2023c). Guidance on assessing national AI ethics readiness.

UNESCO Recommendation on the Ethics of AI. (2022). Values and principles for ethical AI.

United States Equal Employment Opportunity Commission. (1978). *Uniform Guidelines on Employee Selection Procedures* (29 C.F.R. Part 1607). https://www.ecfr.gov/current/title-29/subtitle-B/chapter-XIV/part-1607 (alt. access: https://www.law.cornell.edu/cfr/text/29/part-1607).

Yurdakul, B., & Naranjo, J. D. (2020). Statistical properties of the population stability index. *Journal of Risk Model Validation*, 14(3), 89–100. https://doi.org/10.21314/JRMV.2020.227

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union

AI-ERIT: AI Integration Framework for Higher Education: Ensuring HEIs' Readiness for Inescapable Transformation

Project Code: 2024-1-LT01-KA220-HED-000251565

Discover our resources
on the project website

www.ai-erit.eu

Created by: SMK College of Applied Sciences, Fachhochschule des Mittelstands, Turība University, Akademia WSEI, Aspira University of Applied Sciences, Universidade da Maia, Batman Üniversitesi

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-220-HED-000251565

Co-funded by the European Union