



AI PIONEERS
Artificial Intelligence in education & training

Handbook on Policy and Ethics in AI Education

June 2025

Lead Beneficiary of the deliverable: Universidad del País Vasco/Euskal
Herriko Unibertsitatea (UPV/EHU)

Work Package 5

Deliverable D5.2 "Handbook on Policy and Ethics in AI Education"

Project number: 101087261

Project name: AI and the future of Education

Project acronym: AI Pioneers

Project coordinator: University of Bremen, Institute Technology and
Education (ITB) Call: ERASMUS-EDU-2022-PI-FORWARD

Topic: ERASMUS-EDU-2022-PI-FORWARD-LOT1

Type of action: ERASMUS Lump Sum Grants

Granting authority: European Education and Culture

Executive Agency Project starting date: 01/01/2023

Project end date: 31/12/2025

Project duration: 36 months



Co-funded by
the European Union



AI PIONEERS
Artificial intelligence in education & training



AI PIONEERS
Artificial Intelligence in education & training

English Version of the Deliverable

June 2025



Handbook on Policy and Ethics in AI Education

Gorka Roman Etxebarrieta (Author)

Maria Orcasitas-Vicandi (Author)

Natalia Louleli (Author)

Asier León Nuñez (Author)

Itsaso Biota Piñeiro (Author)

Ana Boyano Murillo (Author)

Raquel Justo Blanco (Author)

Josune Rodriguez Negro (Author)



Table of Contents

Handbook on Policy and Ethics in AI Education	3
1. Introduction	5
1.1 Purpose and scope of the handbook.....	6
1.2 The importance of ethical considerations in AI adoption	6
1.3 A handbook for educators, administrators and policymakers	8
2. Understanding AI in VET and Adult Education	9
2.1. Role of AI in Education	10
2.2. AI in Adult and Vocational Education and Training.....	11
3. Developing Ethical guidelines.....	12
3.1 Importance of ethical policies in AI education.....	12
3.2 Key requirements for trustworthy AI systems	14
3.3 Considering Legal Aspects of AI Implementation in VET Schools	16
3.4 Case Studies and Examples of Effective AI Policies in VET Institutions.....	23
4. Ethical Skills	30
4.1 Ethical Skills definition	30
4.2 Ethical Skills Importance	30
4.3 Ethical Skills areas and examples.....	31
5. Conclusion.....	38
5.1 Key takeaways.....	38
5.2 Final reflections.....	40
Acknowledgement.....	40
References.....	41



1. Introduction

The AI Pioneers project focuses on advancing the integration of Artificial Intelligence (AI) in education, particularly in Adult Education and Vocational Education and Training (VET). This ERASMUS+ Forward Looking Project began in January 2023 and is funded until December 2025.

With AI's growing influence in education and the increasing demand for professional support in AI use and skill development, the project aims to sustain a comprehensive, global network of AI Pioneers across partner countries and beyond. The establishment of the Reference Network of AI Pioneers is a key element, targeting educators, policymakers, educational planners, AI developers, and other relevant stakeholders. By engaging those most invested in AI's use in education, the project aims to build a central hub to assist AI integration in Adult Education and VET, fostering future initiatives and educational AI projects. More information about the network and how to participate can be found via the [project website](#).

Another major goal was to develop a Supplement to the DigCompEDU Framework (Bekiaridis, 2024). This document outlines the necessary skills and competencies educators need to understand, evaluate, and to use AI technologies in education, with a specific focus on Adult Education and VET. This supplement aligns AI competencies with the six core areas of the DigCompEdu Framework (Redecker, 2017): Professional Engagement, Digital Resources, Teaching and Learning, Assessment, Empowering Learners, and Facilitating Learner's Digital Competence. For each area, the supplement:

1. Analyses AI applications within the area
2. Proposes activities to develop relevant skills
3. Suggests progression levels for building competencies
4. Identifies challenges and provides strategies for overcoming them

The project continues to develop and share a range of resources via the AI Pioneers Toolkit, available on the project's website. This toolkit offers curriculum recommendations for AI integration in education, highlights the benefits and risks of AI use in educational settings, and provides best practice guidelines for AI in education (Tommasi & Perini, 2024). It also includes brief overviews of tools for different teaching purposes, such as learning assistants, tools for creating educational materials, and those supporting peer-to-peer learning.

The project is doing important work in addressing the ethical challenges AI poses in educational settings. This document accompanies an evaluation schema and



both are designed to guide institutions and educators in implementing AI responsibly while considering privacy and ethical implications.

The *Evaluation Schema for AI in Education* focuses on practical guidelines for adopting AI tools, ensuring they align with ethical principles and EU values. It is a great resource for educational institutions to make informed decisions about AI integration.

And this, the *Handbook on Policy and Ethics in AI Education* is a more comprehensive guide, offering policies and ethical frameworks to tackle current concerns.

1.1 Purpose and scope of the handbook

The *Handbook on Policy and Ethics in AI Education* serves as a foundational resource for educators, policymakers, and institutions aiming to navigate the complexities of AI in educational settings. It helps create a framework that can ensure ethical considerations are front and centre, protecting students' privacy, promoting fairness, and encouraging responsible usage.

A key benefit is the actionable guidance it provides for educators, empowering them to use AI responsibly, and to help students understand its capabilities, limitations, and ethical implications. This will be critical in fostering a generation of learners who are not only tech-savvy, but also ethically aware when engaging with technology. For policymakers, the consistent framework that the handbook offers guides users to more standardized approaches across institutions, which is essential in mitigating the risks of AI, such as biases in algorithms or violations of privacy. It can also help ensure that AI isn't just integrated for its own sake but rather in a way that enhances the educational experience without compromising ethical standards.

1.2 The importance of ethical considerations in AI adoption

Ethical considerations are crucial in the adoption of AI, especially in fields like adult education and vocational training (VET), where the impact of AI on learners, educators, and institutions can be significant. These considerations help ensure that AI is used responsibly, inclusively, and transparently, and that it supports the broader goals of education without compromising fundamental values.

Some key considerations, which are covered by the schema include;

Diversity and Inclusion: AI systems and tools must be designed to cater to a wide range of learners, including those with different backgrounds, abilities, and



learning needs. In adult education and VET, learners often come from diverse life experiences, such as varying levels of prior education, age groups, and work experience. AI should be flexible and inclusive, ensuring that its use supports rather than excludes any particular group. Ethical adopters of AI must be aware of the ways in which the technology could perpetuate biases or create digital divides in order to mitigate against them.

Transparency: For AI to be ethically adopted in educational settings, transparency is essential. Institutions must be clear about how AI systems work, the data they collect, and the decision-making processes behind them. Learners and educators should understand how AI tools are used in the classroom or training environment, and how the systems can affect learning outcomes. This transparency fosters trust in AI technologies and ensures that there is accountability in their use.

Privacy and Data Protection: Adult education and VET institutions often deal with sensitive personal data, such as learner profiles, assessment results, and career history. The ethical use of AI in this context must prioritize privacy and data protection. Institutions must ensure compliance with data protection laws (e.g., GDPR) and ensure that learners' data is collected, stored, and used responsibly. Consent for data collection should be informed and voluntary, and learners should have control over their personal information.

Sustainability: AI should be integrated in a way that aligns with the long-term sustainability of educational systems. This means considering both the environmental impact of AI technologies and their role in fostering sustainable practices within education and training programs. In the context of VET, AI can also support sustainability in industries by preparing learners with skills that promote green technologies and eco-friendly practices.

AI Literacy skills: As AI becomes more integrated into education, it is vital to build AI literacy among both educators and learners. In adult education and VET, this involves equipping learners with the knowledge and skills to understand AI technologies and use them effectively in their careers. This empowerment allows learners to navigate the evolving job market and ensure they are not left behind. For educators, AI literacy helps them critically assess AI tools and incorporate them in ways that enhance learning without compromising ethical standards.

Fairness and Equity: AI systems should be designed to treat all learners fairly, ensuring that no group is disadvantaged by automated decisions or algorithms. In adult education and VET, this is especially important, as these fields often serve learners who may already face social or economic challenges. It is critical



that AI does not reinforce existing inequalities or create new barriers to access, but rather supports equitable opportunities for all learners.

1.3 A handbook for educators, administrators and policymakers

This handbook is designed for educators, administrators, and policymakers in the context of AI adoption in adult education and vocational training (VET). It is known that AI adoption and its ethical considerations impact each of the aforementioned groups in different ways:

Educators:

For educators, the ethical integration of AI in adult education and VET is about enhancing teaching and learning without undermining the principles of fairness, privacy, and transparency. AI tools can offer personalised learning experiences, streamline administrative tasks, and provide insights into student progress. However, educators must ensure that AI supports diverse learning styles and needs, fostering an inclusive learning environment.

There are plenty of key considerations for educators regarding AI literacy, ethical use and transparency with learners.

AI Literacy: Educators need to be equipped with the skills to understand, evaluate, and effectively integrate AI into their teaching methods. This includes knowledge of how AI tools work and their ethical implications in the classroom.

Ethical Use: Educators must ensure that AI is used to enhance learning, not replace human interaction. They should be prepared to address any issues of bias or discrimination that may arise from AI systems.

Transparency with Learners: Educators should be transparent with students about how AI is being used in the learning process and the data it collects. This helps to build trust and ensures students feel their privacy is respected.

Administrators:

For administrators in adult education and VET institutions, AI adoption involves overseeing the integration of these technologies into the institution's systems and ensuring that the infrastructure is in place to support AI tools. They must also prioritize policies that ensure the ethical use of AI while maintaining institutional integrity.



There are plenty of key considerations for administrators, regarding data security and privacy, equity and access and ongoing training and support.

Data Security and Privacy: Administrators must ensure robust data protection measures are in place, as AI relies on collecting large amounts of learner data. Compliance with data protection regulations (such as GDPR) is essential.

Equity and Access: AI should be deployed in ways that ensure equitable access to all learners, regardless of their socio-economic background. Administrators should advocate for policies that prevent AI tools from exacerbating existing inequalities.

Ongoing Training and Support: Administrators need to allocate resources for continuous professional development so educators can stay updated on AI advancements and ethical best practices.

Policymakers:

Policymakers play a crucial role in creating the frameworks and regulations that govern AI use in education. They must ensure that AI adoption aligns with broader educational goals, respects human rights, and enhances opportunities for lifelong learning.

There is a variety of key considerations for policymakers regarding ethical guidelines and regulation, promoting AI literacy and ensuring inclusivity and accessibility.

Ethical Guidelines and Regulation: Policymakers need to establish ethical standards and regulations that guide the use of AI in education. This includes addressing issues like transparency, data privacy, and fairness in AI-driven educational systems.

Promoting AI Literacy: Policymakers should promote AI literacy across the education system, ensuring both educators and learners have the necessary tools and understanding to navigate AI technologies.

Ensuring Inclusivity and Accessibility: It is vital that AI policies prioritize inclusive education, making sure that AI technologies do not leave certain groups behind. This may involve incentivizing research and development of AI solutions tailored for diverse learners.

2. Understanding AI in VET and Adult Education



The integration of Artificial Intelligence (AI) into education presents transformative opportunities alongside significant challenges. Research underscores AI's potential to enhance teaching through automated tasks, personalised learning, and data-driven pedagogical innovations, while generative AI tools expand possibilities for interactive and reflective education. However, ethical concerns around data privacy, commercial influence, and the digital divide demand robust governance. In adult and vocational education, AI not only informs curriculum design and skill prediction for workforce readiness but also necessitates fostering digital literacy and critical engagement with AI's societal implications. Balancing AI's benefits with equitable, ethical policies remains crucial to shaping inclusive, future-oriented educational ecosystems.

2.1. Role of AI in Education

A recently published policy report from the Joint Research Centre (JRC) of the European Commission, titled [On the Futures of Technology in Education: Emerging Trends and Policy Implications](#), (Tuomi, Cachia & Villar-Onrubia, 2023), provides a comprehensive overview of how technological advancements, particularly in Artificial Intelligence (AI), are transforming educational practices, processes, and institutions. The document highlights the significant and diverse impacts of AI on education, affecting teaching and learning processes, assessment methods, and organizational structures.

AI is seen as a potential solution for reducing teachers' workloads, especially in tasks like marking student homework. By automating these repetitive and time-consuming tasks, AI can make teaching more appealing and allow teachers to concentrate more on instructional activities rather than administrative duties (Li, 2021; Tang & Hai, 2021; Villegas-Ch et al., 2021). This automation could revolutionize formative assessment (Kumar & Boulanger, 2020), enabling continuous feedback for students and facilitating real-time assessment (Porter & Grippa, 2020) and personalised or adaptive learning (Kochmar et al., 2020; Wongvorachan et al., 2022).

Moreover, generative AI and foundational models are expected to have a substantial impact on education. Technologies such as automatic video captioning, translation, and video summarizing offer innovative opportunities for enhancing pedagogy (Chew & Chua, 2020; Vazquez-Cano et al., 2021). These technologies can help develop 'smart companions' or 'learning partners,' shifting the focus from rote learning to metacognition and reflection. This necessitates a careful redistribution of responsibilities between humans and machines, determining which activities to delegate to AI.



We should also emphasize the role of AI in the datafication of education, where large-scale data on educational activities is used to reinforce and inform new pedagogical approaches, learning technologies, and education policies (Williamson et al., 2023). The goal is to record learning processes, not just outcomes, to enhance student engagement and achievement (Porter & Grippa, 2020).

However, ethical concerns about personal data processing, potential misuse, reliance on infrastructures largely controlled by commercial entities and some other challenges are critical issues to address (Seufert, 2024; Wongvorachan et al., 2022). The integration of AI into education prompts significant questions about governance models and policymaking, necessitating a re-evaluation of educational policies to incorporate AI's capabilities and limitations (Attard-Frost et al., 2024; Boninger et al., 2020; Williamson & Eynon, 2020). We must also address the importance of critically evaluating the risks of exacerbating the digital divide and the dominance of commercial interests in the sector (Bulathwela et al., 2024; Celik, 2023). Regulating technological innovation to ensure diverse educational visions and aligning industrial policy with learning theories are crucial aspects we need to consider.

AI's role in education is a transformative and multifaceted tool offering opportunities for enhanced efficiency, personalisation, and innovation in teaching and learning. Nevertheless, it also presents significant challenges and ethical considerations that must be addressed to ensure the equitable, effective, and responsible use of AI in educational settings.

2.2. AI in Adult and Vocational Education and Training

Integrating AI education into curricula involves not only teaching the fundamentals of AI and its applications, but also exploring its societal impacts. This includes fostering critical thinking about AI's capabilities, limitations, and ethical dimensions, such as concerns around bias and privacy (Akgun & Greenhow, 2022; Aldosari, 2020; Attwell et al., 2020; Delcker, et al., 2024; Seufert, 2024; Wongvorachan et al., 2022). Enhancing digital literacy is also essential, as understanding how to interact with and interpret AI systems is becoming a key skill in the modern workforce, including teachers (Bekiaridis & Attwell, 2024; Delcker, et al., 2024; Roppertz, 2020; Rott et al., 2022; Seufert, 2024). As AI continues to evolve, the focus is shifting towards not only using AI tools but also understanding and critically evaluating them, preparing learners for a future where AI is integral to daily life and work (Rott et al., 2022).

Consequently, preparing the workforce for new skills in the age of AI is another crucial aspect of adult and VET education (Attwell et al., 2020; Roppertz, 2020;



Rott et al., 2022). AI's role in identifying emerging trends and predicting future skill requirements is invaluable. This insight allows educational programs to adapt their curricula to meet the evolving needs of the job market (Seufert, 2024). Upskilling and reskilling become central themes, with AI-guided programs enabling adults and professionals to acquire new skills or update existing ones, thus staying in step with technological advancements (Attwell et al., 2020; Emeršič et al., 2024; Delcker, et al., 2024). Moreover, AI-driven training ensures that learners are not just technically adept, but are also prepared to work in AI-integrated environments. This preparation is key to ensuring that the workforce is ready for the challenges and opportunities presented by the rapidly advancing field of AI, making adult and VET education more relevant, effective, and aligned with future employment landscapes.

3. Developing Ethical guidelines

3.1 Importance of ethical policies in AI education

Ethical policies in AI education are essential for ensuring that AI technologies are used in ways that promote fairness, transparency, and responsibility while protecting the rights and well-being of learners and educators. As AI becomes increasingly integrated into educational settings, having clear ethical guidelines is crucial to manage its potential risks and harness its benefits. Below are key reasons why ethical policies are vital in AI education:

Protection of Learners' Rights and Privacy

AI systems in education often require the collection and processing of personal data, including academic records, learning behaviors, and other sensitive information. Ethical policies ensure that data is handled with respect and in compliance with privacy regulations (e.g., GDPR). These policies help protect learners' privacy by ensuring transparency in data collection, giving learners control over their data, and ensuring it is used responsibly and securely.

Ensuring Fairness and Equity

One of the most critical ethical concerns with AI in education is the potential for biased algorithms that may disadvantage certain groups of learners. AI systems can inadvertently perpetuate existing inequalities based on race, gender, socio-economic status, or other factors. Ethical policies are necessary to guide the development and deployment of AI tools that are inclusive and equitable,



ensuring that all learners have equal access to opportunities regardless of their background or personal characteristics.

Promoting Transparency and Accountability

AI technologies can be complex and opaque, making it difficult for educators, learners, and administrators to fully understand how decisions are made. Ethical policies help ensure that AI systems are transparent, meaning stakeholders know how AI tools work, what data is being used, and how decisions are being made. This transparency builds trust and accountability, allowing educators and learners to better understand and interact with AI technologies.

Supporting AI Literacy and Responsible Use

As AI becomes more integrated into the learning process, it is important that educators and learners alike understand how AI works and its potential implications. Ethical policies in AI education should encourage AI literacy, ensuring that both educators and learners are equipped with the knowledge and skills to use AI responsibly and critically. This understanding allows users to maximize the benefits of AI while being aware of its limitations and potential harms.

Fostering Ethical Development and Innovation

AI technologies in education should be developed with a strong ethical foundation to ensure that they are used for the benefit of all learners. Ethical policies provide a framework for developers, researchers, and educational institutions to design AI systems that align with educational values such as fairness, empowerment, and respect for diversity. This also ensures that AI technologies are designed and used in ways that support lifelong learning and the broader goals of education.

Addressing the Risk of Displacement

As AI systems automate certain aspects of education (e.g., grading, tutoring, administrative tasks), there is a risk that they may displace educators or limit human interaction. Ethical policies can help guide the responsible use of AI by ensuring that automation complements human educators, rather than replacing them. Policies can encourage the use of AI to enhance teaching effectiveness and administrative efficiency while maintaining the value of human presence in the educational experience.

Building Trust in AI Systems



For AI to be successfully adopted in educational settings, educators, students, and administrators must trust the technology. Ethical policies play a significant role in establishing and maintaining this trust by providing clear guidelines on how AI should be used and what ethical standards should be followed. When AI systems are developed and deployed in line with these policies, stakeholders are more likely to embrace and use them effectively.

Guiding the Future of AI in Education

As AI continues to evolve, it is crucial that ethical policies provide long-term guidance for its integration into educational systems. These policies help create a framework for the future development of AI technologies that aligns with societal values, ensuring that AI is used to enhance learning outcomes, foster innovation, and support lifelong learning in ways that are both beneficial and ethical.

In general, ethical policies in AI education are necessary to guide the responsible development and use of AI technologies. These policies ensure that AI is used in ways that respect learner privacy, promote fairness, encourage transparency, and build trust. Moreover, they help educators, learners, and institutions navigate the ethical challenges that arise as AI continues to transform educational environments. By establishing these ethical guidelines, educational systems can harness the power of AI while safeguarding the values that are fundamental to effective and inclusive education.

3.2 Key requirements for trustworthy AI systems

The integration of Artificial Intelligence (AI) into education offers immense opportunities for enhancing teaching, learning, and administration. However, to ensure that AI is used ethically and responsibly, educational institutions must develop clear guidelines for the ethical implementation of AI systems. Numerous published frameworks, guidelines, and audits (e.g., UNESCO's AI Audit 2023a and Guidance for Generative AI in Education and Research 2023b) provide valuable insights into what constitutes trustworthy AI. These frameworks emphasize key principles such as **transparency, fairness, accountability, and respect for human rights**. However, the task of developing and applying these principles can be daunting for institutions, especially in a fast-evolving technological landscape.

As part of AI Pioneers, we have developed an evaluation schema (Roman Etxebarrieta et al., 2024) designed to simplify and contextualize these principles for educational settings. Our schema provides a structured approach to help institutions navigate the complexities of AI implementation while ensuring adherence to ethical standards. As long as there are several guidelines available,



our schema aims to offer a practical, user-friendly tool that will guide educational institutions in developing their own ethical AI guidelines, considering their unique contexts and needs.

One of the central insights from the EU AI Act (European Parliament and Council of European Union, 2024) is that the primary responsibility for ensuring AI systems are developed ethically lies with the developers and companies that create them, and that the guidelines to be followed will depend on the level of risk associated with the AI systems. However, educational centers are responsible for ensuring that these systems are implemented and used ethically within their educational environments. Centers must be aware of key ethical issues to understand both the limitations and potential of AI systems they adopt. This involves not only complying with existing policies and legal requirements but also anticipating and addressing challenges that may arise as these technologies evolve.

Specifically, the focus of our evaluation schema is on identifying ethical concerns and developing a policy or strategy that can support and promote AI use in educational centers while adhering to key ethical guidelines. These guidelines include: **diversity and equity, transparency and accountability, privacy and data protection, safety and security, sustainability and social well-being, empowerment of teachers and learners, democratic participation in educational policy planning and AI practices, autonomy, ethical design (related to sustainability and accessibility), commercialization** (Şenocak et al., 2024), **teacher and student agency, and AI literacy.**

The evaluation schema consists of several sections, based on Chan's (2023) dimensions, adjusted with Şenocak et al. 's (2024) review and the European Commission's (2022) ethical guidelines, each focused on different aspects of AI implementation in educational settings. These sections guide institutions in developing ethical frameworks and ensuring AI is used responsibly. The first section focuses on **Assessing Current AI Use and Ethical Concerns**, helping institutions assess their current AI maturity thus allowing them to estimate how prepared the educational center is to implement AI or expand its use of AI (JISC, 2022).

The second section is focused on **AI Governance and Monitoring** and thus addresses issues related to compliance with regional and international policies, covering privacy, data protection, and fairness. It addresses complex ethical issues like non-discrimination and transparency that are also key considerations to ethical AI. The third section, **Operational Considerations**, provides guidance on training teachers, learners, and staff, promoting AI literacy, ensuring human oversight, and supporting democratic participation in AI-related decision-making.



Ensuring all users are adequately trained and aware of AIs capabilities and limitations is key to successful AI implementation and ensures equal access for all.

The fourth section focuses on **Pedagogical Considerations**, considering important issues such as ethical dilemmas in AI-driven assessments, student agency, and competency development, particularly in Adult and Vocational Education and Training (Attwell et al., 2021; UNESCO, 2019). The fifth and final section focuses on **Other Considerations: Sustainability and Ethical Design**, offering guidance for addressing long-term impacts and ensuring accessible AI use.

The evaluation schema provides a practical tool for educational institutions to navigate AI implementation, identify ethical challenges, and develop policies aligned with international standards. For more detailed information, please refer to the full evaluation schema [Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)] and Genially presentation that can be found on the [AI Pioneers website](#).

3.3 Considering Legal Aspects of AI Implementation in VET Schools

3.3.1 The EU AI Act Regulatory Framework

The European Union's AI Act establishes a comprehensive regulatory framework that classifies artificial intelligence systems according to their potential risks, providing essential guidance for vocational education and training (VET) institutions implementing these technologies. This risk-based approach ensures proportionate regulatory oversight while fostering innovation in educational contexts (European Parliament, 2024).

As Zawacki-Richter et al. (2023) articulate in their systematic review of ethical considerations in AI-enhanced education, regulatory frameworks must balance technological innovation with protection of fundamental rights. Their analysis identifies the EU AI Act as a pioneering attempt to establish "clear parameters for responsible AI deployment in learning environments while acknowledging the transformative potential of these technologies" (p. 8).

Risk Classification for Educational AI



AI systems deployed in VET settings predominantly fall within the **high-risk** category due to their significant impact on students' educational experiences and future career opportunities. The EU AI Act categorizes AI systems as follows:

Unacceptable Risk: Systems posing clear threats to fundamental rights or safety are prohibited outright

High-Risk: Applications in critical sectors including education that require rigorous regulatory compliance

Limited Risk: Systems with transparency obligations requiring disclosure of AI interaction

Minimal Risk: Applications with limited regulatory requirements beyond existing frameworks

High-risk classification implications for VET institutions:

- Mandatory risk assessment and mitigation protocols
- Rigorous documentation requirements
- Human oversight mechanisms
- Transparency obligations to affected stakeholders
- Regular conformity assessments.

Research by Ramirez-Montoya et al. (2023) emphasizes that high-risk classification for educational AI systems requires "rigorous governance structures that transcend technical compliance to address broader societal concerns" (p. 142). Their analysis of implementation challenges across European vocational institutions highlights the need for contextualized interpretation of regulatory requirements.

3.3.2 Key Legal Considerations for VET Implementations

Transparency and Accountability Requirements

The EU AI Act mandates comprehensive transparency in AI deployment within educational settings:

- VET institutions must inform students, staff, and relevant stakeholders when AI systems are utilized in assessment, learning recommendations, or administrative decision-making
- Documentation must articulate the purpose, functionality, and limitations of AI applications



- AI-driven decisions affecting student progression or certification must be explainable and justifiable
- Algorithmic processes must be auditable to enable validation of outcomes.

Prinsloo and Knox (2023) identify transparency as a cornerstone ethical principle in educational AI implementation, noting that "algorithmic opacity presents unique challenges in learning contexts where trust between educators and students forms the foundation of effective pedagogy" (p. 76). Their framework for educational AI transparency extends beyond technical disclosure to encompass contextual understanding accessible to diverse stakeholders.

Data Protection and Privacy Compliance

AI implementation in VET contexts requires meticulous adherence to data protection frameworks:

- All AI systems must comply with General Data Protection Regulation (GDPR) provisions
- Clear consent mechanisms must be established for personal data processing
- Data minimization principles must govern collection and retention practices
- Purpose limitation ensures data utilization solely for specified educational objectives
- Technical and organizational safeguards must protect student data integrity and confidentiality.

Holmes et al. (2023) observe that data protection considerations in educational contexts extend beyond legal compliance to encompass "broader ethical questions regarding student agency, power dynamics, and potential educational surveillance" (p. 12). Their work emphasizes the importance of contextualizing data protection frameworks within educational-specific ethical considerations.

Bias Prevention and Fairness Assurance

The EU AI Act establishes explicit requirements to mitigate algorithmic bias in educational contexts:

- VET institutions must conduct regular bias assessments of AI systems to prevent discrimination
- Assessment methodologies must evaluate potential disparate impacts across demographic groups



- Algorithmic training data must be representative of student diversity
- Continuous monitoring protocols must identify and remediate emergent bias in operational systems
- Regular auditing of outcomes across student populations is required to verify equitable impact.

Research by Hodgkin and Kumar (2023) examining algorithmic bias in vocational guidance systems identifies "systematic methodologies for evaluating disparate impact across diverse student populations" (p. 217) as essential for equitable AI implementation. Their framework for comprehensive bias assessment provides practical guidance for VET institutions navigating fairness requirements.

AI in Classroom Tools (Limited Risk)

Numerous AI applications such as: AI-powered tutoring assistants, Plagiarism detection systems, Recommendation engines for learning resources, administrative automation tools, and Virtual laboratory simulations within VET settings may fall under limited risk classification:

These applications must provide clear disclosure to users regarding:

- The AI nature of the system
- Limitations of algorithmic capabilities
- Human oversight provisions
- Mechanisms for contesting outputs.

3.3.3 Implementation Timeline and Compliance Planning

VET institutions must develop strategic implementation plans aligned with the EU AI Act's phased enforcement timeline:

2nd February, 2025: Ban on AI systems posing unacceptable risks

2nd August, 2025: Application of codes of good practice and designation of national regulatory authorities

2nd August, 2026: General enforcement of the AI Act, including rules for high-risk AI systems

2nd August, 2027: Full enforcement of specific obligations for high-risk AI systems.



Prudent compliance planning requires:

- Comprehensive inventory of existing and planned AI applications
- Risk classification assessment of educational technology portfolio
- Gap analysis against regulatory requirements
- Structured implementation roadmap with milestone objectives
- Resource allocation for compliance documentation and monitoring.

3.3.4 Legal Responsibilities and Liability Considerations

VET institutions implementing AI systems assume significant legal responsibilities:

Accountability structures must designate clear responsibility for AI system governance

Liability frameworks must address potential harms arising from algorithmic errors or biases

Contractual agreements with technology providers must explicitly delineate responsibility allocation

Indemnification protocols should address potential claims arising from AI-influenced decisions

Insurance considerations should reflect emerging liability exposures.

Educational institutions must recognize that deploying AI systems, even those procured from third parties, does not absolve them of legal responsibility for outcomes affecting students.

Rodriguez-Abitia and Bribiesca-Correa (2021) identify liability considerations as an often overlooked dimension of educational AI governance, noting that "educational institutions face unique legal exposures when delegating traditionally human functions to algorithmic systems" (p. 94). Their analysis of emerging case law provides valuable precedent for VET institutions developing comprehensive liability frameworks.

3.3.5 Assessment List for Trustworthy AI (ALTAI) in VET Implementation

The Assessment List for Trustworthy Artificial Intelligence (ALTAI) provides VET institutions with a structured framework for ethical implementation and regulatory



compliance. This practical tool operationalizes the Ethics Guidelines for Trustworthy AI developed by the EU High-Level Expert Group on Artificial Intelligence.

In their comparative analysis of ethical assessment frameworks, Zawacki-Richter et al. (2023) identify ALTAI as particularly valuable in educational contexts due to its "comprehensive coverage of both technical and socio-ethical dimensions relevant to learning environments" (p. 14). Their research demonstrates ALTAI's adaptability to diverse educational contexts including vocational training.

Key Components of ALTAI Relevant to VET Implementation

ALTAI structures assessment across seven fundamental dimensions particularly relevant in vocational education contexts:

Human Agency and Oversight

- VET applications must support instructor autonomy rather than supplanting professional judgment
- Student autonomy must be preserved through informed consent and decision transparency
- Meaningful human oversight must be maintained in consequential decisions affecting student progression.

Technical Robustness and Safety

- AI systems must function reliably in diverse vocational training environments
- Security protocols must protect against potential manipulation or unauthorized access
- Fallback procedures must ensure educational continuity during system failures.

Privacy and Data Governance

- Student data utilized in AI training and operation must be protected through comprehensive governance
- Data minimization principles must guide collection and retention practices
- Robust security measures must safeguard sensitive information from unauthorized access.

Transparency



- AI capabilities and limitations must be clearly communicated to educational stakeholders
- Decision processes must be explainable in non-technical language
- Documentation must enable verification of system performance and compliance.

Diversity, Non-discrimination and Fairness

- VET-specific implementations must account for diverse student backgrounds and learning needs
- Assessment methodologies must evaluate equity impact across demographic dimensions
- Bias mitigation strategies must address vocational education's unique diversity considerations.

Societal and Environmental Well-being

- Broad social impacts of AI in vocational training must be systematically evaluated
- Environmental consequences of computational infrastructure must be assessed
- Potential labour market implications of AI skill certification must be considered.

Accountability

- Clear governance structures must establish responsibility for AI systems
- Audit mechanisms must enable verification of regulatory compliance
- Remediation procedures must address harmful outcomes or unintended consequences.

Application of ALTAI in VET Contexts

VET institutions can operationalize ALTAI through systematic assessment protocols:

Initial implementation assessment evaluates proposed AI applications against trustworthiness criteria

Pre-deployment verification ensures compliance with technical and ethical requirements

Post-implementation auditing validates performance across diverse student populations



Periodic reassessment identifies emerging ethical concerns as systems evolve.

By integrating ALTAI into governance frameworks, VET institutions can demonstrate due diligence in regulatory compliance while enhancing stakeholder trust in AI implementation.

3.4 Case Studies and Examples of Effective AI Policies in VET Institutions

3.4.1 Foundational European Policy Frameworks

Effective institutional policies build upon foundational frameworks established at the European level:

High-Level Expert Group on Artificial Intelligence (AI HLEG)

The European Commission's AI HLEG has developed essential guidance through:

Ethics Guidelines for Trustworthy AI: Establishing seven requirements for ethical AI implementation

Policy and Investment Recommendations: Providing strategic direction for sustainable AI development

Assessment List for Trustworthy AI: Offering practical implementation tools for organizations.

These foundational frameworks have informed concrete institutional policies across European VET providers.

As Holmes et al. (2023) observe in their analysis of AI governance frameworks, the AI HLEG guidelines represent "a critical bridge between abstract ethical principles and concrete implementation practices" (p. 18). Their research traces the influence of these guidelines across diverse educational contexts, demonstrating their adaptability to vocational training environments.

3.4.2 Copenhagen Technical College: Comprehensive AI Governance Framework



Copenhagen Technical College has established a multilayered governance structure demonstrating exemplary practice in VET AI implementation.

Policy Architecture and Governance Structure

The institution's approach features:

- Multi-stakeholder AI Ethics Committee including administrators, educators, students, industry representatives, and technical experts
- Tiered approval protocols based on algorithmic risk classification
- Mandatory AI literacy training for all faculty and administrative staff
- Formalized student rights regarding algorithmic transparency and decision contestability
- Structured assessment protocols measuring both educational and ethical dimensions.

This approach aligns with Prinsloo and Knox's (2023) stakeholder engagement framework, which emphasizes the importance of "structured representation across institutional hierarchies and affected communities" (p. 82). Their research demonstrates how multi-stakeholder governance enhances both ethical rigor and implementation effectiveness.

Implementation Methodology

Key implementation elements include:

- Phased deployment beginning with low-risk administrative applications
- Controlled pedagogical pilots with rigorous evaluation protocols
- Structured documentation of both technical performance and ethical impact
- Regular stakeholder feedback mechanisms to inform iterative improvement
- Independent auditing of algorithmic outcomes across demographic dimensions.

Documented Outcomes

The College has reported significant benefits through structured implementation:

- 27% reduction in administrative workload through ethical AI automation
- Improved educational personalisation without compromising instructional quality
- Enhanced transparency in assessment practices



- Increased student satisfaction with technological integration
- Successful compliance with emerging regulatory requirements

3.4.3 Bordeaux Institute of Technology: Transparent AI Assessment Policies

The Bordeaux Institute has developed model policies addressing one of the most ethically complex aspects of educational AI: student assessment and evaluation.

Assessment-Specific Policy Framework

Key policy elements include:

- Prohibition of fully automated high-stakes evaluation
- Mandatory human review of all AI-influenced assessment decisions
- Transparent documentation of algorithmic contribution to evaluations
- Structured appeals mechanism for contesting AI-influenced outcomes
- Regular bias audits comparing assessment results across student populations
- Student representation in assessment governance structures

This approach exemplifies what Ramirez-Montoya et al. (2023) identify as "human-centred assessment governance" characterized by "algorithmic transparency, stakeholder participation, and formalized oversight mechanisms" (p. 149). Their research demonstrates how such frameworks enhance both assessment validity and stakeholder trust.

Implementation Approach

The implementation strategy emphasizes:

- Clear delineation between appropriate and inappropriate assessment automation
- Graduated implementation based on assessment stakes and consequences
- Faculty development in ethical AI assessment methodologies
- Technical infrastructure ensuring explainability of algorithmic recommendations
- Documentation protocols enabling verification of assessment equity.

Measurable Impact

The Institute has documented significant benefits:



- Enhanced assessment consistency while preserving instructor autonomy
- Improved student trust in evaluation methodologies
- Reduced assessment workload without compromising quality
- Better identification of students requiring additional support
- Successful regulatory compliance with AI transparency requirements.

3.4.4 Munich Vocational School Consortium: Collaborative Data Governance

Eight vocational institutions in Munich have established a collaborative approach to AI data governance, demonstrating how resource pooling can enhance ethical implementation.

Shared Governance Architecture

Key structural elements include:

- Common ethical framework for data collection and utilization
- Shared technical infrastructure for data management and security
- Collaborative vendor assessment and contracting protocols
- Joint implementation of federated learning methodologies
- Pooled expertise for regulatory compliance verification.

This collaborative approach aligns with what Rodriguez-Abitia and Bribiesca-Correa (2021) term "resource-optimized governance" characterized by "shared expertise, unified policies, and collective implementation capacity" (p. 98). Their research demonstrates how such approaches can enable smaller institutions to implement sophisticated governance frameworks.

Implementation Methodology

The Consortium's approach features:

- Standardized data protection protocols across participating institutions
- Centralized expertise in data minimization and anonymization techniques
- Shared resources for technical vulnerability assessment
- Collaborative development of privacy-preserving analytical methodologies
- Joint stakeholder engagement in governance development

Demonstrated Benefits

This collaborative approach has yielded:



- More sophisticated data governance than individual institutions could achieve independently
- Enhanced negotiating position with technology vendors
- Improved data security protocols through shared expertise
- Cost-effective regulatory compliance across multiple institutions
- More comprehensive bias mitigation through larger and more diverse datasets.

3.4.5 Dublin Technical Institute: Equity-Focused AI Policy Framework

The Dublin Technical Institute has developed policies specifically prioritizing algorithmic equity in vocational education contexts.

Equity-Centred Policy Framework

Key policy elements include:

- Mandatory equity impact assessments before AI deployment
- Regular algorithmic auditing across demographic dimensions
- Prioritization of explainable AI methodologies
- Requirement for non-technological alternatives to AI systems
- Formalized inclusion of underrepresented populations in governance structures.

This framework exemplifies what Hodgkin and Kumar (2023) identify as "proactive equity governance" characterized by "systematic bias assessment, inclusive stakeholder representation, and continuous monitoring protocols" (p. 219). Their research demonstrates how equity-centred governance enhances both ethical implementation and educational outcomes.

Implementation Strategy

The implementation approach emphasizes:

- Systematic bias testing in vocational guidance algorithms
- Demographic monitoring of AI-influenced educational outcomes
- Development of debiasing methodologies for training data
- Faculty development in equity-conscious AI integration
- Regular stakeholder feedback from diverse student populations.

Documented Impact



The Institute has evidenced:

- Identification and remediation of potential biases in career guidance systems
- Improved educational outcomes for historically underrepresented groups
- Enhanced stakeholder trust in technological integration
- Successful compliance with non-discrimination requirements
- Development of transferable equity-focused methodologies

3.4.6 Application of ALTAI in VET-Specific Contexts

The Assessment List for Trustworthy Artificial Intelligence (ALTAI) has been successfully applied in numerous vocational education contexts, demonstrating its practical utility for ethical implementation.

Research by Zawacki-Richter et al. (2023) identifies ALTAI implementation in vocational education as particularly valuable due to the "concrete professional contexts that characterize vocational training and the explicit career implications of AI-influenced decisions" (p. 15). Their case study analysis demonstrates ALTAI's adaptability to diverse vocational training environments.

AI-Driven Skill Assessment Applications

VET institutions have applied ALTAI to ensure ethical implementation of automated assessment systems:

Human Oversight Integration: Ensuring instructor validation of algorithmic evaluations

Fairness Verification: Testing assessment equity across diverse student populations

Transparency Requirements: Making assessment criteria explicit and understandable

Technical Robustness: Ensuring reliable operation across varying student performance patterns.

AI in Career Guidance and Vocational Counseling

ALTAI application in career recommendation systems has addressed:



- Potential reinforcement of occupational stereotypes
- Transparency in recommendation methodologies
- Data protection in predictive career modeling
- Accessibility of guidance systems across diverse populations.

AI-Powered Adaptive Learning in Vocational Training

Implementation of adaptive learning platforms has benefited from ALTAI assessment through:

- Evaluation of algorithmic personalisation equity
- Verification of learning path diversification
- Assessment of potential filter bubble effects
- Validation of instructor oversight capabilities.

3.4.7 Common Elements of Effective Policies

Analysis of successful implementation across European VET institutions reveals consistent policy patterns:

Multi-stakeholder governance structures involving diverse perspectives in development and oversight

Tiered risk assessment protocols applying proportionate scrutiny based on potential impact

Continuous monitoring mechanisms rather than one-time approval processes

Robust transparency requirements ensuring stakeholder understanding of AI capabilities and limitations

Meaningful human oversight especially in consequential educational decisions

Regular policy review cycles adapting to evolving technologies and regulations

Comprehensive educational components building AI literacy across all stakeholders

Clear accountability structures establishing responsibility for AI-influenced outcomes



Holmes et al. (2023) identify these consistent elements as representing an "emerging consensus regarding foundational governance principles for educational AI implementation" (p. 21). Their comparative analysis across diverse institutional contexts demonstrates the adaptability of these core principles to varied educational environments.

Institutions demonstrating exemplary practice have balanced innovation potential with appropriate safeguards, recognizing that overly restrictive approaches may limit educational benefits while insufficient governance creates significant risks. The most successful frameworks have integrated ethical considerations throughout the implementation lifecycle rather than treating them as post-hoc considerations.

4. Ethical Skills

4.1 Ethical Skills definition

Ethical skills refer to the ability to evaluate whether an AI is used ethically in a specific context. In the educational context, for example, ethical skills is whether the policies of an educational center are ethical, if the data security and privacy is maintained etc.

Ethical skills refer to the critical abilities required to evaluate and ensure the ethical use of AI and data within educational settings, based on principles such as human agency, equity, humanity, and justified choice. Human agency involves enabling individuals to make informed choices and take responsibility for their actions, fostering autonomy and accountability. Equity ensures fair treatment and equal access to opportunities for all, promoting inclusion and non-discrimination. Humanity emphasizes respecting human dignity, identity, and well-being, prioritizing meaningful human connections over instrumental use of data. Justified choice requires transparent decision-making processes grounded in ethical principles and stakeholder participation, ensuring AI systems in education are deployed responsibly and benefit all involved parties.

4.2 Ethical skills Importance

These ethical considerations are crucial for guiding educators and educational institutions in navigating the complexities of AI implementation. By adhering to ethical requirements such as those proposed by the AI Ethics Guidelines, educational stakeholders can effectively mitigate risks, promote fairness, and uphold the well-being of students and staff alike.



4.3 Ethical Skills areas and examples

Area 1: Professional Engagement. Using digital technologies for communication, collaboration, and professional development.	
Competence element	Potential indicators
Is able to critically describe positive and negative impacts of AI and data use in education	Takes an active part in continuous professional learning on AI and learning analytics and their ethical use.
	Able to give examples of AI systems and describe their relevance.
	Knows how the ethical impact of AI systems is assessed in the school.
	Knows how to initiate and promote strategies across the school and its wider community that promote ethical and responsible use of AI and data.
Understand the basics of AI and learning analytics	Aware that AI algorithms work in ways that are usually not visible or easily understood by users.
	Able to interact and give feedback to the AI system to influence what it recommends next.
	Aware that sensors used in many digital technologies and applications generate large amounts of data, including personal data, that can be used to train an AI system.
	Aware of EU AI ethics guidelines and self-assessment instruments.



Example 1: Ana is a teacher that wants to teach a lesson of fluid mechanics, and she wants to use the AI to generate some pictures and include them in the subject powerpoint. Before including the AI in the lesson, Ana has read the EU AI ethics guidelines and evaluated if it is ethical to include the use of it or not.

Example 2: Lucas teaches in the Environmental Education and Control VET. Several teachers of his educational centre use AI, so he has proposed a meeting with the entire educational centre to analyse the use of AI by the teachers. Lucas' goal is to ensure that all teachers are doing an ethical and responsible use of AI and the data.

Area 2: Digital Resources: Sourcing, creating and sharing digital resources

Competence element	Potential indicators
Data governance	Aware of the various forms of personal data used in education and training.
	Aware of responsibilities in maintaining data security and privacy.
	Knows that the processing of personal data is subject to national and EU regulation including General Data Protection regulation (GDPR).
	Knows who has access to student data, how access is monitored, and how long data is retained.
	Knows that all EU citizens have the right to not be subject to fully automated decision making.



	Able to give examples of sensitive data, including biometric data.
	Able to weigh the benefits and risks before allowing third parties to process personal data especially when using AI systems.
AI governance	Knows that AI systems are subject to national and EU regulation (notably AI Act to be adopted).
	Able to explain the risk-based approach of the AI Act (to be adopted).
	Knows the high-risk AI use cases in education and the associated requirements under the AI Act (when adopted).
	Knows how to incorporate AI edited/manipulated digital content in one's own work and how that work should be credited.
	Able to explain key principles of data quality in AI systems.
Example 1: Maria is a teacher at the Energy Efficiency and Solar Thermal Energy VET school. She is creating notes for her course to give to the students, and she has incorporated some AI edited digital content into the notes using the Copilot tool. Maria knows that this content needs to be credited, and she makes sure to credit it correctly.	
Example 2: Jaime teaches at the Mechanical Manufacturing VET school. He uses ChatGPT sometimes to develop some contents or activities of his subject. At some point he assesses the possibility to include his students personal data (name, surname, date of birth and address) to create a document that includes all this information. He is aware of his responsibility of maintaining data security and privacy, so he decided not to include the personal data.	

Area 3: Using digital technologies and strategies to enhance assessment



Competence element	Potential indicators
Models of learning	Knows that AI systems implement designer's understanding of what learning is and how learning can be measured; can explain key pedagogical assumptions that underpin a given digital learning system.
Objectives of education	Knows how a given digital system addresses the different social objectives of education (qualification, socialisation, subjectification).
Human agency	Able to consider the AI system impact on teacher autonomy, professional development, and educational innovation.
	Consider the sources of unacceptable bias in data-driven AI.
Fairness	Considers risks related to emotional dependency and student self-image when using interactive AI systems and learning analytics.
Humanity	Able to consider the impact of AI and data use on the student community.
	Confident in discussing the ethical aspects of AI, and how they influence the way technology is used.
Participates in the development of learning practices that use AI and data	Can explain how ethical principles and values are considered and negotiated in co-design and co-creation of learning practices that use AI and data (linked to learning design).
Example 1: Luisa, a teacher of Electromechanical Maintenance at a training company, implements AI software that utilises adaptive learning models to personalise teaching. The software assesses each apprentice's initial level and adjusts the content and tasks based on their progress. This allows every	



apprentice to advance at their own pace, receiving immediate feedback and specific resources to overcome difficulties. Additionally, Luisa can access detailed reports to identify areas where her students need additional support, ensuring inclusive and effective teaching.

Example 2: Javi, a teacher of Graphic Arts in the medium-level Graphic Printing program of a VET school, integrates AI software that enhances human capacity by assisting with complex design tasks. The software helps students by automating repetitive tasks, such as colour correction and layout adjustments, allowing them to focus on creative and strategic aspects of their projects. This empowers students to develop their artistic skills and critical thinking while the AI handles technical details.

Area 4: Assessment: Managing and orchestrating the use of digital technologies in teaching and learning.

Competence element	Potential indicators
Personal differences	Aware that students react in different ways to automated feedback.
Algorithmic bias	Consider the sources of unacceptable bias in AI systems and how it can be mitigated.
Cognitive focus	Aware that AI systems assess student progress based on predefined domain-specific models of knowledge.
	Aware that most AI systems do not assess collaboration, social competences, or creativity.
	Aware of common ways to manipulate AI-based assessment.



New ways to misuse technology	Aware that teachers may rely too much on AI-generated feedback, which sometimes can feel impersonal and/or fail to address specific student needs or misunderstandings."
Example 1: Jorge teaches a subject in the Installations and Maintenance VET school. After finishing the 1st topic of the subject, he uses an AI tool to assess the acquisition of the contents by his students. Furthermore, this AI tool will automatically provide feedback and adapt the level of the questions based on previous answers. Jorge is aware of how students react to automated feedback, so he makes sure to provide them personalised feedback and to check the students' progress himself. Therefore, he uses the AI only as a tool to give an initial assessment, which is checked and revised by him subsequently.	
Example 2: Natalia is a trainer at a VET service provider for renewable energies. She encourages students to use Windows Copilot in her training lessons for certain activities. Even if this tool works as an AI assistant that can be helpful with productivity and creativity, she is aware that other kinds of activities and assessment are needed to improve collaboration, social competences and creativity, as AI is not able to assess those aspects. She continues to take on the relevant activities and assessment as a trainer herself and uses AI as a tool during her lessons and for training preparation.	

Area 5: Empowering Learners: Using digital technologies to enhance inclusion, personalisation and learners' active engagement	
Competence element	Potential indicators
AI addressing learners' diverse learning needs	Knows the different ways personalised learning systems can adapt their behaviour (content, learning path, pedagogical approach).



	Able to explain how a given system can benefit all students, independent of their cognitive, cultural, economic, or physical differences
	Aware that digital learning systems treat different student groups differently
	Able to consider impact on the development of student self-efficiency, self-image, mindset, and cognitive and affective self-regulation skills.
Justified choice	Knows that AI and data use may benefit some learners more than others.
	Able to explain what evidence has been used to justify the deployment of a given AI system in the classroom.
	Recognises the need for constant monitoring of the outcomes of AI use.
Example 1: Teresa, teaching Advanced Vocational Training in Administration and Finance at an adult education provider uses AI to tailor learning experiences. The AI system assesses each student's needs and adjusts resources and activities accordingly. For instance, visual learners might receive interactive tools, while analytical students get detailed case studies. This approach ensures personalised learning that meets diverse learning styles and maximizes student engagement.	
Example 2: Óscar, teaching Advanced Vocational Training in Mechanical Manufacturing Design, employs AI software for designing mechanical parts. The software offers multiple design solutions and simulates performance across different conditions. Students analyse results to select the most efficient design, guided by Óscar to develop critical and analytical skills crucial for their careers.	



Area 6: Facilitating learners' digital competence enabling learners to creatively and responsibly use digital technologies for information, communication, content creation, wellbeing and problem-solving

Competence element	Potential indicators
AI and Learning Analytics ethics	Able to use AI projects and deployments to help students learn about ethics of AI and data use in education and training.

Example 1: Professor Sara teaches advanced marketing analytics to her vocational teacher students at a university. She discusses a case study where AI analyses social media data for targeted advertising. Sara prompts a debate on ethical concerns like privacy, consent, and algorithmic bias. She guides students to develop ethical guidelines for AI use in marketing, fostering their understanding of responsible AI deployment in the industry.

Example 2: Professor Jesús teaches culinary management, focusing on AI's role in personalising restaurant menus. He discusses a case study where AI analyzes dietary preferences to offer personalised menu recommendations. Jesús engages trainees in a debate on ethical issues such as data privacy, ensuring menu diversity, and AI's impact on customer satisfaction. He guides the trainees in developing ethical frameworks for using AI in menu customisation, preparing them to balance innovation with ethical considerations in restaurant operations.

5. Conclusion

As AI continues to shape educational landscapes, it is essential that the adoption of AI technologies is done in a way that upholds ethical values, promotes fairness, and safeguards students' privacy. This handbook has outlined the purpose and scope of AI implementation, underscoring the importance of ethical considerations to ensure that AI technologies are used responsibly across educational contexts.

5.1 Key takeaways



The integration of AI in education brings immense potential, but it must be guided by ethical principles that prioritize diversity, inclusion, and transparency. AI tools should be designed to accommodate learners from diverse backgrounds, taking into account cultural, linguistic, and socio-economic differences to promote equity in educational outcomes. Transparency is equally vital—educators and students must understand how AI systems make decisions, especially in areas like grading, feedback, or personalized learning paths. Ensuring privacy and data protection is essential, requiring full compliance with regulations such as the GDPR and adherence to ethical data management practices. AI systems must also be sustainable, minimizing their environmental impact and ensuring their long-term viability in educational ecosystems.

Equally important is the promotion of AI literacy among all stakeholders. Educators need targeted training to ethically and effectively integrate AI tools into their pedagogy, while students should be equipped with a critical understanding of AI's capabilities and limitations. Institutional administrators play a key role in aligning school or university policies with ethical AI practices, fostering an environment of accountability and fairness. Policymakers must step in to create robust regulatory frameworks that guide responsible AI adoption, mitigate algorithmic bias, and prevent discriminatory outcomes. Ultimately, a multi-stakeholder approach is crucial to ensure that AI enhances education in a manner that is fair, inclusive, and sustainable.

It is also very important to consider some key legal and regulatory frameworks, such as the EU AI Act (2024), which introduces a comprehensive legal framework that categorizes AI systems based on their risk to fundamental rights, with significant implications for educational institutions.

Based on that there are the High-risk AI applications, such as those used in student assessments or admissions, that are subject to stringent compliance requirements, including rigorous documentation, human oversight, and robust risk management protocols.

Moreover, transparency and accountability are not an option. Institutions must clearly disclose how AI systems function and ensure that users can understand and challenge automated decisions. Furthermore, the Act mandates proactive measures to audit and mitigate algorithmic bias, ensuring fairness and preventing discriminatory outcomes. To support these efforts, the Assessment List for Trustworthy AI (ALTAI) offers a practical tool that helps institutions evaluate the ethical impact of their AI systems, fostering a culture of responsible and trustworthy AI deployment in education.



5.2 Final reflections

We should all be aware of the fact that as AI continues to reshape university and vocational education, this handbook emphasizes the importance of maintaining a careful equilibrium between innovative practices and ethical responsibility. The benefits of AI starting from personalized learning pathways to improved administrative efficiency are significant, yet they come with risks that must not be overlooked or ignored. Issues such as algorithmic bias, data privacy breaches, and the diminishing role of human judgement demand constant attention. To navigate these complexities, there must be constant collaboration among educators, administrators, and policymakers. Of course, ethical principles must be deeply embedded in the development, deployment, and evaluation of AI tools to ensure they support inclusion, respect autonomy, and reinforce trust in the educational process.

Looking ahead in the future, the successful integration of AI in education will rely not only on technological readiness but also on a shared vision grounded in transparency, fairness, and lifelong learning. The sector must foster AI literacy among all stakeholders and create robust systems for accountability and oversight. This is not a one-time adjustment but an evolving journey that calls for continuous reflection and dialogue. By championing human-centered approaches and ethical innovation, the education community can transform AI into a powerful ally that enhances learning opportunities while preserving the values for quality education. For comprehensive guidance, stakeholders are encouraged to consult the AI Pioneers project resources and the Evaluation Schema for AI in Education, which are accessible in multiple languages.

Acknowledgements

We would like to thank all those who supported the development of this document through their constructive criticism and invaluable contributions. We are deeply thankful to all our project partners, who contributed to the content as well as to the AI Pioneers Reference Network and participants in our consultation seminars. Their shared concerns and opinions offered valuable guidance throughout this process.

Lastly, we acknowledge the significance of the referenced literature, along with various other resources such as scientific papers, reports, seminars, guidelines from educational institutions, and blog posts by researchers and educators. Collectively, these contributions have greatly informed our approach to addressing ethical considerations in the use of AI in education at all levels.



References

- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.
- Aldosari, S. A. M. (2020). The future of higher education in the light of artificial intelligence transformations. *International Journal of Higher Education*, 9(3), 145-151.
- Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, 41(2), 101929.
- Attwell, G., Bekiaridis, G., Deitmer, L., Perini, M., Roppertz, S., & Tütlys, V. (2020). Artificial intelligence in policies, processes and practices of vocational education and training.
- Attwell, G., Deitmer, L., Tütlys, V., Roppertz, S., & Perini, M. (2020). Digitalisation, artificial intelligence and vocational occupations and skills: what are the needs for training teachers and trainers?. In *Trends in vocational education and training research, Vol. III. Proceedings of the European Conference on Educational Research (ECER), Vocational Education and Training Network (VETNET)* (pp. 30-42).
- Bekiaridis, G., & Attwell, G. (2024). Integrating Artificial Intelligence in Vocational and Adult Education: A Supplement to the DigCompEdu Framework. *Ubiquity Proceedings*, 4(1).
- Boninger, F., Molnar, A., & Saldaña, C. (2020). Big Claims, Little Evidence, Lots of Money: The Reality behind the Summit Learning Program and the Push to Adopt Digital Personalized Learning Programs. *Commercialism in Education Research Unit*.
- Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational inequality, techno-solutionism and inclusive tools. *Sustainability*, 16(781). <https://doi.org/10.3390/su16020781>
- Celik, I. (2023). Exploring the determinants of artificial intelligence (Ai) literacy: Digital divide, computational thinking, cognitive absorption. *Telematics and Informatics*, 83, 102026.
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology*



in *Higher Education*, 20(1), 38. <https://link.springer.com/article/10.1186/s41239-023-00408-3>

Chew, E., & Chua, X. N. (2020). Robotic Chinese language tutor: Personalising progress assessment and feedback or taking over your job? *On the Horizon*, 28(3), 113–124. <https://doi.org/10.1108/OTH-04-2020-0015>

Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). First-year students AI-competence as a predictor for intended and de facto use of AI-tools for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 18.

European Commission, Directorate-General for Education, Youth, Sport, and Culture. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/153756>

European Parliament. (2024). EU AI Act: First regulation on artificial intelligence. Retrieved from <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. Official Journal of the European Union, L series, 1–144.

Emeršič, Ž., Hrastnik, G., Meh Peer, N., & Peer, P. (2025). AIM@VET-inspired university level education strategies for teaching computer vision and biometrics. *ROSUS 2025 – Računalniška obdelava slik in njena uporaba v Sloveniji 2025: Zbornik 19. strokovne konference*. <https://doi.org/10.18690/um.ferj.2.2025.4>

Hodgkin, E., & Kumar, A. (2023). Algorithmic fairness in vocational guidance systems: A comparative analysis of implementation frameworks. *Educational Technology Research and Development*, 71(3), 215-231.

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign



JISC (2022). *AI in tertiary Education. A summary of the current state of play*. JISC Repository. Retrieved from <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020). Automated personalized feedback improves learning gains in an intelligent tutoring system. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II* 21 (pp. 140-146). Springer International Publishing.

Kumar, V., & Boulanger, D. (2020, October). Explainable automated essay scoring: Deep learning really has pedagogical value. In *Frontiers in education* (Vol. 5, p. 572367). Frontiers Media SA.

Li, Q. (2021). The use of artificial intelligence combined with cloud computing in the design of education information management platform. *International Journal of Emerging Technologies in Learning*, 16(5), 32–44. <https://doi.org/10.3991/ijet.v16i05.20309>

Porter, B., & Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability*, 12(24), 10243.

Prinsloo, P., & Knox, J. (2023). Exploring the ethical principles for the implementation of artificial intelligence in education: Towards a future agenda. *Research and Practice in Technology Enhanced Learning*, 18(1), 73-94. Retrieved from https://www.researchgate.net/publication/368740229_Exploring-the-Ethical-Principles-for-the-Implementation-of-Artificial-Intelligence-in-Education-Towards-a-Future-Agenda

Ramirez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., & Miranda, J. (2023). Artificial intelligence in vocational education: Ethical challenges and governance frameworks. *International Journal of Educational Technology in Higher Education*, 20(1), 138-155. <https://doi.org/10.1186/s41239-023-00408-3>

Rodriguez-Abitia, G., & Bribiesca-Correa, G. (2021). Assessing the ethical impact of artificial intelligence in education. *RED. Revista de Educación a Distancia*, 6(1). <https://doi.org/10.34627/redvol6iss1e202303>

Roman Etxebarrieta, G., Orcasitas-Vicandi, M., & Antzaka, A. (2024). *Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)*. AIPioneers.org. Retrieved from: <https://aipioneers.org/evaluation-schema-for-ai-in-education-on-data-privacy-ethics-and-eu-values-wp5/>



Roppertz, S. (2020). Artificial Intelligence and Vocational Education and Training—Perspective of German Vet Teachers. In *European Distance and E-Learning Network (EDEN) Conference Proceedings* (No. 2, pp. 207-216). European Distance and E-Learning Network.

Rott, K. J., Lao, L., Petridou, E., & Schmidt-Hertha, B. (2022). Needs and requirements for an additional AI qualification during dual vocational training: Results from studies of apprentices and teachers. *Computers and education: Artificial intelligence*, 3, 100102.

Şenocak, D., Bozkurt, A., & Koçdar, S. (2024). Exploring the Ethical Principles for the Implementation of Artificial Intelligence in Education: Towards a Future Agenda. In *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 200-213). IGI Global.

Seufert, S. (2024). Artificial Intelligence in Vocational Education and Training (VET): Evaluating VET Leaders' Acceptance of AI in Switzerland.

Tang, J., & Hai, L. (2021). Construction and exploration of an intelligent evaluation system for educational app through artificial intelligence technology. *International Journal of Emerging Technologies in Learning*, 16(5), 17–31. <https://doi.org/10.3991/ijet.v16i05.20293>

Tuomi, I., Cachia, R., & Villar-Onrubia, D. (2023). On the futures of technology in education: Emerging trends and policy implications. *Publications Office of the European Union, Luxembourg*.

UNESCO (2023a). *ChatGPT and artificial intelligence in higher education*. Retrieved from: https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

UNESCO (2023b). *Guidance for generative AI in education and research*. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18, 1-20.

Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S., & Roman-Cañizares, M. (2021). Implementation of a virtual assistant for the academic management of a university with the use of artificial intelligence. *Future Internet*, 13(4), 97.



Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235.

Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1), 1–5. <https://doi.org/10.1080/17439884.2023.2167830>

Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology*, 95-116.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2023). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00408-3>

CONTACT

www.aipioneers.org

gorka.roman@ehu.es

maria.orcasitas@ehu.es

natalia.louleli@ehu.eus



AI PIONEERS
Artificial Intelligence in education & training

Handbuch zu Politik und Ethik in der KI-Bildung

Juni 2025

Zuständiger Projektpartner: Universidad del País Vasco/Euskal Herriko
Unibertsitatea (UPV/EHU)

Arbeitspaket 5

Projektergebnis D5.2 "Handbuch zu Politik und Ethik in der KI-Bildung"

Projektnummer: 101087261

Projektname: AI and the future of Education

Akronym des Projekts: AI Pioneers

Projektkoordinator: Universität Bremen, Institut Technik und Bildung (ITB)

Förderschwerpunkt: ERASMUS-EDU-2022-PI-FORWARD

Thema: ERASMUS-EDU-2022-PI-FORWARD-LOT1

Aktionstyp: ERASMUS Lump Sum Grants

Bewilligungsbehörde: European Education and Culture

Projektbeginn: 01/01/2023

Projektende: 31/12/2025

Projektdauer: 36 Monate



Kofinanziert von der
Europäischen Union



AI PIONEERS
Artificial Intelligence in education & training



AI PIONEERS
Artificial Intelligence in education & training

Deutsche Version des Projektergebnisses

Juni 2025



Kofinanziert von der
Europäischen Union



AI PIONEERS
Artificial Intelligence in education & training

Handbuch zu Politik und Ethik in der KI-Bildung

Gorka Roman Etxebarrieta (Autor)

Maria Orcasitas-Vicandi (Autorin)

Natalia Louleli (Autorin)

Asier León Nuñez (Autor)

Itsaso Biota Piñeiro (Autorin)

Ana Boyano Murillo (Autorin)

Raquel Justo Blanco (Autorin)

Josune Rodriguez Negro (Autorin)



Inhaltsverzeichnis

Handbuch zu Politik und Ethik in der KI-Bildung	3
1. Einleitung.....	5
1.1 Zweck und Umfang des Handbuchs.....	6
1.2 Die Bedeutung ethischer Überlegungen bei der Einführung von KI	7
1.3 Ein Handbuch für Lehrende, Administrator:innen und politische Entscheidungsträger:innen	8
2. Verständnis von KI in der Berufs- und Erwachsenenbildung	11
2.1. Die Rolle von KI im Bildungswesen	11
2.2 KI in der Erwachsenen- und Berufsbildung	13
3. Entwicklung ethischer Leitlinien.....	14
3.1 Die Bedeutung ethischer Grundsätze in der KI-Bildung	14
3.2 Schlüsselanforderungen für vertrauenswürdige KI-Systeme.....	16
3.3 Berücksichtigung rechtlicher Aspekte der KI-Implementierung in berufsbildenden Schulen.....	19
3.4 Fallstudien und Beispiele für wirksame KI-Politiken in Berufsbildungsinstitutionen	26
4. Ethische Fähigkeiten.....	35
4.1 Definition ethischer Fähigkeiten	35
4.2 Bedeutung ethischer Fähigkeiten.....	35
4.3 Bereiche und Beispiele für ethische Fähigkeiten	36
5. Schlussfolgerung	45
5.1 Wichtigste Erkenntnisse.....	45
5.2 Abschließende Reflexionen.....	46
Danksagung.....	47
Referenzen.....	47



1. Einleitung

Das Projekt AI Pioneers konzentriert sich darauf, die Integration von Künstlicher Intelligenz (KI) in die Bildung voranzutreiben, insbesondere in der Erwachsenenbildung und der beruflichen Aus- und Weiterbildung. Dieses zukunftsorientierte ERASMUS+-Projekt begann im Januar 2023 und wird bis Dezember 2025 finanziert.

Angesichts des wachsenden Einflusses von KI im Bildungsbereich und der steigenden Nachfrage nach professioneller Unterstützung bei der Nutzung von KI und der Entwicklung von Fähigkeiten zielt das Projekt darauf ab, ein umfassendes, globales Netzwerk von KI-Pionieren in den Partnerländern und darüber hinaus aufzubauen. Die Einrichtung des Referenznetzwerks der KI-Pioniere ist ein Schlüsselement, das sich an Lehrende, politische Entscheidungsträger:innen, Bildungsplanende, KI-Entwickelnde und andere relevante Interessengruppen richtet. Durch die Einbindung derjenigen, die am intensivsten am Einsatz von KI in der Bildung interessiert sind, soll ein zentraler Knotenpunkt geschaffen werden, der die Integration von KI in die Erwachsenenbildung und die Berufsbildung unterstützt und künftige Initiativen und KI-Projekte im Bildungsbereich fördert. Weitere Informationen über das Netzwerk und die Möglichkeit zur Teilnahme finden Sie auf der [Projektwebsite](#).

Ein weiteres wichtiges Ziel war die Entwicklung einer Ergänzung zum DigCompEdu- Rahmenwerk (Bekiaridis, 2024). Dieses Dokument umreißt die notwendigen Fähigkeiten und Kompetenzen, die Lehrende benötigen, um KI-Technologien in der Bildung zu verstehen, zu bewerten und zu nutzen, wobei ein besonderer Schwerpunkt auf der Erwachsenenbildung und der Berufsbildung liegt. In dieser Ergänzung werden die KI-Kompetenzen mit den sechs Kernbereichen des DigCompEdu-Rahmenwerks (Redecker, 2017) in Einklang gebracht: Professionelles Engagement, digitale Ressourcen, Lehren und Lernen, Bewertung, Befähigung der Lernenden und Förderung der digitalen Kompetenz der Lernenden. Für jeden Bereich, die Ergänzung:

1. Analysiert KI-Anwendungen in diesem Bereich
2. Schlägt Aktivitäten zur Entwicklung relevanter Fähigkeiten vor
3. Schlägt Professionsstufen für den Aufbau von Kompetenzen vor
4. Identifiziert Herausforderungen und bietet Strategien zu deren Bewältigung

Im Rahmen des Projekts wird weiterhin eine Reihe von Ressourcen entwickelt und über das AI Pioneers Toolkit, das auf der Website des Projekts zur Verfügung



steht, verbreitet. Dieses Toolkit bietet Lehrplanempfehlungen für die Integration von KI in die Bildung, hebt die Vorteile und Risiken des Einsatzes von KI in Bildungsumgebungen hervor und bietet Best-Practice-Leitlinien für KI in der Bildung (Tommasi & Perini, 2024). Es enthält auch kurze Übersichten über Tools für verschiedene Unterrichtszwecke, wie z. B. Lernassistenten, Tools zur Erstellung von Unterrichtsmaterialien und solche, die Peer-to-Peer-Lernen unterstützen.

Das Projekt leistet wichtige Arbeit bei der Bewältigung der ethischen Herausforderungen, die KI im Bildungsbereich mit sich bringt. Dieses Dokument wird von einem Bewertungsschema begleitet. Beide sollen Institutionen und Lehrende bei der verantwortungsvollen Implementierung von KI unter Berücksichtigung des Datenschutzes und ethischer Implikationen helfen.

Das *Bewertungsschema für KI im Bildungswesen* konzentriert sich auf praktische Leitlinien für die Einführung von KI-Tools, die sicherstellen, dass sie mit ethischen Grundsätzen und EU-Werten übereinstimmen. Es ist eine hervorragende Ressource für Bildungseinrichtungen, um fundierte Entscheidungen über die Integration von KI zu treffen.

Dieses *Handbuch zu Politik und Ethik in der KI-Bildung* ist ein umfassenderer Leitfaden, der politische und ethische Rahmenbedingungen für die Bewältigung aktueller Probleme bietet.

1.1 Zweck und Umfang des Handbuchs

Das *Handbuch zu Politik und Ethik in der KI-Bildung* dient als grundlegende Ressource für Lehrende, politische Entscheidungsträger:innen und Institutionen, die sich mit der Komplexität von KI im Bildungsbereich auseinandersetzen wollen. Es hilft dabei, einen Rahmen zu schaffen, der sicherstellt, dass ethische Überlegungen im Vordergrund stehen, die Privatsphäre der Lernenden geschützt, Fairness gefördert und eine verantwortungsvolle Nutzung unterstützt wird.

Ein entscheidender Vorteil ist die handlungsorientierte Hilfestellung für Lehrende, die ihnen einen verantwortungsvollen Umgang mit KI ermöglicht und den Lernenden hilft, deren Fähigkeiten, Grenzen und ethische Implikationen zu verstehen. Dies ist entscheidend für die Förderung einer Generation von Lernenden, die nicht nur technisch versiert sind, sondern auch ethisch bewusst mit Technologie umgehen. Für politische Entscheidungsträger ist der konsistente Rahmen, den das Handbuch bietet, ein Wegweiser zu mehr standardisierten Ansätzen in allen Institutionen, was für die Abschwächung der Risiken von KI,



wie etwa Verzerrungen in Algorithmen oder Verletzungen der Privatsphäre, von entscheidender Bedeutung ist. Es kann auch dazu beitragen, dass KI nicht nur um ihrer selbst willen integriert wird, sondern in einer Weise, die das Bildungserlebnis verbessert, ohne ethische Standards zu gefährden.

1.2 Die Bedeutung ethischer Überlegungen bei der Einführung von KI

Ethische Überlegungen sind bei der Einführung von KI von entscheidender Bedeutung, insbesondere in Bereichen wie der Erwachsenenbildung und der beruflichen Bildung, wo die Auswirkungen von KI auf Lernende, Lehrende und Institutionen groß sein können. Diese Überlegungen tragen dazu bei, dass KI verantwortungsvoll, inklusiv und transparent eingesetzt wird und dass sie die allgemeinen Ziele der Bildung unterstützt, ohne grundlegende Werte zu gefährden.

Wichtige Aspekte, die das Schema abdeckt, sind unter anderem:

Diversität und Inklusion: KI-Systeme und -Werkzeuge müssen so konzipiert sein, dass sie für ein breites Spektrum von Lernenden geeignet sind, einschließlich solcher mit unterschiedlichem Hintergrund, Fähigkeiten und Lernbedarf. In der Erwachsenen- und Berufsbildung bringen die Lernenden oft unterschiedliche Lebenserfahrungen mit, z. B. unterschiedliche Bildungsniveaus, Altersgruppen und Berufserfahrungen. Künstliche Intelligenz sollte flexibel und integrativ sein und sicherstellen, dass ihr Einsatz bestimmte Gruppen unterstützt und nicht ausschließt. Ethische Verfechter:innen von KI müssen sich darüber im Klaren sein, auf welche Art und Weise die Technologie Vorurteile aufrechterhalten oder digitale Klüfte schaffen könnte, um diese zu mildern.

Transparenz: Damit KI im Bildungsbereich ethisch vertretbar eingesetzt werden kann, ist Transparenz unerlässlich. Die Einrichtungen müssen sich darüber im Klaren sein, wie KI-Systeme funktionieren, welche Daten sie sammeln und welche Entscheidungsprozesse dahinterstehen. Lernende und Lehrende sollten verstehen, wie KI-Tools im Klassenzimmer oder in der Ausbildungsumgebung eingesetzt werden und wie die Systeme die Lernergebnisse beeinflussen können. Diese Transparenz fördert das Vertrauen in KI-Technologien und stellt sicher, dass ihr Einsatz verantwortungsvoll erfolgt.

Privatsphäre und Datenschutz: Einrichtungen der Erwachsenenbildung und der beruflichen Bildung haben oft mit sensiblen personenbezogenen Daten zu tun, z. B. mit den Profilen der Lernenden, den Bewertungsergebnissen und dem beruflichen Werdegang. Der ethische Einsatz von KI in diesem Kontext muss den Schutz der Privatsphäre und den Datenschutz in den Vordergrund stellen. Die



Einrichtungen müssen die Einhaltung der Datenschutzgesetze (z. B. DSGVO) gewährleisten und sicherstellen, dass die Daten der Lernenden auf verantwortungsvolle Weise erhoben, gespeichert und verwendet werden. Die Zustimmung zur Datenerhebung sollte in Kenntnis der Sachlage und freiwillig erfolgen, und die Lernenden sollten die Kontrolle über ihre persönlichen Daten haben.

Nachhaltigkeit: KI sollte in einer Weise integriert werden, die mit der langfristigen Nachhaltigkeit von Bildungssystemen in Einklang steht. Dies bedeutet, dass sowohl die Umweltauswirkungen von KI-Technologien als auch ihre Rolle bei der Förderung nachhaltiger Praktiken in Bildungs- und Ausbildungsprogrammen berücksichtigt werden müssen. Im Kontext der Berufsbildung kann KI auch die Nachhaltigkeit in der Industrie unterstützen, indem sie die Lernenden mit Fähigkeiten ausstattet, die grüne Technologien und umweltfreundliche Praktiken fördern

KI-Kenntnisse/ Fähigkeiten: Mit der zunehmenden Integration von KI in die Bildung ist es von entscheidender Bedeutung, KI-Kenntnisse sowohl bei Lehrenden als auch bei Lernenden aufzubauen. In der Erwachsenen- und Berufsbildung bedeutet dies, die Lernenden mit dem Wissen und den Fähigkeiten auszustatten, die sie benötigen, um KI-Technologien zu verstehen und sie in ihrer beruflichen Laufbahn effektiv einzusetzen. Diese Befähigung ermöglicht es den Lernenden, sich auf dem sich wandelnden Arbeitsmarkt zurechtzufinden und sicherzustellen, dass sie nicht zurückbleiben. Für Lehrende bedeutet KI-Kompetenz, dass sie KI-Tools kritisch bewerten und sie so einsetzen können, dass sie das Lernen verbessern, ohne ethische Standards zu verletzen.

Fairness und Gerechtigkeit: KI-Systeme sollten so konzipiert sein, dass sie alle Lernenden fair behandeln und sicherstellen, dass keine Gruppe durch automatische Entscheidungen oder Algorithmen benachteiligt wird. In der Erwachsenen- und Berufsbildung ist dies besonders wichtig, da diese Bereiche häufig Lernende bedienen, die bereits mit sozialen oder wirtschaftlichen Herausforderungen konfrontiert sind. Es ist von entscheidender Bedeutung, dass KI bestehende Ungleichheiten nicht verstärkt oder neue Zugangshindernisse schafft, sondern vielmehr gerechte Chancen für alle Lernenden unterstützt.

1.3 Ein Handbuch für Lehrende, Administrator:innen und politische Entscheidungsträger:innen

Dieses Handbuch richtet sich an Lehrende, Administrator:innen und politische Entscheidungsträger:innen im Zusammenhang mit dem Einsatz von KI in der



Erwachsenenbildung und der beruflichen Bildung. Es ist bekannt, dass der Einsatz von KI und die damit verbundenen ethischen Überlegungen jede der genannten Gruppen auf unterschiedliche Weise betreffen:

Lehrende:

Für Lehrende geht es bei der ethischen Integration von KI in der Erwachsenen- und Berufsbildung darum, das Lehren und Lernen zu verbessern, ohne die Grundsätze der Fairness, der Privatsphäre und der Transparenz zu missachten. KI-Tools können personalisierte Lernerfahrungen bieten, Verwaltungsaufgaben rationalisieren und Einblicke in die Fortschritte der Lernenden geben. Lehrende müssen jedoch sicherstellen, dass KI unterschiedliche Lernstile und -bedürfnisse unterstützt und ein integratives Lernumfeld fördert.

Es gibt viele wichtige Überlegungen für Lehrende in Bezug auf KI-Kenntnisse, ethische Nutzung und Transparenz gegenüber den Lernenden.

KI-Kenntnisse: Lehrende müssen mit den Fähigkeiten ausgestattet sein, KI zu verstehen, zu bewerten und effektiv in ihre Lehrmethoden zu integrieren. Dazu gehört auch das Wissen über die Funktionsweise von KI-Tools und ihre ethischen Auswirkungen im Klassenzimmer.

Ethische Nutzung: Lehrende müssen sicherstellen, dass KI zur Verbesserung des Lernens eingesetzt wird und nicht die menschliche Interaktion ersetzt. Sie sollten darauf vorbereitet sein, Fragen der Voreingenommenheit oder Diskriminierung anzusprechen, die durch KI-Systeme entstehen können.

Transparenz gegenüber Lernenden: Lehrende sollten den Lernenden gegenüber transparent sein, wie KI im Lernprozess eingesetzt wird und welche Daten sie sammelt. Dies trägt dazu zum Aufbau von Vertrauen bei und gibt den Lernenden das Gefühl, dass ihre Privatsphäre respektiert wird.

Administrator:innen:

Für Administrator:innen in Erwachsenenbildungs- und Berufsbildungseinrichtungen bedeutet die Einführung von KI, dass sie die Integration dieser Technologien in die Systeme der Einrichtung beaufsichtigen und sicherstellen müssen, dass die Infrastruktur für die Unterstützung von KI-Tools vorhanden ist. Sie müssen auch Richtlinien festlegen, die den ethischen Einsatz von KI sicherstellen und gleichzeitig die institutionelle Integrität wahren.



Für die Administrator:innen gibt es viele wichtige Überlegungen zum Datenschutz und zur Privatsphäre, zur Gleichberechtigung und zum Zugang sowie zur laufenden Schulung und Unterstützung.

Datenschutz und Privatsphäre: Die Administrator:innen müssen sicherstellen, dass robuste Datenschutzmaßnahmen vorhanden sind, da KI auf der Erfassung großer Mengen von Daten der Lernenden beruht. Die Einhaltung von Datenschutzbestimmungen (wie der DSGVO) ist unerlässlich.

Gerechtigkeit und Zugang: KI sollte so eingesetzt werden, dass alle Lernenden unabhängig von ihrem sozioökonomischen Hintergrund gleichberechtigten Zugang erhalten. Administrator:innen sollten sich für Maßnahmen einsetzen, mit denen verhindert wird, dass KI-Tools bestehende Ungleichheiten verschärfen.

Laufende Schulung und Unterstützung: Die Administrator:innen müssen Ressourcen für die kontinuierliche berufliche Weiterbildung bereitstellen, damit die Lehrenden über KI-Fortschritte und ethische Best Practices auf dem Laufenden bleiben können.

Politische Entscheidungsträger:innen:

Politische Entscheidungsträger:innen spielen eine entscheidende Rolle bei der Erstellung der Rahmenbedingungen und Vorschriften, die den Einsatz von KI im Bildungswesen regeln. Sie müssen sicherstellen, dass der Einsatz von KI mit den allgemeinen Bildungszielen übereinstimmt, die Menschenrechte respektiert und die Möglichkeiten für lebenslanges Lernen verbessert werden.

Für die politischen Entscheidungsträger:innen gibt es eine Reihe wichtiger Überlegungen zu ethischen Richtlinien und Vorschriften, zur Förderung der KI-Kompetenz und zur Gewährleistung von Inklusion und Zugänglichkeit.

Ethische Richtlinien und Vorschriften: Die politischen Entscheidungsträger:innen müssen ethische Standards und Vorschriften für den Einsatz von KI in der Bildung festlegen. Dazu gehört auch die Behandlung von Themen wie Transparenz, Datenschutz und Fairness in KI-gesteuerten Bildungssystemen.

Förderung der KI-Kenntnisse: Politische Entscheidungsträger:innen sollten die KI-Kenntnisse im gesamten Bildungssystem fördern und sicherstellen, dass sowohl Lehrende als auch Lernende über die notwendigen Werkzeuge und das Verständnis für den Umgang mit KI-Technologien verfügen.



Gewährleistung von Inklusion und Zugänglichkeit: Es ist von entscheidender Bedeutung, dass die KI-Politik der inklusiven Bildung einen hohen Stellenwert einräumt und dafür sorgt, dass KI-Technologien bestimmte Gruppen nicht benachteiligen. Dies kann Anreize für die Forschung und Entwicklung von KI-Lösungen beinhalten, die auf unterschiedliche Lernende zugeschnitten sind.

2. Verständnis von KI in der Berufs- und Erwachsenenbildung

Die Integration von Künstlicher Intelligenz (KI) in die Bildung birgt neben großen Herausforderungen auch neue Chancen. Die Forschung unterstreicht das Potenzial der KI, den Unterricht durch automatisierte Aufgaben, personalisiertes Lernen und datengesteuerte pädagogische Innovationen zu verbessern, während generative KI-Tools die Möglichkeiten für interaktive und reflektierende Bildung erweitern. Ethische Bedenken hinsichtlich des Datenschutzes, der kommerziellen Einflussnahme und der digitalen Kluft erfordern jedoch eine stabile Steuerung. In der Erwachsenen- und Berufsbildung beeinflusst KI nicht nur die Gestaltung von Lehrplänen und die Vorhersage von Fähigkeiten, um die Bereitschaft zur Arbeit zu erhöhen, sondern erfordert auch die Förderung digitaler Kompetenzen und eine kritische Auseinandersetzung mit den gesellschaftlichen Auswirkungen von KI. Der Ausgleich zwischen den Vorteilen der KI und einer gerechten, ethischen Politik bleibt entscheidend für die Gestaltung inklusiver, zukunftsorientierter Bildungssysteme.

2.1. Die Rolle von KI im Bildungswesen

Ein kürzlich veröffentlichter Bericht des Joint Research Centre (JRC) der Europäischen Kommission mit dem Titel [On the Futures of Technology in Education: Emerging Trends and Policy Implications](#) (Tuomi, Cachia & Villar-Onrubia, 2023) bietet einen umfassenden Überblick darüber, wie technologische Fortschritte, insbesondere im Bereich der künstlichen Intelligenz (KI), Bildungspraktiken, -prozesse und -einrichtungen verändern. Das Dokument hebt die bedeutenden und vielfältigen Auswirkungen von KI auf die Bildung hervor, die sich auf Lehr- und Lernprozesse, Bewertungsmethoden und Organisationsstrukturen auswirken.

KI wird als potenzielle Lösung zur Verringerung der Arbeitsbelastung von Lehrkräften angesehen, insbesondere bei Aufgaben wie der Benotung von Lernendenhausaufgaben. Durch die Automatisierung dieser sich wiederholenden und zeitaufwändigen Aufgaben kann KI den Unterricht attraktiver



machen und es den Lehrkräften ermöglichen, sich mehr auf den Unterricht als auf administrative Aufgaben zu konzentrieren (Li, 2021; Tang & Hai, 2021; Villegas-Ch et al., 2021). Diese Automatisierung könnte die formative Bewertung revolutionieren (Kumar & Boulanger, 2020), indem sie kontinuierliches Feedback für die Lernenden ermöglicht und eine Bewertung in Echtzeit (Porter & Grippa, 2020) sowie personalisiertes oder adaptives Lernen (Kochmar et al., 2020; Wongvorachan et al., 2022) erleichtert.

Darüber hinaus wird erwartet, dass generative KI und grundlegende Modelle einen erheblichen Einfluss auf die Bildung haben werden. Technologien wie automatische Videountertitelung, Übersetzung und Videozusammenfassung bieten innovative Möglichkeiten zur Verbesserung der Pädagogik (Chew & Chua, 2020; Vazquez-Cano et al., 2021). Diese Technologien können dazu beitragen, „intelligente Begleiter“ oder „Lernpartner“ zu entwickeln, die den Schwerpunkt vom Auswendiglernen auf Metakognition und Reflexion verlagern. Dies erfordert eine sorgfältige Neuverteilung der Zuständigkeiten zwischen Menschen und Maschinen, wobei festzulegen ist, welche Tätigkeiten an die KI delegiert werden sollen.

Wir sollten auch die Rolle der KI bei der Datafizierung der Bildung hervorheben, bei der umfangreiche Daten zu Bildungsaktivitäten genutzt werden, um neue pädagogische Ansätze, Lerntechnologien und bildungspolitische Maßnahmen zu fördern und zu informieren (Williamson et al., 2023). Das Ziel ist die Erfassung von Lernprozessen, und nicht nur von Ergebnissen, um das Engagement und die Leistungen der Lernenden zu verbessern (Porter & Grippa, 2020).

Allerdings sind ethische Bedenken hinsichtlich der Verarbeitung personenbezogener Daten, des potenziellen Missbrauchs, der Abhängigkeit von Infrastrukturen, die weitgehend von kommerziellen Unternehmen kontrolliert werden, und einiger anderer Herausforderungen kritische Fragen, die es zu klären gilt (Seufert, 2024; Wongvorachan et al., 2022). Die Integration von KI in die Bildung wirft wichtige Fragen zu Governance-Modellen und politischen Entscheidungen auf und erfordert eine Neubewertung der Bildungspolitik, um die Fähigkeiten und Grenzen von KI zu berücksichtigen (Attard-Frost et al., 2024; Boninger et al., 2020; Williamson & Eynon, 2020). Wir müssen auch die Risiken einer Verschärfung der digitalen Kluft und der Dominanz kommerzieller Interessen in diesem Sektor kritisch bewerten (Bulathwela et al., 2024; Celik, 2023). Die Regulierung der technologischen Innovation zur Gewährleistung vielfältiger Bildungsvisionen und die Abstimmung der Industriepolitik auf Lerntheorien sind entscheidende Aspekte, die wir berücksichtigen müssen.



Die Rolle der KI in der Bildung ist ein transformatives und vielschichtiges Instrument, das Möglichkeiten für mehr Effizienz, Personalisierung und Innovation beim Lehren und Lernen bietet. Allerdings bringt sie auch erhebliche Herausforderungen und ethische Überlegungen mit sich, die angegangen werden müssen, um einen gerechten, effektiven und verantwortungsvollen Einsatz von KI im Bildungsbereich zu gewährleisten.

2.2 KI in der Erwachsenen- und Berufsbildung

Die Integration der KI- Bildung in die Lehrpläne beinhaltet nicht nur die Vermittlung der Grundlagen der KI und ihrer Anwendungen, sondern auch die Erforschung ihrer gesellschaftlichen Auswirkungen. Dazu gehört auch die Förderung des kritischen Denkens über die Fähigkeiten, Grenzen und ethischen Dimensionen von KI, wie z. B. Bedenken hinsichtlich Voreingenommenheit und Datenschutz (Akgun & Greenhow, 2022; Aldosari, 2020; Attwell et al., 2020; Delcker, et al., 2024; Seufert, 2024; Wongvorachan et al., 2022). Die Verbesserung der digitalen Kompetenz ist ebenfalls von entscheidender Bedeutung, da das Verständnis für die Interaktion mit und die Interpretation von KI-Systemen zu einer Schlüsselqualifikation für moderne Arbeitskräfte, einschließlich Lehrkräften, wird (Bekiaridis & Attwell, 2024; Delcker, et al., 2024; Roppertz, 2020; Rott et al., 2022; Seufert, 2024). Mit der Weiterentwicklung der KI verlagert sich der Schwerpunkt darauf, KI-Tools nicht nur zu nutzen, sondern auch zu verstehen und kritisch zu bewerten, um die Lernenden auf eine Zukunft vorzubereiten, in der KI integraler Bestandteil des täglichen Lebens und der Arbeit ist (Rott et al., 2022).

Folglich ist die Vorbereitung der Arbeitskräfte auf neue Fähigkeiten im Zeitalter der KI ein weiterer wichtiger Aspekt der Erwachsenen- und Berufsbildung (Attwell et al., 2020; Roppertz, 2020; Rott et al., 2022). Die Rolle der KI bei der Erkennung neuer Trends und der Vorhersage künftiger Qualifikationsanforderungen ist von unschätzbarem Wert. Diese Erkenntnisse ermöglichen Bildungsprogrammen die Anpassung ihrer Lehrpläne an die sich wandelnden Anforderungen des Arbeitsmarktes (Seufert, 2024). Höherqualifizierung und Umschulung werden zu zentralen Themen, wobei KI-gesteuerte Programme Erwachsenen und Fachkräften ermöglichen, neue Fähigkeiten zu erwerben oder bestehende zu aktualisieren und so mit dem technologischen Fortschritt Schritt zu halten (Attwell et al., 2020; Emeršič et al., 2024; Delcker, et al., 2024). Darüber hinaus stellt eine KI-gesteuerte Ausbildung sicher, dass die Lernenden nicht nur technisch versiert sind, sondern auch auf die Arbeit in KI-integrierten Umgebungen vorbereitet werden. Diese Vorbereitung ist der Schlüssel, um sicherzustellen, dass die Arbeitskräfte auf die Herausforderungen und Chancen des sich rasch



entwickelnden Bereichs der KI vorbereitet sind, wodurch die Erwachsenen- und Berufsbildung relevanter und effektiver wird und sich an der künftigen Beschäftigungslandschaft orientiert.

3. Entwicklung ethischer Leitlinien

3.1 Die Bedeutung ethischer Grundsätze in der KI-Bildung

Ethische Grundsätze in der KI-Bildung sind unerlässlich, damit sichergestellt werden kann, dass KI-Technologien in einer Weise eingesetzt werden, die Fairness, Transparenz und Verantwortung fördert und gleichzeitig die Rechte und das Wohlergehen von Lernenden und Lehrenden schützt. Da KI zunehmend in den Bildungsbereich integriert wird, sind klare ethische Richtlinien von entscheidender Bedeutung, um die potenziellen Risiken zu beherrschen und die Vorteile zu nutzen. Im Folgenden werden die wichtigsten Gründe genannt, warum ethische Richtlinien in der KI-Bildung unerlässlich sind:

Schutz der Rechte der Lernenden und der Privatsphäre

KI-Systeme im Bildungswesen erfordern häufig die Erfassung und Verarbeitung personenbezogener Daten, einschließlich akademischer Aufzeichnungen, Lernverhalten und anderer sensibler Informationen. Ethische Richtlinien stellen sicher, dass die Daten mit Respekt und in Übereinstimmung mit den Datenschutzbestimmungen (z. B. DSGVO) behandelt werden. Diese Richtlinien tragen dazu bei, die Privatsphäre der Lernenden zu schützen, indem sie für Transparenz bei der Datenerfassung sorgen, den Lernenden die Kontrolle über ihre Daten geben und sicherstellen, dass diese verantwortungsvoll und sicher verwendet werden.

Gewährleistung von Fairness und Gleichberechtigung

Eines der größten ethischen Probleme im Zusammenhang mit KI im Bildungsbereich ist das Potenzial für voreingenommene Algorithmen, die bestimmte Gruppen von Lernenden benachteiligen könnten. KI-Systeme können unbeabsichtigt bestehende Ungleichheiten aufgrund von Ethnie, Geschlecht, sozioökonomischem Status oder anderen Faktoren aufrechterhalten. Ethische Richtlinien sind notwendig, um die Entwicklung und den Einsatz von KI-Tools zu lenken, die inklusiv und gerecht sind und sicherstellen, dass alle Lernenden unabhängig von ihrem Hintergrund oder ihren persönlichen Merkmalen gleichen Zugang zu den Möglichkeiten haben.



Förderung von Transparenz und Rechenschaftspflicht

KI-Technologien können komplex und undurchsichtig sein, so dass es für Lehrende, Lernende und Administrator:innen schwierig ist zu verstehen, wie Entscheidungen getroffen werden. Ethische Richtlinien tragen dazu bei, dass KI-Systeme transparent sind, d. h. die Beteiligten wissen, wie KI-Tools funktionieren, welche Daten verwendet werden und wie Entscheidungen getroffen werden. Diese Transparenz schafft Vertrauen und Verantwortlichkeit und ermöglicht es Lehrenden und Lernenden, KI-Technologien besser zu verstehen und mit ihnen zu interagieren.

Förderung von KI-Kenntnissen und verantwortungsvoller Nutzung

Da KI zunehmend in den Lernprozess integriert wird, ist es wichtig, dass Lehrende und Lernende gleichermaßen verstehen, wie KI funktioniert und welche Auswirkungen sie haben kann. Ethische Maßnahmen in der KI-Bildung sollten die KI-Kompetenz fördern und sicherstellen, dass sowohl Lehrende als auch Lernende mit dem Wissen und den Fähigkeiten ausgestattet sind, KI verantwortungsvoll und kritisch zu nutzen. Dieses Verständnis ermöglicht es den Nutzenden, die Vorteile der KI zu maximieren und sich gleichzeitig ihrer Grenzen und potenziellen Schäden bewusst zu sein.

Förderung von ethischer Entwicklung und Innovation

KI-Technologien im Bildungsbereich sollten auf einer soliden ethischen Grundlage entwickelt werden, um sicherzustellen, dass sie zum Nutzen aller Lernenden eingesetzt werden. Ethische Richtlinien bieten Entwickelnden, Forschenden und Bildungseinrichtungen einen Rahmen für die Gestaltung von KI-Systemen, die mit pädagogischen Werten wie Fairness, Befähigung und Respekt für Vielfalt in Einklang stehen. Dadurch wird auch sichergestellt, dass KI-Technologien so konzipiert und eingesetzt werden, dass sie das lebenslange Lernen und die allgemeinen Ziele der Bildung unterstützen.

Umgang mit dem Verlagerungsrisiko

In dem Maße, wie KI-Systeme bestimmte Aspekte der Bildung automatisieren (z. B. Benotung, Nachhilfe, Verwaltungsaufgaben), besteht die Gefahr, dass sie Lehrende verdrängen oder die menschliche Interaktion einschränken. Ethische Richtlinien können dazu beitragen, den verantwortungsvollen Einsatz von KI zu steuern, indem gewährleistet wird, dass KI die menschlichen Lehrenden ergänzt, anstatt sie zu ersetzen. Richtlinien können den Einsatz von KI fördern, um die



Effektivität des Unterrichts und die Effizienz der Verwaltung zu verbessern und gleichzeitig den Wert der menschlichen Präsenz in der Bildung zu erhalten.

Vertrauen in die KI-Systeme schaffen

Damit KI im Bildungsbereich erfolgreich eingesetzt werden kann, müssen Lehrende, Lernende und Administrator:innen der Technologie vertrauen. Ethische Richtlinien spielen eine wichtige Rolle bei der Schaffung und Aufrechterhaltung dieses Vertrauens, indem sie klare Leitlinien für den Einsatz von KI und die Einhaltung ethischer Standards vorgeben. Wenn KI-Systeme im Einklang mit diesen Richtlinien entwickelt und eingesetzt werden, ist es wahrscheinlicher, dass die Beteiligten sie annehmen und effektiv nutzen.

Die Zukunft der KI im Bildungswesen steuern

Da sich die KI weiterentwickelt, ist es von entscheidender Bedeutung, dass ethische Richtlinien eine langfristige Orientierung für ihre Integration in Bildungssysteme bieten. Diese Richtlinien tragen dazu bei, einen Rahmen für die künftige Entwicklung von KI-Technologien zu schaffen, der mit den gesellschaftlichen Werten in Einklang steht und sicherstellt, dass KI zur Verbesserung von Lernergebnissen, zur Förderung von Innovationen und zur Unterstützung des lebenslangen Lernens auf eine Weise eingesetzt wird, die sowohl nützlich als auch ethisch vertretbar ist.

Generell sind ethische Richtlinien in der KI-Bildung notwendig, um die verantwortungsvolle Entwicklung und Nutzung von KI-Technologien zu steuern. Diese Richtlinien stellen sicher, dass KI in einer Weise eingesetzt wird, die die Privatsphäre der Lernenden respektiert, Fairness fördert, Transparenz unterstützt und Vertrauen schafft. Darüber hinaus helfen sie Lehrenden, Lernenden und Institutionen bei der Bewältigung der ethischen Herausforderungen, die sich aus der fortschreitenden Veränderung des Bildungsumfelds durch KI ergeben. Durch die Festlegung dieser ethischen Richtlinien können Bildungssysteme die Möglichkeiten der KI nutzen und gleichzeitig die Werte schützen, die für eine effektive und inklusive Bildung grundlegend sind.

3.2 Schlüsselanforderungen für vertrauenswürdige KI-Systeme

Die Integration von Künstlicher Intelligenz (KI) in die Bildung bietet immense Möglichkeiten zur Verbesserung von Lehre, Lernen und Verwaltung. Um jedoch sicherzustellen, dass KI ethisch und verantwortungsvoll eingesetzt wird, müssen Bildungseinrichtungen klare Richtlinien für die ethische Implementierung von KI-



Systemen entwickeln. Zahlreiche veröffentlichte Rahmenwerke, Richtlinien und Audits (z. B. AI Audit 2023a der UNESCO und Guidance for Generative AI in Education and Research 2023b) bieten wertvolle Einblicke in die Frage, was vertrauenswürdige KI ausmacht. Diese Rahmenwerke betonen Schlüsselprinzipien wie **Transparenz, Fairness, Rechenschaftspflicht und die Achtung der Menschenrechte**. Die Aufgabe, diese Grundsätze zu entwickeln und anzuwenden, kann jedoch für Institutionen entmutigend sein, insbesondere in einer sich schnell entwickelnden technologischen Landschaft.

Im Rahmen von AI Pioneers haben wir ein Bewertungsschema entwickelt (Roman Etxebarrieta et al., 2024), das diese Prinzipien für den Bildungsbereich vereinfachen und kontextualisieren soll. Unser Schema bietet einen strukturierten Ansatz, der Institutionen dabei hilft, die Komplexität der KI-Implementierung zu bewältigen und gleichzeitig die Einhaltung ethischer Standards zu gewährleisten. Solange es mehrere Richtlinien gibt, soll unser Schema ein praktisches, benutzendenfreundliches Werkzeug bieten, das Bildungseinrichtungen bei der Entwicklung ihrer eigenen ethischen KI-Richtlinien unter Berücksichtigung ihres einzigartigen Kontexts und ihrer Bedürfnisse helfen kann.

Eine der zentralen Erkenntnisse des EU-KI-Gesetzes (Europäisches Parlament und Rat der Europäischen Union, 2024) ist, dass die Hauptverantwortung für die ethische Entwicklung von KI-Systemen bei den Entwickler:innen und Unternehmen liegt, die sie herstellen, und dass die zu befolgenden Richtlinien vom Risikoniveau der KI-Systeme abhängen. Die Bildungseinrichtungen sind jedoch dafür verantwortlich, dass diese Systeme in ihrem Bildungsumfeld auf ethische Weise implementiert und genutzt werden. Die Zentren müssen sich der wichtigsten ethischen Fragen bewusst sein, um sowohl die Grenzen als auch das Potenzial der von ihnen eingesetzten KI-Systeme zu verstehen. Dazu gehört nicht nur die Einhaltung bestehender Richtlinien und rechtlicher Anforderungen, sondern auch die Vorwegnahme und Bewältigung von Herausforderungen, die sich aus der Weiterentwicklung dieser Technologien ergeben können.

Der Schwerpunkt unseres Bewertungsschemas liegt auf der Identifizierung ethischer Bedenken und der Entwicklung einer Politik oder Strategie, die den Einsatz von KI in Bildungszentren unterstützen und fördern kann und dabei die wichtigsten ethischen Leitlinien beachtet. Zu diesen Leitlinien gehören: **Vielfalt und Gerechtigkeit, Transparenz und Rechenschaftspflicht, Privatsphäre und Datenschutz, Sicherheit, Nachhaltigkeit und soziales Wohlergehen, Befähigung von Lehrenden und Lernenden, demokratische Beteiligung an der bildungspolitischen Planung und an KI-Praktiken, Autonomie, ethisches Design (in Bezug auf Nachhaltigkeit und Zugänglichkeit),**



Kommerzialisierung (Şenocak et al., 2024), Handlungskompetenz von Lehrenden und Lernenden sowie KI-Kompetenz.

Das Bewertungsschema besteht aus mehreren Abschnitten, die auf den Dimensionen von Chan (2023) beruhen und mit der Überprüfung von Şenocak et al. (2024) und den ethischen Leitlinien der Europäischen Kommission (2022) abgestimmt sind, wobei jeder Abschnitt auf verschiedene Aspekte der KI-Implementierung im Bildungsbereich ausgerichtet ist. Diese Abschnitte geben den Institutionen eine Anleitung für die Entwicklung ethischer Rahmenbedingungen und die Gewährleistung eines verantwortungsvollen Umgangs mit KI. Der erste Abschnitt befasst sich mit der **Bewertung des aktuellen KI-Einsatzes und ethischer Bedenken** und bietet Einrichtungen eine Hilfestellung bei der Bewertung ihres aktuellen KI-Reifegrads, so dass sie beurteilen können, inwieweit die Bildungseinrichtung bereit ist, KI zu implementieren oder ihren KI-Einsatz zu erweitern (JISC, 2022).

Der zweite Abschnitt konzentriert sich auf **KI-Governance und -Überwachung** und befasst sich daher mit Fragen der Einhaltung regionaler und internationaler Richtlinien, die den Schutz der Privatsphäre, den Datenschutz und die Fairness betreffen. Er befasst sich mit komplexen ethischen Fragen wie Antidiskriminierung und Transparenz, die ebenfalls zu den wichtigsten Überlegungen für ethische KI gehören. Der dritte Abschnitt, **Operative Überlegungen**, enthält Hinweise zur Schulung von Lehrenden, Lernenden und Mitarbeitenden, zur Förderung von KI-Kenntnissen, zur Gewährleistung der menschlichen Aufsicht und zur Unterstützung der demokratischen Beteiligung an KI-bezogenen Entscheidungen. Der Schlüssel zu einer erfolgreichen KI-Implementierung liegt darin, alle Nutzenden angemessen zu schulen und sie über die Möglichkeiten und Grenzen von KI aufzuklären, um einen gleichberechtigten Zugang für alle zu gewährleisten.

Der vierte Abschnitt befasst sich mit **pädagogischen Überlegungen**, wobei wichtige Themen wie ethische Dilemmata bei KI-gesteuerten Bewertungen, die Handlungskompetenz der Lernenden und die Kompetenzentwicklung, insbesondere in der Erwachsenenbildung und der beruflichen Bildung, berücksichtigt werden (Attwell et al., 2021; UNESCO, 2019). Der fünfte und letzte Abschnitt befasst sich mit weiteren Überlegungen: **Nachhaltigkeit und ethisches Design**. Hier werden Leitlinien für den Umgang mit langfristigen Auswirkungen und die Gewährleistung einer zugänglichen KI-Nutzung angeboten.

Das Bewertungsschema ist ein praktisches Instrument für Bildungseinrichtungen, um die KI-Implementierung zu steuern, ethische Herausforderungen zu



identifizieren und Richtlinien zu entwickeln, die mit internationalen Standards übereinstimmen. Ausführlichere Informationen finden Sie im vollständigen Bewertungsschema [Bewertungsschema für KI in der Bildung zu Daten, Datenschutz, Ethik und EU-Werten (WP5)] und in der Präsentation von Genially, die auf der [AI Pioneers-Website](#) zu finden ist.

3.3 Berücksichtigung rechtlicher Aspekte der KI-Implementierung in berufsbildenden Schulen

3.3.1 Der Rechtsrahmen des EU-KI-Gesetzes

Mit dem KI-Gesetz der Europäischen Union wurde ein umfassender Rechtsrahmen geschaffen, der Systeme der künstlichen Intelligenz nach ihren potenziellen Risiken einstuft und den Einrichtungen der beruflichen Aus- und Weiterbildung, die diese Technologien einsetzen, eine wichtige Orientierungshilfe bietet. Dieser risikobasierte Ansatz gewährleistet eine verhältnismäßige Regulierungsaufsicht und fördert gleichzeitig Innovationen im Bildungskontext (Europäisches Parlament, 2024).

Wie Zawacki-Richter et al. (2023) in ihrem systematischen Überblick über ethische Erwägungen in der KI-gestützten Bildung darlegen, müssen die rechtlichen Rahmenbedingungen ein Gleichgewicht zwischen technologischer Innovation und dem Schutz der Grundrechte herstellen. In ihrer Analyse wird das EU-KI-Gesetz als wegweisender Versuch bezeichnet, „klare Parameter für den verantwortungsvollen Einsatz von KI in Lernumgebungen festzulegen und gleichzeitig das transformative Potenzial dieser Technologien anzuerkennen“ (S. 8).

Risikoklassifizierung für KI im Bildungsbereich

KI-Systeme, die in der beruflichen Bildung eingesetzt werden, fallen aufgrund ihrer erheblichen Auswirkungen auf die Bildungserfahrungen und die künftigen Karrierechancen der Lernenden überwiegend in die Kategorie „**hohes Risiko**“. Der EU AI Act kategorisiert AI-Systeme wie folgt:

Inakzeptables Risiko: Systeme, die eine eindeutige Bedrohung für die Grundrechte oder die Sicherheit darstellen, sind gänzlich verboten.

Hohes Risiko: Anwendungen in kritischen Sektoren, wie dem Bildungswesen, die eine strenge Einhaltung von Vorschriften erfordern.



Begrenztes Risiko: Systeme mit Transparenzpflichten, die eine Offenlegung der KI-Interaktion erfordern.

Minimales Risiko: Anwendungen mit begrenzten regulatorischen Anforderungen, die über den bestehenden Rahmen hinausgehen.

Auswirkungen der Hochrisikoklassifizierung auf Berufsbildungseinrichtungen:

- Obligatorische Risikobewertung und Protokolle zur Risikominimierung
- Strenge Dokumentationsanforderungen
- Mechanismen für die menschliche Aufsicht
- Transparenzverpflichtungen gegenüber den betroffenen Akteuren
- Regelmäßige Konformitätsbewertungen.

Ramirez-Montoya et al. (2023) betonen, dass die Einstufung von KI-Systemen im Bildungsbereich als risikoreich „rigorose Governance-Strukturen erfordert, die über die technische Einhaltung hinausgehen und umfassendere gesellschaftliche Anliegen berücksichtigen“ (S. 142). Ihre Analyse der Umsetzungs Herausforderungen in europäischen Berufsbildungseinrichtungen unterstreicht die Notwendigkeit einer kontextabhängigen Auslegung der rechtlichen Anforderungen.

3.3.2 Wichtige rechtliche Überlegungen für die Umsetzung der Berufsbildung

Anforderungen an Transparenz und Rechenschaftspflicht

Das EU-KI-Gesetz schreibt eine umfassende Transparenz beim Einsatz von KI im Bildungsbereich vor:

- Berufsbildungseinrichtungen müssen Lernende, Mitarbeitende und relevante Stakeholder informieren, wenn KI-Systeme zur Bewertung, für Lernempfehlungen oder für administrative Entscheidungen eingesetzt werden.
- Die Dokumentation muss den Zweck, die Funktionalität und die Grenzen der KI-Anwendungen darlegen.
- KI-gesteuerte Entscheidungen, die sich auf den Lernfortschritt oder die Zertifizierung auswirken, müssen erklärbar und begründbar sein.
- Algorithmische Prozesse müssen überprüfbar sein, um eine Validierung der Ergebnisse zu ermöglichen.



Prinsloo und Knox (2023) bezeichnen Transparenz als ein ethisches Grundprinzip bei der Implementierung von KI im Bildungsbereich und stellen fest, dass „die Undurchsichtigkeit von Algorithmen eine besondere Herausforderung in Lernkontexten darstellt, in denen das Vertrauen zwischen Lehrenden und Lernenden die Grundlage für eine effektive Pädagogik bildet“ (S. 76). Ihr Rahmen für die Transparenz von KI im Bildungsbereich geht über die technische Offenlegung hinaus und umfasst ein kontextuelles Verständnis, das für verschiedene Interessengruppen zugänglich ist.

Datenschutz und Einhaltung der Privatsphäre

Die Implementierung von KI im Berufsbildungskontext erfordert eine genaue Einhaltung der Datenschutzbestimmungen:

- Alle KI-Systeme müssen die Bestimmungen der Datenschutzgrundverordnung (DSGVO) einhalten.
- Es müssen klare Zustimmungsmechanismen für die Verarbeitung personenbezogener Daten festgelegt werden.
- Die Grundsätze der Datenminimierung müssen die Praktiken der Datenerhebung und -speicherung bestimmen.
- Die Zweckbindung stellt sicher, dass die Daten ausschließlich für bestimmte Bildungsziele verwendet werden.
- Technische und organisatorische Sicherheitsvorkehrungen müssen die Integrität und Vertraulichkeit der Lernendendaten schützen.

Holmes et al. (2023) stellen fest, dass Datenschutzerwägungen in Bildungskontexten über die Einhaltung von Rechtsvorschriften hinausgehen und „umfassendere ethische Fragen in Bezug auf die Handlungsfähigkeit von Lernenden, Machtdynamik und potenzielle pädagogische Überwachung“ (S. 12) einschließen. Ihre Arbeit unterstreicht die Bedeutung der Kontextualisierung von Datenschutzrahmen innerhalb bildungsspezifischer ethischer Überlegungen.

Bias-Prävention und Fairness-Garantie

Das EU-KI-Gesetz legt ausdrückliche Anforderungen fest, um algorithmische Voreingenommenheit in Bildungskontexten abzuschwächen:

- Berufsbildungseinrichtungen müssen regelmäßige Bewertungen von KI-Systemen durchführen, um Diskriminierung zu verhindern
- Die Bewertungsmethoden müssen potenziell unterschiedliche Auswirkungen auf verschiedene demografische Gruppen bewerten.



- Algorithmische Trainingsdaten müssen repräsentativ für die Vielfalt der Lernenden sein.
- Kontinuierliche Überwachungsprotokolle müssen aufkommende Verzerrungen in operativen Systemen identifizieren und beheben.
- Regelmäßige Überprüfungen der Ergebnisse für alle Lernendengruppen sind erforderlich, um eine gerechte Wirkung zu überprüfen.

Die Forschung von Hodgkin und Kumar (2023), die algorithmische Verzerrungen in Berufsberatungssystemen untersucht, identifiziert „systematische Methoden zur Bewertung ungleicher Auswirkungen auf unterschiedliche Schülerpopulationen“ (S. 217) als wesentlich für eine gerechte Implementierung von KI. Ihr Rahmen für eine umfassende Bewertung von Vorurteilen bietet praktische Anleitungen für Berufsbildungseinrichtungen, die Fairnessanforderungen erfüllen müssen.

KI im Klassenzimmer (begrenztes Risiko)

Zahlreiche KI-Anwendungen wie z. B.: KI-gestützte Nachhilfeassistenten, Plagiatserkennungssysteme, Empfehlungsmaschinen für Lernressourcen, Automatisierungswerkzeuge für die Verwaltung und virtuelle Laborsimulationen in der beruflichen Bildung können als eingeschränktes Risiko eingestuft werden:

Diese Anwendungen müssen den Nutzenden klare Informationen liefern:

- Die KI-Natur des Systems
- Beschränkungen der algorithmischen Fähigkeiten
- Bestimmungen über die menschliche Aufsicht
- Mechanismen zur Anfechtung der Ergebnisse.

3.3.3 Zeitplan für die Implementierung und Planung der Einhaltung von Vorschriften

Berufsbildungseinrichtungen müssen strategische Umsetzungspläne entwickeln, die mit dem Zeitplan für die schrittweise Umsetzung des EU-KI-Gesetzes übereinstimmen:

2. Februar 2025: Verbot von KI-Systemen, die unannehmbare Risiken darstellen

2. August 2025: Anwendungen der Verhaltenskodizes und Benennung der nationalen Regulierungsbehörden



2. August 2026: Allgemeine Durchsetzung des KI-Gesetzes, einschließlich Vorschriften für KI-Systeme mit hohem Risiko

2. August 2027: Vollständige Durchsetzung der spezifischen Verpflichtungen für KI-Systeme mit hohem Risiko.

Eine umsichtige Planung der Einhaltung der Vorschriften erfordert:

- Umfassende Bestandsaufnahme der bestehenden und geplanten KI-Anwendungen
- Bewertung der Risikoklassifizierung des Portfolios an Bildungstechnologien
- Lückenanalyse im Vergleich zu den regulatorischen Anforderungen
- Strukturierter Implementierungsfahrplan mit Etappenzielen
- Ressourcenzuweisung für die Dokumentation und Überwachung der Einhaltung der Vorschriften.

3.3.4 Rechtliche Verantwortlichkeiten und Haftungsüberlegungen

Berufsbildungseinrichtungen, die KI-Systeme einführen, übernehmen erhebliche rechtliche Verantwortung:

Rechenschaftsstrukturen müssen klare Verantwortlichkeiten für die Steuerung von KI-Systemen festlegen

Haftungsrahmen müssen sich mit potenziellen Schäden befassen, die durch algorithmische Fehler oder Verzerrungen entstehen

Vertragliche Vereinbarungen mit Technologieanbietern müssen die Verantwortungszuweisung ausdrücklich regeln

Entschädigungsprotokolle sollten sich mit potenziellen Ansprüchen aufgrund von KI-beeinflussten Entscheidungen befassen

Versicherungsüberlegungen sollten neue Haftungsrisiken widerspiegeln.

Bildungseinrichtungen müssen sich darüber im Klaren sein, dass der Einsatz von KI-Systemen, auch wenn sie von Dritten beschafft wurden, sie nicht von der rechtlichen Verantwortung für die Ergebnisse entbindet, die die Lernenden betreffen.



Rodriguez-Abitia und Bribiesca-Correa (2021) identifizieren Haftungsüberlegungen als eine oft übersehene Dimension der KI-Governance im Bildungsbereich und stellen fest, dass „Bildungseinrichtungen einzigartigen rechtlichen Risiken ausgesetzt sind, wenn sie traditionell menschliche Funktionen an algorithmische Systeme delegieren“ (S. 94). Ihre Analyse der aufkommenden Rechtsprechung liefert wertvolle Präzedenzfälle für Berufsbildungseinrichtungen, die umfassende Haftungsrahmen entwickeln.

3.3.5 Bewertungsliste für vertrauenswürdige KI (ALTAI) bei der Implementierung in die Berufsbildung

Die Bewertungsliste für vertrauenswürdige künstliche Intelligenz (ALTAI) bietet Berufsbildungseinrichtungen einen strukturierten Rahmen für die ethische Umsetzung und die Einhaltung von Vorschriften. Dieses praktische Instrument operationalisiert die Ethik-Leitlinien für vertrauenswürdige KI, die von der hochrangigen EU-Sachverständigengruppe für künstliche Intelligenz entwickelt wurden.

In ihrer vergleichenden Analyse ethischer Bewertungsrahmen identifizieren Zawacki-Richter et al. (2023) ALTAI als besonders wertvoll in Bildungskontexten aufgrund seiner „umfassenden Abdeckung sowohl technischer als auch sozio-ethischer Dimensionen, die für Lernumgebungen relevant sind“ (S. 14). Ihre Forschung zeigt die Anpassungsfähigkeit von ALTAI an verschiedene Bildungskontexte, einschließlich der Berufsausbildung.

Schlüsselkomponenten von ALTAI, die für die Umsetzung der Berufsbildung relevant sind

ALTAI strukturiert die Bewertung anhand von sieben grundlegenden Dimensionen, die besonders im Kontext der beruflichen Bildung von Bedeutung sind:

Menschliches Handeln und Aufsichtspflicht

- Berufsbildungsanwendungen müssen die Autonomie der Auszubildenden unterstützen und dürfen nicht an die Stelle des professionellen Urteils treten.
- Die Autonomie der Lernenden muss durch informierte Zustimmung und transparente Entscheidungen gewahrt werden.
- Bei Entscheidungen, die sich auf den Lernfortschritt der Lernenden auswirken, muss eine sinnvolle menschliche Kontrolle gewährleistet sein.



Technische Robustheit und Sicherheit

- KI-Systeme müssen in verschiedenen Berufsbildungsumgebungen zuverlässig funktionieren
- Sicherheitsprotokolle müssen vor möglicher Manipulation oder unberechtigtem Zugriff schützen
- Fallback-Verfahren müssen die Kontinuität der Ausbildung bei Systemausfällen gewährleisten.

Privatsphäre und Datenverwaltung

- Die im KI-Training und Betrieb genutzten Lernendendaten müssen durch umfassende Governance geschützt werden
- Grundsätze der Datenminimierung müssen die Erfassungs- und Aufbewahrungspraktiken leiten
- Robuste Sicherheitsmaßnahmen müssen sensible Informationen vor unbefugtem Zugriff schützen.

Transparenz

- KI-Fähigkeiten und -Grenzen müssen den Akteuren im Bildungswesen klar kommuniziert werden
- Entscheidungsprozesse müssen in nicht-technischer Sprache erklärbar sein
- Die Dokumentation muss die Überprüfung der Systemleistung und -konformität ermöglichen.

Diversität, Anti-Diskriminierung und Fairness

- Berufsbildungsspezifische Implementierungen müssen den unterschiedlichen Hintergründen und Lernbedürfnissen der Lernenden Rechnung tragen
- Bewertungsmethoden müssen die Auswirkungen auf die Chancengleichheit in allen demografischen Dimensionen bewerten
- Strategien zur Abschwächung von Bias müssen die besonderen Aspekte der Vielfalt in der beruflichen Bildung berücksichtigen

Gesellschaftliches und ökologisches Wohlergehen

- Die allgemeinen sozialen Auswirkungen von KI in der Berufsbildung müssen systematisch bewertet werden.
- Die allgemeinen sozialen Auswirkungen von KI in der Berufsbildung müssen systematisch bewertet werden.



- Potenzielle Auswirkungen der Zertifizierung von KI-Fähigkeiten auf dem Arbeitsmarkt müssen berücksichtigt werden

Rechenschaftspflicht

- Klare Governance-Strukturen müssen die Verantwortung für KI-Systeme festlegen
- Audit-Mechanismen müssen die Überprüfung der Einhaltung von Vorschriften ermöglichen
- Abhilfemaßnahmen müssen sich mit schädlichen Ergebnissen oder unbeabsichtigten Folgen befassen.

Anwendung von ALTAI im Berufsbildungskontext

Berufsbildungseinrichtungen können ALTAI durch systematische Bewertungsprotokolle operationalisieren:

Die **Erstbewertung der Implementierung** bewertet die vorgeschlagenen KI-Anwendungen anhand von Vertrauenswürdigkeitskriterien

Überprüfung vor dem Einsatz zur Sicherstellung der Einhaltung der technischen und ethischen Anforderungen

Prüfung nach der Implementierung zur Validierung der Leistung bei unterschiedlichen Lernendengruppen

Regelmäßige Neubewertung identifiziert aufkommende ethische Bedenken, wenn sich die Systeme weiterentwickeln.

Durch die Integration von ALTAI in den Governance-Rahmen können Berufsbildungseinrichtungen ihre Sorgfaltspflicht bei der Einhaltung von Vorschriften nachweisen und gleichzeitig das Vertrauen der Stakeholder in die KI-Implementierung stärken.

3.4 Fallstudien und Beispiele für wirksame KI-Politiken in Berufsbildungsinstitutionen

3.4.1 Grundlegende europäische politische Rahmenwerke

Eine wirksame institutionelle Politik stützt sich auf die auf europäischer Ebene geschaffenen Rahmenbedingungen:



Hochrangige Expertengruppe für Künstliche Intelligenz (AI HLEG)

Die Hochrangige Expertengruppe für Künstliche Intelligenz (AI HLEG) der Europäischen Kommission hat wichtige Leitlinien entwickelt:

Ethische Leitlinien für vertrauenswürdige KI: Festlegung von sieben Anforderungen für die ethische Implementierung von KI

Politik- und Investitionsempfehlungen: Strategische Leitlinien für eine nachhaltige KI-Entwicklung

Assessment List for Trustworthy AI (Bewertungsliste für vertrauenswürdige KI): Bietet praktische Umsetzungsinstrumente für Organisationen.

Diese grundlegenden Rahmenwerke haben die konkrete institutionelle Politik der europäischen Berufsbildungsanbieter beeinflusst.

Wie Holmes et al. (2023) in ihrer Analyse von KI-Governance-Rahmenwerken feststellen, stellen die KI-HLEG-Leitlinien „eine entscheidende Brücke zwischen abstrakten ethischen Grundsätzen und konkreten Umsetzungspraktiken“ dar (S. 18). Ihre Forschung zeichnet den Einfluss dieser Leitlinien in verschiedenen Bildungskontexten nach und zeigt ihre Anpassungsfähigkeit an Berufsbildungsumgebungen.

3.4.2 Technische Hochschule Kopenhagen: Umfassender KI-Governance-Rahmen

Die Technische Hochschule Kopenhagen hat eine vielschichtige Governance-Struktur eingerichtet, die beispielhaft für die Umsetzung von KI in der Berufsbildung ist.

Politische Architektur und Governance-Struktur

Der Ansatz der Hochschule zeichnet sich durch folgende Merkmale aus:

- KI-Ethikkomitee mit mehreren Interessenvertretern, darunter Verwaltungsangestellte, Lehrkräfte, Studenten, Industrievertreter und technische Experten
- Gestaffelte Genehmigungsprotokolle auf der Grundlage einer algorithmischen Risikoklassifizierung



- Verpflichtende KI-Schulungen für alle Lehrkräfte und Verwaltungsmitarbeiter
- Formalisierte Rechte der Studierenden hinsichtlich der Transparenz von Algorithmen und der Anfechtbarkeit von Entscheidungen
- Strukturierte Bewertungsprotokolle, die sowohl pädagogische als auch ethische Dimensionen messen.

Dieser Ansatz steht im Einklang mit dem Stakeholder-Engagement-Rahmen von Prinsloo und Knox (2023), der die Bedeutung einer „strukturierten Vertretung über institutionelle Hierarchien und betroffene Gemeinschaften hinweg“ (S. 82) hervorhebt. Ihre Forschung zeigt, wie Multi-Stakeholder-Governance sowohl die ethische Strenge als auch die Wirksamkeit der Umsetzung verbessert.

Implementierungsmethodik

Die wichtigsten Elemente der Implementierung sind:

- Schrittweise Einführung, beginnend mit risikoarmen Verwaltungsanwendungen
- Kontrollierte pädagogische Pilotprojekte mit strengen Bewertungsprotokollen
- Strukturierte Dokumentation sowohl der technischen Leistung als auch der ethischen Auswirkungen
- Regelmäßige Feedback-Mechanismen für die Stakeholder, um iterative Verbesserungen zu ermöglichen
- Unabhängige Prüfung der algorithmischen Ergebnisse in allen demografischen Dimensionen.

Dokumentierte Ergebnisse

Die Hochschule hat erhebliche Vorteile durch die strukturierte Umsetzung festgestellt:

- 27%ige Reduzierung des Verwaltungsaufwands durch ethische KI-Automatisierung
- Verbesserte Personalisierung des Unterrichts ohne Beeinträchtigung der Unterrichtsqualität
- Verbesserte Transparenz in der Bewertungspraxis
- Erhöhte Zufriedenheit der Lernenden mit der technologischen Integration
- Erfolgreiche Einhaltung neuer regulatorischer Anforderungen



3.4.3 Technisches Institut von Bordeaux: Transparente AI-Bewertungsrichtlinien

Das Bordeaux-Institut hat Modellrichtlinien entwickelt, die sich mit einem der ethisch komplexesten Aspekte der KI im Bildungsbereich befassen: der Beurteilung und Bewertung von Lernenden.

Bewertungsspezifischer politischer Rahmen

Zu den Schlüsselementen der Politik gehören:

- Verbot der vollautomatischen Bewertung mit hohem Schwierigkeitsgrad
- Obligatorische menschliche Überprüfung alle KI-beeinflussten Bewertungsentscheidungen
- Transparente Dokumentation des algorithmischen Beitrags zu Bewertungen
- Strukturierte Berufungsmechanismen für die Anfechtung von KI-beeinflussten Ergebnissen
- Regelmäßige Bias-Audits zum Vergleich von Bewertungsergebnissen für verschiedene Lernendengruppen
- Vertretung der Lernenden in den Entscheidungsstrukturen für die Bewertung

Dieser Ansatz veranschaulicht, was Ramirez-Montoya et al. (2023) als „human-centred assessment governance“ bezeichnen, die sich durch „algorithmische Transparenz, Beteiligung von Interessengruppen und formalisierte Aufsichtsmechanismen“ auszeichnet (S. 149). Ihre Forschung zeigt, wie solche Rahmenwerke sowohl die Validität der Bewertung als auch das Vertrauen der Stakeholder verbessern.

Umsetzungsansatz

Die Umsetzungsstrategie legt den Schwerpunkt auf:

- Klare Abgrenzung zwischen angemessener und unangemessener Bewertungsautomatisierung
- Abgestufte Implementierung auf der Grundlage von Bewertungsanteilen und Konsequenzen
- Schulung der Lehrkräfte in ethischen Bewertungsmethoden
- Technische Infrastruktur, die die Erklärbarkeit der algorithmischen Empfehlungen gewährleistet



- Dokumentationsprotokolle, die eine Überprüfung der Bewertungsgerechtigkeit ermöglichen.

Messbare Auswirkungen

Das Institut hat erhebliche Vorteile dokumentiert:

- Verbesserte Konsistenz der Bewertung bei gleichzeitiger Wahrung der Autonomie der Lehrkräfte
- Verbessertes Vertrauen der Lernenden in die Bewertungsmethoden
- Geringerer Arbeitsaufwand für die Bewertung ohne Qualitätseinbußen
- Bessere Identifizierung von Lernenden, die zusätzliche Unterstützung benötigen
- Erfolgreiche Einhaltung der regulatorischen Anforderungen an die KI-Transparenz.

3.4.4 Berufsschulkonsortium München: Kollaborative Datenverwaltung

Acht Berufsbildungseinrichtungen in München haben einen kollaborativen Ansatz für die Verwaltung von KI-Daten entwickelt, der zeigt, wie die Bündelung von Ressourcen die ethische Umsetzung verbessern kann.

Gemeinsame Governance-Architektur

Zu den wichtigsten Strukturelementen gehören:

- Gemeinsamer ethischer Rahmen für die Datenerhebung und -nutzung
- Gemeinsame technische Infrastruktur für Datenmanagement und -struktur
- Gemeinsame Protokolle für Anbieterbewertung und Vertragsabschluss
- Gemeinsame Implementierung von förderierten Lernmethoden
- Gebündeltes Fachwissen für die Überprüfung der Einhaltung gesetzlicher Vorschriften.

Dieser kollaborative Ansatz entspricht dem, was Rodriguez-Abitia und Bribiesca-Correa (2021) als „ressourcenoptimierte Governance“ bezeichnen, die durch „gemeinsames Fachwissen, einheitliche Strategien und kollektive Umsetzungskapazitäten“ gekennzeichnet ist (S. 98). Ihre Forschung zeigt, wie solche Ansätze kleinere Institutionen in die Lage versetzen können, anspruchsvolle Governance-Rahmenwerke umzusetzen.



Umsetzungsmethodik

Der Ansatz des Konsortiums zeichnet sich durch folgende Merkmale aus:

- Standardisierte Datenschutzprotokolle zwischen den teilnehmenden Institutionen
- Zentralisiertes Fachwissen über Datenminimierung und Anonymisierungstechniken
- Gemeinsame Ressourcen für die technische Schwachstellenanalyse
- Gemeinsame Entwicklung von Analysemethoden zur Wahrung der Privatsphäre
- Gemeinsames Engagement der Stakeholder bei der Entwicklung von Governance

Nachgewiesene Vorteile

Dieser kollaborative Ansatz hat zu folgenden Ergebnissen geführt:

- Eine ausgereifere Datenverwaltung, als die einzelnen Institutionen unabhängig voneinander erreichen könnten
- Verbesserte Verhandlungspositionen gegenüber Technologieanbietenden
- Verbesserte Datensicherheitsprotokolle durch gemeinsames Fachwissen
- Kosteneffiziente Einhaltung von Vorschriften über mehrere Institutionen hinweg
- Umfassendere Abschwächung von Verzerrungen durch größere und vielfältigere Datensätze.

3.4.5 Technisches Institut Dublin: Gleichstellungsorientierte KI-Politikrahmen

Das Technische Institut in Dublin hat Strategien entwickelt, die speziell auf algorithmische Gerechtigkeit in der beruflichen Bildung abzielen.

Gleichstellungsorientierter politischer Rahmen

Zu den wichtigsten politischen Elementen gehören:

- Obligatorische Folgenabschätzungen für Gerechtigkeit vor dem Einsatz von KI
- Regelmäßige Überprüfung von Algorithmen nach demografischen Gesichtspunkten
- Priorisierung von erklärbaren KI-Methoden



- Forderung nach nicht-technologischen Alternativen zu KI-Systemen
- Formalisierte Einbeziehung von unterrepräsentierten Bevölkerungsgruppen in die Governance-Strukturen.

Dieser Rahmen veranschaulicht, was Hodgkin und Kumar (2023) als „proaktive Equity Governance“ bezeichnen, die durch „systematische Bewertung von Vorurteilen, integrative Vertretung von Interessengruppen und kontinuierliche Überwachungsprotokolle“ gekennzeichnet ist (S. 219). Ihre Forschung zeigt, wie eine auf Gleichberechtigung ausgerichtete Governance sowohl die ethische Umsetzung als auch die Bildungsergebnisse verbessert.

Implementierungsstrategie

Der Implementierungsansatz legt den Schwerpunkt auf:

- Systematisches Testen von Berufsberatungsalgorithmen auf Verzerrungen
- Demografisches Monitoring von KI-beeinflussten Bildungsergebnissen
- Entwicklung von Entschärfungsmethoden für Ausbildungsdaten
- Entwicklung von Lehrkräften zur gleichstellungsbewussten Integration von KI
- Regelmäßiges Stakeholder-Feedback aus unterschiedlichen Lernendengruppen

Dokumentierte Wirkung

Das Institut hat nachgewiesen:

- Identifizierung und Beseitigung potenzieller Vorurteile in Berufsberatungssystemen
- Verbesserte Bildungsergebnisse für historisch unterrepräsentierte Gruppen
- Verbessertes Vertrauen der Beteiligten in die technologische Integration
- Erfolgreiche Einhaltung von Antidiskriminierungsvorschriften
- Entwicklung von übertragbaren, auf Gleichberechtigung ausgerichteten Methoden



3.4.6 Anwendung von ALTAI in berufsbildungsspezifischen Kontexten

Die Bewertungsliste für vertrauenswürdige künstliche Intelligenz (ALTAI) wurde in zahlreichen Kontexten der beruflichen Bildung erfolgreich angewandt und hat ihren praktischen Nutzen für die ethische Umsetzung bewiesen.

Die Forschung von Zawacki-Richter et al. (2023) identifiziert die ALTAI-Implementierung in der beruflichen Bildung als besonders wertvoll aufgrund der „konkreten beruflichen Kontexte, die die berufliche Bildung charakterisieren, und der expliziten Karriereimplikationen von KI-beeinflussten Entscheidungen“ (S. 15). Ihre Fallstudienanalyse demonstriert die Anpassungsfähigkeit von ALTAI an verschiedene Berufsbildungsumgebungen.

AI-gesteuerte Anwendungen zur Bewertung von Fähigkeiten

Berufsbildungseinrichtungen haben ALTAI angewandt, um die ethische Implementierung von automatisierten Bewertungssystemen sicherzustellen:

Integration menschlicher Aufsicht: Sicherstellung der Validierung von algorithmischen Bewertungen durch Auszubildende

Überprüfung der Fairness: Prüfung der Bewertungsgerechtigkeit bei unterschiedlichen Lernendenpopulationen

Transparenzanforderungen: Bewertungskriterien explizit und verständlich machen

Technische Robustheit: Sicherstellung eines zuverlässigen Betriebs bei unterschiedlichen Leistungsmustern der Lernenden.

KI in der Berufsberatung und Berufsorientierung

ALTAI-Anwendung in Berufsempfehlungssystemen wurde angesprochen:

- Mögliche Verstärkung von Berufsstereotypen
- Transparenz in Empfehlungsmethoden
- Datenschutz bei der prädiktiven Karrieremodellierung
- Zugänglichkeit von Beratungssystemen für unterschiedliche Bevölkerungsgruppen.

KI-gestütztes adaptives Lernen in der Berufsausbildung



Die Umsetzung von adaptiven Lernplattformen hat von der ALTAI-Bewertung profitiert:

- Bewertung der algorithmischen Personalisierungsgerechtigkeit
- Überprüfung der Diversifizierung der Lernpfade
- Bewertung potenzieller Filterblaseneffekte
- Validierung der Aufsichtskompetenz der Auszubildenden.

3.4.5 Gemeinsame Elemente wirksamer Politiken

Die Analyse der erfolgreichen Umsetzung in den europäischen Berufsbildungseinrichtungen zeigt konsistente politische Muster:

Multi-Stakeholder-Governance-Strukturen, die verschiedene Perspektiven in die Entwicklung und Überwachung einbeziehen

Abgestufte Risikobewertungsprotokolle, die eine verhältnismäßige Prüfung auf der Grundlage der potenziellen Auswirkungen vorsehen

Kontinuierliche Überwachungsmechanismen anstelle von einmaligen Genehmigungsprozessen

Robuste Transparenzanforderungen, die sicherstellen, dass die Beteiligten die Fähigkeiten und Grenzen der KI verstehen

Sinnvolle menschliche Aufsicht, insbesondere bei folgenreichen Bildungsentscheidungen

Regelmäßige Überprüfung der Richtlinien zur Anpassung an sich entwickelnde Technologien und Vorschriften

Umfassende Bildungskomponenten zum Aufbau von KI-Kenntnissen bei allen Beteiligten

Klare Strukturen der Rechenschaftspflicht, die die Verantwortung für KI-beeinflusste Ergebnisse festlegen

Holmes et al. (2023) stellen fest, dass diese konsistenten Elemente einen „sich abzeichnenden Konsens hinsichtlich grundlegender Governance-Prinzipien für die Implementierung von KI im Bildungsbereich“ darstellen (S. 21). Ihre vergleichende Analyse über verschiedene institutionelle Kontexte hinweg zeigt



die Anpassungsfähigkeit dieser Kernprinzipien an unterschiedliche Bildungsumgebungen.

Einrichtungen mit vorbildlichen Praktiken haben ein Gleichgewicht zwischen Innovationspotenzial und angemessenen Sicherheitsvorkehrungen gefunden, da sie erkannt haben, dass zu restriktive Ansätze den pädagogischen Nutzen einschränken können, während eine unzureichende Governance erhebliche Risiken birgt. Die erfolgreichsten Rahmenwerke haben ethische Überlegungen in den gesamten Umsetzungszyklus integriert, anstatt sie als nachträgliche Überlegungen zu behandeln.

4. Ethische Fähigkeiten

4.1 Definition ethischer Fähigkeiten

Ethische Fähigkeiten beziehen sich auf die Fähigkeit zu beurteilen, ob eine KI in einem bestimmten Kontext ethisch korrekt eingesetzt wird. Im Bildungskontext geht es beispielsweise darum, ob die Richtlinien eines Bildungszentrums ethisch vertretbar sind, ob die Datensicherheit und der Datenschutz gewahrt werden usw.

Ethische Fähigkeiten beziehen sich auf die kritischen Fähigkeiten, die erforderlich sind, um die ethische Nutzung von KI und Daten im Bildungsbereich auf der Grundlage von Prinzipien wie menschliches Handeln, Gerechtigkeit, Menschlichkeit und gerechtfertigte Entscheidungen zu bewerten und sicherzustellen. Menschliches Handeln bedeutet, dass Individuen in die Lage versetzt werden, fundierte Entscheidungen zu treffen und die Verantwortung für ihr Handeln zu übernehmen, wodurch Autonomie und Verantwortlichkeit gefördert werden. Gleichberechtigung gewährleistet eine faire Behandlung und gleichen Zugang zu Chancen für alle und fördert Integration und Antidiskriminierung. Menschlichkeit betont die Achtung der Würde, der Identität und des Wohlbefindens des Menschen und gibt sinnvollen menschlichen Beziehungen Vorrang vor der instrumentellen Nutzung von Daten. Gerechtfertigte Entscheidungen erfordern transparente Entscheidungsprozesse, die auf ethischen Grundsätzen und der Beteiligung von Stakeholdern beruhen, um sicherzustellen, dass KI-Systeme im Bildungswesen verantwortungsbewusst eingesetzt werden und allen Beteiligten zugutekommen.

4.2 Bedeutung ethischer Fähigkeiten

Diese ethischen Überlegungen sind entscheidend, um Lehrende und Bildungseinrichtungen bei der Bewältigung der Komplexität der KI-Implementierung zu unterstützen. Durch die Einhaltung ethischer Anforderungen,



wie sie in den KI-Ethikrichtlinien vorgeschlagen werden, können Bildungsakteure Risiken wirksam mindern, Fairness fördern und das Wohlergehen von Lernenden und Mitarbeitenden gleichermaßen aufrechterhalten.

4.3 Bereiche und Beispiele für ethische Fähigkeiten

Bereich 1: Berufliches Engagement: Nutzung digitaler Technologien für Kommunikation, Zusammenarbeit und berufliche Entwicklung	
Kompetenz-Element	Potenzielle Indikatoren
Ist in der Lage, positive und negative Auswirkungen von KI und Datennutzung im Bildungsbereich kritisch zu beschreiben	Nimmt aktiv an der kontinuierlichen beruflichen Weiterbildung zu KI und Lernanalysen und deren ethischer Nutzung teil.
	Ist in der Lage, Beispiele für KI-Systeme zu nennen und ihre Relevanz zu beschreiben.
	Weiß, wie die ethischen Auswirkungen von KI-Systemen in der Schule bewertet werden.
	Weiß, wie man Strategien in der Schule und in der weiteren Gemeinschaft initiiert und fördert, die eine ethische und verantwortungsvolle Nutzung von KI und Daten fördern.
Verstehen der Grundlagen der KI und der lernenden Analytik	Bewusstsein dafür, dass KI-Algorithmen auf eine Art und Weise arbeiten, die für die Nutzer normalerweise nicht sichtbar oder leicht verständlich ist.
	Kann mit dem KI-System interagieren und ihm Feedback geben, um seine nächsten Empfehlungen zu beeinflussen.
	weiß, dass Sensoren, die in vielen digitalen Technologien und Anwendungen eingesetzt werden, große Mengen an Daten, einschließlich



	personenbezogener Daten, erzeugen, die zum Trainieren eines KI-Systems verwendet werden können.
	kennt die EU-Ethikrichtlinien für KI und die Selbstbewertungsinstrumente.
Beispiel 1: Ana ist eine Lehrerin, die eine Lektion über Strömungsmechanik unterrichten möchte. Sie möchte die KI nutzen, um einige Bilder zu generieren und sie in die Powerpoint-Präsentation zum Thema einzubinden. Bevor sie die KI in den Unterricht einbaut, hat Ana die EU-Ethikrichtlinien für KI gelesen und bewertet, ob der Einsatz der KI ethisch vertretbar ist oder nicht.	
Beispiel 2: Lucas unterrichtet in der Berufsschule für Umwelterziehung und -kontrolle. Mehrere Lehrende seines Bildungszentrums verwenden KI, daher hat er ein Treffen mit dem gesamten Bildungszentrum vorgeschlagen, um die Verwendung von KI durch die Lehrenden zu analysieren. Lucas' Ziel ist es, sicherzustellen, dass alle Lehrende die KI und die Daten ethisch und verantwortungsvoll nutzen.	

Bereich 2: Digitale Ressourcen: Beschaffung, Erstellung und gemeinsame Nutzung digitaler Ressourcen	
Kompetenz-Element	Potenzielle Indikatoren
Daten-Governance	Kennt die verschiedenen Formen von personenbezogenen Daten, die in der allgemeinen und beruflichen Bildung verwendet werden.
	Ist sich der Verantwortung für die Wahrung der Datensicherheit und des Datenschutzes bewusst.



	weiß, dass die Verarbeitung personenbezogener Daten nationalen und EU-Vorschriften unterliegt, einschließlich der allgemeinen Datenschutzverordnung (DSGVO).
	weiß, wer Zugriff auf Lernendendaten hat, wie der Zugriff überwacht wird und wie lange die Daten gespeichert werden.
	weiß, dass alle EU-Bürger:innen das Recht haben, keiner vollautomatischen Entscheidungsfindung unterworfen zu werden.
	Kann Beispiele für sensible Daten, einschließlich biometrischer Daten, nennen.
	Kann die Vorteile und Risiken abwägen, bevor er Dritten die Verarbeitung personenbezogener Daten erlaubt, insbesondere wenn er KI-Systeme einsetzt.
KI-Governance	Weiß, dass KI-Systeme nationalen und EU-Vorschriften unterliegen (insbesondere dem noch zu verabschiedenden KI-Gesetz).
	Kann den risikobasierten Ansatz des (noch zu verabschiedenden) KI-Gesetzes erläutern.
	Kennt die risikoreichen KI-Anwendungsfälle im Bildungsbereich und die damit verbundenen Anforderungen des KI-Gesetzes (wenn es verabschiedet ist).
	Weiß, wie man mit KI bearbeitete/manipulierte digitale Inhalte in die eigene Arbeit einbezieht und wie diese Arbeit gutgeschrieben werden sollte
	Kann die wichtigsten Grundsätze der Datenqualität in KI-Systemen erklären.



Beispiel 1: Maria ist Lehrerin an der Berufsfachschule für Energieeffizienz und Solarthermie. Sie erstellt Notizen für ihren Kurs, die sie an die Lernenden weitergibt, und hat mit Hilfe des Copilot-Tools einige von KI bearbeitete digitale Inhalte in die Notizen integriert. Maria weiß, dass dieser Inhalt mit einem Vermerk versehen werden muss, und sie stellt sicher, dass dieser Vermerk korrekt ist.

Beispiel 2: Jaime unterrichtet an der Berufsschule für mechanische Fertigung. Er verwendet ChatGPT manchmal, um einige Inhalte oder Aktivitäten seines Faches zu entwickeln. Irgendwann prüft er die Möglichkeit, die persönlichen Daten seiner Lernenden (Name, Nachname, Geburtsdatum und Adresse) in ein Dokument aufzunehmen, das all diese Informationen enthält. Da er sich seiner Verantwortung für die Wahrung der Datensicherheit und des Datenschutzes bewusst ist, entschied er sich, die persönlichen Daten nicht aufzunehmen.

Bereich 3: Einsatz digitaler Technologien und Strategien zur Verbesserung der Bewertung

Kompetenz-Element	Potenzielle Indikatoren
Lernmodelle	weiß, dass KI-Systeme die Vorstellungen der Designer darüber umsetzen, was Lernen ist und wie Lernen gemessen werden kann; kann die wichtigsten pädagogischen Annahmen erklären, die einem bestimmten digitalen Lernsystem zugrunde liegen.
Bildungsziele	weiß, wie ein bestimmtes digitales System die verschiedenen sozialen Ziele der Bildung (Qualifikation, Sozialisierung, Subjektivierung) anspricht.
Menschliches Handeln	Ist in der Lage, die Auswirkungen von KI-Systemen auf die Autonomie der Lehrkräfte, die berufliche Entwicklung und die Bildungsinnovation zu berücksichtigen.



	Berücksichtigt die Quellen inakzeptabler Verzerrungen bei datengesteuerter KI.
Fairness	Berücksichtigt die Risiken im Zusammenhang mit der emotionalen Abhängigkeit und dem Selbstverständnis der Lernenden beim Einsatz von interaktiven KI-Systemen und Lernanalyseverfahren.
Menschlichkeit	Ist in der Lage, die Auswirkungen von KI und Datennutzung auf die Lernendengemeinschaft zu berücksichtigen.
	Kann die ethischen Aspekte von KI und deren Einfluss auf den Einsatz von Technologie diskutieren.
Beteiligt sich an der Entwicklung von Lernpraktiken, die KI und Daten nutzen.	Kann erklären, wie ethische Grundsätze und Werte bei der Mitgestaltung von Lernpraktiken, die KI und Daten nutzen, berücksichtigt und verhandelt werden (in Verbindung mit Lerndesign).
Beispiel 1: Luisa, eine Ausbilderin für elektromechanische Instandhaltung in einem Ausbildungsunternehmen, setzt eine KI-Software ein, die adaptive Lernmodelle zur Personalisierung des Unterrichts nutzt. Die Software bewertet das Ausgangsniveau jedes Auszubildenden und passt den Inhalt und die Aufgaben auf der Grundlage seiner Fortschritte an. So kann jeder Auszubildende in seinem eigenen Tempo vorankommen und erhält sofortiges Feedback und spezifische Ressourcen, um Schwierigkeiten zu überwinden. Außerdem kann Luisa auf detaillierte Berichte zugreifen, um Bereiche zu ermitteln, in denen ihre Lernenden zusätzliche Unterstützung benötigen, und so einen integrativen und effektiven Unterricht gewährleisten.	
Beispiel 2: Javi, ein Lehrer für grafische Künste im Mittelstufenprogramm für Grafikdruck an einer berufsbildenden Schule, integriert eine KI-Software, die die menschlichen Fähigkeiten durch Unterstützung bei komplexen Designaufgaben verbessert. Die Software hilft den Lernenden durch die Automatisierung sich wiederholender Aufgaben wie Farbkorrekturen und Layoutanpassungen, so dass sie sich auf die kreativen und strategischen Aspekte ihrer Projekte konzentrieren können. So können die Lernenden ihre	



künstlerischen Fähigkeiten und ihr kritisches Denken weiterentwickeln, während sich die KI um technische Details kümmert.

Bereich 4: Bewertung: Verwaltung und Steuerung des Einsatzes digitaler Technologien beim Lehren und Lernen

Kompetenz-Elemente	Potenzielle Indikatoren
Persönliche Unterschiede	ist sich bewusst, dass Lernende unterschiedlich auf automatisches Feedback reagieren.
Algorithmische Bias	Berücksichtigt die Quellen inakzeptabler Voreingenommenheit in KI-Systemen und wie sie gemildert werden können.
Kognitiver Fokus	Ist sich bewusst, dass KI-Systeme den Fortschritt der Lernenden auf der Grundlage vordefinierter domänenspezifischer Wissensmodelle bewerten.
	ist sich bewusst, dass die meisten KI-Systeme keine Zusammenarbeit, sozialen Kompetenzen oder Kreativität bewerten.
	Kennt die gängigen Methoden zur Manipulation von KI-basierten Beurteilungen..
Neue Möglichkeiten des Missbrauchs von Technologie	ist sich bewusst, dass sich Lehrkräfte möglicherweise zu sehr auf KI-generiertes Feedback verlassen, das manchmal unpersönlich wirkt und/oder nicht auf spezifische Bedürfnisse oder Missverständnisse der Schüler eingeht.

Beispiel 1: Jorge unterrichtet ein Fach in der Berufsschule für Installation und Wartung. Nachdem er das erste Thema des Fachs abgeschlossen hat, verwendet er ein KI-Tool, um die Aneignung der Inhalte durch seine Lernende



zu bewerten. Außerdem gibt dieses KI-Tool automatisch Feedback und passt das Niveau der Fragen auf der Grundlage früherer Antworten an. Jorge ist sich bewusst, wie die Lernenden auf automatisches Feedback reagieren, daher achtet er darauf, ihnen ein individuelles Feedback zu geben und die Fortschritte der Lernenden selbst zu überprüfen. Daher verwendet er die KI nur als Hilfsmittel für eine erste Bewertung, die anschließend von ihm überprüft und überarbeitet wird.

Beispiel 2: Natalia ist Ausbilderin bei einem Berufsbildungsdienstleister für erneuerbare Energien. Sie ermutigt die Lernenden, Windows Copilot in ihrem Unterricht für bestimmte Aktivitäten zu nutzen. Auch wenn dieses Tool als KI-Assistent funktioniert, der bei der Produktivität und Kreativität hilfreich sein kann, ist sie sich bewusst, dass andere Arten von Aktivitäten und Bewertungen erforderlich sind, um die Zusammenarbeit, die sozialen Kompetenzen und die Kreativität zu verbessern, da die KI nicht in der Lage ist, diese Aspekte zu bewerten. Sie übernimmt die entsprechenden Aktivitäten und Beurteilungen als Ausbilderin weiterhin selbst und nutzt KI als Hilfsmittel während ihres Unterrichts und für die Trainingsvorbereitung.

Bereich 5: Befähigung der Lernenden: Einsatz digitaler Technologien zur Förderung der Integration, Personalisierung und des aktiven Engagements der Lernenden

Kompetenz-Element	Potenzielle Indikatoren
KI geht auf die unterschiedlichen Lernbedürfnisse der Lernenden ein	Kennt die verschiedenen Möglichkeiten, wie personalisierte Lernsysteme ihr Verhalten anpassen können (Inhalt, Lernpfad, pädagogischer Ansatz).
	Kann erklären, wie ein bestimmtes System allen Lernenden zugutekommen kann, unabhängig von ihren kognitiven, kulturellen, wirtschaftlichen oder physischen Unterschieden



	ist sich bewusst, dass digitale Lernsysteme verschiedene Lernendengruppen unterschiedlich behandeln Kann Auswirkungen auf die Entwicklung der Selbstwirksamkeit, des Selbstbildes, der Denkweise sowie der kognitiven und affektiven Selbstregulierungsfähigkeiten der Lernenden berücksichtigen.
Gerechtfertigte Wahl	Weiß, dass KI und die Nutzung von Daten für einige Lernende vorteilhafter sein können als für andere.
	Ist in der Lage zu erklären, welche Beweise verwendet wurden, um den Einsatz eines bestimmten KI-Systems im Klassenzimmer zu rechtfertigen.
	erkennt die Notwendigkeit einer ständigen Überwachung der Ergebnisse des KI-Einsatzes.
Beispiel 1: Teresa, die bei einem Anbieter von Erwachsenenbildung eine fortgeschrittene Berufsausbildung in Verwaltung und Finanzen unterrichtet, nutzt KI, um Lernerfahrungen maßgeschneidert zu gestalten. Das KI-System bewertet die Bedürfnisse jedes Lernenden und passt die Ressourcen und Aktivitäten entsprechend an. So können beispielsweise visuelle Lernende interaktive Tools erhalten, während analytische Lernende detaillierte Fallstudien bekommen. Dieser Ansatz gewährleistet ein personalisiertes Lernen, das den unterschiedlichen Lernstilen gerecht wird und das Engagement der Lernenden maximiert.	
Beispiel 2: Óscar, der eine fortgeschrittene Berufsausbildung in mechanischem Fertigungsdesign unterrichtet, setzt KI-Software für die Konstruktion mechanischer Teile ein. Die Software bietet mehrere Konstruktionslösungen und simuliert die Leistung unter verschiedenen Bedingungen. Die Lernenden analysieren die Ergebnisse, um das effizienteste Design auszuwählen. Dabei werden sie von Óscar angeleitet, kritische und	



analytische Fähigkeiten zu entwickeln, die für ihre berufliche Laufbahn entscheidend sind.

Bereich 6: Förderung der digitalen Kompetenz der Lernenden, um sie in die Lage zu versetzen, digitale Technologien kreativ und verantwortungsbewusst für Information, Kommunikation, Erstellung von Inhalten, Wohlbefinden und Problemlösung zu nutzen

Kompetenz-Element	Potenzielle Indikatoren
KI und Lernanalyse-Ethik	Kann anhand von KI-Projekten und -Einsätzen Studenten helfen, die Ethik der KI und der Datennutzung in Bildung und Ausbildung zu verstehen.

Beispiel 1: Professorin Sara unterrichtet ihre Berufsschullehramtstudierenden an einer Universität in fortgeschrittener Marketing-Analytik. Sie erörtert eine Fallstudie, in der KI Daten aus sozialen Medien für gezielte Werbung analysiert. Sara regt eine Debatte über ethische Bedenken wie Datenschutz, Zustimmung und algorithmische Verzerrungen an. Sie leitet die Studierenden an, ethische Richtlinien für den Einsatz von KI im Marketing zu entwickeln und fördert so ihr Verständnis für einen verantwortungsvollen Einsatz von KI in der Branche.

Beispiel 2: Professor Jesús lehrt kulinarisches Management und konzentriert sich dabei auf die Rolle der KI bei der Personalisierung von Restaurantmenüs. Er erörtert eine Fallstudie, in der KI Ernährungspräferenzen analysiert, um personalisierte Menüempfehlungen anzubieten. Jesús verwickelt die Auszubildenden in eine Debatte über ethische Fragen wie den Datenschutz, die Gewährleistung der Menüvielfalt und die Auswirkungen der KI auf die Kundenzufriedenheit. Er leitet die Teilnehmer bei der Entwicklung eines ethischen Rahmens für den Einsatz von KI bei der Menüanpassung an und bereitet sie darauf vor, Innovation und ethische Überlegungen im Restaurantbetrieb in Einklang zu bringen.



5. Schlussfolgerung

Da KI weiterhin die Bildungslandschaft prägt, ist es wichtig, die Einführung von KI-Technologien so zu gestalten, dass ethische Werte gewahrt werden, Fairness gefördert wird und die Privatsphäre der Lernenden geschützt wird. In diesem Handbuch werden Zweck und Umfang der KI-Implementierung umrissen und die Bedeutung ethischer Überlegungen hervorgehoben, um sicherzustellen, dass KI-Technologien in allen Bildungskontexten verantwortungsvoll eingesetzt werden.

5.1 Wichtigste Erkenntnisse

Die Integration von KI in die Bildung birgt ein immenses Potenzial, muss aber von ethischen Grundsätzen geleitet werden, die Vielfalt, Inklusion und Transparenz in den Vordergrund stellen. KI-Tools sollten so konzipiert sein, dass sie Lernenden mit unterschiedlichem Hintergrund entgegenkommen und kulturelle, sprachliche und sozioökonomische Unterschiede berücksichtigen, um die Chancengleichheit bei den Bildungsergebnissen zu fördern. Transparenz ist ebenso wichtig - Lehrende und Lernende müssen verstehen, wie KI-Systeme Entscheidungen treffen, insbesondere in Bereichen wie Benotung, Feedback oder personalisierte Lernpfade. Die Gewährleistung von Privatsphäre und Datenschutz ist von entscheidender Bedeutung und erfordert die vollständige Einhaltung von Vorschriften wie der Datenschutz-Grundverordnung (GDPR) und die Einhaltung ethischer Datenverwaltungspraktiken. KI-Systeme müssen auch nachhaltig sein, ihre Umweltauswirkungen minimieren und ihre langfristige Lebensfähigkeit in Bildungsökosystemen sicherstellen.

Ebenso wichtig ist die Förderung von KI-Kenntnissen bei allen Beteiligten. Lehrende müssen gezielt geschult werden, um KI-Tools ethisch korrekt und effektiv in ihre Pädagogik zu integrieren, während Lernende mit einem kritischen Verständnis für die Möglichkeiten und Grenzen von KI ausgestattet werden sollten. Die institutionelle Verwaltung spielt eine Schlüsselrolle, wenn es darum geht, die Schul- oder Universitätspolitik mit ethischen KI-Praktiken in Einklang zu bringen und ein Umfeld der Verantwortlichkeit und Fairness zu fördern. Politische Entscheidungsträger:innen müssen eingreifen, um robuste rechtliche Rahmenbedingungen zu schaffen, die einen verantwortungsvollen Einsatz von KI ermöglichen, algorithmische Verzerrungen abmildern und diskriminierende Ergebnisse verhindern. Letztlich ist ein Multi-Stakeholder-Ansatz entscheidend, um sicherzustellen, dass KI die Bildung auf faire, inklusive und nachhaltige Weise verbessert.



Es ist auch sehr wichtig, einige wichtige rechtliche und regulatorische Rahmenbedingungen zu berücksichtigen, wie z. B. das EU-KI-Gesetz (2024), das einen umfassenden Rechtsrahmen einführt, der KI-Systeme auf der Grundlage ihres Risikos für die Grundrechte kategorisiert, was erhebliche Auswirkungen auf Bildungseinrichtungen hat.

Darauf aufbauend gibt es KI-Anwendungen mit hohem Risiko, wie z. B. solche, die bei der Beurteilung von Lernenden oder bei der Zulassung eingesetzt werden, für die strenge Compliance-Anforderungen gelten, einschließlich strenger Dokumentation, menschlicher Aufsicht und solider Risikomanagement-Protokolle.

Außerdem sind Transparenz und Rechenschaftspflicht keine optionale Voraussetzung. Die Institute müssen klar offenlegen, wie KI-Systeme funktionieren, und sicherstellen, dass die Nutzenden automatisierte Entscheidungen verstehen und anfechten können. Darüber hinaus schreibt das Gesetz proaktive Maßnahmen zur Prüfung und Milderung algorithmischer Verzerrungen vor, um Fairness zu gewährleisten und diskriminierende Ergebnisse zu verhindern. Zur Unterstützung dieser Bemühungen bietet die Bewertungsliste für vertrauenswürdige KI (Assessment List for Trustworthy AI, ALTAI) ein praktisches Instrument, das Institutionen hilft, die ethischen Auswirkungen ihrer KI-Systeme zu bewerten und eine Kultur des verantwortungsvollen und vertrauenswürdigen KI-Einsatzes im Bildungswesen zu fördern.

5.2 Abschließende Reflexionen

Wir alle sollten uns der Tatsache bewusst sein, dass KI die Hochschul- und Berufsbildung immer weiter umgestaltet. Dieses Handbuch betont, wie wichtig es ist, ein sorgfältiges Gleichgewicht zwischen innovativen Praktiken und ethischer Verantwortung zu wahren. Die Vorteile der KI, angefangen von personalisierten Lernwegen bis hin zu einer verbesserten Verwaltungseffizienz, sind beträchtlich, doch sie bergen auch Risiken, die nicht übersehen oder ignoriert werden dürfen. Themen wie algorithmische Verzerrungen, Datenschutzverletzungen und die schwindende Rolle des menschlichen Urteilsvermögens erfordern ständige Aufmerksamkeit. Um diese komplexen Zusammenhänge zu bewältigen, ist eine ständige Zusammenarbeit zwischen Lehrenden, Administrator:innen und politischen Entscheidungsträger:innen erforderlich. Natürlich müssen ethische Grundsätze in die Entwicklung, den Einsatz und die Bewertung von KI-Tools einfließen, um sicherzustellen, dass sie die Inklusion fördern, die Autonomie respektieren und das Vertrauen in den Bildungsprozess stärken.



Mit Blick auf die Zukunft wird die erfolgreiche Integration von KI in die Bildung nicht nur von der technologischen Bereitschaft abhängen, sondern auch von einer gemeinsamen Vision, die auf Transparenz, Fairness und lebenslangem Lernen beruht. Der Sektor muss die KI-Kompetenz aller Beteiligten fördern und robuste Systeme für die Rechenschaftspflicht und Aufsicht schaffen. Dabei handelt es sich nicht um eine einmalige Anpassung, sondern um eine sich entwickelnde Reise, die kontinuierliche Reflexion und Dialog erfordert. Indem sie sich für menschenzentrierte Ansätze und ethische Innovationen einsetzt, kann die Bildungsgemeinschaft KI in einen mächtigen Verbündeten verwandeln, der die Lernmöglichkeiten verbessert und gleichzeitig die Werte für eine hochwertige Bildung bewahrt. Umfassende Orientierungshilfen finden die Beteiligten in den Ressourcen des AI Pioneers-Projekts und im Evaluierungsschema für KI im Bildungswesen, die in mehreren Sprachen verfügbar sind.

Danksagung

Wir möchten all jenen danken, die die Entwicklung dieses Dokuments durch ihre konstruktive Kritik und unschätzbaren Beiträge unterstützt haben. Wir sind allen unseren Projektpartnern, die zum Inhalt beigetragen haben, sowie dem Referenznetzwerk der KI-Pioniere und den Teilnehmenden an unseren Konsultationsseminaren zu großem Dank verpflichtet. Ihre gemeinsamen Anliegen und Meinungen haben uns während des gesamten Prozesses wertvolle Hinweise gegeben.

Schließlich würdigen wir die Bedeutung der referenzierten Literatur sowie verschiedener anderer Ressourcen wie wissenschaftliche Abhandlungen, Berichte, Seminare, Richtlinien von Bildungseinrichtungen und Blogbeiträge von Forschenden und Lehrenden. Insgesamt haben diese Beiträge unseren Ansatz, ethische Erwägungen beim Einsatz von KI in der Bildung auf allen Ebenen anzusprechen, stark beeinflusst.

Referenzen

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.

Aldosari, S. A. M. (2020). The future of higher education in the light of artificial intelligence transformations. *International Journal of Higher Education*, 9(3), 145-151.



Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, 41(2), 101929.

Attwell, G., Bekiaridis, G., Deitmer, L., Perini, M., Roppertz, S., & Tütlys, V. (2020). Artificial intelligence in policies, processes and practices of vocational education and training.

Attwell, G., Deitmer, L., Tütlys, V., Roppertz, S., & Perini, M. (2020). Digitalisation, artificial intelligence and vocational occupations and skills: what are the needs for training teachers and trainers?. In *Trends in vocational education and training research, Vol. III. Proceedings of the European Conference on Educational Research (ECER), Vocational Education and Training Network (VETNET)* (pp. 30-42).

Bekiaridis, G., & Attwell, G. (2024). Integrating Artificial Intelligence in Vocational and Adult Education: A Supplement to the DigCompEdu Framework. *Ubiquity Proceedings*, 4(1).

Boninger, F., Molnar, A., & Saldaña, C. (2020). Big Claims, Little Evidence, Lots of Money: The Reality behind the Summit Learning Program and the Push to Adopt Digital Personalized Learning Programs. *Commercialism in Education Research Unit*.

Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational inequality, techno-solutionism and inclusive tools. *Sustainability*, 16(781). <https://doi.org/10.3390/su16020781>

Celik, I. (2023). Exploring the determinants of artificial intelligence (Ai) literacy: Digital divide, computational thinking, cognitive absorption. *Telematics and Informatics*, 83, 102026.

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://link.springer.com/article/10.1186/s41239-023-00408-3>

Chew, E., & Chua, X. N. (2020). Robotic Chinese language tutor: Personalising progress assessment and feedback or taking over your job? *On the Horizon*, 28(3), 113–124. <https://doi.org/10.1108/OTH-04-2020-0015>

Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). First-year students AI-competence as a predictor for intended and de facto use of AI-tools



for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 18.

European Commission, Directorate-General for Education, Youth, Sport, and Culture. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/153756>

European Parliament. (2024). EU AI Act: First regulation on artificial intelligence. Retrieved from <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. Official Journal of the European Union, L series, 1–144.

Emeršič, Ž., Hrastnik, G., Meh Peer, N., & Peer, P. (2025). AIM@VET-inspired university level education strategies for teaching computer vision and biometrics. *ROSUS 2025 – Računalniška obdelava slik in njena uporaba v Sloveniji 2025: Zbornik 19. strokovne konference*. <https://doi.org/10.18690/um.ferj.2.2025.4>

Hodgkin, E., & Kumar, A. (2023). Algorithmic fairness in vocational guidance systems: A comparative analysis of implementation frameworks. *Educational Technology Research and Development*, 71(3), 215-231.

Holmes, W., Bialik, M., & Fadel, C. (2023). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign

JISC (2022). *AI in tertiary Education. A summary of the current state of play*. JISC Repository. Retrieved from <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020). Automated personalized feedback improves learning gains in an intelligent tutoring system. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II* 21 (pp. 140-146). Springer International Publishing.



Kumar, V., & Boulanger, D. (2020, October). Explainable automated essay scoring: Deep learning really has pedagogical value. In *Frontiers in education* (Vol. 5, p. 572367). Frontiers Media SA.

Li, Q. (2021). The use of artificial intelligence combined with cloud computing in the design of education information management platform. *International Journal of Emerging Technologies in Learning*, 16(5), 32–44. <https://doi.org/10.3991/ijet.v16i05.20309>

Porter, B., & Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability*, 12(24), 10243.

Prinsloo, P., & Knox, J. (2023). Exploring the ethical principles for the implementation of artificial intelligence in education: Towards a future agenda. *Research and Practice in Technology Enhanced Learning*, 18(1), 73-94. Retrieved from https://www.researchgate.net/publication/368740229_Exploring-the-Ethical-Principles-for-the-Implementation-of-Artificial-Intelligence-in-Education-Towards-a-Future-Agenda

Ramirez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., & Miranda, J. (2023). Artificial intelligence in vocational education: Ethical challenges and governance frameworks. *International Journal of Educational Technology in Higher Education*, 20(1), 138-155. <https://doi.org/10.1186/s41239-023-00408-3>

Rodriguez-Abitia, G., & Bribiesca-Correa, G. (2021). Assessing the ethical impact of artificial intelligence in education. *RED. Revista de Educación a Distancia*, 6(1). <https://doi.org/10.34627/redvol6iss1e202303>

Roman Etxebarrieta, G., Orcasitas-Vicandi, M., & Antzaka, A. (2024). *Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)*. AIPioneers.org. Retrieved from: <https://aipioneers.org/evaluation-schema-for-ai-in-education-on-data-privacy-ethics-and-eu-values-wp5/>

Roppertz, S. (2020). Artificial Intelligence and Vocational Education and Training—Perspective of German Vet Teachers. In *European Distance and E-Learning Network (EDEN) Conference Proceedings* (No. 2, pp. 207-216). European Distance and E-Learning Network.

Rott, K. J., Lao, L., Petridou, E., & Schmidt-Hertha, B. (2022). Needs and requirements for an additional AI qualification during dual vocational training: Results from studies of apprentices and teachers. *Computers and education: Artificial intelligence*, 3, 100102.



Şenocak, D., Bozkurt, A., & Koçdar, S. (2024). Exploring the Ethical Principles for the Implementation of Artificial Intelligence in Education: Towards a Future Agenda. In *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 200-213). IGI Global.

Seufert, S. (2024). Artificial Intelligence in Vocational Education and Training (VET): Evaluating VET Leaders' Acceptance of AI in Switzerland.

Tang, J., & Hai, L. (2021). Construction and exploration of an intelligent evaluation system for educational app through artificial intelligence technology. *International Journal of Emerging Technologies in Learning*, 16(5), 17–31. <https://doi.org/10.3991/ijet.v16i05.20293>

Tuomi, I., Cachia, R., & Villar-Onrubia, D. (2023). On the futures of technology in education: Emerging trends and policy implications. *Publications Office of the European Union, Luxembourg*.

UNESCO (2023a). *ChatGPT and artificial intelligence in higher education*. Retrieved from: https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

UNESCO (2023b). *Guidance for generative AI in education and research*. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18, 1-20.

Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S., & Roman-Cañizares, M. (2021). Implementation of a virtual assistant for the academic management of a university with the use of artificial intelligence. *Future Internet*, 13(4), 97.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235.

Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1), 1–5. <https://doi.org/10.1080/17439884.2023.2167830>

Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology*, 95-116.



Kofinanziert von der
Europäischen Union



AI PIONEERS
Artificial Intelligence in education & training

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2023). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 20(1). <https://doi.org/10.1186/s41239-023-00408-3>

KONTAKT

www.aipioneers.org

gorka.roman@ehu.es

maria.orcasitas@ehu.es

natalia.louleli@ehu.eus



AI PIONEERS
Artificial Intelligence in education & training

Manual de política e ética no ensino da IA

Junho 2025

Beneficiário principal do entregável: Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

Work Package 5

Entregável 5.2 “Manual de política e ética no ensino da IA”

N.º do projeto: 101087261

Nome do projeto: AI and the future of Education

Acrónimo do projeto: AI Pioneers

Coordenação do projeto: Universidade de Bremen, Instituto de Tecnologia e Educação (ITB)

Call: ERASMUS-EDU-2022-PI-FORWARD

Tópico: ERASMUS-EDU-2022-PI-FORWARD-LOT1

Tipo de Ação: ERASMUS Lump Sum Grants

Autoridade outorgante: Agência de Execução Europeia da Educação e da Cultura

Data de início do projeto: 01/01/2023

Data do final do projeto: 31/12/2025

Duração do projeto: 36 meses



Cofinanciado pela
União Europeia



AI PIONEERS
Artificial Intelligence in education & training



AI PIONEERS
Artificial Intelligence in education & training

Versão Portuguesa do Entregável

Junho 2025

Tradução e adaptação para língua portuguesa:
Fátima Correia - TecMinho



Manual de política e ética no ensino da IA

Gorka Roman Etxebarrieta (Autor)

Maria Orcasitas-Vicandi (Autor)

Natalia Louleli (Autor)

Asier León Nuñez (Autor)

Itsaso Biota Piñeiro (Autor)

Ana Boyano Murillo (Autor)

Raquel Justo Blanco (Autor)

Josune Rodriguez Negro (Autor)



Índice

Manual de política e ética no ensino da IA.....	3
1. Introdução.....	5
1.1 Objetivo e âmbito do manual	6
1.2 A importância das considerações éticas na adoção da IA.....	7
1.3 Um manual para educadores, administradores e decisores políticos	8
2. Compreender a IA no EFP e na educação de adultos	11
2.1. O Papel da IA na Educação	11
2.2. IA no ensino e formação de adultos e profissional	13
3. Desenvolvimento de orientações éticas	14
3.1 Importância das políticas éticas no ensino da IA.....	14
3.2 Requisitos essenciais para sistemas fiáveis de IA	16
3.3 Consideração dos aspetos jurídicos da aplicação da IA nas instituições de ensino e formação profissional	18
3.4 Estudos de casos e exemplos de políticas de IA eficazes em instituições de ensino e formação profissional	26
4. Competências éticas.....	34
4.1 Definição de competências éticas.....	34
4.2 Importância das competências éticas	35
4.3 Áreas e exemplos de competências éticas	35
5. Conclusão.....	45
5.1 Principais conclusões	45
5.2 Reflexões finais	46
Agradecimentos	47
Referências.....	47



1. Introdução

O projeto AI Pioneers está centrado na promoção e apoio à integração da Inteligência Artificial (IA) na educação, com ênfase na Educação de Adultos e na Educação e Formação Profissional (EFP). Trata-se de um projeto prospetivo ERASMUS+, com início em janeiro de 2023 e financiado até dezembro de 2025.

Com a crescente influência da IA na educação e a crescente procura de apoio profissional na utilização da IA e no desenvolvimento de competências, o projeto visa manter uma rede global e abrangente de Pioneiros da IA nos países parceiros e não só. A criação da Rede de Referência de Pioneiros da IA é um elemento-chave, visando educadores, decisores políticos, planeadores educativos, criadores de IA e outras partes interessadas relevantes. Ao envolver os que mais investem na utilização da IA na educação, o projeto visa criar um núcleo central para apoiar a integração da IA na Educação de Adultos e no Ensino e Formação Profissional (EFP), promovendo iniciativas futuras e projetos de IA educativa. Para mais informações sobre a rede e a forma de participar, poderá ser consultado o [sítio web do projeto](#).

Outro objetivo importante foi o desenvolvimento de um suplemento ao quadro DigCompEDU (Bekiaridis, 2024). Este documento descreve as aptidões e competências necessárias para que os educadores compreendam, avaliem e utilizem as tecnologias de IA na educação, com um enfoque específico na Educação de Adultos e no EFP. Este suplemento alinha as competências de IA com as seis áreas principais do Quadro DigCompEdu (Redecker, 2017): Envolvimento Profissional, Recursos Digitais, Ensino e Aprendizagem, Avaliação, Capacitação dos Aprendentes e Promoção da Competência Digital dos Aprendentes. Para cada área, o suplemento:

1. Analisa a forma como a IA pode ser aplicada no domínio específico
2. Propõe atividades de apoio ao desenvolvimento das competências pertinentes
3. Sugere níveis de progressão para o desenvolvimento de competências
4. Apresenta potenciais desafios e estratégias para os ultrapassar

O projeto continua a desenvolver e a partilhar uma série de recursos através do [Toolkit AI Pioneers](#), disponível no [sítio web do projeto](#). Este conjunto de ferramentas inclui recomendações curriculares para a integração da IA na educação, uma panorâmica dos benefícios e riscos da utilização da IA em contextos educativos e recursos sobre as melhores práticas para a utilização



da IA na educação (Tommasi & Perini, 2024). O *toolkit* inclui, também, breves apresentações de ferramentas utilizadas para diferentes fins no ensino e formação (por exemplo, assistentes de aprendizagem, ferramentas para criar materiais didáticos ou para apoiar a aprendizagem entre pares).

O projeto está a realizar um trabalho relevante na abordagem dos desafios éticos que a IA coloca em contextos educativos. Este documento acompanha um esquema de avaliação e ambos foram concebidos para orientar as instituições e os educadores na implementação responsável da IA, tendo em conta a privacidade e as implicações éticas.

O *Esquema de avaliação da IA na educação* centra-se em orientações práticas para a adoção de ferramentas de IA, garantindo o seu alinhamento com os princípios éticos e os valores da UE. Trata-se de um excelente recurso para as instituições de ensino tomarem decisões informadas sobre a integração da IA.

O *Manual de Política e Ética no Ensino da IA* é um guia mais completo, que oferece políticas e quadros éticos para responder às preocupações atuais.

1.1 Objetivo e âmbito do manual

O *Manual de Política e Ética no Ensino da IA* é um recurso fundamental para educadores, decisores políticos e instituições que pretendem navegar pelas complexidades da IA em contextos educativos. Ajuda a criar um quadro que pode assegurar que as considerações éticas estão no centro das atenções, protegendo a privacidade dos estudantes, promovendo a justiça e encorajando uma utilização responsável.

Uma das principais vantagens é a orientação prática que fornece aos educadores, capacitando-os a utilizar a IA de forma responsável e a ajudar os alunos a compreender as suas capacidades, limitações e implicações éticas. Isto será fundamental para fomentar uma geração de alunos que não só sejam proficientes em tecnologia, mas que também estejam eticamente conscientes quando se envolvem com a tecnologia. Para os decisores políticos, a estrutura consistente que o manual oferece orienta os utilizadores para abordagens mais padronizadas em todas as instituições, o que é essencial para mitigar os riscos da IA, tais como preconceitos em algoritmos ou violações de privacidade. Também pode ajudar a garantir que a IA não seja integrada apenas por si mesma, mas de uma forma que melhore a experiência educacional sem comprometer os padrões éticos.



1.2 A importância das considerações éticas na adoção da IA

As considerações éticas são cruciais na adoção da IA, especialmente em domínios como a educação de adultos e a formação profissional (EFP), onde o impacto da IA nos alunos, educadores e instituições pode ser significativo. Estas considerações ajudam a garantir que a IA é utilizada de forma responsável, inclusiva e transparente, e que apoia os objetivos mais amplos da educação sem comprometer os valores fundamentais.

Algumas considerações fundamentais, que são abrangidas pelo esquema, incluem:

Diversidade e inclusão: Os sistemas e ferramentas de IA devem ser concebidos para responder a um vasto leque de aprendentes, incluindo aqueles com diferentes antecedentes, capacidades e necessidades de aprendizagem. Na educação de adultos e no ensino e formação profissionais, os aprendentes provêm frequentemente de experiências de vida diversas, tais como diferentes níveis de educação prévia, grupos etários e experiência profissional. A IA deve ser flexível e inclusiva, assegurando que a sua utilização apoia e não exclui qualquer grupo em particular. Os utilizadores éticos da IA devem estar conscientes das formas como a tecnologia pode perpetuar preconceitos ou criar fossos digitais, a fim de os atenuar.

Transparência: Para que a IA seja eticamente adotada em contextos educativos, a transparência é essencial. As instituições devem ser claras quanto à forma como os sistemas de IA funcionam, aos dados que recolhem e aos processos de decisão que lhes estão subjacentes. Os alunos e os educadores devem compreender como as ferramentas de IA são utilizadas na sala de aula ou no ambiente de formação e como os sistemas podem afetar os resultados da aprendizagem. Esta transparência promove a confiança nas tecnologias de IA e garante a responsabilização pela sua utilização.

Privacidade e proteção de dados: As instituições de educação de adultos e de ensino e formação profissional lidam frequentemente com dados pessoais sensíveis, como os perfis dos alunos, os resultados das avaliações e o historial profissional. A utilização ética da IA neste contexto deve dar prioridade à privacidade e à proteção de dados. As instituições devem garantir a conformidade com as leis de proteção de dados (por exemplo, o RGPD) e assegurar que os dados dos alunos são recolhidos, armazenados e utilizados de forma responsável. O consentimento para a recolha de dados deve ser informado e voluntário, e os alunos devem ter controlo sobre as suas informações pessoais.



Sustentabilidade: A IA deve ser integrada de uma forma que se alinhe com a sustentabilidade a longo prazo dos sistemas educativos. Isto significa considerar tanto o impacto ambiental das tecnologias de IA como o seu papel na promoção de práticas sustentáveis no âmbito dos programas de educação e formação. No contexto do ensino e formação profissionais, a IA pode também apoiar a sustentabilidade nas indústrias, preparando os alunos com competências que promovam tecnologias ecológicas e práticas respeitadoras do ambiente.

Competências de literacia em IA: À medida que a IA se torna mais integrada na educação, é vital desenvolver a literacia em IA tanto entre os educadores como entre os aprendentes. Na educação de adultos e no EFP, isto implica dotar os alunos de conhecimentos e competências para compreender as tecnologias de IA e utilizá-las eficazmente nas suas carreiras. Esta capacitação permite aos aprendentes navegar no mercado de trabalho em evolução e garantir que não são deixados para trás. Para os educadores, a literacia em IA ajuda-os a avaliar criticamente as ferramentas de IA e a incorporá-las de forma a melhorar a aprendizagem sem comprometer as normas éticas.

Imparcialidade e equidade: Os sistemas de IA devem ser concebidos para tratar todos os alunos de forma justa, garantindo que nenhum grupo seja prejudicado por decisões ou algoritmos automatizados. Na educação de adultos e no EFP, isto é especialmente importante, uma vez que estes domínios servem, com frequência, aprendentes que podem já enfrentar desafios sociais ou económicos. É fundamental que a IA não reforce as desigualdades existentes ou crie novas barreiras ao acesso, mas que apoie oportunidades equitativas para todos os alunos.

1.3 Um manual para educadores, administradores e decisores políticos

Este manual foi concebido para educadores, administradores e decisores políticos no contexto da adoção da IA na educação de adultos e na formação profissional (EFP). Sabe-se que a adoção da IA e as suas considerações éticas têm um impacto diferente em cada um dos grupos acima mencionados:

Educadores:

Para os educadores, a integração ética da IA na educação de adultos e no EFP consiste em melhorar o ensino e a aprendizagem sem pôr em causa os



princípios da equidade, da privacidade e da transparência. As ferramentas de IA podem oferecer experiências de aprendizagem personalizadas, simplificar as tarefas administrativas e fornecer informações sobre o progresso dos alunos. No entanto, os educadores devem garantir que a IA apoia diversos estilos e necessidades de aprendizagem, promovendo um ambiente de aprendizagem inclusivo.

Existem muitas considerações importantes para os educadores relativamente à literacia em IA, à utilização ética e à transparência com os alunos.

Literacia em IA: Os educadores deverão estar equipados com as competências necessárias para compreender, avaliar e integrar eficazmente a IA nos seus métodos de ensino. Isto inclui o conhecimento do funcionamento das ferramentas de IA e das suas implicações éticas na sala de aula.

Utilização ética: Os educadores devem garantir que a IA é utilizada para melhorar a aprendizagem e não para substituir a interação humana. Devem estar preparados para lidar com quaisquer questões de preconceito ou discriminação que possam surgir dos sistemas de IA.

Transparência para com os alunos: Os educadores devem ser transparentes com os alunos sobre a forma como a IA está a ser utilizada no processo de aprendizagem e os dados que recolhe. Isto ajuda a criar confiança e garante que os alunos sintam que a sua privacidade é respeitada.

Administradores:

Para os gestores das instituições de educação de adultos e de ensino e formação profissional, a adoção da IA implica supervisionar a integração destas tecnologias nos sistemas da instituição e garantir a existência de infraestruturas que suportem as ferramentas de IA. Devem, igualmente, dar prioridade a políticas que garantam a utilização ética da IA, mantendo a integridade institucional.

Existem muitas considerações-chave para os administradores relativas à segurança e à privacidade dos dados, à equidade e ao acesso e à formação e apoio contínuos.

Segurança e privacidade dos dados: Os administradores devem garantir a existência de medidas robustas de proteção de dados, uma vez que a IA se baseia na recolha de grandes quantidades de dados dos alunos. A



conformidade com os regulamentos de proteção de dados (como o RGPD) é essencial.

Equidade e acesso: A IA deve ser utilizada de forma a garantir um acesso equitativo a todos os alunos, independentemente da sua origem socioeconómica. Os administradores devem defender políticas que impeçam as ferramentas de IA de exacerbar as desigualdades existentes.

Formação e apoio contínuos: Os administradores precisam de atribuir recursos para o desenvolvimento profissional contínuo, para que os educadores possam manter-se atualizados sobre os avanços da IA e as melhores práticas éticas.

Decisores políticos:

Os decisores políticos desempenham um papel crucial na criação dos quadros e regulamentos que regem a utilização da IA na educação. Têm de garantir que a adoção da IA se alinha com objetivos educativos mais amplos, respeita os direitos humanos e aumenta as oportunidades de aprendizagem ao longo da vida.

Há uma variedade de considerações fundamentais para os decisores políticos no que respeita às orientações éticas e à regulamentação, promovendo a literacia em IA e garantindo a inclusão e a acessibilidade.

Orientações éticas e regulamentação: Os decisores políticos têm de estabelecer normas e regulamentos éticos que orientem a utilização da IA na educação. Isto inclui a abordagem de questões como a transparência, a privacidade dos dados e a equidade nos sistemas educativos orientados para a IA.

Promover a literacia em IA: Os decisores políticos devem promover a literacia em IA em todo o sistema educativo, garantindo que tanto os educadores como os alunos têm as ferramentas e a compreensão necessárias para navegar nas tecnologias de IA.

Garantir a inclusão e a acessibilidade: É vital que as políticas de IA deem prioridade à educação inclusiva, assegurando que as tecnologias de IA não deixem determinados grupos em desvantagem. Isto pode implicar o incentivo à



investigação e ao desenvolvimento de soluções de IA adaptadas a diversos alunos.

2. Compreender a IA no EFP e na educação de adultos

A integração da Inteligência Artificial (IA) na educação apresenta oportunidades transformadoras e desafios significativos. A investigação sublinha o potencial da IA para melhorar o ensino através de tarefas automatizadas, da aprendizagem personalizada e de inovações pedagógicas baseadas em dados, enquanto as ferramentas de IA generativa expandem as possibilidades de uma educação interativa e reflexiva. No entanto, as preocupações éticas em torno da privacidade dos dados, da influência comercial e do fosso digital exigem uma governação sólida. Na educação de adultos e no ensino profissional, a IA não só contribui para a conceção de programas curriculares e a previsão de competências para a preparação da força de trabalho, como também exige a promoção da literacia digital e o envolvimento crítico com as implicações sociais da IA. Equilibrar os benefícios da IA com políticas éticas e equitativas continua a ser crucial para a criação de ecossistemas educativos inclusivos e orientados para o futuro.

2.1. O Papel da IA na Educação

Um relatório de políticas recentemente publicado pelo Joint Research Centre (JRC) da Comissão Europeia, intitulado [On the Futures of Technology in Education: Emerging Trends and Policy Implications](#), (Tuomi, Cachia & Villar-Onrubia, 2023), fornece uma visão abrangente da forma como os avanços tecnológicos, em particular na Inteligência Artificial (IA), estão a transformar as práticas, os processos e as instituições educativas. O documento destaca os impactos significativos e diversos da IA na educação, afetando os processos de ensino e aprendizagem, os métodos de avaliação e as estruturas organizacionais.

A IA é vista como uma solução potencial para reduzir a carga de trabalho dos professores, especialmente em tarefas como a correção dos trabalhos de casa dos alunos. Ao automatizar estas tarefas repetitivas e morosas, a IA pode tornar o ensino mais apelativo e permitir que os professores se concentrem mais nas atividades de ensino do que nas tarefas administrativas (Li, 2021; Tang & Hai, 2021; Villegas-Ch et al., 2021). Esta automatização poderá



revolucionar a avaliação formativa (Kumar & Boulanger, 2020), permitindo um *feedback* contínuo aos alunos e facilitando a avaliação em tempo real (Porter & Grippa, 2020) e a aprendizagem personalizada ou adaptativa (Kochmar et al., 2020; Wongvorachan et al., 2022).

Além disso, espera-se que a IA generativa e os modelos fundamentais tenham um impacto substancial na educação. Tecnologias como a legendagem automática de vídeos, a tradução e o resumo de vídeos oferecem oportunidades inovadoras para melhorar a pedagogia (Chew & Chua, 2020; Vazquez-Cano et al., 2021). Estas tecnologias podem ajudar a desenvolver “companheiros inteligentes” ou “parceiros de aprendizagem”, deslocando o foco da aprendizagem mecânica para a metacognição e a reflexão. Isto exige uma redistribuição cuidadosa das responsabilidades entre humanos e máquinas, determinando quais as atividades a delegar na IA.

Também devemos enfatizar o papel da IA na dataficação da educação, onde dados em grande escala sobre atividades educacionais são usados para reforçar e informar novas abordagens pedagógicas, tecnologias de aprendizagem e políticas educacionais (Williamson et al., 2023). O objetivo é registar os processos de aprendizagem, e não apenas os resultados, para melhorar o envolvimento e o desempenho dos alunos (Porter & Grippa, 2020).

No entanto, as preocupações éticas sobre o tratamento de dados pessoais, a potencial utilização indevida, a dependência de infra-estruturas em grande parte controladas por entidades comerciais e alguns outros desafios são questões críticas a abordar (Seufert, 2024; Wongvorachan et al., 2022). A integração da IA na educação suscita questões significativas sobre os modelos de governação e a elaboração de políticas, exigindo uma reavaliação das políticas educativas para incorporar as capacidades e limitações da IA (Attard-Frost et al., 2024; Boninger et al., 2020; Williamson & Eynon, 2020). Devemos, também, abordar a importância de avaliar criticamente os riscos de exacerbar o fosso digital e o domínio dos interesses comerciais no setor (Bulathwela et al., 2024; Celik, 2023). Regulamentar a inovação tecnológica para garantir visões educativas diversificadas e alinhar a política industrial com as teorias da aprendizagem são aspetos cruciais que temos de considerar.

O papel da IA na educação é uma ferramenta transformadora e multifacetada que oferece oportunidades para uma maior eficiência, personalização e inovação no ensino e na aprendizagem. No entanto, também apresenta desafios significativos e considerações éticas que devem ser abordadas para



garantir a utilização equitativa, eficaz e responsável da IA em contextos educativos.

2.2. IA no ensino e formação de adultos e profissional

A integração do ensino da IA nos programas curriculares implica não só o ensino dos fundamentos da IA e das suas aplicações, como também a exploração dos seus impactos sociais. Isto inclui a promoção do pensamento crítico sobre as capacidades, limitações e dimensões éticas da IA, como as preocupações em torno do preconceito e da privacidade (Akgun & Greenhow, 2022; Aldosari, 2020; Attwell et al., 2020; Delcker, et al., 2024; Seufert, 2024; Wongvorachan et al., 2022). O reforço da literacia digital é, também essencial, uma vez que a compreensão de como interagir e interpretar os sistemas de IA está a tornar-se uma competência fundamental para a força de trabalho moderna, incluindo os educadores (Bekiaridis & Attwell, 2024; Delcker, et al., 2024; Roppertz, 2020; Rott et al., 2022; Seufert, 2024). À medida que a IA continua a evoluir, o foco está a mudar não só para a utilização de ferramentas de IA, mas também para a sua compreensão e avaliação crítica, preparando os alunos para um futuro em que a IA é parte integrante da vida quotidiana e do trabalho (Rott et al., 2022).

Como tal, a preparação da mão de obra para novas competências na era da IA é outro aspeto crucial da educação de adultos e do EFP (Attwell et al., 2020; Roppertz, 2020; Rott et al., 2022). O papel da IA na identificação de tendências emergentes e na previsão de futuras necessidades de competências é inestimável. Esta visão permite que os programas educativos adaptem os seus currículos para satisfazer as necessidades em evolução do mercado de trabalho (Seufert, 2024). O *upskilling* e o *reskilling* tornam-se temas centrais, com programas guiados por IA que permitem que adultos e profissionais adquiram novas competências ou atualizem as existentes, mantendo-se assim a par dos avanços tecnológicos (Attwell et al., 2020; Emeršič et al., 2024; Delcker, et al., 2024). Para além disso, a formação orientada para a IA garante que os alunos não são apenas tecnicamente competentes, mas estão também preparados para trabalhar em ambientes integrados na IA. Esta preparação é fundamental para garantir que a mão de obra esteja pronta para os desafios e oportunidades apresentados pelo domínio da IA, que avança rapidamente, tornando a educação de adultos e o EFP mais relevantes, eficazes e alinhados com os futuros cenários de emprego.



3. Desenvolvimento de orientações éticas

3.1 Importância das políticas éticas no ensino da IA

As políticas éticas na educação sobre IA são essenciais para garantir que as tecnologias de IA são utilizadas de forma a promover a equidade, a transparência e a responsabilidade, protegendo simultaneamente os direitos e o bem-estar dos alunos e dos educadores. À medida que a IA se vai integrando cada vez mais nos contextos educativos, é crucial dispor de orientações éticas claras para gerir os seus potenciais riscos e aproveitar os seus benefícios. Seguem-se as principais razões pelas quais as políticas éticas são vitais no ensino da IA:

Proteção dos direitos e da privacidade dos alunos

Os sistemas de IA na educação exigem frequentemente a recolha e o tratamento de dados pessoais, incluindo registos académicos, comportamentos de aprendizagem e outras informações sensíveis. As políticas éticas garantem que os dados são tratados com respeito e em conformidade com os regulamentos de privacidade (por exemplo, RGPD). Estas políticas ajudam a proteger a privacidade dos alunos, garantindo a transparência na recolha de dados, dando aos alunos controlo sobre os seus dados e assegurando que são utilizados de forma responsável e segura.

Garantir a imparcialidade e a equidade

Uma das preocupações éticas mais importantes com a IA na educação é o potencial para algoritmos tendenciosos que possam prejudicar determinados grupos de alunos. Os sistemas de IA podem, inadvertidamente, perpetuar as desigualdades existentes com base na raça, no género, no estatuto socioeconómico ou noutros fatores. São necessárias políticas éticas para orientar o desenvolvimento e a implantação de ferramentas de IA que sejam inclusivas e equitativas, garantindo que todos os alunos tenham igual acesso a oportunidades, independentemente da sua origem ou características pessoais.

Promover a transparência e a responsabilização

As tecnologias de IA podem ser complexas e opacas, o que torna difícil para os educadores, aprendentes e administradores compreenderem plenamente a forma como as decisões são tomadas. As políticas éticas ajudam a garantir que os sistemas de IA são transparentes, o que significa que as partes interessadas sabem como as ferramentas de IA funcionam, que dados estão a ser utilizados



e como as decisões estão a ser tomadas. Esta transparência gera confiança e responsabilidade, permitindo que os educadores e os alunos compreendam e interajam melhor com as tecnologias de IA.

Apoiar a literacia e a utilização responsável da IA

À medida que a IA se integra cada vez mais no processo de aprendizagem, é importante que tanto os educadores como os alunos compreendam como funciona a IA e as suas potenciais implicações. As políticas éticas no ensino da IA devem incentivar a literacia em IA, assegurando que tanto os educadores como os alunos estão equipados com os conhecimentos e as competências para utilizar a IA de forma responsável e crítica. Esta compreensão permite que os utilizadores maximizem os benefícios da IA, enquanto estão conscientes das suas limitações e potenciais danos.

Promover o desenvolvimento ético e a inovação

As tecnologias de IA na educação devem ser desenvolvidas com uma forte base ética para garantir que são utilizadas em benefício de todos os alunos. As políticas éticas fornecem um quadro para os criadores, investigadores e instituições educativas conceberem sistemas de IA que se alinhem com valores educativos como a justiça, a capacitação e o respeito pela diversidade. Isto também garante que as tecnologias de IA são concebidas e utilizadas de forma a apoiar a aprendizagem ao longo da vida e os objetivos mais amplos da educação.

Abordar o risco de substituição

Dado que os sistemas de IA automatizam certos aspetos da educação (por exemplo, classificação, tutoria, tarefas administrativas), existe o risco de substituírem os educadores ou limitarem a interação humana. As políticas éticas podem ajudar a orientar a utilização responsável da IA, garantindo que a automatização complementa os educadores humanos, em vez de os substituir. As políticas podem incentivar a utilização da IA para melhorar a eficácia do ensino e a eficiência administrativa, mantendo simultaneamente o valor da presença humana na experiência educativa.

Criar confiança nos sistemas de IA

Para que a IA seja adotada com êxito em contextos educativos, os educadores, os alunos e os administradores devem confiar na tecnologia. As políticas éticas desempenham um papel importante no estabelecimento e manutenção desta



confiança, fornecendo orientações claras sobre a forma como a IA deve ser utilizada e quais as normas éticas que devem ser seguidas. Quando os sistemas de IA são desenvolvidos e implementados de acordo com estas políticas, é mais provável que as partes interessadas os adotem e utilizem eficazmente.

Orientar o futuro da IA na educação

À medida que a IA continua a evoluir, é crucial que as políticas éticas forneçam orientações a longo prazo para a sua integração nos sistemas educativos. Estas políticas ajudam a criar um quadro para o futuro desenvolvimento de tecnologias de IA que se alinhe com os valores sociais, assegurando que a IA é utilizada para melhorar os resultados da aprendizagem, promover a inovação e apoiar a aprendizagem ao longo da vida de forma benéfica e ética.

De um modo geral, as políticas éticas no domínio da educação em matéria de IA são necessárias para orientar o desenvolvimento e a utilização responsáveis das tecnologias de IA. Estas políticas garantem que a IA é utilizada de forma a respeitar a privacidade dos alunos, promover a equidade, incentivar a transparência e criar confiança. Além disso, ajudam os educadores, os alunos e as instituições a enfrentar os desafios éticos que surgem à medida que a IA continua a transformar os ambientes educativos. Ao estabelecer estas orientações éticas, os sistemas educativos podem aproveitar o poder da IA, salvaguardando os valores que são fundamentais para uma educação eficaz e inclusiva.

3.2 Requisitos essenciais para sistemas fiáveis de IA

A integração da Inteligência Artificial (IA) na educação oferece imensas oportunidades para melhorar o ensino, a aprendizagem e a administração. No entanto, para garantir que a IA é utilizada de forma ética e responsável, as instituições de ensino devem desenvolver orientações claras para a implementação ética dos sistemas de IA. Numerosas abordagens, orientações e auditorias publicadas (por exemplo, UNESCO's AI Audit 2023a and Guidance for Generative AI in Education and Research 2023b) fornecem informações valiosas sobre o que constitui uma IA fiável. Estes quadros enfatizam princípios-chave como a **transparência, a justiça, a responsabilidade e o respeito pelos direitos humanos**. No entanto, a tarefa de desenvolver e aplicar esses princípios pode ser assustadora para as instituições, especialmente num cenário tecnológico em rápida evolução.



Como parte dos AI Pioneers, desenvolvemos um esquema de avaliação (Roman Etxebarrieta et al., 2024) concebido para simplificar e contextualizar estes princípios para contextos educativos. O nosso esquema fornece uma abordagem estruturada para ajudar as instituições a navegar pelas complexidades da implementação da IA, assegurando simultaneamente a adesão a normas éticas. Enquanto existirem várias orientações disponíveis, o nosso esquema visa oferecer uma ferramenta prática e de fácil utilização que guiará as instituições de ensino e formação no desenvolvimento das suas próprias orientações éticas em matéria de IA, tendo em conta os seus contextos e necessidades específicos.

Uma das ideias centrais do Regulamento da Inteligência Artificial da UE – AI Act (Parlamento Europeu e Conselho da União Europeia, 2024) é que a principal responsabilidade por garantir que os sistemas de IA são desenvolvidos de forma ética cabe aos programadores e às empresas que os criam, e que as orientações a seguir dependerão do nível de risco associado aos sistemas de IA. No entanto, os centros educativos são responsáveis por garantir que estes sistemas sejam implementados e utilizados de forma ética nos seus ambientes educativos. Os centros devem estar conscientes das principais questões éticas para compreenderem as limitações e o potencial dos sistemas de IA que adotam. Isto implica não só o cumprimento das políticas e dos requisitos legais existentes, como também a antecipação e a resolução dos desafios que possam surgir à medida que estas tecnologias evoluem.

Especificamente, o nosso esquema de avaliação centra-se na identificação de preocupações éticas e no desenvolvimento de uma política ou estratégia que possa apoiar e promover a utilização da IA em instituições de ensino, respeitando as principais orientações éticas. Estas orientações incluem: **diversidade e equidade, transparência e responsabilidade, privacidade e proteção de dados, segurança e proteção, sustentabilidade e bem-estar social, capacitação de professores e alunos, participação democrática no planeamento de políticas educativas e práticas de IA, autonomia, conceção ética (relacionada com a sustentabilidade e a acessibilidade), comercialização (Şenocak et al., 2024), agência de professores e alunos e literacia em IA.**

O esquema de avaliação é composto por várias secções, baseadas nas dimensões de Chan (2023), ajustadas com a revisão de Şenocak et al. (2024) e as orientações éticas da Comissão Europeia (2022), cada uma centrada em diferentes aspetos da implementação da IA em contextos educativos. Estas secções orientam as instituições no desenvolvimento de quadros éticos e na garantia de que a IA é utilizada de forma responsável. A primeira secção



centra-se na **Avaliação da Utilização Atual da IA e das Preocupações Éticas**, ajudando as instituições a avaliar a sua maturidade atual em matéria de IA, permitindo-lhes assim estimar o grau de preparação do centro educativo para implementar a IA ou expandir a sua utilização da IA (JISC, 2022).

A segunda secção centra-se na **Governança e Monitorização da IA** e, por conseguinte, aborda questões relacionadas com a conformidade com as políticas regionais e internacionais, abrangendo a privacidade, a proteção de dados e a equidade. A terceira secção, **Considerações Operacionais**, fornece orientações sobre a formação de professores, alunos e pessoal, promovendo a literacia em IA, assegurando a supervisão humana e apoiando a participação democrática na tomada de decisões relacionadas com a IA. Assegurar que todos os utilizadores são adequadamente formados e estão conscientes das capacidades e limitações da IA é fundamental para uma implementação bem-sucedida da IA e garante a igualdade de acesso para todos.

A quarta secção centra-se nas **Considerações Pedagógicas**, considerando questões importantes como os dilemas éticos nas avaliações baseadas em IA, a agência do aluno e o desenvolvimento de competências, particularmente na Educação e Formação de Adultos e Profissional (Attwell et al., 2021; UNESCO, 2019). A quinta e última secção centra-se em **Outras Considerações: Sustentabilidade e Conceção Ética**, oferecendo orientações para abordar os impactos a longo prazo e garantir uma utilização acessível da IA.

O esquema de avaliação fornece uma ferramenta prática para as instituições de ensino navegarem na implementação da IA, identificarem desafios éticos e desenvolverem políticas alinhadas com as normas internacionais. Para obter informações mais pormenorizadas, consulte o esquema de avaliação completo [Esquema de avaliação da IA na educação sobre dados, privacidade, ética e valores da UE (WP5)] e a apresentação Genially que pode ser encontrada no [website do projeto AI Pioneers](#).

3.3 Consideração dos aspetos jurídicos da aplicação da IA nas instituições de ensino e formação profissional

3.3.1 O Quadro Normativo do Regulamento da Inteligência Artificial da EU

O regulamento da IA da União Europeia (AI Act) estabelece um quadro regulamentar abrangente que classifica os sistemas de inteligência artificial de acordo com os seus riscos potenciais, fornecendo orientações essenciais para as instituições de ensino e formação profissional (EFP) que implementam estas



tecnologias. Esta abordagem baseada no risco assegura uma supervisão regulamentar proporcionada, enquanto promove a inovação em contextos educativos (Parlamento Europeu, 2024).

Como Zawacki-Richter et al. (2023) articulam na sua revisão sistemática das considerações éticas na educação melhorada pela IA, os quadros regulamentares devem equilibrar a inovação tecnológica com a proteção dos direitos fundamentais. A sua análise identifica o Regulamento da IA da UE como uma tentativa pioneira de estabelecer “parâmetros claros para a implantação responsável da IA em ambientes de aprendizagem, reconhecendo simultaneamente o potencial transformador destas tecnologias” (p. 8).

Classificação de riscos para a IA educativa

Os sistemas de IA implantados em contextos de EFP inserem-se predominantemente na categoria de alto risco devido ao seu impacto significativo nas experiências educativas dos estudantes e nas suas futuras oportunidades de carreira. O Regulamento da UE sobre a IA classifica os sistemas de IA da seguinte forma:

Risco inaceitável: São totalmente proibidos os sistemas que representem ameaças claras aos direitos fundamentais ou à segurança

Alto risco: Aplicações em setores críticos, incluindo a educação, que exigem uma rigorosa conformidade regulamentar

Risco limitado: Sistemas com obrigações de transparência que exigem a divulgação da interação com a IA

Risco mínimo: Aplicações com requisitos regulamentares limitados para além dos quadros existentes.

Implicações da classificação de alto risco para os estabelecimentos de ensino e formação profissional:

- Protocolos obrigatórios de avaliação e mitigação de riscos
- Requisitos rigorosos de documentação
- Mecanismos de supervisão humana
- Obrigações de transparência para as partes interessadas afetadas
- Avaliações regulares de conformidade.



A investigação realizada por Ramirez-Montoya et al. (2023) enfatiza que a classificação de alto risco para sistemas de IA educacionais requer «estruturas de governança rigorosas que transcendam a conformidade técnica para abordar questões sociais mais amplas» (p. 142). A sua análise dos desafios de implementação em instituições de ensino profissional europeias destaca a necessidade de uma interpretação contextualizada dos requisitos regulatórios.

3.3.2 Principais considerações legais para a implementação do EFP

Requisitos de transparência e de responsabilidade

O Regulamento da IA da União Europeia (AI Act) exige uma ampla transparência na implantação da IA em contextos educativos:

- Os estabelecimentos de ensino e formação profissional devem informar os estudantes, o pessoal e as partes interessadas quando os sistemas de IA são utilizados na avaliação, nas recomendações de aprendizagem ou na tomada de decisões administrativas
- A documentação deve articular o objetivo, a funcionalidade e as limitações das aplicações de IA
- As decisões baseadas na IA que afetam a progressão ou a certificação dos estudantes devem ser explicáveis e justificáveis
- Os processos algorítmicos devem ser auditáveis para permitir a validação dos resultados.

Prinsloo e Knox (2023) identificam a transparência como um princípio ético fundamental na implementação da IA educacional, observando que "a opacidade algorítmica apresenta desafios únicos em contextos de aprendizagem em que a confiança entre educadores e alunos constitui a base de uma pedagogia eficaz" (p. 76). O seu enquadramento para a transparência da IA educativa vai para além da divulgação técnica, abrangendo a compreensão contextual acessível a diversas partes interessadas.

Proteção de dados e conformidade com a privacidade

A implementação da IA em contextos de EFP exige uma adesão meticulosa aos quadros de proteção de dados:

- Todos os sistemas de IA devem cumprir as disposições do Regulamento Geral sobre a Proteção de Dados (RGPD)
- Devem ser estabelecidos mecanismos claros de consentimento para o tratamento de dados pessoais



- Os princípios de minimização de dados devem reger as práticas de recolha e conservação
- A limitação da finalidade garante a utilização dos dados apenas para objetivos educativos específicos
- As salvaguardas técnicas e organizacionais devem proteger a integridade e a confidencialidade dos dados dos alunos.

Holmes et al. (2023) observam que as considerações sobre a proteção de dados em contextos educativos vão além da conformidade legal para abranger "questões éticas mais amplas relativas à agência do aluno, à dinâmica de poder e à potencial vigilância educativa" (p. 12). O seu trabalho enfatiza a importância de contextualizar os quadros de proteção de dados no âmbito de considerações éticas específicas da educação.

Prevenção de preconceitos e garantia de equidade

O regulamento da IA da União Europeia estabelece requisitos explícitos para atenuar o enviesamento algorítmico em contextos educativos:

- As instituições de ensino e formação profissional devem efetuar avaliações regulares dos sistemas de IA para evitar a discriminação
- As metodologias de avaliação devem avaliar potenciais impactos díspares entre grupos demográficos
- Os dados de treino algorítmico devem ser representativos da diversidade dos alunos
- Os protocolos de monitorização contínua devem identificar e corrigir os desvios emergentes nos sistemas operacionais
- Para verificar o impacto equitativo, é necessário proceder a uma auditoria regular dos resultados obtidos junto das populações estudantis.

A investigação levada a cabo por Hodgkin e Kumar (2023), que examinou o enviesamento algorítmico em sistemas de orientação profissional, identifica "metodologias sistemáticas para avaliar o impacto díspar em diversas populações de estudantes" (p. 217) como essenciais para uma implementação equitativa da IA. A sua estrutura para uma avaliação abrangente do enviesamento fornece orientações práticas para as instituições de EFP que navegam pelos requisitos de equidade.

Ferramentas de IA na sala de aula (risco limitado)

Podem ser classificados como de risco limitado numerosas aplicações de IA, tais como assistentes de tutoria alimentados por IA, sistemas de deteção de



plágio, motores de recomendação para recursos de aprendizagem, ferramentas de automatização administrativa e simulações de laboratórios virtuais em contextos de EFP.

Estas aplicações devem informar claramente os utilizadores sobre:

- A natureza da IA do sistema
- Limitações das capacidades algorítmicas
- Disposições relativas à supervisão humana
- Mecanismos de contestação dos resultados.

3.3.3 Cronograma de implementação e planeamento da conformidade

As instituições de ensino e formação profissional devem desenvolver planos estratégicos de implementação alinhados com o calendário de aplicação faseada do regulamento da IA da União Europeia:

2 de fevereiro de 2025: Proibição de sistemas de IA que representem riscos inaceitáveis

2 de agosto de 2025: Aplicação de códigos de boas práticas e designação de autoridades reguladoras nacionais

2 de agosto de 2026: Aplicação geral do regulamento da IA, incluindo regras para sistemas de IA de alto risco

2 de agosto de 2027: Aplicação integral de obrigações específicas para sistemas de IA de alto risco.

Um planeamento prudente da conformidade exige:

- Inventário exaustivo das aplicações de IA existentes e planeadas
- Avaliação da classificação de risco da carteira de tecnologias educativas
- Análise de lacunas em relação aos requisitos regulamentares
- Roteiro de implementação estruturado com objetivos de etapas
- Afetação de recursos para a documentação e o controlo da conformidade.

3.3.4 Responsabilidades legais e considerações sobre a responsabilidade

As instituições de EFP que implementam sistemas de IA assumem responsabilidades jurídicas significativas:



As estruturas de responsabilização devem designar claramente a responsabilidade pela governação do sistema de IA

Os quadros de responsabilidade devem ter em conta os potenciais danos resultantes de erros ou enviesamentos dos algoritmos

Os acordos contratuais com os fornecedores de tecnologia devem definir explicitamente a atribuição de responsabilidades

Os protocolos de indemnização devem contemplar potenciais reclamações decorrentes de decisões influenciadas pela IA

As considerações em matéria de seguros devem refletir os riscos de responsabilidade emergentes.

As instituições de ensino devem reconhecer que a implantação de sistemas de IA, mesmo os adquiridos a terceiros, não as isenta da responsabilidade legal pelos resultados que afetam os alunos.

Rodriguez-Abitia e Bribiesca-Correa (2021) identificam considerações de responsabilidade como uma dimensão frequentemente negligenciada da governação da IA educacional, observando que "as instituições educacionais enfrentam exposições legais únicas ao delegar funções tradicionalmente humanas em sistemas algorítmicos" (p. 94). A sua análise da jurisprudência emergente fornece um precedente valioso para as instituições de EFP que desenvolvem quadros de responsabilidade abrangentes.

3.3.5 Lista de avaliação para a implementação de IA fiável no EFP

A Lista de Avaliação para uma Inteligência Artificial Fiável (ALTAI) fornece às instituições de EFP um quadro estruturado para a implementação ética e a conformidade regulamentar. Esta ferramenta prática operacionaliza as Orientações Éticas para uma IA Fiável desenvolvidas pelo Grupo de Peritos de Alto Nível da UE em Inteligência Artificial.

Na sua análise comparativa dos quadros de avaliação ética, Zawacki-Richter et al. (2023) identificam a ALTAI como particularmente valioso em contextos educativos devido à sua "cobertura abrangente das dimensões técnicas e sócio éticas relevantes para os ambientes de aprendizagem" (p. 14). A sua investigação demonstra a adaptabilidade da ALTAI a diversos contextos educativos, incluindo a formação profissional.



Componentes-chave da ALTAI relevantes para a implementação no EFP

A ALTAI estrutura a avaliação em sete dimensões fundamentais, particularmente relevantes em contextos de ensino profissional:

Agência humana e supervisão

- As aplicações de ensino e formação profissional devem apoiar a autonomia dos formadores, em vez de se sobrepor ao juízo profissional
- A autonomia dos estudantes deve ser preservada através do consentimento informado e da transparência das decisões
- Deve ser mantida uma supervisão humana significativa nas decisões consequentes que afetam a progressão dos alunos.

Robustez técnica e segurança

- Os sistemas de IA devem funcionar de forma fiável em diversos ambientes de formação profissional
- Os protocolos de segurança devem proteger contra potenciais manipulações ou acessos não autorizados
- Os procedimentos de recurso devem assegurar a continuidade do ensino em caso de falha do sistema.

Privacidade e governação de dados

- Os dados dos alunos utilizados na formação e funcionamento da IA devem ser protegidos através de uma governação abrangente
- Os princípios de minimização de dados devem orientar as práticas de recolha e conservação
- Medidas de segurança robustas devem proteger as informações sensíveis contra o acesso não autorizado.

Transparência

- As capacidades e limitações da IA devem ser claramente comunicadas às partes interessadas do sector da educação
- Os processos de decisão devem ser explicáveis numa linguagem não técnica
- A documentação deve permitir a verificação do desempenho e da conformidade do sistema.



Diversidade, não-discriminação e equidade

- As implementações específicas do EFP devem ter em conta a diversidade dos contextos dos estudantes e das necessidades de aprendizagem
- As metodologias de avaliação devem avaliar o impacto da equidade em todas as dimensões demográficas
- As estratégias de atenuação dos preconceitos devem ter em conta as considerações de diversidade específicas do ensino profissional.

Bem-estar social e ambiental

- Os impactos sociais gerais da IA na formação profissional devem ser avaliados de forma sistemática
- As consequências ambientais das infraestruturas informáticas devem ser avaliadas
- Devem ser consideradas as potenciais implicações da certificação de competências em IA para o mercado de trabalho.

Responsabilidade

- As estruturas de governação claras devem estabelecer a responsabilidade pelos sistemas de IA
- Os mecanismos de auditoria devem permitir a verificação do cumprimento da regulamentação
- Os procedimentos de correção devem abordar os resultados prejudiciais ou as consequências não intencionais.

Aplicação da ALTAI em contextos de ensino e formação profissional

As instituições de ensino e formação profissional podem operacionalizar a ALTAI através de protocolos de avaliação sistemáticos:

- **A avaliação inicial da implementação** avalia as aplicações de IA propostas em função de critérios de fiabilidade
- **A verificação antes da implementação** garante a conformidade com os requisitos técnicos e éticos
- **A auditoria pós-implementação** valida o desempenho em diversas populações de estudantes
- **A reavaliação periódica** identifica as preocupações éticas emergentes à medida que os sistemas evoluem.



Ao integrar a ALTAI nos quadros de governação, as instituições de ensino e formação profissional podem demonstrar a devida diligência no cumprimento da regulamentação, reforçando simultaneamente a confiança das partes interessadas na implementação da IA.

3.4 Estudos de casos e exemplos de políticas de IA eficazes em instituições de ensino e formação profissional

3.4.1 Quadros políticos europeus fundamentais

Políticas institucionais eficazes baseiam-se em quadros fundamentais estabelecidos a nível europeu:

Grupo de Peritos de Alto Nível em Inteligência Artificial (High-Level Expert Group on Artificial Intelligence - AI HLEG)

O Grupo de Alto Nível para a IA da Comissão Europeia desenvolveu orientações essenciais através de:

Orientações éticas para uma IA fiável: Estabelecer sete requisitos para uma implementação ética da IA

Recomendações políticas e de investimento: Proporcionar uma direcção estratégica para o desenvolvimento sustentável da IA

Lista de avaliação para uma IA fiável: Oferece ferramentas de implementação práticas para as organizações.

Estes quadros de referência serviram de base a políticas institucionais concretas em muitos das entidades de EFP a nível europeu.

Como Holmes et al. (2023) observam na sua análise dos quadros de governação da IA, as orientações do AI HLEG representam "uma ponte crítica entre princípios éticos abstratos e práticas de implementação concretas" (p. 18). A sua investigação traça a influência destas orientações em diversos contextos educativos, demonstrando a sua adaptabilidade a ambientes de formação profissional.



3.4.2 Escola Técnica de Copenhaga: Quadro abrangente de governação da IA

O Copenhagen Technical College criou uma estrutura de governação a vários níveis, demonstrando uma prática exemplar na implementação da IA no EFP.

Arquitetura política e estrutura de governação

A abordagem da instituição caracteriza-se por:

- Comité de Ética da IA com várias partes interessadas, incluindo administradores, educadores, estudantes, representantes da indústria e peritos técnicos
- Protocolos de aprovação escalonados com base na classificação algorítmica do risco
- Formação obrigatória em literacia em IA para todos os docentes e pessoal administrativo
- Formalização dos direitos dos alunos relativamente à transparência algorítmica e à contestabilidade das decisões
- Protocolos de avaliação estruturados que medem as dimensões educativa e ética.

Esta abordagem alinha-se com o quadro de envolvimento das partes interessadas de Prinsloo e Knox (2023), que salienta a importância da "representação estruturada através das hierarquias institucionais e das comunidades afetadas" (p. 82). A sua investigação demonstra como a governação de múltiplas partes interessadas melhora o rigor ético e a eficácia da implementação.

Metodologia de implementação

Os principais elementos de implementação incluem:

- Implementação faseada, começando com aplicações administrativas de baixo risco
- Pilotos pedagógicos controlados com protocolos rigorosos de avaliação
- Documentação estruturada do desempenho técnico e do impacto ético
- Mecanismos regulares de *feedback* das partes interessadas para informar a melhoria iterativa
- Auditoria independente dos resultados algorítmicos em todas as dimensões demográficas.



Resultados documentados

A Escola registou benefícios significativos através da implementação estruturada:

- 27% de redução do volume de trabalho administrativo através da automatização ética da IA
- Melhoria da personalização do ensino sem comprometer a qualidade do ensino
- Maior transparência nas práticas de avaliação
- Aumento da satisfação dos alunos com a integração tecnológica
- Conformidade bem-sucedida com os requisitos regulamentares emergentes.

3.4.3 Instituto de Tecnologia de Bordéus: Políticas transparentes de avaliação da IA

O Instituto desenvolveu modelos de políticas que abordam um dos aspetos eticamente mais complexos da IA educativa: a avaliação dos alunos.

Quadro político específico da avaliação

Os principais elementos políticos incluem:

- Proibição da avaliação totalmente automatizada em provas importantes
- Revisão humana obrigatória de todas as decisões de avaliação influenciadas pela IA
- Documentação transparente da contribuição algorítmica para as avaliações
- Mecanismo de recurso estruturado para contestar resultados influenciados pela IA
- Auditorias regulares sobre preconceitos, comparando os resultados da avaliação entre as populações de estudantes
- Representação dos estudantes nas estruturas de governação da avaliação

Esta abordagem exemplifica o que Ramirez-Montoya et al. (2023) identificam como "governação da avaliação centrada no ser humano", caracterizada por "transparência algorítmica, participação das partes interessadas e mecanismos de supervisão formalizados" (p. 149). A sua investigação demonstra como



estes quadros melhoram tanto a validade da avaliação como a confiança das partes interessadas.

Abordagem de implementação

A estratégia de implementação dá ênfase a:

- Delimitação clara entre automatismos de avaliação adequados e inadequados
- Implementação gradual com base na avaliação de riscos e consequências
- Desenvolvimento do corpo docente em metodologias de avaliação ética da IA
- Infraestruturas técnicas que garantem a explicabilidade das recomendações algorítmicas
- Protocolos de documentação que permitam a verificação da equidade da avaliação.

Impacto mensurável

O Instituto documentou benefícios significativos:

- Maior consistência na avaliação, preservando a autonomia do instrutor
- Maior confiança dos alunos nas metodologias de avaliação
- Redução do volume de trabalho de avaliação sem comprometer a qualidade
- Melhor identificação dos alunos que necessitam de apoio adicional
- Conformidade regulamentar bem-sucedida com os requisitos de transparência da IA.

3.4.4 Consórcio de escolas profissionais de Munique: Governança colaborativa de dados

Oito instituições de ensino profissional em Munique estabeleceram uma abordagem colaborativa à governança dos dados de IA, demonstrando como a partilha de recursos pode melhorar a implementação ética.

Arquitetura de governança partilhada

Os principais elementos estruturais incluem:



- Quadro ético comum para a recolha e utilização de dados
- Infraestrutura técnica partilhada para a gestão e segurança dos dados
- Protocolos partilhados de avaliação e contratação de fornecedores
- Aplicação conjunta de metodologias de aprendizagem federada
- Conhecimentos especializados comuns para a verificação da conformidade regulamentar.

Esta abordagem colaborativa alinha-se com o que Rodriguez-Abitia e Bribiesca-Correa (2021) designam por "governança otimizada em termos de recursos", caracterizada por "conhecimentos especializados partilhados, políticas unificadas e capacidade de implementação coletiva" (p. 98). A sua pesquisa demonstra como essas abordagens podem permitir que instituições menores implementem estruturas de governança sofisticadas.

Metodologia de implementação

A abordagem do Consórcio inclui:

- Protocolos normalizados de proteção de dados nas instituições participantes
- Experiência centralizada em técnicas de minimização e anonimização de dados
- Recursos partilhados para a avaliação da vulnerabilidade técnica
- Desenvolvimento colaborativo de metodologias analíticas que preservam a privacidade
- Envolvimento conjunto das partes interessadas no desenvolvimento da governação

Benefícios demonstrados

Esta abordagem de colaboração produziu resultados:

- Governação de dados mais sofisticada do que as instituições individuais poderiam alcançar de forma independente
- Melhoria da posição negocial com os fornecedores de tecnologia
- Melhoria dos protocolos de segurança dos dados através da partilha de conhecimentos especializados
- Conformidade regulamentar rentável em várias instituições
- Atenuação mais abrangente dos enviesamentos através de conjuntos de dados maiores e mais diversificados.



3.4.5 Instituto Técnico de Dublin: Quadro de política de IA centrado na equidade

O Dublin Technical Institute desenvolveu políticas que dão especificamente prioridade à equidade algorítmica em contextos de ensino profissional.

Quadro político centrado na equidade

Os principais elementos políticos incluem:

- Avaliações obrigatórias do impacto na equidade antes da implementação da IA
- Auditoria algorítmica regular em todas as dimensões demográficas
- Priorização de metodologias de IA explicáveis
- Necessidade de alternativas não tecnológicas aos sistemas de IA
- Inclusão formalizada de populações sub-representadas nas estruturas de governação.

Este quadro exemplifica o que Hodgkin e Kumar (2023) identificam como "governação proactiva da equidade" caracterizada por "avaliação sistemática de preconceitos, representação inclusiva das partes interessadas e protocolos de monitorização contínua" (p. 219). A sua investigação demonstra como a governação centrada na equidade melhora tanto a implementação ética como os resultados educativos.

Estratégia de implementação

A abordagem de implementação enfatiza:

- Teste sistemático do enviesamento em algoritmos de orientação profissional
- Monitorização demográfica dos resultados educativos influenciados pela IA
- Desenvolvimento de metodologias de desenviesamento para dados de treino
- Desenvolvimento do corpo docente na integração da IA consciente da equidade
- Reações regulares das partes interessadas provenientes de diversas populações de estudantes.



Impacto documentado

O Instituto evidenciou:

- Identificação e correção de potenciais enviesamentos nos sistemas de orientação profissional
- Melhores resultados educativos para grupos historicamente sub-representados
- Reforço da confiança das partes interessadas na integração tecnológica
- Cumprimento com êxito dos requisitos de não-discriminação
- Desenvolvimento de metodologias transferíveis centradas na equidade.

3.4.6 Aplicação da ALTAI em contextos específicos de EFP

A Lista de Avaliação para uma Inteligência Artificial Fiável (ALTAI) tem sido aplicada com sucesso em numerosos contextos de ensino profissional, demonstrando a sua utilidade prática para a implementação ética.

A investigação de Zawacki-Richter et al. (2023) identifica a implementação da ALTAI no ensino profissional como particularmente valiosa devido aos "contextos profissionais concretos que caracterizam a formação profissional e as implicações explícitas na carreira das decisões influenciadas pela IA" (p. 15). A análise do seu estudo de caso demonstra a adaptabilidade da ALTAI a diversos ambientes de formação profissional.

Aplicações de avaliação de competências baseadas em IA

As instituições de ensino e formação profissional aplicaram a ALTAI para garantir a implementação ética de sistemas de avaliação automatizados:

Integração da supervisão humana: Garantir a validação das avaliações algorítmicas pelo instrutor

Verificação da equidade: Testar a equidade da avaliação em diversas populações de estudantes

Requisitos de transparência: Tornar os critérios de avaliação explícitos e compreensíveis

Robustez técnica: Assegurar um funcionamento fiável em diferentes padrões de desempenho dos alunos.



IA em Orientação Profissional e Aconselhamento Vocacional

A aplicação da ALTAI em sistemas de recomendação de carreira abordou:

- Potencial reforço de estereótipos profissionais
- Transparência nas metodologias de recomendação
- Proteção de dados na modelação preditiva de carreiras
- Acessibilidade dos sistemas de orientação para as diferentes populações.

Aprendizagem adaptativa com recurso à IA na formação profissional

A implementação de plataformas de aprendizagem adaptativa beneficiou da avaliação ALTAI através de:

- Avaliação da equidade da personalização algorítmica
- Verificação da diversificação do percurso de aprendizagem
- Avaliação dos potenciais efeitos das bolhas de filtragem
- Validação das capacidades de supervisão do instrutor.

3.4.7 Elementos comuns de políticas eficazes

A análise da implementação bem-sucedida nas instituições europeias de EFP revela padrões de política consistentes:

Estruturas de governação multiparticipadas que envolvem diversas perspetivas no desenvolvimento e na supervisão

Protocolos de avaliação de risco por níveis que aplicam um escrutínio proporcional com base no impacto potencial

Mecanismos de controlo contínuo em vez de processos de aprovação pontuais

Requisitos de transparência robustos que asseguram a compreensão das capacidades e limitações da IA pelas partes interessadas

Controlo humano significativo, especialmente nas decisões educativas consequentes

Ciclos regulares de revisão das políticas, adaptando-se à evolução das tecnologias e da regulamentação



Componentes educativos abrangentes para desenvolver a literacia em IA em todas as partes interessadas

Estruturas de responsabilização claras que estabelecem a responsabilidade pelos resultados influenciados pela IA.

Holmes et al. (2023) identificam estes elementos consistentes como representando um "consenso emergente relativamente aos princípios fundamentais de governação para a implementação da IA educacional" (p. 21). A sua análise comparativa em diversos contextos institucionais demonstra a adaptabilidade destes princípios fundamentais a ambientes educativos variados.

As instituições que demonstram uma prática exemplar equilibraram o potencial de inovação com salvaguardas adequadas, reconhecendo que abordagens demasiado restritivas podem limitar os benefícios educativos, enquanto uma governação insuficiente cria riscos significativos. Os quadros mais bem-sucedidos integraram considerações éticas ao longo do ciclo de vida da implementação, em vez de as tratarem como considerações *post-hoc*.

4. Competências éticas

4.1 Definição de competências éticas

As competências éticas referem-se à capacidade de avaliar se uma IA é utilizada de forma ética num contexto específico. No contexto educativo, por exemplo, as competências éticas consistem em saber se as políticas de um centro educativo são éticas, se a segurança e a privacidade dos dados são mantidas, etc.

As competências éticas referem-se às capacidades críticas necessárias para avaliar e garantir a utilização ética da IA e dos dados em contextos educativos, com base em princípios como a agência humana, a equidade, a humanidade e a escolha justificada. A agência humana implica permitir que os indivíduos façam escolhas informadas e assumam a responsabilidade pelas suas ações, promovendo a autonomia e a responsabilização. A equidade assegura um tratamento justo e a igualdade de acesso às oportunidades para todos, promovendo a inclusão e a não discriminação. A humanidade sublinha o respeito pela dignidade, identidade e bem-estar humanos, dando prioridade às ligações humanas significativas em detrimento da utilização instrumental dos dados. A escolha justificada requer processos transparentes de tomada de



decisão baseados em princípios éticos e na participação das partes interessadas, garantindo que os sistemas de IA na educação sejam implantados de forma responsável e beneficiem todas as partes envolvidas.

4.2 Importância das competências éticas

Estas considerações éticas são cruciais para orientar os educadores e as instituições educativas na navegação pelas complexidades da implementação da IA. Ao aderir a requisitos éticos como os propostos pelas Orientações de Ética da IA, as partes interessadas do setor educativo podem mitigar eficazmente os riscos, promover a justiça e defender o bem-estar dos alunos e do pessoal.

4.3 Áreas e exemplos de competências éticas

Área 1: Envolvimento profissional. Utilização de tecnologias digitais para comunicação, colaboração e desenvolvimento profissional.	
Elemento de competência	Indicadores potenciais
É capaz de descrever de forma crítica os impactos positivos e negativos da IA e da utilização de dados na educação	Participa ativamente na aprendizagem profissional contínua sobre a IA e a análise da aprendizagem e a sua utilização ética.
	É capaz de dar exemplos de sistemas de IA e descrever a sua relevância.
	Sabe como o impacto ético dos sistemas de IA é avaliado na escola.
	Sabe como iniciar e promover estratégias em toda a escola e na sua comunidade mais alargada que promovam uma utilização ética e responsável da IA e dos dados.
	Está consciente de que os algoritmos de IA funcionam de formas que normalmente não



Compreende os conceitos básicos da IA e da análise de aprendizagem	são visíveis ou facilmente compreendidas pelos utilizadores.
	É capaz de interagir e dar <i>feedback</i> ao sistema de IA para influenciar o que este recomenda a seguir.
	Está consciente de que os sensores utilizados em muitas tecnologias e aplicações digitais geram grandes quantidades de dados, incluindo dados pessoais, que podem ser utilizados para treinar um sistema de IA.
	Conhece as orientações éticas da UE em matéria de IA e os instrumentos de autoavaliação.
Exemplo 1: Ana é uma professora que quer dar uma aula de mecânica dos fluidos e quer utilizar a IA para gerar algumas imagens e incluí-las no powerpoint da disciplina. Antes de incluir a IA na aula, Ana leu as orientações éticas da UE sobre a IA e avaliou se é ético incluir ou não a sua utilização.	
Exemplo 2: Lucas leciona no curso de Educação e Controlo Ambiental. Vários professores da sua instituição educativa utilizam a IA, pelo que propôs uma reunião com toda a instituição para analisar a utilização da IA pelos professores. O objetivo de Lucas é garantir que todos os docentes fazem uma utilização ética e responsável da IA e dos dados.	



Área 2: Recursos digitais: Obter, criar e partilhar recursos digitais		
Elemento de competência	de	Elemento de competência
Governação de dados		Está consciente das diferentes formas de dados pessoais utilizados na educação e na formação.
		Está consciente das responsabilidades na manutenção da segurança e da privacidade dos dados.
		Sabe que o tratamento de dados pessoais está sujeito à regulamentação nacional e comunitária, incluindo o Regulamento Geral sobre a Proteção de Dados (RGPD).
		Sabe quem tem acesso aos dados dos alunos, como o acesso é monitorizado e durante quanto tempo os dados são conservados.
		Sabe que todos os cidadãos da UE têm o direito de não serem sujeitos a uma tomada de decisão totalmente automatizada.
		É capaz de dar exemplos de dados sensíveis, incluindo dados biométricos.
		É capaz de ponderar os benefícios e os riscos antes de permitir que terceiros tratem dados pessoais, especialmente quando utilizam sistemas de IA.
Governação da IA		Sabe que os sistemas de IA estão sujeitos a regulamentação nacional e comunitária (nomeadamente o Regulamento da Inteligência Artificial da UE a ser adotado).



	Capaz de explicar a abordagem baseada no risco do Regulamento da Inteligência Artificial da UE (a ser adotado).
	Conhece os casos de utilização de IA de alto risco na educação e os requisitos associados ao abrigo do Regulamento da Inteligência Artificial da UE (quando adotado).
	Sabe como incorporar conteúdos digitais editados/manipulados por IA no seu próprio trabalho e como esse trabalho deve ser creditado.
	Capacidade para explicar os princípios fundamentais da qualidade dos dados em sistemas de IA.
Exemplo 1: Maria é professora na escola de EFP de Eficiência Energética e Energia Solar Térmica. Está a criar notas para o seu curso, para entregar aos alunos, e incorporou alguns conteúdos digitais editados por IA nas notas, utilizando a ferramenta Copilot. Maria sabe que este conteúdo tem de ser creditado e certifica-se de que o faz corretamente.	
Exemplo 2: O Jaime dá aulas na escola de formação profissional de Fabricação Mecânica. Por vezes, utiliza o ChatGPT para desenvolver alguns conteúdos ou atividades da sua disciplina. A dada altura, avalia a possibilidade de incluir os dados pessoais dos alunos (nome, apelido, data de nascimento e morada) para criar um documento que inclua todas estas informações. Consciente da sua responsabilidade de manter a segurança e a privacidade dos dados, decidiu não incluir os dados pessoais.	



Área 3: Utilização de tecnologias e estratégias digitais para melhorar a avaliação		
Elemento de competência	de	Elemento de competência
Modelos de aprendizagem	de	Sabe que os sistemas de IA implementam o entendimento do designer sobre o que é a aprendizagem e como esta pode ser medida; pode explicar os principais pressupostos pedagógicos subjacentes a um determinado sistema de aprendizagem digital.
Objetivos da educação		Sabe de que forma um determinado sistema digital responde aos diferentes objetivos sociais da educação (qualificação, socialização, subjetivação).
Agência humana		Capaz de considerar o impacto do sistema de IA na autonomia dos professores, no desenvolvimento profissional e na inovação educativa.
		Considerar as fontes de enviesamento inaceitável na IA baseada em dados.
Equidade		Considera os riscos relacionados com a dependência emocional e a autoimagem dos alunos ao utilizar sistemas interativos de IA e análises de aprendizagem.
Humanidade		Capaz de considerar o impacto da IA e da utilização de dados na comunidade estudantil.
		Confiança para discutir os aspetos éticos da IA e a forma como estes influenciam a utilização da tecnologia.



Participa no desenvolvimento de práticas de aprendizagem que utilizam IA e dados	É capaz de explicar de que forma os princípios e valores éticos são considerados e negociados na co-conceção e co-criação de práticas de aprendizagem que utilizam a IA e os dados (associados à conceção da aprendizagem).
Exemplo 1: Luísa, professora de Manutenção Eletromecânica numa entidade de formação, implementa um <i>software</i> de IA que utiliza modelos de aprendizagem adaptativa para personalizar o ensino. O <i>software</i> avalia o nível inicial de cada formando e ajusta o conteúdo e as tarefas com base no seu progresso. Isto permite que cada formando avance ao seu próprio ritmo, recebendo <i>feedback</i> imediato e recursos específicos para ultrapassar as dificuldades. Além disso, Luísa pode aceder a relatórios detalhados para identificar as áreas em que os seus formandos necessitam de apoio adicional, assegurando um ensino inclusivo e eficaz.	
Exemplo 2: Javi, um professor de artes gráficas no curso de impressão gráfica de nível médio de uma escola de formação profissional, integra <i>software</i> de IA que melhora a capacidade humana, ajudando em tarefas de <i>design</i> complexas. O <i>software</i> ajuda os alunos automatizando tarefas repetitivas, como a correção de cores e os ajustes de <i>layout</i> , permitindo-lhes concentrar-se nos aspetos criativos e estratégicos dos seus projetos. Isto permite que os estudantes desenvolvam as suas capacidades artísticas e pensamento crítico, enquanto a IA trata dos pormenores técnicos.	



Área 4: Avaliação: Gerir e orquestrar a utilização das tecnologias digitais no ensino e na aprendizagem.

Elemento de competência	Indicadores potenciais
Diferenças pessoais	Consciente de que os alunos reagem de formas diferentes ao <i>feedback</i> automático.
Enviesamento algorítmico	Analisar as fontes de enviesamento inaceitável nos sistemas de IA e a forma como pode ser atenuado.
Foco cognitivo	Conscientes de que os sistemas de IA avaliam os progressos dos alunos com base em modelos de conhecimentos específicos de um domínio predefinido.
	Consciente de que a maioria dos sistemas de IA não avalia a colaboração, as competências sociais ou a criatividade.
	Consciente das formas comuns de manipular a avaliação baseada em IA.
Novas formas de utilização abusiva da tecnologia	Consciente de que os alunos reagem de formas diferentes ao <i>feedback</i> automático.

Exemplo 1: O Jorge leciona uma disciplina na escola de formação profissional de Instalações e Manutenção. Depois de terminar o 1.º tópico da disciplina, utiliza uma ferramenta de IA para avaliar a aquisição dos conteúdos pelos seus alunos. Além disso, esta ferramenta de IA fornece automaticamente *feedback* e adapta o nível das perguntas com base nas respostas anteriores. O Jorge está ciente da forma como os alunos reagem ao *feedback* automático, pelo que se certifica de que lhes dá *feedback* personalizado e verifica ele próprio o progresso dos alunos. Por conseguinte, utiliza a IA apenas como uma



ferramenta para fazer uma avaliação inicial, que é posteriormente verificada e revista por ele.

Exemplo 2: Natalia é formadora numa entidade de EFP na área das energias renováveis. Incentiva os alunos a utilizar o Windows Copilot nas suas aulas de formação para determinadas atividades. Embora esta ferramenta funcione como um assistente de IA que pode ser útil em termos de produtividade e criatividade, Natalia está consciente de que são necessários outros tipos de atividades e de avaliação para melhorar a colaboração, as competências sociais e a criatividade, uma vez que a IA não é capaz de avaliar esses aspetos. Continua a assumir as atividades e a avaliação relevantes como formadora e utiliza a IA como ferramenta durante as suas aulas e na preparação da formação.

Area 5: Capacitação dos Aprendentes: Utilizar tecnologias digitais para melhorar a inclusão, a personalização e o envolvimento ativo dos alunos

Elemento de competência	Indicadores potenciais
A IA responde às diferentes necessidades de aprendizagem dos alunos	Conhece as diferentes formas como os sistemas de aprendizagem personalizados podem adaptar o seu comportamento (conteúdo, percurso de aprendizagem, abordagem pedagógica).
	É capaz de explicar como um determinado sistema pode beneficiar todos os alunos, independentemente das suas diferenças cognitivas, culturais, económicas ou físicas.
	Está consciente de que os sistemas de aprendizagem digital tratam os diferentes grupos de alunos de forma diferente.



	É capaz de considerar o impacto no desenvolvimento da autoeficiência, da autoimagem, da mentalidade e das competências cognitivas e afetivas de autorregulação dos alunos.
Escolha justificada	Sabe que a IA e a utilização de dados podem beneficiar mais alguns alunos do que outros.
	É capaz de explicar que provas foram utilizadas para justificar a utilização de um determinado sistema de IA na sala de aula.
	Reconhece a necessidade de um controlo constante dos resultados da utilização da IA.
Exemplo 1: Teresa, professora de Formação Profissional Avançada em Administração e Finanças num estabelecimento de ensino para adultos, utiliza a IA para adaptar as experiências de aprendizagem. O sistema de IA avalia as necessidades de cada aluno e ajusta os recursos e as atividades em conformidade. Por exemplo, os alunos visuais podem receber ferramentas interativas, enquanto os alunos analíticos recebem estudos de caso detalhados. Esta abordagem garante uma aprendizagem personalizada que vai ao encontro de diversos estilos de aprendizagem e maximiza a participação dos alunos.	
Exemplo 2: Óscar, professor de Formação Profissional Avançada em <i>Design</i> de Fabrico Mecânico, utiliza <i>software</i> de IA para desenhar peças mecânicas. O <i>software</i> oferece várias soluções de conceção e simula o desempenho em diferentes condições. Os alunos analisam os resultados para selecionar o desenho mais eficiente, orientados por Óscar para desenvolverem competências críticas e analíticas cruciais para as suas carreiras.	



Área 6: Promoção da competência digital dos aprendentes, permitindo-lhes utilizar de forma criativa e responsável as tecnologias digitais para fins de informação, comunicação, criação de conteúdos, bem-estar e resolução de problemas

Elemento de competência	Indicadores potenciais
Ética da IA e da análise da aprendizagem	É capaz de utilizar projetos e implantações de IA para ajudar os alunos a aprender sobre a ética da IA e a utilização de dados na educação e na formação.

Exemplo 1: A professora Sara ensina análise avançada de marketing aos seus alunos do ensino profissional numa entidade de ensino superior. Ela discute um estudo de caso em que a IA analisa dados de redes sociais para publicidade direcionada. Sara promove um debate sobre questões éticas como a privacidade, o consentimento e o enviesamento algorítmico. Orienta os alunos no sentido de desenvolverem orientações éticas para a utilização da IA no marketing, promovendo a sua compreensão da utilização responsável da IA no setor.

Exemplo 2: O Professor Jesús ensina gestão culinária, centrando-se no papel da IA na personalização dos menus dos restaurantes. Discute um estudo de caso em que a IA analisa as preferências alimentares para oferecer recomendações de menus personalizados. Jesús envolve os formandos num debate sobre questões éticas como a privacidade dos dados, a garantia da diversidade dos menus e o impacto da IA na satisfação do cliente. Orienta os formandos no desenvolvimento de quadros éticos para a utilização da IA na personalização de menus, preparando-os para equilibrar a inovação com considerações éticas nas operações dos restaurantes.



5. Conclusão

À medida que a IA continua a moldar os cenários educativos, é essencial que a adoção de tecnologias de IA seja feita de uma forma que defenda os valores éticos, promova a justiça e salvguarde a privacidade dos alunos. Este manual delineou o objetivo e o âmbito da implementação da IA, sublinhando a importância das considerações éticas para garantir que as tecnologias de IA são utilizadas de forma responsável em todos os contextos educativos.

5.1 Principais conclusões

A integração da IA na educação tem um potencial imenso, mas deve ser orientada por princípios éticos que dão prioridade à diversidade, à inclusão e à transparência. As ferramentas de IA devem ser concebidas para acomodar alunos de diversas origens, tendo em conta as diferenças culturais, linguísticas e socioeconómicas para promover a equidade nos resultados educativos. A transparência é igualmente vital - educadores e alunos devem entender como os sistemas de IA tomam decisões, especialmente em áreas como classificação, feedback ou caminhos de aprendizagem personalizados. Garantir a privacidade e a proteção de dados é essencial, exigindo total conformidade com regulamentos como o RGPD e a adesão a práticas éticas de gerenciamento de dados. Os sistemas de IA também devem ser sustentáveis, minimizando o seu impacto ambiental e garantindo a sua viabilidade a longo prazo nos ecossistemas educativos.

Igualmente importante é a promoção da literacia em IA entre todas as partes interessadas. Os educadores precisam de formação específica para integrar de forma ética e eficaz as ferramentas de IA na sua pedagogia, enquanto os estudantes devem estar equipados com uma compreensão crítica das capacidades e limitações da IA. Os administradores institucionais desempenham um papel fundamental no alinhamento das políticas escolares ou universitárias com práticas éticas de IA, promovendo um ambiente de responsabilidade e justiça. Os decisores políticos devem intervir para criar quadros regulamentares robustos que orientem a adoção responsável da IA, mitiguem o enviesamento algorítmico e evitem resultados discriminatórios. Em última análise, é crucial uma abordagem de várias partes interessadas para garantir que a IA melhore a educação de uma forma justa, inclusiva e sustentável.

É, também, muito importante ter em conta alguns quadros jurídicos e regulamentares fundamentais, como o Regulamento da Inteligência Artificial da



UE (2024), que introduz um quadro jurídico abrangente que categoriza os sistemas de IA com base no seu risco para os direitos fundamentais, com implicações significativas para as instituições de ensino.

Com base nisso, existem as aplicações de IA de alto risco, como as utilizadas em avaliações ou admissões de estudantes, que estão sujeitas a requisitos de conformidade rigorosos, incluindo documentação rigorosa, supervisão humana e protocolos sólidos de gestão de riscos.

Além disso, a transparência e a responsabilização não são uma opção. As instituições devem divulgar claramente o funcionamento dos sistemas de IA e garantir que os utilizadores possam compreender e contestar as decisões automatizadas. Além disso, a lei impõe medidas proactivas para auditar e atenuar o enviesamento algorítmico, garantindo a equidade e evitando resultados discriminatórios. Para apoiar estes esforços, a Lista de Avaliação para uma IA Fiável (ALTAI) oferece uma ferramenta prática que ajuda as instituições a avaliar o impacto ético dos seus sistemas de IA, promovendo uma cultura de implantação de IA responsável e fiável na educação.

5.2 Reflexões finais

Todos devemos estar conscientes do facto de que, à medida que a IA continua a remodelar o ensino universitário e profissional, este manual sublinha a importância de manter um equilíbrio cuidadoso entre práticas inovadoras e responsabilidade ética. Os benefícios da IA, que vão desde percursos de aprendizagem personalizados até à melhoria da eficiência administrativa, são significativos, mas comportam riscos que não devem ser negligenciados ou ignorados. Questões como o enviesamento algorítmico, as violações da privacidade dos dados e a diminuição do papel do julgamento humano exigem uma atenção constante. Para navegar nestas complexidades, deve haver uma colaboração constante entre educadores, administradores e decisores políticos. Naturalmente, os princípios éticos devem ser profundamente integrados no desenvolvimento, na implementação e na avaliação das ferramentas de IA para garantir que apoiam a inclusão, respeitam a autonomia e reforçam a confiança no processo educativo.

Olhando para o futuro, a integração bem-sucedida da IA na educação dependerá não só da preparação tecnológica, mas também de uma visão partilhada assente na transparência, na equidade e na aprendizagem ao longo da vida. O setor deve promover a literacia em IA entre todas as partes interessadas e criar sistemas sólidos de responsabilização e supervisão. Não se trata de um ajustamento pontual, mas de um percurso evolutivo que exige



reflexão e diálogo contínuos. Ao defender abordagens centradas no ser humano e a inovação ética, a comunidade educativa pode transformar a IA num poderoso aliado que melhora as oportunidades de aprendizagem, preservando simultaneamente os valores da educação de qualidade. Para uma orientação abrangente, as partes interessadas são encorajadas a consultar os recursos do projeto AI Pioneers e o Esquema de avaliação da IA na educação, que estão acessíveis em várias línguas.

Agradecimentos

Olhando para o futuro, a integração bem-sucedida da IA na educação dependerá não só da preparação tecnológica, mas também de uma visão partilhada assente na transparência, na equidade e na aprendizagem ao longo da vida. O setor deve promover a literacia em IA entre todas as partes interessadas e criar sistemas sólidos de responsabilização e supervisão. Não se trata de um ajustamento pontual, mas de um percurso evolutivo que exige reflexão e diálogo contínuos. Ao defender abordagens centradas no ser humano e a inovação ética, a comunidade educativa pode transformar a IA num poderoso aliado que melhora as oportunidades de aprendizagem, preservando simultaneamente os valores da educação de qualidade. Para uma orientação abrangente, as partes interessadas são encorajadas a consultar os recursos do projeto AI Pioneers e o Esquema de avaliação da IA na educação, que estão acessíveis em várias línguas.

Referências

- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.
- Aldosari, S. A. M. (2020). The future of higher education in the light of artificial intelligence transformations. *International Journal of Higher Education*, 9(3), 145-151.
- Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, 41(2), 101929.
- Attwell, G., Bekiaridis, G., Deitmer, L., Perini, M., Roppertz, S., & Tütlys, V. (2020). Artificial intelligence in policies, processes and practices of vocational education and training.



- Attwell, G., Deitmer, L., Tütlys, V., Roppertz, S., & Perini, M. (2020). Digitalisation, artificial intelligence and vocational occupations and skills: what are the needs for training teachers and trainers?. In *Trends in vocational education and training research, Vol. III. Proceedings of the European Conference on Educational Research (ECER), Vocational Education and Training Network (VETNET)* (pp. 30-42).
- Bekiaridis, G., & Attwell, G. (2024). Integrating Artificial Intelligence in Vocational and Adult Education: A Supplement to the DigCompEdu Framework. *Ubiquity Proceedings*, 4(1).
- Boninger, F., Molnar, A., & Saldaña, C. (2020). Big Claims, Little Evidence, Lots of Money: The Reality behind the Summit Learning Program and the Push to Adopt Digital Personalized Learning Programs. *Commercialism in Education Research Unit*.
- Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational inequality, techno-solutionism and inclusive tools. *Sustainability*, 16(781). <https://doi.org/10.3390/su16020781>
- Celik, I. (2023). Exploring the determinants of artificial intelligence (Ai) literacy: Digital divide, computational thinking, cognitive absorption. *Telematics and Informatics*, 83, 102026.
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://link.springer.com/article/10.1186/s41239-023-00408-3>
- Chew, E., & Chua, X. N. (2020). Robotic Chinese language tutor: Personalising progress assessment and feedback or taking over your job? *On the Horizon*, 28(3), 113–124. <https://doi.org/10.1108/OTH-04-2020-0015>
- Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). First-year students AI-competence as a predictor for intended and de facto use of AI-tools for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 18.
- European Commission, Directorate-General for Education, Youth, Sport, and Culture. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/153756>



European Parliament. (2024). EU AI Act: First regulation on artificial intelligence. Retrieved from <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. Official Journal of the European Union, L series, 1–144.

Emeršič, Ž., Hrastnik, G., Meh Peer, N., & Peer, P. (2025). AIM@VET-inspired university level education strategies for teaching computer vision and biometrics. *ROSUS 2025 – Računalniška obdelava slik in njena uporaba v Sloveniji 2025: Zbornik 19. strokovne konference*. <https://doi.org/10.18690/um.feri.2.2025.4>

Hodgkin, E., & Kumar, A. (2023). Algorithmic fairness in vocational guidance systems: A comparative analysis of implementation frameworks. *Educational Technology Research and Development*, 71(3), 215-231.

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign

JISC (2022). *AI in tertiary Education. A summary of the current state of play*. JISC Repository. Retrieved from <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020). Automated personalized feedback improves learning gains in an intelligent tutoring system. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II* 21 (pp. 140-146). Springer International Publishing.

Kumar, V., & Boulanger, D. (2020, October). Explainable automated essay scoring: Deep learning really has pedagogical value. In *Frontiers in education* (Vol. 5, p. 572367). Frontiers Media SA.

Li, Q. (2021). The use of artificial intelligence combined with cloud computing in the design of education information management platform. *International Journal*



of Emerging Technologies in Learning, 16(5), 32–44.

<https://doi.org/10.3991/ijet.v16i05.20309>

Porter, B., & Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability*, 12(24), 10243.

Prinsloo, P., & Knox, J. (2023). Exploring the ethical principles for the implementation of artificial intelligence in education: Towards a future agenda. *Research and Practice in Technology Enhanced Learning*, 18(1), 73-94. Retrieved from https://www.researchgate.net/publication/368740229_Exploring-the-Ethical-Principles-for-the-Implementation-of-Artificial-Intelligence-in-Education-Towards-a-Future-Agenda

Ramirez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., & Miranda, J. (2023). Artificial intelligence in vocational education: Ethical challenges and governance frameworks. *International Journal of Educational Technology in Higher Education*, 20(1), 138-155. <https://doi.org/10.1186/s41239-023-00408-3>

Rodriguez-Abitia, G., & Bribiesca-Correa, G. (2021). Assessing the ethical impact of artificial intelligence in education. *RED. Revista de Educación a Distancia*, 6(1). <https://doi.org/10.34627/redvol6iss1e202303>

Roman Etxebarrieta, G., Orcasitas-Vicandi, M., & Antzaka, A. (2024). *Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)*. AIPioneers.org. Retrieved from: <https://aipioneers.org/evaluation-schema-for-ai-in-education-on-data-privacy-ethics-and-eu-values-wp5/>

Roppertz, S. (2020). Artificial Intelligence and Vocational Education and Training—Perspective of German Vet Teachers. In *European Distance and E-Learning Network (EDEN) Conference Proceedings* (No. 2, pp. 207-216). European Distance and E-Learning Network.

Rott, K. J., Lao, L., Petridou, E., & Schmidt-Hertha, B. (2022). Needs and requirements for an additional AI qualification during dual vocational training: Results from studies of apprentices and teachers. *Computers and education: Artificial intelligence*, 3, 100102.

Şenocak, D., Bozkurt, A., & Koçdar, S. (2024). Exploring the Ethical Principles for the Implementation of Artificial Intelligence in Education: Towards a Future Agenda. In *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 200-213). IGI Global.



Seufert, S. (2024). Artificial Intelligence in Vocational Education and Training (VET): Evaluating VET Leaders' Acceptance of AI in Switzerland.

Tang, J., & Hai, L. (2021). Construction and exploration of an intelligent evaluation system for educational app through artificial intelligence technology. *International Journal of Emerging Technologies in Learning*, 16(5), 17–31. <https://doi.org/10.3991/ijet.v16i05.20293>

Tuomi, I., Cachia, R., & Villar-Onrubia, D. (2023). On the futures of technology in education: Emerging trends and policy implications. *Publications Office of the European Union, Luxembourg*.

UNESCO (2023a). *ChatGPT and artificial intelligence in higher education*. Retrieved from: https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

UNESCO (2023b). *Guidance for generative AI in education and research*. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18, 1-20.

Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S., & Roman-Cañizares, M. (2021). Implementation of a virtual assistant for the academic management of a university with the use of artificial intelligence. *Future Internet*, 13(4), 97.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235.

Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1), 1–5. <https://doi.org/10.1080/17439884.2023.2167830>

Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology*, 95-116.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2023). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational*



Cofinanciado pela
União Europeia



AI PIONEERS
Artificial intelligence in education & training

Technology in Higher Education, 20(1). <https://doi.org/10.1186/s41239-023-00408-3>

C O N T A C T O S

www.aipioneers.org

gorka.roman@ehu.es

maria.orcasitas@ehu.es

natalia.louleli@ehu.eus



AI PIONEERS
Artificial Intelligence in education & training

Εγχειρίδιο για την Πολιτική και την Ηθική στην Εκπαίδευση στην Τεχνητή Νοημοσύνη

Ιούνιος 2025

Κύριος δικαιούχος του παραδοτέου: Universidad del País Vasco/Euskal
Herriko Πανεπιστήμιο (UPV/EHU)

Πακέτο Εργασίας 5

Παραδοτέο D5.1 "Εγχειρίδιο για την Πολιτική και τη Δεοντολογία στην
Εκπαίδευση στην Τεχνητή Νοημοσύνη"

Αριθμός έργου: 101087261

Όνομα έργου: Τεχνητή Νοημοσύνη και το
μέλλον της εκπαίδευσης Ακρωνύμιο έργου:

Πρωτοπόροι της Τεχνητής Νοημοσύνης

Συντονιστής έργου: Πανεπιστήμιο της Βρέμης, Ινστιτούτο Τεχνολογίας και
Εκπαίδευσης (ITB) Πρόσκληση: ERASMUS-EDU-2022-PI-FORWARD

Θέμα: ERASMUS-EDU-2022-PI-FORWARD-LOT1

Είδος δράσης: Επιχορηγήσεις εφάπαξ ποσού ERASMUS

Χορηγούσα αρχή: Εκτελεστικός Οργανισμός Ευρωπαϊκής
Εκπαίδευσης και Πολιτισμού Ημερομηνία έναρξης έργου:

01/01/2023

Ημερομηνία λήξης έργου:

31/12/2025 Διάρκεια έργου:

36 μήνες



Co-funded by
the European Union



AI PIONEERS
Artificial intelligence in education & training



AI PIONEERS
Artificial Intelligence in education & training

Ελληνική έκδοση του παραδοτέου

Ιούνιος 2025



Εγχειρίδιο για την Πολιτική και την Ηθική στην Εκπαίδευση στην Τεχνητή Νοημοσύνη

Gorka Roman Etxebarrieta (Συγγραφέας)

Maria Orcasitas-Vicandi (Συγγραφέας)

Natalia Louleli (Συγγραφέας)

Asier León Nuñez (Συγγραφέας)

Itsaso Biota Piñeiro (Συγγραφέας)

Ana Boyano Murillo (Συγγραφέας)

Raquel Justo Blanco (Συγγραφέας)

Josune Rodriguez Negro (Συγγραφέας)



Πίνακας περιεχομένων

Εγχειρίδιο για την Πολιτική και την Ηθική στην Εκπαίδευση στην Τεχνητή Νοημοσύνη	3
1. Εισαγωγή.....	5
1.1 Σκοπός και πεδίο εφαρμογής του εγχειριδίου	6
1.2 Η σημασία των ηθικών παραμέτρων στην υιοθέτηση της Τεχνητής Νοημοσύνης	7
1.3 Ένα εγχειρίδιο για εκπαιδευτικούς, διοικητικούς υπαλλήλους και υπεύθυνους χάραξης πολιτικής	9
2. Κατανόηση της Τεχνητής Νοημοσύνης στην ΕΕΚ και την Εκπαίδευση Ενηλίκων	12
2.1. Ο ρόλος της Τεχνητής Νοημοσύνης στην Εκπαίδευση	12
2.2. Τεχνητή Νοημοσύνη στην Εκπαίδευση και Κατάρτιση Ενηλίκων και Επαγγελματικής Κατάρτισης.....	14
3. Ανάπτυξη ηθικών κατευθυντήριων γραμμών	15
3.1 Σημασία των ηθικών πολιτικών στην εκπαίδευση στην Τεχνητή Νοημοσύνη..	15
3.2 Βασικές απαιτήσεις για αξιόπιστα συστήματα Τεχνητής Νοημοσύνης.....	18
3.3 Λαμβάνοντας υπόψη τις νομικές πτυχές της εφαρμογής της Τεχνητής Νοημοσύνης στις Σχολές Επαγγελματικής Εκπαίδευσης και Κατάρτισης.....	21
3.4 Μελέτες Περιπτώσεων και Παραδείγματα Αποτελεσματικών Πολιτικών Τεχνητής Νοημοσύνης σε Ιδρύματα Επαγγελματικής Εκπαίδευσης και Κατάρτισης	29
4. Ηθικές Δεξιότητες	38
4.1 Ορισμός Ηθικών Δεξιοτήτων	38
4.2 Σημασία ηθικών δεξιοτήτων	39
4.3 Ηθικά Δεξιότητες περιοχές και παραδείγματα	39
5. Συμπέρασμα	49
5.1 Βασικά συμπεράσματα	50
5.2 Τελικές σκέψεις.....	51
Ευχαριστίες	52
Αναφορές.....	53
	4

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



1. Εισαγωγή

Το έργο AI Pioneers επικεντρώνεται στην προώθηση της ενσωμάτωσης της Τεχνητής Νοημοσύνης (TN) στην εκπαίδευση, ιδίως στην Εκπαίδευση Ενηλίκων και την Επαγγελματική Εκπαίδευση και Κατάρτιση (EEK). Αυτό το Προοδευτικό Έργο ERASMUS+ ξεκίνησε τον Ιανουάριο του 2023 και χρηματοδοτείται έως τον Δεκέμβριο του 2025.

Με την αυξανόμενη επιρροή της Τεχνητής Νοημοσύνης στην εκπαίδευση και την αυξανόμενη ζήτηση για επαγγελματική υποστήριξη στη χρήση της Τεχνητής Νοημοσύνης και την ανάπτυξη δεξιοτήτων, το έργο στοχεύει στη διατήρηση ενός ολοκληρωμένου, παγκόσμιου δικτύου Πρωτοπόρων της Τεχνητής Νοημοσύνης σε όλες τις χώρες εταίρους και πέραν αυτών. Η δημιουργία του Δικτύου Αναφοράς Πρωτοπόρων της Τεχνητής Νοημοσύνης είναι ένα βασικό στοιχείο, που απευθύνεται σε εκπαιδευτικούς, υπεύθυνους χάραξης πολιτικής, εκπαιδευτικούς σχεδιαστές, προγραμματιστές Τεχνητής Νοημοσύνης και άλλα σχετικά ενδιαφερόμενα μέρη. Με την εμπλοκή όσων έχουν επενδύσει περισσότερο στη χρήση της Τεχνητής Νοημοσύνης στην εκπαίδευση, το έργο στοχεύει στη δημιουργία ενός κεντρικού κόμβου για την υποστήριξη της ενσωμάτωσης της Τεχνητής Νοημοσύνης στην Εκπαίδευση Ενηλίκων και την EEK, ενθαρρύνοντας μελλοντικές πρωτοβουλίες και εκπαιδευτικά έργα Τεχνητής Νοημοσύνης. Περισσότερες πληροφορίες σχετικά με το δίκτυο και τον τρόπο συμμετοχής μπορείτε να βρείτε μέσω της [ιστοσελίδας του έργου](#).

Ένας άλλος σημαντικός στόχος ήταν η ανάπτυξη ενός Συμπληρώματος στο Πλαίσιο DigCompEDU (Bekiaridis , 2024). Το παρόν έγγραφο περιγράφει τις απαραίτητες δεξιότητες και ικανότητες που χρειάζονται οι εκπαιδευτικοί για να κατανοήσουν, να αξιολογήσουν και να χρησιμοποιήσουν τις τεχνολογίες Τεχνητής Νοημοσύνης στην εκπαίδευση, με ιδιαίτερη έμφαση στην Εκπαίδευση Ενηλίκων και την EEK. Αυτό το συμπλήρωμα ευθυγραμμίζει τις ικανότητες Τεχνητής Νοημοσύνης με τους έξι βασικούς τομείς του Πλαισίου DigCompEdu (Redecker , 2017): Επαγγελματική Συμμετοχή, Ψηφιακοί Πόροι, Διδασκαλία και Μάθηση, Αξιολόγηση, Ενδυνάμωση των Μαθητών και Διευκόλυνση της Ψηφιακής Ικανότητας των Μαθητών. Για κάθε περιοχή , η συμπλήρωμα :

1. Αναλύει εφαρμογές Τεχνητής Νοημοσύνης στην περιοχή

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



2. Προτείνει δραστηριότητες για την ανάπτυξη σχετικών δεξιοτήτων
3. Προτείνει επίπεδα προόδου για την ανάπτυξη ικανοτήτων
4. Εντοπίζει προκλήσεις και παρέχει στρατηγικές για την αντιμετώπισή τους

Το έργο συνεχίζει να αναπτύσσει και να μοιράζεται μια σειρά πόρων μέσω του AI Pioneers Toolkit, το οποίο είναι διαθέσιμο στον ιστότοπο του έργου. Αυτό το κιτ εργαλείων προσφέρει συστάσεις για το πρόγραμμα σπουδών για την ενσωμάτωση της Τεχνητής Νοημοσύνης στην εκπαίδευση, επισημαίνει τα οφέλη και τους κινδύνους της χρήσης της Τεχνητής Νοημοσύνης σε εκπαιδευτικά περιβάλλοντα και παρέχει κατευθυντήριες γραμμές βέλτιστων πρακτικών για την Τεχνητή Νοημοσύνη στην εκπαίδευση (Tommasi & Perini, 2024). Περιλαμβάνει επίσης σύντομες επισκοπήσεις εργαλείων για διαφορετικούς διδακτικούς σκοπούς, όπως βοηθούς μάθησης, εργαλεία για τη δημιουργία εκπαιδευτικού υλικού και εργαλεία που υποστηρίζουν τη μάθηση από ομοτίμους.

Το έργο επιτελεί σημαντικό έργο στην αντιμετώπιση των ηθικών προκλήσεων που θέτει η Τεχνητή Νοημοσύνη σε εκπαιδευτικά περιβάλλοντα. Το παρόν έγγραφο συνοδεύει ένα σχήμα αξιολόγησης και αμφότερα έχουν σχεδιαστεί για να καθοδηγήσουν τα ιδρύματα και τους εκπαιδευτικούς στην υπεύθυνη εφαρμογή της Τεχνητής Νοημοσύνης, λαμβάνοντας παράλληλα υπόψη τις επιπτώσεις στην ιδιωτικότητα και τις ηθικές επιπτώσεις.

Το *Σχήμα Αξιολόγησης για την Τεχνητή Νοημοσύνη στην Εκπαίδευση* εστιάζει σε πρακτικές κατευθυντήριες γραμμές για την υιοθέτηση εργαλείων Τεχνητής Νοημοσύνης, διασφαλίζοντας ότι ευθυγραμμίζονται με τις ηθικές αρχές και τις αξίες της ΕΕ. Αποτελεί έναν εξαιρετικό πόρο για τα εκπαιδευτικά ιδρύματα ώστε να λαμβάνουν τεκμηριωμένες αποφάσεις σχετικά με την ενσωμάτωση της Τεχνητής Νοημοσύνης.

Και αυτό, το *Εγχειρίδιο για την Πολιτική και την Ηθική στην Εκπαίδευση στην Τεχνητή Νοημοσύνη*, είναι ένας πιο ολοκληρωμένος οδηγός, που προσφέρει πολιτικές και ηθικά πλαίσια για την αντιμετώπιση των τρεχόντων ανησυχιών.

1.1 Σκοπός και πεδίο εφαρμογής του εγχειριδίου

Το *Εγχειρίδιο για την Πολιτική και την Ηθική στην Εκπαίδευση στην Τεχνητή Νοημοσύνη* χρησιμεύει ως θεμελιώδης πηγή για εκπαιδευτικούς, υπεύθυνους χάραξης πολιτικής και ιδρύματα που στοχεύουν στην πλοήγηση στην πολυπλοκότητα της Τεχνητής Νοημοσύνης σε εκπαιδευτικά περιβάλλοντα. Βοηθά στη δημιουργία ενός πλαισίου που μπορεί να διασφαλίσει ότι οι ηθικές



παράμετροι βρίσκονται στο επίκεντρο , προστατεύοντας το απόρρητο των μαθητών, προωθώντας τη δικαιοσύνη και ενθαρρύνοντας την υπεύθυνη χρήση.

Ένα βασικό πλεονέκτημα είναι η εφαρμόσιμη καθοδήγηση που παρέχει στους εκπαιδευτικούς, δίνοντάς τους τη δυνατότητα να χρησιμοποιούν την Τεχνητή Νοημοσύνη υπεύθυνα και βοηθώντας τους μαθητές να κατανοήσουν τις δυνατότητές της, τους περιορισμούς και τις ηθικές επιπτώσεις της. Αυτό θα είναι κρίσιμο για την καλλιέργεια μιας γενιάς μαθητών που δεν είναι μόνο τεχνολογικά καταρτισμένοι, αλλά και ηθικά συνειδητοποιημένοι όταν ασχολούνται με την τεχνολογία. Για τους υπεύθυνους χάραξης πολιτικής, το συνεπές πλαίσιο που προσφέρει το εγχειρίδιο καθοδηγεί τους χρήστες σε πιο τυποποιημένες προσεγγίσεις σε όλα τα ιδρύματα, κάτι που είναι απαραίτητο για τον μετριασμό των κινδύνων της Τεχνητής Νοημοσύνης, όπως οι προκαταλήψεις σε αλγόριθμους ή οι παραβιάσεις της ιδιωτικής ζωής. Μπορεί επίσης να βοηθήσει να διασφαλιστεί ότι η Τεχνητή Νοημοσύνη δεν ενσωματώνεται μόνο για τον εαυτό της, αλλά μάλλον με τρόπο που βελτιώνει την εκπαιδευτική εμπειρία χωρίς να διακυβεύονται τα ηθικά πρότυπα.

1.2 Η σημασία των ηθικών παραμέτρων στην υιοθέτηση της Τεχνητής Νοημοσύνης

Οι ηθικές παράμετροι είναι κρίσιμες για την υιοθέτηση της Τεχνητής Νοημοσύνης, ιδίως σε τομείς όπως η εκπαίδευση ενηλίκων και η επαγγελματική κατάρτιση (ΕΕΚ), όπου ο αντίκτυπος της Τεχνητής Νοημοσύνης στους μαθητές, τους εκπαιδευτικούς και τα ιδρύματα μπορεί να είναι σημαντικός. Αυτές οι παράμετροι συμβάλλουν στη διασφάλιση ότι η Τεχνητή Νοημοσύνη χρησιμοποιείται με υπευθυνότητα, χωρίς αποκλεισμούς και διαφάνεια, και ότι υποστηρίζει τους ευρύτερους στόχους της εκπαίδευσης χωρίς να θέτει σε κίνδυνο τις θεμελιώδεις αξίες.

Ορισμένες βασικές παραμέτρους που καλύπτονται από το σχήμα περιλαμβάνουν:

Ποικιλομορφία και Ένταξη : Τα συστήματα και τα εργαλεία Τεχνητής Νοημοσύνης πρέπει να σχεδιάζονται ώστε να καλύπτουν ένα ευρύ φάσμα μαθητών, συμπεριλαμβανομένων εκείνων με διαφορετικό υπόβαθρο, ικανότητες και μαθησιακές ανάγκες. Στην εκπαίδευση ενηλίκων και την ΕΕΚ, οι μαθητές συχνά προέρχονται από ποικίλες εμπειρίες ζωής, όπως διαφορετικά επίπεδα προηγούμενης εκπαίδευσης, ηλικιακές ομάδες και εργασιακή εμπειρία. Η Τεχνητή Νοημοσύνη θα πρέπει να είναι ευέλικτη και συμπεριληπτική, διασφαλίζοντας ότι η χρήση της υποστηρίζει και όχι αποκλείει οποιαδήποτε



συγκεκριμένη ομάδα. Οι ηθικοί υιοθετητές της Τεχνητής Νοημοσύνης πρέπει να γνωρίζουν τους τρόπους με τους οποίους η τεχνολογία θα μπορούσε να διαιωνίσει προκαταλήψεις ή να δημιουργήσει ψηφιακά χάσματα, προκειμένου να τα μετριάσει.

Διαφάνεια : Για να υιοθετηθεί ηθικά η Τεχνητή Νοημοσύνη σε εκπαιδευτικά περιβάλλοντα, η διαφάνεια είναι απαραίτητη. Τα ιδρύματα πρέπει να είναι σαφή σχετικά με τον τρόπο λειτουργίας των συστημάτων Τεχνητής Νοημοσύνης, τα δεδομένα που συλλέγουν και τις διαδικασίες λήψης αποφάσεων που τα διέπουν. Οι μαθητές και οι εκπαιδευτικοί θα πρέπει να κατανοούν πώς χρησιμοποιούνται τα εργαλεία Τεχνητής Νοημοσύνης στην τάξη ή στο περιβάλλον εκπαίδευσης και πώς τα συστήματα μπορούν να επηρεάσουν τα μαθησιακά αποτελέσματα. Αυτή η διαφάνεια ενισχύει την εμπιστοσύνη στις τεχνολογίες Τεχνητής Νοημοσύνης και διασφαλίζει ότι υπάρχει λογοδοσία στη χρήση τους.

Ιδιωτικότητα και Προστασία Δεδομένων : Τα ιδρύματα εκπαίδευσης ενηλίκων και επαγγελματικής κατάρτισης συχνά ασχολούνται με ευαίσθητα προσωπικά δεδομένα, όπως προφίλ μαθητών, αποτελέσματα αξιολόγησης και ιστορικό σταδιοδρομίας. Η ηθική χρήση της Τεχνητής Νοημοσύνης σε αυτό το πλαίσιο πρέπει να δίνει προτεραιότητα στην ιδιωτικότητα και την προστασία των δεδομένων. Τα ιδρύματα πρέπει να διασφαλίζουν τη συμμόρφωση με τους νόμους περί προστασίας δεδομένων (π.χ., ΓΚΠΔ) και να διασφαλίζουν ότι τα δεδομένα των μαθητών συλλέγονται, αποθηκεύονται και χρησιμοποιούνται υπεύθυνα. Η συγκατάθεση για τη συλλογή δεδομένων θα πρέπει να είναι ενημερωμένη και εθελοντική, και οι μαθητές θα πρέπει να έχουν τον έλεγχο των προσωπικών τους πληροφοριών.

Βιωσιμότητα : Η Τεχνητή Νοημοσύνη θα πρέπει να ενσωματωθεί με τρόπο που να ευθυγραμμίζεται με τη μακροπρόθεσμη βιωσιμότητα των εκπαιδευτικών συστημάτων. Αυτό σημαίνει ότι πρέπει να λαμβάνεται υπόψη τόσο ο περιβαλλοντικός αντίκτυπος των τεχνολογιών Τεχνητής Νοημοσύνης όσο και ο ρόλος τους στην προώθηση βιώσιμων πρακτικών στο πλαίσιο των προγραμμάτων εκπαίδευσης και κατάρτισης. Στο πλαίσιο της ΕΕΚ, η Τεχνητή Νοημοσύνη μπορεί επίσης να υποστηρίξει τη βιωσιμότητα στις βιομηχανίες, προετοιμάζοντας τους μαθητές με δεξιότητες που προωθούν τις πράσινες τεχνολογίες και τις φιλικές προς το περιβάλλον πρακτικές.

Δεξιότητες Γραμματισμού στην Τεχνητή Νοημοσύνη : Καθώς η Τεχνητή Νοημοσύνη ενσωματώνεται όλο και περισσότερο στην εκπαίδευση, είναι ζωτικής σημασίας να ενισχυθεί ο γραμματισμός στην Τεχνητή Νοημοσύνη τόσο μεταξύ



των εκπαιδευτικών όσο και των εκπαιδευόμενων. Στην εκπαίδευση ενηλίκων και την ΕΕΚ, αυτό περιλαμβάνει τον εξοπλισμό των εκπαιδευόμενων με τις γνώσεις και τις δεξιότητες για να κατανοήσουν τις τεχνολογίες Τεχνητής Νοημοσύνης και να τις χρησιμοποιήσουν αποτελεσματικά στην καριέρα τους. Αυτή η ενδυνάμωση επιτρέπει στους εκπαιδευόμενους να πλοηγούνται στην εξελισσόμενη αγορά εργασίας και να διασφαλίζουν ότι δεν θα μείνουν πίσω. Για τους εκπαιδευτικούς, ο γραμματισμός στην Τεχνητή Νοημοσύνη τους βοηθά να αξιολογούν κριτικά τα εργαλεία Τεχνητής Νοημοσύνης και να τα ενσωματώνουν με τρόπους που ενισχύουν τη μάθηση χωρίς να διακυβεύονται τα ηθικά πρότυπα.

Δικαιοσύνη και Ισότητα : Τα συστήματα Τεχνητής Νοημοσύνης θα πρέπει να σχεδιάζονται έτσι ώστε να αντιμετωπίζουν όλους τους μαθητές δίκαια, διασφαλίζοντας ότι καμία ομάδα δεν θα μειονεκτεί από αυτοματοποιημένες αποφάσεις ή αλγόριθμους. Στην εκπαίδευση ενηλίκων και την ΕΕΚ, αυτό είναι ιδιαίτερα σημαντικό, καθώς αυτοί οι τομείς συχνά εξυπηρετούν μαθητές που ενδέχεται να αντιμετωπίζουν ήδη κοινωνικές ή οικονομικές προκλήσεις. Είναι κρίσιμο η Τεχνητή Νοημοσύνη να μην ενισχύει τις υπάρχουσες ανισότητες ή να δημιουργεί νέα εμπόδια στην πρόσβαση, αλλά μάλλον να υποστηρίζει ίσες ευκαιρίες για όλους τους μαθητές.

1.3 Ένα εγχειρίδιο για εκπαιδευτικούς, διοικητικούς υπαλλήλους και υπεύθυνους χάραξης πολιτικής

Αυτό το εγχειρίδιο έχει σχεδιαστεί για εκπαιδευτικούς, διοικητικούς υπαλλήλους και υπεύθυνους χάραξης πολιτικής στο πλαίσιο της υιοθέτησης της Τεχνητής Νοημοσύνης στην εκπαίδευση ενηλίκων και την επαγγελματική κατάρτιση (ΕΕΚ). Είναι γνωστό ότι η υιοθέτηση της Τεχνητής Νοημοσύνης και οι ηθικές της παραμέτρους επηρεάζουν καθεμία από τις προαναφερθείσες ομάδες με διαφορετικούς τρόπους:

Εκπαιδευτικοί :

Για τους εκπαιδευτικούς, η ηθική ενσωμάτωση της Τεχνητής Νοημοσύνης στην εκπαίδευση ενηλίκων και την ΕΕΚ αφορά την ενίσχυση της διδασκαλίας και της μάθησης χωρίς να υπονομεύονται οι αρχές της δικαιοσύνης, της ιδιωτικότητας και της διαφάνειας. Τα εργαλεία Τεχνητής Νοημοσύνης μπορούν να προσφέρουν εξατομικευμένες μαθησιακές εμπειρίες, να βελτιστοποιήσουν τις διοικητικές εργασίες και να παρέχουν πληροφορίες για την πρόοδο των μαθητών. Ωστόσο, οι εκπαιδευτικοί πρέπει να διασφαλίζουν ότι η Τεχνητή Νοημοσύνη υποστηρίζει



ποικίλα στυλ και ανάγκες μάθησης, προωθώντας ένα περιβάλλον μάθησης χωρίς αποκλεισμούς.

Υπάρχουν πολλές βασικές παραμέτρους που πρέπει να λάβουν υπόψη οι εκπαιδευτικοί σχετικά με τον γραμματισμό στην Τεχνητή Νοημοσύνη, την ηθική χρήση και τη διαφάνεια με τους μαθητές.

Γραμματισμός στην Τεχνητή Νοημοσύνη : Οι εκπαιδευτικοί πρέπει να είναι εξοπλισμένοι με τις δεξιότητες για να κατανοούν, να αξιολογούν και να ενσωματώνουν αποτελεσματικά την Τεχνητή Νοημοσύνη στις διδακτικές τους μεθόδους. Αυτό περιλαμβάνει τη γνώση του τρόπου λειτουργίας των εργαλείων Τεχνητής Νοημοσύνης και τις ηθικές τους επιπτώσεις στην τάξη.

Ηθική Χρήση : Οι εκπαιδευτικοί πρέπει να διασφαλίζουν ότι η Τεχνητή Νοημοσύνη χρησιμοποιείται για την ενίσχυση της μάθησης και όχι για την αντικατάσταση της ανθρώπινης αλληλεπίδρασης. Θα πρέπει να είναι προετοιμασμένοι να αντιμετωπίσουν τυχόν ζητήματα προκατάληψης ή διακρίσεων που ενδέχεται να προκύψουν από συστήματα Τεχνητής Νοημοσύνης.

Διαφάνεια με τους μαθητές : Οι εκπαιδευτικοί θα πρέπει να είναι διαφανείς με τους μαθητές σχετικά με τον τρόπο χρήσης της Τεχνητής Νοημοσύνης στη μαθησιακή διαδικασία και τα δεδομένα που συλλέγει. Αυτό βοηθά στην οικοδόμηση εμπιστοσύνης και διασφαλίζει ότι οι μαθητές αισθάνονται ότι το απόρρητό τους γίνεται σεβαστό.

Διαχειριστές :

Για τους διοικητικούς υπαλλήλους σε ιδρύματα εκπαίδευσης ενηλίκων και επαγγελματικής κατάρτισης, η υιοθέτηση της Τεχνητής Νοημοσύνης περιλαμβάνει την επίβλεψη της ενσωμάτωσης αυτών των τεχνολογιών στα συστήματα του ιδρύματος και τη διασφάλιση ότι υπάρχει η κατάλληλη υποδομή για την υποστήριξη εργαλείων Τεχνητής Νοημοσύνης. Πρέπει επίσης να δώσουν προτεραιότητα σε πολιτικές που διασφαλίζουν την ηθική χρήση της Τεχνητής Νοημοσύνης, διατηρώντας παράλληλα την ακεραιότητα του ιδρύματος.

Υπάρχουν πολλά βασικά ζητήματα που πρέπει να λάβουν υπόψη οι διαχειριστές, όσον αφορά την ασφάλεια και το απόρρητο των δεδομένων, την ισότητα και την πρόσβαση, καθώς και τη συνεχή εκπαίδευση και υποστήριξη.



Ασφάλεια δεδομένων και ιδιωτικό απόρρητο : Οι διαχειριστές πρέπει να διασφαλίζουν ότι εφαρμόζονται ισχυρά μέτρα προστασίας δεδομένων, καθώς η Τεχνητή Νοημοσύνη βασίζεται στη συλλογή μεγάλων ποσοτήτων δεδομένων μαθητών. Η συμμόρφωση με τους κανονισμούς προστασίας δεδομένων (όπως ο ΓΚΠΔ) είναι απαραίτητη.

Ισότητα και Πρόσβαση : Η Τεχνητή Νοημοσύνη θα πρέπει να αναπτύσσεται με τρόπους που να διασφαλίζουν την ισότιμη πρόσβαση σε όλους τους μαθητές, ανεξάρτητα από το κοινωνικοοικονομικό τους υπόβαθρο. Οι διαχειριστές θα πρέπει να υποστηρίζουν πολιτικές που εμποδίζουν τα εργαλεία Τεχνητής Νοημοσύνης να επιδεινώνουν τις υπάρχουσες ανισότητες.

Συνεχής Εκπαίδευση και Υποστήριξη : Οι διαχειριστές πρέπει να διαθέτουν πόρους για συνεχή επαγγελματική ανάπτυξη, ώστε οι εκπαιδευτικοί να μπορούν να παραμένουν ενημερωμένοι σχετικά με τις εξελίξεις στην Τεχνητή Νοημοσύνη και τις βέλτιστες πρακτικές δεοντολογίας.

Υπεύθυνοι χάραξης πολιτικής :

Οι υπεύθυνοι χάραξης πολιτικής διαδραματίζουν κρίσιμο ρόλο στη δημιουργία των πλαισίων και των κανονισμών που διέπουν τη χρήση της Τεχνητής Νοημοσύνης στην εκπαίδευση. Πρέπει να διασφαλίσουν ότι η υιοθέτηση της Τεχνητής Νοημοσύνης ευθυγραμμίζεται με τους ευρύτερους εκπαιδευτικούς στόχους, σέβεται τα ανθρώπινα δικαιώματα και ενισχύει τις ευκαιρίες για δια βίου μάθηση.

Υπάρχουν ποικίλα βασικά ζητήματα που πρέπει να λάβουν υπόψη οι υπεύθυνοι χάραξης πολιτικής σχετικά με τις δεοντολογικές κατευθυντήριες γραμμές και τους κανονισμούς, την προώθηση του γραμματισμού στην Τεχνητή Νοημοσύνη και τη διασφάλιση της συμπερίληψης και της προσβασιμότητας.

Ηθικές Κατευθυντήριες Γραμμές και Κανονισμοί : Οι υπεύθυνοι χάραξης πολιτικής πρέπει να θεσπίσουν ηθικά πρότυπα και κανονισμούς που να καθοδηγούν τη χρήση της Τεχνητής Νοημοσύνης στην εκπαίδευση. Αυτό περιλαμβάνει την αντιμετώπιση ζητημάτων όπως η διαφάνεια, το απόρρητο των δεδομένων και η δικαιοσύνη στα εκπαιδευτικά συστήματα που βασίζονται στην Τεχνητή Νοημοσύνη.

Προώθηση της παιδείας στην Τεχνητή Νοημοσύνη : Οι υπεύθυνοι χάραξης πολιτικής θα πρέπει να προωθήσουν την παιδεία στην Τεχνητή Νοημοσύνη σε όλο το εκπαιδευτικό σύστημα, διασφαλίζοντας ότι τόσο οι εκπαιδευτικοί όσο και



οι μαθητές διαθέτουν τα απαραίτητα εργαλεία και την κατανόηση για την πλοήγηση στις τεχνολογίες Τεχνητής Νοημοσύνης.

Διασφάλιση της Συμπερίληψης και της Προσβασιμότητας : Είναι ζωτικής σημασίας οι πολιτικές για την Τεχνητή Νοημοσύνη να δώσουν προτεραιότητα στην συμπεριληπτική εκπαίδευση, διασφαλίζοντας ότι οι τεχνολογίες Τεχνητής Νοημοσύνης δεν αφήνουν πίσω ορισμένες ομάδες. Αυτό μπορεί να περιλαμβάνει την παροχή κινήτρων για την έρευνα και την ανάπτυξη λύσεων Τεχνητής Νοημοσύνης προσαρμοσμένων σε διαφορετικούς μαθητές.

2. Κατανόηση της Τεχνητής Νοημοσύνης στην ΕΕΚ και την Εκπαίδευση Ενηλίκων

Η ενσωμάτωση της Τεχνητής Νοημοσύνης (TN) στην εκπαίδευση παρουσιάζει μετασχηματιστικές ευκαιρίες παράλληλα με σημαντικές προκλήσεις. Η έρευνα υπογραμμίζει τις δυνατότητες της TN να βελτιώσει τη διδασκαλία μέσω αυτοματοποιημένων εργασιών, εξατομικευμένης μάθησης και παιδαγωγικών καινοτομιών που βασίζονται σε δεδομένα, ενώ τα εργαλεία γενετικής TN διευρύνουν τις δυνατότητες για διαδραστική και αναστοχαστική εκπαίδευση. Ωστόσο, οι ηθικές ανησυχίες σχετικά με το απόρρητο των δεδομένων, την εμπορική επιρροή και το ψηφιακό χάσμα απαιτούν ισχυρή διακυβέρνηση. Στην εκπαίδευση ενηλίκων και την επαγγελματική εκπαίδευση, η TN όχι μόνο επηρεάζει τον σχεδιασμό των προγραμμάτων σπουδών και την πρόβλεψη δεξιοτήτων για την ετοιμότητα του εργατικού δυναμικού, αλλά απαιτεί επίσης την ενίσχυση του ψηφιακού γραμματισμού και της κριτικής εμπλοκής με τις κοινωνικές επιπτώσεις της TN. Η εξισορρόπηση των οφελών της TN με δίκαιες, ηθικές πολιτικές παραμένει κρίσιμη για τη διαμόρφωση συμπεριληπτικών, μελλοντοστρεφών εκπαιδευτικών οικοσυστημάτων.

2.1. Ο ρόλος της Τεχνητής Νοημοσύνης στην Εκπαίδευση

Μια πρόσφατα δημοσιευμένη έκθεση πολιτικής από το Κοινό Κέντρο Ερευνών (JRC) της Ευρωπαϊκής Επιτροπής, με τίτλο « Σχετικά με το Μέλλον της Τεχνολογίας στην Εκπαίδευση: Αναδυόμενες Τάσεις και Επιπτώσεις Πολιτικής » (Tuomi , Cachia & Villar- Onrubia , 2023), παρέχει μια ολοκληρωμένη επισκόπηση του πώς οι τεχνολογικές εξελίξεις, ιδίως στην Τεχνητή Νοημοσύνη (TN), μετασχηματίζουν τις εκπαιδευτικές πρακτικές, τις διαδικασίες και τους θεσμούς. Το έγγραφο υπογραμμίζει τις σημαντικές και ποικίλες επιπτώσεις της



TN στην εκπαίδευση, επηρεάζοντας τις διαδικασίες διδασκαλίας και μάθησης, τις μεθόδους αξιολόγησης και τις οργανωτικές δομές.

Η Τεχνητή Νοημοσύνη θεωρείται ως μια πιθανή λύση για τη μείωση του φόρτου εργασίας των εκπαιδευτικών, ειδικά σε εργασίες όπως η διόρθωση των εργασιών των μαθητών για το σπίτι. Αυτοματοποιώντας αυτές τις επαναλαμβανόμενες και χρονοβόρες εργασίες, η Τεχνητή Νοημοσύνη μπορεί να κάνει τη διδασκαλία πιο ελκυστική και να επιτρέψει στους εκπαιδευτικούς να επικεντρωθούν περισσότερο σε διδακτικές δραστηριότητες παρά σε διοικητικά καθήκοντα (Li, 2021; Tang & Hai, 2021; Villegas-Ch et al., 2021). Αυτή η αυτοματοποίηση θα μπορούσε να φέρει επανάσταση στη διαμορφωτική αξιολόγηση (Kumar & Boulanger, 2020), επιτρέποντας τη συνεχή ανατροφοδότηση για τους μαθητές και διευκολύνοντας την αξιολόγηση σε πραγματικό χρόνο (Porter & Grippa, 2020) και την εξατομικευμένη ή προσαρμοστική μάθηση (Kochmar et al., 2020; Wongvorachan et al., 2022).

Επιπλέον, η παραγωγική Τεχνητή Νοημοσύνη και τα θεμελιώδη μοντέλα αναμένεται να έχουν σημαντικό αντίκτυπο στην εκπαίδευση. Τεχνολογίες όπως η αυτόματη δημιουργία υποτίτλων σε βίντεο, η μετάφραση και η σύνοψη βίντεο προσφέρουν καινοτόμες ευκαιρίες για την ενίσχυση της παιδαγωγικής (Chew & Chua, 2020; Vazquez- Cano et al., 2021). Αυτές οι τεχνολογίες μπορούν να βοηθήσουν στην ανάπτυξη «έξυπνων συντρόφων» ή «συνεργατών μάθησης», μετατοπίζοντας την εστίαση από την απομνημόνευση στη μεταγνώση και τον αναστοχασμό. Αυτό απαιτεί μια προσεκτική ανακατανομή των ευθυνών μεταξύ ανθρώπων και μηχανών, καθορίζοντας ποιες δραστηριότητες θα ανατεθούν στην Τεχνητή Νοημοσύνη.

Θα πρέπει επίσης να τονίσουμε τον ρόλο της Τεχνητής Νοημοσύνης (TN) στην επεξεργασία δεδομένων στην εκπαίδευση, όπου δεδομένα μεγάλης κλίμακας σχετικά με εκπαιδευτικές δραστηριότητες χρησιμοποιούνται για την ενίσχυση και την ενημέρωση νέων παιδαγωγικών προσεγγίσεων, τεχνολογιών μάθησης και εκπαιδευτικών πολιτικών (Williamson et al., 2023). Στόχος είναι η καταγραφή των μαθησιακών διαδικασιών, όχι μόνο των αποτελεσμάτων, για την ενίσχυση της εμπλοκής και των επιδόσεων των μαθητών (Porter & Grippa, 2020).

Ωστόσο, οι ηθικές ανησυχίες σχετικά με την επεξεργασία προσωπικών δεδομένων, η πιθανή κακή χρήση, η εξάρτηση από υποδομές που ελέγχονται σε μεγάλο βαθμό από εμπορικές οντότητες και ορισμένες άλλες προκλήσεις αποτελούν κρίσιμα ζητήματα που πρέπει να αντιμετωπιστούν (Seufert, 2024; Wongvorachan et al., 2022). Η ενσωμάτωση της Τεχνητής Νοημοσύνης στην εκπαίδευση εγείρει σημαντικά ερωτήματα σχετικά με τα



μοντέλα διακυβέρνησης και τη χάραξη πολιτικής, καθιστώντας απαραίτητη την επαναξιολόγηση των εκπαιδευτικών πολιτικών ώστε να ενσωματώνονται οι δυνατότητες και οι περιορισμοί της Τεχνητής Νοημοσύνης (Attard-Frost et al., 2024; Boninger et al., 2020; Williamson & Eynon, 2020). Πρέπει επίσης να αντιμετωπίσουμε τη σημασία της κριτικής αξιολόγησης των κινδύνων επιδείνωσης του ψηφιακού χάσματος και της κυριαρχίας των εμπορικών συμφερόντων στον τομέα (Bulathwela et al., 2024; Celik, 2023). Η ρύθμιση της τεχνολογικής καινοτομίας για τη διασφάλιση ποικίλων εκπαιδευτικών οραμάτων και η ευθυγράμμιση της βιομηχανικής πολιτικής με τις θεωρίες μάθησης είναι κρίσιμες πτυχές που πρέπει να λάβουμε υπόψη.

Ο ρόλος της Τεχνητής Νοημοσύνης στην εκπαίδευση είναι ένα μετασχηματιστικό και πολύπλευρο εργαλείο που προσφέρει ευκαιρίες για βελτιωμένη αποτελεσματικότητα, εξατομίκευση και καινοτομία στη διδασκαλία και τη μάθηση. Παρ' όλα αυτά, παρουσιάζει επίσης σημαντικές προκλήσεις και ηθικά ζητήματα που πρέπει να αντιμετωπιστούν για να διασφαλιστεί η δίκαιη, αποτελεσματική και υπεύθυνη χρήση της Τεχνητής Νοημοσύνης σε εκπαιδευτικά περιβάλλοντα.

2.2. Τεχνητή Νοημοσύνη στην Εκπαίδευση και Κατάρτιση Ενηλίκων και Επαγγελματικής Κατάρτισης

Η ενσωμάτωση της εκπαίδευσης στην Τεχνητή Νοημοσύνη στα προγράμματα σπουδών περιλαμβάνει όχι μόνο τη διδασκαλία των βασικών αρχών της Τεχνητής Νοημοσύνης και των εφαρμογών της, αλλά και τη διερεύνηση των κοινωνικών επιπτώσεών της. Αυτό περιλαμβάνει την καλλιέργεια κριτικής σκέψης σχετικά με τις δυνατότητες, τους περιορισμούς και τις ηθικές διαστάσεις της Τεχνητής Νοημοσύνης, όπως οι ανησυχίες σχετικά με την προκατάληψη και την ιδιωτικότητα (Akgun & Greenhow, 2022; Aldosari, 2020; Attwell et al., 2020; Delcker, et al., 2024; Seufert, 2024; Wongvorachan et al., 2022). Η ενίσχυση του ψηφιακού γραμματισμού είναι επίσης απαραίτητη, καθώς η κατανόηση του τρόπου αλληλεπίδρασης και ερμηνείας των συστημάτων Τεχνητής Νοημοσύνης γίνεται βασική δεξιότητα στο σύγχρονο εργατικό δυναμικό, συμπεριλαμβανομένων των εκπαιδευτικών (Bekiaridis & Attwell, 2024; Delcker, et al., 2024; Roppertz, 2020; Rott et al., 2022; Seufert, 2024). Καθώς η Τεχνητή Νοημοσύνη συνεχίζει να εξελίσσεται, η εστίαση μετατοπίζεται όχι μόνο στη χρήση εργαλείων Τεχνητής Νοημοσύνης, αλλά και στην κατανόηση και την κριτική αξιολόγησή τους, προετοιμάζοντας τους μαθητές για ένα μέλλον όπου η Τεχνητή Νοημοσύνη είναι αναπόσπαστο κομμάτι της καθημερινής ζωής και εργασίας (Rott et al., 2022).



Συνεπώς, η προετοιμασία του εργατικού δυναμικού για νέες δεξιότητες στην εποχή της Τεχνητής Νοημοσύνης (TN) αποτελεί μια ακόμη κρίσιμη πτυχή της εκπαίδευσης ενηλίκων και της ΕΕΚ (Attwell et al., 2020; Roppertz , 2020; Rott et al., 2022). Ο ρόλος της TN στον εντοπισμό αναδυόμενων τάσεων και στην πρόβλεψη μελλοντικών απαιτήσεων σε δεξιότητες είναι ανεκτίμητος. Αυτή η γνώση επιτρέπει στα εκπαιδευτικά προγράμματα να προσαρμόζουν τα προγράμματα σπουδών τους ώστε να ανταποκρίνονται στις εξελισσόμενες ανάγκες της αγοράς εργασίας (Seufert, 2024). Η αναβάθμιση και η επανειδίκευση γίνονται κεντρικά θέματα, με προγράμματα καθοδηγούμενα από την TN να επιτρέπουν σε ενήλικες και επαγγελματίες να αποκτούν νέες δεξιότητες ή να ενημερώνουν τις υπάρχουσες, παραμένοντας έτσι σε επαφή με τις τεχνολογικές εξελίξεις (Attwell et al., 2020; Emeršič et al., 2024; Delcker , et al., 2024). Επιπλέον, η εκπαίδευση που βασίζεται στην TN διασφαλίζει ότι οι εκπαιδευόμενοι δεν είναι μόνο τεχνικά καταρτισμένοι, αλλά είναι επίσης προετοιμασμένοι να εργαστούν σε περιβάλλοντα ενσωματωμένα στην TN. Αυτή η προετοιμασία είναι το κλειδί για να διασφαλιστεί ότι το εργατικό δυναμικό είναι έτοιμο για τις προκλήσεις και τις ευκαιρίες που παρουσιάζει ο ταχέως εξελισσόμενος τομέας της Τεχνητής Νοημοσύνης, καθιστώντας την εκπαίδευση ενηλίκων και την επαγγελματική εκπαίδευση και κατάρτιση πιο σχετική, αποτελεσματική και ευθυγραμμισμένη με τα μελλοντικά τοπία απασχόλησης.

3. Ανάπτυξη ηθικών κατευθυντήριων γραμμών

3.1 Σημασία των ηθικών πολιτικών στην εκπαίδευση στην Τεχνητή Νοημοσύνη

Οι ηθικές πολιτικές στην εκπαίδευση στην Τεχνητή Νοημοσύνη είναι απαραίτητες για να διασφαλιστεί ότι οι τεχνολογίες Τεχνητής Νοημοσύνης χρησιμοποιούνται με τρόπους που προωθούν τη δικαιοσύνη, τη διαφάνεια και την υπευθυνότητα, προστατεύοντας παράλληλα τα δικαιώματα και την ευημερία των μαθητών και των εκπαιδευτικών. Καθώς η Τεχνητή Νοημοσύνη ενσωματώνεται ολοένα και περισσότερο στα εκπαιδευτικά περιβάλλοντα, η ύπαρξη σαφών ηθικών κατευθυντήριων γραμμών είναι ζωτικής σημασίας για τη διαχείριση των πιθανών κινδύνων της και την αξιοποίηση των οφελών της. Παρακάτω παρατίθενται οι βασικοί λόγοι για τους οποίους οι ηθικές πολιτικές είναι ζωτικής σημασίας στην εκπαίδευση στην Τεχνητή Νοημοσύνη:

Προστασία των Δικαιωμάτων και της Ιδιωτικότητας των Μαθητών



Τα συστήματα Τεχνητής Νοημοσύνης στην εκπαίδευση συχνά απαιτούν τη συλλογή και επεξεργασία προσωπικών δεδομένων, συμπεριλαμβανομένων ακαδημαϊκών αρχείων, μαθησιακών συμπεριφορών και άλλων ευαίσθητων πληροφοριών. Οι ηθικές πολιτικές διασφαλίζουν ότι τα δεδομένα αντιμετωπίζονται με σεβασμό και σύμφωνα με τους κανονισμούς περί απορρήτου (π.χ., GDPR). Αυτές οι πολιτικές βοηθούν στην προστασία του απορρήτου των μαθητών διασφαλίζοντας τη διαφάνεια στη συλλογή δεδομένων, δίνοντας στους μαθητές τον έλεγχο των δεδομένων τους και διασφαλίζοντας ότι χρησιμοποιούνται υπεύθυνα και με ασφάλεια.

Διασφάλιση Δικαιοσύνης και Ισότητας

Μία από τις πιο κρίσιμες ηθικές ανησυχίες σχετικά με την Τεχνητή Νοημοσύνη στην εκπαίδευση είναι η πιθανότητα μεροληπτικών αλγορίθμων που ενδέχεται να θέσουν σε μειονεκτική θέση ορισμένες ομάδες μαθητών. Τα συστήματα Τεχνητής Νοημοσύνης μπορούν ακούσια να διαιωνίσουν τις υπάρχουσες ανισότητες με βάση τη φυλή, το φύλο, την κοινωνικοοικονομική κατάσταση ή άλλους παράγοντες. Οι ηθικές πολιτικές είναι απαραίτητες για την καθοδήγηση της ανάπτυξης και της εφαρμογής εργαλείων Τεχνητής Νοημοσύνης που είναι χωρίς αποκλεισμούς και δίκαια, διασφαλίζοντας ότι όλοι οι μαθητές έχουν ίση πρόσβαση σε ευκαιρίες ανεξάρτητα από το υπόβαθρό τους ή τα προσωπικά τους χαρακτηριστικά.

Προώθηση της διαφάνειας και της λογοδοσίας

Οι τεχνολογίες τεχνητής νοημοσύνης μπορεί να είναι πολύπλοκες και αδιαφανείς, γεγονός που καθιστά δύσκολο για τους εκπαιδευτικούς, τους μαθητές και τους διαχειριστές να κατανοήσουν πλήρως τον τρόπο λήψης των αποφάσεων. Οι ηθικές πολιτικές συμβάλλουν στη διασφάλιση της διαφάνειας των συστημάτων τεχνητής νοημοσύνης, που σημαίνει ότι τα ενδιαφερόμενα μέρη γνωρίζουν πώς λειτουργούν τα εργαλεία τεχνητής νοημοσύνης, ποια δεδομένα χρησιμοποιούνται και πώς λαμβάνονται οι αποφάσεις. Αυτή η διαφάνεια χτίζει εμπιστοσύνη και λογοδοσία, επιτρέποντας στους εκπαιδευτικούς και τους μαθητές να κατανοούν καλύτερα και να αλληλεπιδρούν με τις τεχνολογίες τεχνητής νοημοσύνης.

Υποστήριξη της Τεχνητής Νοημοσύνης και της Υπεύθυνης Χρήσης

Καθώς η Τεχνητή Νοημοσύνη ενσωματώνεται όλο και περισσότερο στη μαθησιακή διαδικασία, είναι σημαντικό τόσο οι εκπαιδευτικοί όσο και οι μαθητές να κατανοήσουν πώς λειτουργεί η Τεχνητή Νοημοσύνη και τις πιθανές επιπτώσεις της. Οι ηθικές πολιτικές στην εκπαίδευση στην Τεχνητή Νοημοσύνη



θα πρέπει να ενθαρρύνουν τον αλφαριθμητισμό στην Τεχνητή Νοημοσύνη, διασφαλίζοντας ότι τόσο οι εκπαιδευτικοί όσο και οι μαθητές είναι εξοπλισμένοι με τις γνώσεις και τις δεξιότητες για να χρησιμοποιούν την Τεχνητή Νοημοσύνη υπεύθυνα και κριτικά. Αυτή η κατανόηση επιτρέπει στους χρήστες να μεγιστοποιήσουν τα οφέλη της, έχοντας παράλληλα επίγνωση των περιορισμών και των πιθανών βλαβών της.

Ενθάρρυνση της Ηθικής Ανάπτυξης και της Καινοτομίας

Οι τεχνολογίες τεχνητής νοημοσύνης στην εκπαίδευση θα πρέπει να αναπτύσσονται με ισχυρή ηθική βάση, ώστε να διασφαλίζεται ότι χρησιμοποιούνται προς όφελος όλων των μαθητών. Οι ηθικές πολιτικές παρέχουν ένα πλαίσιο για τους προγραμματιστές, τους ερευνητές και τα εκπαιδευτικά ιδρύματα, ώστε να σχεδιάζουν συστήματα τεχνητής νοημοσύνης που ευθυγραμμίζονται με εκπαιδευτικές αξίες, όπως η δικαιοσύνη, η ενδυνάμωση και ο σεβασμός της διαφορετικότητας. Αυτό διασφαλίζει επίσης ότι οι τεχνολογίες τεχνητής νοημοσύνης σχεδιάζονται και χρησιμοποιούνται με τρόπους που υποστηρίζουν τη δια βίου μάθηση και τους ευρύτερους στόχους της εκπαίδευσης.

Αντιμετώπιση του Κινδύνου Εκτοπισμού

Καθώς τα συστήματα Τεχνητής Νοημοσύνης αυτοματοποιούν ορισμένες πτυχές της εκπαίδευσης (π.χ. βαθμολόγηση, διδασκαλία, διοικητικά καθήκοντα), υπάρχει ο κίνδυνος να εκτοπίσουν τους εκπαιδευτικούς ή να περιορίσουν την ανθρώπινη αλληλεπίδραση. Οι ηθικές πολιτικές μπορούν να βοηθήσουν στην καθοδήγηση της υπεύθυνης χρήσης της Τεχνητής Νοημοσύνης διασφαλίζοντας ότι ο αυτοματισμός συμπληρώνει τους ανθρώπινους εκπαιδευτικούς, αντί να τους αντικαθιστά. Οι πολιτικές μπορούν να ενθαρρύνουν τη χρήση της Τεχνητής Νοημοσύνης για την ενίσχυση της διδακτικής αποτελεσματικότητας και της διοικητικής αποδοτικότητας, διατηρώντας παράλληλα την αξία της ανθρώπινης παρουσίας στην εκπαιδευτική εμπειρία.

Οικοδόμηση εμπιστοσύνης στα συστήματα τεχνητής νοημοσύνης

Για να υιοθετηθεί με επιτυχία η Τεχνητή Νοημοσύνη σε εκπαιδευτικά περιβάλλοντα, οι εκπαιδευτικοί, οι μαθητές και οι διοικητικοί υπάλληλοι πρέπει να εμπιστεύονται την τεχνολογία. Οι ηθικές πολιτικές διαδραματίζουν σημαντικό ρόλο στην εδραίωση και διατήρηση αυτής της εμπιστοσύνης, παρέχοντας σαφείς οδηγίες σχετικά με τον τρόπο χρήσης της Τεχνητής Νοημοσύνης και ποια ηθικά πρότυπα πρέπει να ακολουθούνται. Όταν τα συστήματα Τεχνητής Νοημοσύνης αναπτύσσονται και αναπτύσσονται σύμφωνα με αυτές τις πολιτικές, τα



ενδιαφερόμενα μέρη είναι πιο πιθανό να τα υιοθετήσουν και να τα χρησιμοποιήσουν αποτελεσματικά.

Καθοδηγώντας το μέλλον της Τεχνητής Νοημοσύνης στην Εκπαίδευση

Καθώς η Τεχνητή Νοημοσύνη συνεχίζει να εξελίσσεται, είναι ζωτικής σημασίας οι ηθικές πολιτικές να παρέχουν μακροπρόθεσμη καθοδήγηση για την ενσωμάτωση της στα εκπαιδευτικά συστήματα. Αυτές οι πολιτικές συμβάλλουν στη δημιουργία ενός πλαισίου για τη μελλοντική ανάπτυξη τεχνολογιών Τεχνητής Νοημοσύνης που ευθυγραμμίζεται με τις κοινωνικές αξίες, διασφαλίζοντας ότι η Τεχνητή Νοημοσύνη χρησιμοποιείται για τη βελτίωση των μαθησιακών αποτελεσμάτων, την ενίσχυση της καινοτομίας και την υποστήριξη της δια βίου μάθησης με τρόπους που είναι τόσο ωφέλιμοι όσο και ηθικοί.

Γενικά, οι ηθικές πολιτικές στην εκπαίδευση στην Τεχνητή Νοημοσύνη είναι απαραίτητες για να καθοδηγήσουν την υπεύθυνη ανάπτυξη και χρήση τεχνολογιών Τεχνητής Νοημοσύνης. Αυτές οι πολιτικές διασφαλίζουν ότι η Τεχνητή Νοημοσύνη χρησιμοποιείται με τρόπους που σέβονται το απόρρητο των μαθητών, προωθούν τη δικαιοσύνη, ενθαρρύνουν τη διαφάνεια και χτίζουν εμπιστοσύνη. Επιπλέον, βοηθούν τους εκπαιδευτικούς, τους μαθητές και τα ιδρύματα να αντιμετωπίσουν τις ηθικές προκλήσεις που προκύπτουν καθώς η Τεχνητή Νοημοσύνη συνεχίζει να μεταμορφώνει τα εκπαιδευτικά περιβάλλοντα. Με τη θέσπιση αυτών των ηθικών κατευθυντήριων γραμμών, τα εκπαιδευτικά συστήματα μπορούν να αξιοποιήσουν τη δύναμη της Τεχνητής Νοημοσύνης, διαφυλάσσοντας παράλληλα τις αξίες που είναι θεμελιώδεις για την αποτελεσματική και χωρίς αποκλεισμούς εκπαίδευση.

3.2 Βασικές απαιτήσεις για αξιόπιστα συστήματα Τεχνητής Νοημοσύνης

Η ενσωμάτωση της Τεχνητής Νοημοσύνης (TN) στην εκπαίδευση προσφέρει τεράστιες ευκαιρίες για την ενίσχυση της διδασκαλίας, της μάθησης και της διοίκησης. Ωστόσο, για να διασφαλιστεί ότι η TN χρησιμοποιείται ηθικά και υπεύθυνα, τα εκπαιδευτικά ιδρύματα πρέπει να αναπτύξουν σαφείς κατευθυντήριες γραμμές για την ηθική εφαρμογή των συστημάτων TN. Πολυάριθμα δημοσιευμένα πλαίσια, κατευθυντήριες γραμμές και έλεγχοι (π.χ., ο Έλεγχος TN 2023α της UNESCO και οι Οδηγίες για την Γενετική TN στην Εκπαίδευση και την Έρευνα 2023β) παρέχουν πολύτιμες πληροφορίες για το τι συνιστά αξιόπιστη TN. Αυτά τα πλαίσια δίνουν έμφαση σε βασικές αρχές όπως η διαφάνεια, η δικαιοσύνη, η λογοδοσία και ο σεβασμός των ανθρωπίνων



δικαιωμάτων . Ωστόσο, το έργο της ανάπτυξης και εφαρμογής αυτών των αρχών μπορεί να είναι τρομακτικό για τα ιδρύματα, ειδικά σε ένα ταχέως εξελισσόμενο τεχνολογικό τοπίο.

Ως μέρος του προγράμματος AI Pioneers, έχουμε αναπτύξει ένα σχήμα αξιολόγησης (Roman Etxebarrieta et al., 2024) που έχει σχεδιαστεί για να απλοποιήσει και να εντάξει αυτές τις αρχές σε ένα εκπαιδευτικό περιβάλλον. Το σχήμα μας παρέχει μια δομημένη προσέγγιση για να βοηθήσει τα ιδρύματα να πλοηγηθούν στις πολυπλοκότητες της εφαρμογής της Τεχνητής Νοημοσύνης, διασφαλίζοντας παράλληλα την τήρηση των ηθικών προτύπων. Εφόσον υπάρχουν αρκετές διαθέσιμες κατευθυντήριες γραμμές, το σχήμα μας στοχεύει να προσφέρει ένα πρακτικό, φιλικό προς το χρήστη εργαλείο που θα καθοδηγήσει τα εκπαιδευτικά ιδρύματα στην ανάπτυξη των δικών τους ηθικών κατευθυντήριων γραμμών για την Τεχνητή Νοημοσύνη, λαμβάνοντας υπόψη τα μοναδικά τους πλαίσια και τις ανάγκες τους.

Μία από τις κεντρικές διαπιστώσεις από τον Νόμο της ΕΕ για την Τεχνητή Νοημοσύνη (Ευρωπαϊκό Κοινοβούλιο και Συμβούλιο της Ευρωπαϊκής Ένωσης, 2024) είναι ότι η πρωταρχική ευθύνη για τη διασφάλιση της ηθικής ανάπτυξης των συστημάτων Τεχνητής Νοημοσύνης βαρύνει τους προγραμματιστές και τις εταιρείες που τα δημιουργούν και ότι οι κατευθυντήριες γραμμές που πρέπει να ακολουθούνται θα εξαρτηθούν από το επίπεδο κινδύνου που σχετίζεται με τα συστήματα Τεχνητής Νοημοσύνης. Ωστόσο, τα εκπαιδευτικά κέντρα είναι υπεύθυνα για τη διασφάλιση της ηθικής εφαρμογής και χρήσης αυτών των συστημάτων στο εκπαιδευτικό τους περιβάλλον. Τα κέντρα πρέπει να γνωρίζουν βασικά ηθικά ζητήματα, ώστε να κατανοούν τόσο τους περιορισμούς όσο και τις δυνατότητες των συστημάτων Τεχνητής Νοημοσύνης που υιοθετούν. Αυτό περιλαμβάνει όχι μόνο τη συμμόρφωση με τις υπάρχουσες πολιτικές και τις νομικές απαιτήσεις, αλλά και την πρόβλεψη και αντιμετώπιση των προκλήσεων που ενδέχεται να προκύψουν καθώς εξελίσσονται αυτές οι τεχνολογίες.

Συγκεκριμένα, το σχήμα αξιολόγησής μας εστιάζει στον εντοπισμό ηθικών ζητημάτων και στην ανάπτυξη μιας πολιτικής ή στρατηγικής που μπορεί να υποστηρίξει και να προωθήσει τη χρήση της Τεχνητής Νοημοσύνης σε εκπαιδευτικά κέντρα, τηρώντας παράλληλα βασικές ηθικές κατευθυντήριες γραμμές. Αυτές οι κατευθυντήριες γραμμές περιλαμβάνουν: **ποικιλομορφία και ισότητα, διαφάνεια και λογοδοσία, ιδιωτικότητα και προστασία δεδομένων, ασφάλεια, βιωσιμότητα και κοινωνική ευημερία, ενδυνάμωση εκπαιδευτικών και μαθητών, δημοκρατική συμμετοχή στον σχεδιασμό εκπαιδευτικής πολιτικής και στις πρακτικές της Τεχνητής Νοημοσύνης, αυτονομία, ηθικός σχεδιασμός (σχετικός με τη βιωσιμότητα και την**



προσβασιμότητα), εμπορευματοποίηση (Şenocak et al., 2024) , εμπλοκή εκπαιδευτικών και μαθητών και γραμματισμός στην Τεχνητή Νοημοσύνη.

Το σχήμα αξιολόγησης αποτελείται από διάφορα τμήματα, βασισμένα στις διαστάσεις του Chan (2023), προσαρμοσμένα με την ανασκόπηση του Şenocak et al. (2024) και τις ηθικές κατευθυντήριες γραμμές της Ευρωπαϊκής Επιτροπής (2022), καθένα από τα οποία επικεντρώνεται σε διαφορετικές πτυχές της εφαρμογής της Τεχνητής Νοημοσύνης σε εκπαιδευτικά περιβάλλοντα. Αυτά τα τμήματα καθοδηγούν τα ιδρύματα στην ανάπτυξη ηθικών πλαισίων και στη διασφάλιση της υπεύθυνης χρήσης της Τεχνητής Νοημοσύνης. Η πρώτη ενότητα εστιάζει στην **Αξιολόγηση της Τρέχουσας Χρήσης της Τεχνητής Νοημοσύνης και των Ηθικών Ζητημάτων**, βοηθώντας τα ιδρύματα να αξιολογήσουν την τρέχουσα ωριμότητά τους στην Τεχνητή Νοημοσύνη, επιτρέποντάς τους έτσι να εκτιμήσουν πόσο προετοιμασμένο είναι το εκπαιδευτικό κέντρο να εφαρμόσει την Τεχνητή Νοημοσύνη ή να επεκτείνει τη χρήση της (JISC, 2022).

Η δεύτερη ενότητα επικεντρώνεται στη **Διακυβέρνηση και την Παρακολούθηση της Τεχνητής Νοημοσύνης** και, ως εκ τούτου, ασχολείται με ζητήματα που σχετίζονται με τη συμμόρφωση με τις περιφερειακές και διεθνείς πολιτικές, καλύπτοντας την ιδιωτικότητα, την προστασία των δεδομένων και τη δικαιοσύνη. Αντιμετωπίζει σύνθετα ηθικά ζητήματα, όπως η μη διάκριση και η διαφάνεια, που αποτελούν επίσης βασικές παραμέτρους για την ηθική Τεχνητή Νοημοσύνη. Η τρίτη ενότητα, **Επιχειρησιακές Παραμέτρους**, παρέχει καθοδήγηση σχετικά με την εκπαίδευση εκπαιδευτικών, μαθητών και προσωπικού, την προώθηση του γραμματισμού στην Τεχνητή Νοημοσύνη, τη διασφάλιση της ανθρώπινης εποπτείας και την υποστήριξη της δημοκρατικής συμμετοχής στη λήψη αποφάσεων που σχετίζονται με την Τεχνητή Νοημοσύνη. Η διασφάλιση ότι όλοι οι χρήστες είναι επαρκώς εκπαιδευμένοι και γνωρίζουν τις δυνατότητες και τους περιορισμούς της Τεχνητής Νοημοσύνης είναι το κλειδί για την επιτυχή εφαρμογή της Τεχνητής Νοημοσύνης και διασφαλίζει ισότιμη πρόσβαση για όλους.

Η τέταρτη ενότητα επικεντρώνεται σε **Παιδαγωγικές Σκέψεις**, λαμβάνοντας υπόψη σημαντικά ζητήματα όπως τα ηθικά διλήμματα στις αξιολογήσεις που βασίζονται στην Τεχνητή Νοημοσύνη, η αυτονομία των μαθητών και η ανάπτυξη ικανοτήτων, ιδίως στην Εκπαίδευση και Κατάρτιση Ενηλίκων και Επαγγελματικής Εκπαίδευσης και Κατάρτισης (Attwell et al., 2021; UNESCO, 2019). Η πέμπτη και τελευταία ενότητα επικεντρώνεται σε **Άλλες Σκέψεις: Βιωσιμότητα και Ηθικός Σχεδιασμός**, προσφέροντας καθοδήγηση για την αντιμετώπιση των



μακροπρόθεσμων επιπτώσεων και τη διασφάλιση της προσβάσιμης χρήσης της Τεχνητής Νοημοσύνης.

Το σχήμα αξιολόγησης παρέχει ένα πρακτικό εργαλείο για τα εκπαιδευτικά ιδρύματα ώστε να πλοηγηθούν στην εφαρμογή της Τεχνητής Νοημοσύνης, να εντοπίσουν ηθικές προκλήσεις και να αναπτύξουν πολιτικές ευθυγραμμισμένες με τα διεθνή πρότυπα. Για πιο λεπτομερείς πληροφορίες, ανατρέξτε στο πλήρες σχήμα αξιολόγησης [Σχέδιο αξιολόγησης για την Τεχνητή Νοημοσύνη στην εκπαίδευση σχετικά με τα δεδομένα, την ιδιωτικότητα, την ηθική και τις αξίες της ΕΕ (WP5)] και στην παρουσίαση Genially που μπορείτε να βρείτε στον [Ιστότοπο AI Pioneers](#).

3.3 Λαμβάνοντας υπόψη τις νομικές πτυχές της εφαρμογής της Τεχνητής Νοημοσύνης στις Σχολές Επαγγελματικής Εκπαίδευσης και Κατάρτισης

3.3.1 Το κανονιστικό πλαίσιο του νόμου της ΕΕ για την τεχνητή νοημοσύνη

Ο νόμος περί τεχνητής νοημοσύνης της Ευρωπαϊκής Ένωσης θεσπίζει ένα ολοκληρωμένο κανονιστικό πλαίσιο που ταξινομεί τα συστήματα τεχνητής νοημοσύνης ανάλογα με τους πιθανούς κινδύνους που ενέχουν, παρέχοντας ουσιαστική καθοδήγηση στα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης (ΕΕΚ) που εφαρμόζουν αυτές τις τεχνολογίες. Αυτή η προσέγγιση που βασίζεται στον κίνδυνο διασφαλίζει αναλογική κανονιστική εποπτεία, ενώ παράλληλα ενθαρρύνει την καινοτομία σε εκπαιδευτικά πλαίσια (Ευρωπαϊκό Κοινοβούλιο, 2024).

Όπως διατυπώνουν οι Zawacki -Richter et al. (2023) στη συστηματική ανασκόπηση των ηθικών ζητημάτων στην εκπαίδευση με την ενίσχυση της Τεχνητής Νοημοσύνης, τα κανονιστικά πλαίσια πρέπει να εξισορροπούν την τεχνολογική καινοτομία με την προστασία των θεμελιωδών δικαιωμάτων. Η ανάλυσή τους προσδιορίζει τον Νόμο της ΕΕ για την Τεχνητή Νοημοσύνη ως μια πρωτοποριακή προσπάθεια για τη θέσπιση «σαφών παραμέτρων για την υπεύθυνη ανάπτυξη της Τεχνητής Νοημοσύνης σε μαθησιακά περιβάλλοντα, αναγνωρίζοντας παράλληλα το μετασχηματιστικό δυναμικό αυτών των τεχνολογιών» (σελ. 8).

Ταξινόμηση Κινδύνου για την Εκπαιδευτική Τεχνητή Νοημοσύνη



Τα συστήματα τεχνητής νοημοσύνης που αναπτύσσονται σε περιβάλλοντα επαγγελματικής εκπαίδευσης και κατάρτισης εμπίπτουν κυρίως στην κατηγορία **υψηλού κινδύνου** λόγω του σημαντικού τους αντίκτυπου στις εκπαιδευτικές εμπειρίες των μαθητών και στις μελλοντικές ευκαιρίες σταδιοδρομίας τους. Ο νόμος της ΕΕ για την τεχνητή νοημοσύνη κατηγοριοποιεί τα συστήματα τεχνητής νοημοσύνης ως εξής:

Μη Αποδεκτός Κίνδυνος : Απαγορεύονται ρητά συστήματα που θέτουν σαφείς απειλές για τα θεμελιώδη δικαιώματα ή την ασφάλεια.

Υψηλού Κινδύνου : Εφαρμογές σε κρίσιμους τομείς, συμπεριλαμβανομένης της εκπαίδευσης, που απαιτούν αυστηρή κανονιστική συμμόρφωση

Περιορισμένος Κίνδυνος : Συστήματα με υποχρεώσεις διαφάνειας που απαιτούν αποκάλυψη της αλληλεπίδρασης με την Τεχνητή Νοημοσύνη

Ελάχιστος Κίνδυνος : Εφαρμογές με περιορισμένες κανονιστικές απαιτήσεις πέραν των υφιστάμενων πλαισίων

Επιπτώσεις ταξινόμησης υψηλού κινδύνου για τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης:

- Υποχρεωτικά πρωτόκολλα αξιολόγησης κινδύνου και μετριάσμού
- Αυστηρός απόδειξη με έγγραφα απαιτήσεις
- Ανθρώπινη εποπτεία μηχανισμοί
- Διαφάνεια υποχρεώσεις να προσβεβλημένος ενδιαφερόμενοι φορείς
- Τακτική συμμόρφωση αξιολογήσεις .

Η έρευνα των Ramirez-Montoya et al. (2023) τονίζει ότι η ταξινόμηση υψηλού κινδύνου για τα εκπαιδευτικά συστήματα Τεχνητής Νοημοσύνης απαιτεί «αυστηρές δομές διακυβέρνησης που υπερβαίνουν την τεχνική συμμόρφωση για την αντιμετώπιση ευρύτερων κοινωνικών ανησυχιών» (σελ. 142). Η ανάλυσή τους σχετικά με τις προκλήσεις εφαρμογής σε όλα τα ευρωπαϊκά επαγγελματικά ιδρύματα υπογραμμίζει την ανάγκη για ερμηνεία των κανονιστικών απαιτήσεων με βάση τα συμφραζόμενα.

3.3.2 Βασικές Νομικές Παραμέτρους για την Υλοποίηση της Επαγγελματικής Εκπαίδευσης και Κατάρτισης

Απαιτήσεις Διαφάνειας και Λογοδοσίας

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



Ο νόμος της ΕΕ για την τεχνητή νοημοσύνη επιβάλλει πλήρη διαφάνεια στην ανάπτυξη της τεχνητής νοημοσύνης σε εκπαιδευτικά περιβάλλοντα:

- Τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης πρέπει να ενημερώνουν τους φοιτητές, το προσωπικό και τα σχετικά ενδιαφερόμενα μέρη όταν χρησιμοποιούνται συστήματα τεχνητής νοημοσύνης στην αξιολόγηση, στις συστάσεις μάθησης ή στη λήψη διοικητικών αποφάσεων.
- Η τεκμηρίωση πρέπει να διατυπώνει τον σκοπό, τη λειτουργικότητα και τους περιορισμούς των εφαρμογών Τεχνητής Νοημοσύνης.
- Οι αποφάσεις που βασίζονται στην Τεχνητή Νοημοσύνη και επηρεάζουν την πρόοδο ή την πιστοποίηση των φοιτητών πρέπει να είναι εξηγήσιμες και δικαιολογημένες.
- Οι αλγοριθμικές διαδικασίες πρέπει να είναι ελέγξιμες για να επιτρέπουν την επικύρωση των αποτελεσμάτων.

Οι Prinsloo και Knox (2023) προσδιορίζουν τη διαφάνεια ως ακρογωνιαίό λίθο της ηθικής αρχής στην εφαρμογή της εκπαιδευτικής Τεχνητής Νοημοσύνης, σημειώνοντας ότι «η αλγοριθμική αδιαφάνεια παρουσιάζει μοναδικές προκλήσεις σε μαθησιακά περιβάλλοντα όπου η εμπιστοσύνη μεταξύ εκπαιδευτικών και μαθητών αποτελεί το θεμέλιο της αποτελεσματικής παιδαγωγικής» (σελ. 76). Το πλαίσιο τους για τη διαφάνεια της εκπαιδευτικής Τεχνητής Νοημοσύνης εκτείνεται πέρα από την τεχνική αποκάλυψη, ώστε να περιλαμβάνει την κατανόηση των συμφραζομένων, η οποία είναι προσβάσιμη σε ποικίλα ενδιαφερόμενα μέρη.

Προστασία Δεδομένων και Συμμόρφωση με το Απόρρητο

Η εφαρμογή της Τεχνητής Νοημοσύνης σε περιβάλλοντα ΕΕΚ απαιτεί σχολαστική τήρηση των πλαισίων προστασίας δεδομένων:

- Όλα τα συστήματα τεχνητής νοημοσύνης πρέπει να συμμορφώνονται με τις διατάξεις του Γενικού Κανονισμού για την Προστασία Δεδομένων (GDPR).
- Πρέπει να θεσπιστούν σαφείς μηχανισμοί συγκατάθεσης για την επεξεργασία προσωπικών δεδομένων
- Οι αρχές ελαχιστοποίησης δεδομένων πρέπει να διέπουν τις πρακτικές συλλογής και διατήρησης
- Ο περιορισμός του σκοπού διασφαλίζει τη χρήση των δεδομένων αποκλειστικά για συγκεκριμένους εκπαιδευτικούς στόχους.
- Οι τεχνικές και οργανωτικές δικλίδες ασφαλείας πρέπει να προστατεύουν την ακεραιότητα και την εμπιστευτικότητα των δεδομένων των φοιτητών.



Οι Holmes et al. (2023) παρατηρούν ότι οι παράμετροι προστασίας δεδομένων σε εκπαιδευτικά πλαίσια εκτείνονται πέρα από τη νομική συμμόρφωση και περιλαμβάνουν «ευρύτερα ηθικά ζητήματα σχετικά με την ενεργό δράση των μαθητών, τη δυναμική της εξουσίας και την πιθανή εκπαιδευτική επιτήρηση» (σελ. 12). Το έργο τους τονίζει τη σημασία της ένταξης των πλαισίων προστασίας δεδομένων στο πλαίσιο ηθικών παραμέτρων που αφορούν συγκεκριμένα την εκπαίδευση.

Πρόληψη Προκαταλήψεων και Διασφάλιση Δικαιοσύνης

Ο νόμος της ΕΕ για την τεχνητή νοημοσύνη θεσπίζει σαφείς απαιτήσεις για τον μετριασμό της αλγοριθμικής προκατάληψης σε εκπαιδευτικά πλαίσια:

- Τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης πρέπει να διενεργούν τακτικές αξιολογήσεις προκατάληψης των συστημάτων τεχνητής νοημοσύνης για την πρόληψη των διακρίσεων.
- Οι μεθοδολογίες αξιολόγησης πρέπει να αξιολογούν τις πιθανές ανομοιόμορφες επιπτώσεις σε όλες τις δημογραφικές ομάδες
- Τα αλγοριθμικά δεδομένα εκπαίδευσης πρέπει να είναι αντιπροσωπευτικά της ποικιλομορφίας των μαθητών
- Τα πρωτόκολλα συνεχούς παρακολούθησης πρέπει να εντοπίζουν και να διορθώνουν αναδυόμενες μεροληψίες σε λειτουργικά συστήματα.
- Απαιτείται τακτικός έλεγχος των αποτελεσμάτων σε όλους τους μαθητικούς πληθυσμούς για την επαλήθευση του δίκαιου αντίκτυπου.

Η έρευνα των Hodgkin και Kumar (2023) που εξετάζει την αλγοριθμική προκατάληψη στα συστήματα επαγγελματικής καθοδήγησης προσδιορίζει «συστηματικές μεθοδολογίες για την αξιολόγηση των ανόμοιων επιπτώσεων σε διαφορετικούς μαθητικούς πληθυσμούς» (σελ. 217) ως απαραίτητες για την ισότιμη εφαρμογή της Τεχνητής Νοημοσύνης. Το πλαίσιό τους για την ολοκληρωμένη αξιολόγηση της προκατάληψης παρέχει πρακτική καθοδήγηση στα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης που πλοηγούνται στις απαιτήσεις δίκαιης μεταχείρισης.

Τεχνητή Νοημοσύνη σε Εργαλεία Τάξης (Περιορισμένου Κινδύνου)

Πολυάριθμες εφαρμογές Τεχνητής Νοημοσύνης, όπως: βοηθοί διδασκαλίας με τεχνητή νοημοσύνη, συστήματα ανίχνευσης λογοκλοπής, μηχανές συστάσεων για μαθησιακούς πόρους, εργαλεία διοικητικού αυτοματισμού και προσομοιώσεις



εικονικού εργαστηρίου σε περιβάλλοντα ΕΕΚ, ενδέχεται να εμπίπτουν σε κατηγορία περιορισμένου κινδύνου:

Αυτές οι εφαρμογές πρέπει να παρέχουν σαφή γνωστοποίηση στους χρήστες σχετικά με:

- Η φύση του συστήματος που βασίζεται στην Τεχνητή Νοημοσύνη
- Περιορισμοί του αλγοριθμικός δυνατότητες
- Ανθρώπινη εποπτεία προμήθειες
- Μηχανισμοί για αμφισβητούμενα αποτελέσματα.

3.3.3 Χρονοδιάγραμμα Υλοποίησης και Συμμόρφωση Σχεδίαση

Τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης πρέπει να αναπτύξουν στρατηγικά σχέδια εφαρμογής ευθυγραμμισμένα με το χρονοδιάγραμμα σταδιακής εφαρμογής του νόμου της ΕΕ για την τεχνητή νοημοσύνη:

2 ^{Φεβρουαρίου} 2025 : Απαγόρευση συστημάτων τεχνητής νοημοσύνης που ενέχουν απαράδεκτους κινδύνους

2 ^{Αυγούστου} 2025 : Εφαρμογή κωδίκων ορθής πρακτικής και ορισμός εθνικών ρυθμιστικών αρχών

2 ^{Αυγούστου} 2026 : Γενική εφαρμογή του νόμου περί τεχνητής νοημοσύνης, συμπεριλαμβανομένων κανόνων για συστήματα τεχνητής νοημοσύνης υψηλού κινδύνου

2 ^{Αυγούστου} 2027 : Πλήρης εφαρμογή συγκεκριμένων υποχρεώσεων για συστήματα τεχνητής νοημοσύνης υψηλού κινδύνου.

Συνετός συμμόρφωση σχεδίαση απαιτεί :

- Πλήρης απογραφή υφιστάμενων και προγραμματισμένων εφαρμογών Τεχνητής Νοημοσύνης
- Αξιολόγηση ταξινόμησης κινδύνου χαρτοφυλακίου εκπαιδευτικής τεχνολογίας
- Ανάλυση κενών κατά ρυθμιστικούς απαιτήσεις
- Δομημένος οδικός χάρτης υλοποίησης με στόχους-ορόσημα
- Κατανομή πόρων για την τεκμηρίωση και την παρακολούθηση της συμμόρφωσης.



3.3.4 Νομικές Ευθύνες και Ζητήματα Ευθύνης

Τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης που εφαρμόζουν συστήματα τεχνητής νοημοσύνης αναλαμβάνουν σημαντικές νομικές ευθύνες:

Οι δομές λογοδοσίας πρέπει να ορίζουν σαφή ευθύνη για τη διακυβέρνηση του συστήματος τεχνητής νοημοσύνης.

Τα πλαίσια ευθύνης πρέπει να αντιμετωπίζουν πιθανές ζημιές που προκύπτουν από αλγοριθμικά σφάλματα ή προκαταλήψεις.

Οι συμβατικές συμφωνίες με τους παρόχους τεχνολογίας πρέπει να οριοθετούν ρητά την κατανομή ευθυνών.

Τα πρωτόκολλα αποζημίωσης θα πρέπει να αντιμετωπίζουν πιθανές αξιώσεις που προκύπτουν από αποφάσεις που επηρεάζονται από την τεχνητή νοημοσύνη.

Οι ασφαλιστικές παραμέτρους θα πρέπει να αντικατοπτρίζουν τις αναδυόμενες εκθέσεις αστικής ευθύνης.

Τα εκπαιδευτικά ιδρύματα πρέπει να αναγνωρίσουν ότι η ανάπτυξη συστημάτων Τεχνητής Νοημοσύνης, ακόμη και εκείνων που προμηθεύονται από τρίτους, δεν τα απαλλάσσει από τη νομική ευθύνη για τα αποτελέσματα που επηρεάζουν τους μαθητές.

Οι Rodriguez- Abitia και Bribiesca -Correa (2021) προσδιορίζουν τις παραμέτρους ευθύνης ως μια συχνά παραβλεπόμενη διάσταση της διακυβέρνησης της εκπαιδευτικής Τεχνητής Νοημοσύνης, σημειώνοντας ότι «τα εκπαιδευτικά ιδρύματα αντιμετωπίζουν μοναδικά νομικά ζητήματα όταν αναθέτουν παραδοσιακά ανθρώπινες λειτουργίες σε αλγοριθμικά συστήματα» (σελ. 94). Η ανάλυσή τους για την αναδυόμενη νομολογία παρέχει πολύτιμο προηγούμενο για τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης που αναπτύσσουν ολοκληρωμένα πλαίσια ευθύνης.

3.3.5 Λίστα Αξιολόγησης για Αξιόπιστη Τεχνητή Νοημοσύνη (ALTAI) στην Εφαρμογή Επαγγελματικής Εκπαίδευσης και Κατάρτισης

Ο Κατάλογος Αξιολόγησης για την Αξιόπιστη Τεχνητή Νοημοσύνη (ALTAI) παρέχει στα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης ένα δομημένο



πλαίσιο για την ηθική εφαρμογή και τη συμμόρφωση με τους κανονισμούς. Αυτό το πρακτικό εργαλείο εφαρμόζει τις Κατευθυντήριες Γραμμές Δεοντολογίας για την Αξιοπίστη Τεχνητή Νοημοσύνη που αναπτύχθηκαν από την Ομάδα Εμπειρογνομώνων Υψηλού Επιπέδου της ΕΕ για την Τεχνητή Νοημοσύνη.

Στη συγκριτική τους ανάλυση των πλαισίων ηθικής αξιολόγησης, οι Zawacki - Richter et al. (2023) προσδιορίζουν την ALTAI ως ιδιαίτερα πολύτιμη σε εκπαιδευτικά πλαίσια λόγω της «ολοκληρωμένης κάλυψης τόσο των τεχνικών όσο και των κοινωνικο-ηθικών διαστάσεων που σχετίζονται με τα μαθησιακά περιβάλλοντα» (σελ. 14). Η έρευνά τους καταδεικνύει την προσαρμοστικότητα της ALTAI σε ποικίλα εκπαιδευτικά πλαίσια, συμπεριλαμβανομένης της επαγγελματικής κατάρτισης.

Βασικά Στοιχεία του ALTAI Σχετικά με την Υλοποίηση της Επαγγελματικής Εκπαίδευσης και Κατάρτισης

Το ALTAI διαρθρώνει την αξιολόγηση σε επτά θεμελιώδεις διαστάσεις, ιδιαίτερα σημαντικές σε περιβάλλοντα επαγγελματικής εκπαίδευσης:

Ανθρώπινη Δράση και Εποπτεία

- Οι αιτήσεις για επαγγελματική εκπαίδευση και κατάρτιση πρέπει να υποστηρίζουν την αυτονομία των εκπαιδευτών αντί να υποκαθιστούν την επαγγελματική κρίση.
- Η αυτονομία των φοιτητών πρέπει να διατηρείται μέσω της ενημερωμένης συναίνεσης και της διαφάνειας των αποφάσεων.
- Πρέπει να διατηρείται ουσιαστική ανθρώπινη εποπτεία στις επακόλουθες αποφάσεις που επηρεάζουν την πρόοδο των μαθητών.

Τεχνικός Ανθεκτικότητα και Ασφάλεια

- Τα συστήματα τεχνητής νοημοσύνης πρέπει να λειτουργούν αξιόπιστα σε ποικίλα περιβάλλοντα επαγγελματικής κατάρτισης
- Τα πρωτόκολλα ασφαλείας πρέπει να προστατεύουν από πιθανή χειραγώγηση ή μη εξουσιοδοτημένη πρόσβαση
- Οι διαδικασίες εφεδρείας πρέπει να διασφαλίζουν τη συνέχεια της εκπαίδευσης κατά τη διάρκεια βλαβών του συστήματος.

Ιδιωτικότητα και Διακυβέρνηση Δεδομένων



- Τα δεδομένα των μαθητών που χρησιμοποιούνται στην εκπαίδευση και λειτουργία της Τεχνητής Νοημοσύνης πρέπει να προστατεύονται μέσω ολοκληρωμένης διακυβέρνησης.
- Οι αρχές ελαχιστοποίησης δεδομένων πρέπει να καθοδηγούν τις πρακτικές συλλογής και διατήρησης
- Τα ισχυρά μέτρα ασφαλείας πρέπει να προστατεύουν τις ευαίσθητες πληροφορίες από μη εξουσιοδοτημένη πρόσβαση.

Διαφάνεια

- Οι δυνατότητες και οι περιορισμοί της Τεχνητής Νοημοσύνης πρέπει να κοινοποιούνται με σαφήνεια στους εκπαιδευτικούς φορείς.
- Οι διαδικασίες λήψης αποφάσεων πρέπει να είναι εξηγήσιμες σε μη τεχνική γλώσσα
- Η τεκμηρίωση πρέπει να επιτρέπει την επαλήθευση της απόδοσης και της συμμόρφωσης του συστήματος.

Ποικιλομορφία , Μη Διακρίσεις και Δικαιοσύνη

- Οι εφαρμογές που αφορούν ειδικά την ΕΕΚ πρέπει να λαμβάνουν υπόψη τα διαφορετικά υπόβαθρα των μαθητών και τις μαθησιακές τους ανάγκες.
- Οι μεθοδολογίες αξιολόγησης πρέπει να αξιολογούν τον αντίκτυπο στην ισότητα σε όλες τις δημογραφικές διαστάσεις
- Οι στρατηγικές μετριασμού των προκαταλήψεων πρέπει να λαμβάνουν υπόψη τις μοναδικές παραμέτρους ποικιλομορφίας της επαγγελματικής εκπαίδευσης.

Κοινωνικό και Περιβαλλοντικό Ευημερία

- Οι ευρείες κοινωνικές επιπτώσεις της Τεχνητής Νοημοσύνης στην επαγγελματική κατάρτιση πρέπει να αξιολογούνται συστηματικά.
- Πρέπει να αξιολογηθούν οι περιβαλλοντικές συνέπειες της υπολογιστικής υποδομής
- Πρέπει να ληφθούν υπόψη οι πιθανές επιπτώσεις στην αγορά εργασίας από την πιστοποίηση δεξιοτήτων Τεχνητής Νοημοσύνης.

Ευθύνη

- Σαφείς δομές διακυβέρνησης πρέπει να καθορίζουν την ευθύνη για τα συστήματα Τεχνητής Νοημοσύνης.
- Οι μηχανισμοί ελέγχου πρέπει να επιτρέπουν την επαλήθευση της κανονιστικής συμμόρφωσης



- Οι διαδικασίες αποκατάστασης πρέπει να αντιμετωπίζουν επιβλαβή αποτελέσματα ή ακούσιες συνέπειες.

Εφαρμογή του ALTAI σε περιβάλλοντα επαγγελματικής εκπαίδευσης και κατάρτισης

Τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης μπορούν να θέσουν σε λειτουργία την ALTAI μέσω συστηματικών πρωτοκόλλων αξιολόγησης:

Η αρχική αξιολόγηση εφαρμογής αξιολογεί τις προτεινόμενες εφαρμογές τεχνητής νοημοσύνης με βάση κριτήρια αξιοπιστίας

Η επαλήθευση πριν από την ανάπτυξη διασφαλίζει τη συμμόρφωση με τις τεχνικές και δεοντολογικές απαιτήσεις

Ο έλεγχος μετά την εφαρμογή επικυρώνει την απόδοση σε ποικίλους μαθητικούς πληθυσμούς

Η περιοδική επαναξιολόγηση εντοπίζει αναδυόμενα ηθικά ζητήματα καθώς τα συστήματα εξελίσσονται.

Ενσωματώνοντας την ALTAI στα πλαίσια διακυβέρνησης, τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης μπορούν να επιδείξουν δέουσα επιμέλεια στη συμμόρφωση με τους κανονισμούς, ενισχύοντας παράλληλα την εμπιστοσύνη των ενδιαφερόμενων μερών στην εφαρμογή της Τεχνητής Νοημοσύνης.

3.4 Μελέτες Περιπτώσεων και Παραδείγματα Αποτελεσματικών Πολιτικών Τεχνητής Νοημοσύνης σε Ιδρύματα Επαγγελματικής Εκπαίδευσης και Κατάρτισης

3.4.1 Θεμελιώδη Ευρωπαϊκά Πλαίσια Πολιτικής

Οι αποτελεσματικές θεσμικές πολιτικές βασίζονται σε θεμελιώδη πλαίσια που έχουν θεσπιστεί σε ευρωπαϊκό επίπεδο:

Ομάδα Εμπειρογνομόνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη (AI HLEG)



Η Ομάδα Υψηλού Επιπέδου Τεχνητής Νοημοσύνης (HLEG) της Ευρωπαϊκής Επιτροπής έχει αναπτύξει ουσιαστική καθοδήγηση μέσω:

Οδηγίες Δεοντολογίας για Αξιοπίστη Τεχνητή Νοημοσύνη : Καθιέρωση επτά απαιτήσεων για την εφαρμογή της Τεχνητής Νοημοσύνης ως ηθικής

Συστάσεις πολιτικής και επενδύσεων : Παροχή στρατηγικής κατεύθυνσης για βιώσιμη ανάπτυξη της τεχνητής νοημοσύνης

Λίστα Αξιολόγησης για Αξιοπίστη Τεχνητή Νοημοσύνη : Προσφορά πρακτικών εργαλείων εφαρμογής για οργανισμούς.

Αυτά τα θεμελιώδη πλαίσια έχουν διαμορφώσει συγκεκριμένες θεσμικές πολιτικές σε όλους τους ευρωπαίους παρόχους επαγγελματικής εκπαίδευσης και κατάρτισης.

Όπως παρατηρούν οι Holmes et al. (2023) στην ανάλυσή τους για τα πλαίσια διακυβέρνησης της Τεχνητής Νοημοσύνης, οι κατευθυντήριες γραμμές AI HLEG αντιπροσωπεύουν «μια κρίσιμη γέφυρα μεταξύ αφηρημένων ηθικών αρχών και συγκεκριμένων πρακτικών εφαρμογής» (σελ. 18). Η έρευνά τους εντοπίζει την επιρροή αυτών των κατευθυντήριων γραμμών σε ποικίλα εκπαιδευτικά πλαίσια, καταδεικνύοντας την προσαρμοστικότητά τους σε περιβάλλοντα επαγγελματικής κατάρτισης.

3.4.2 Τεχνικό Κολλέγιο Κοπεγχάγης: Ολοκληρωμένο Πλαίσιο Διακυβέρνησης Τεχνητής Νοημοσύνης

Το Τεχνικό Κολλέγιο της Κοπεγχάγης έχει δημιουργήσει μια πολυεπίπεδη δομή διακυβέρνησης που επιδεικνύει υποδειγματική πρακτική στην εφαρμογή της Τεχνητής Νοημοσύνης στην ΕΕΚ.

Αρχιτεκτονική Πολιτικής και Δομή Διακυβέρνησης

Η προσέγγιση του ιδρύματος περιλαμβάνει τα εξής χαρακτηριστικά:

- Επιτροπή Δεοντολογίας Τεχνητής Νοημοσύνης από πολλαπλούς φορείς, συμπεριλαμβανομένων διοικητικών στελεχών, εκπαιδευτικών, φοιτητών, εκπροσώπων του κλάδου και τεχνικών εμπειρογνομόνων
- Πρωτόκολλα έγκρισης σε επίπεδα που βασίζονται σε αλγοριθμική ταξινόμηση κινδύνου

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



- Υποχρεωτική εκπαίδευση στην Τεχνητή Νοημοσύνη για όλο το διδακτικό και διοικητικό προσωπικό
- Τυποποιημένα δικαιώματα των φοιτητών σχετικά με την αλγοριθμική διαφάνεια και την αμφισβήτηση αποφάσεων
- Δομημένα πρωτόκολλα αξιολόγησης που μετρούν τόσο τις εκπαιδευτικές όσο και τις ηθικές διαστάσεις.

Αυτή η προσέγγιση ευθυγραμμίζεται με το πλαίσιο εμπλοκής των ενδιαφερόμενων μερών των Prinsloo και Knox (2023), το οποίο τονίζει τη σημασία της «δομημένης εκπροσώπησης σε θεσμικές ιεραρχίες και επηρεαζόμενες κοινότητες» (σελ. 82). Η έρευνά τους καταδεικνύει πώς η διακυβέρνηση πολλαπλών ενδιαφερόμενων μερών ενισχύει τόσο την ηθική αυστηρότητα όσο και την αποτελεσματικότητα της εφαρμογής.

Μεθοδολογία Υλοποίησης

Τα βασικά στοιχεία υλοποίησης περιλαμβάνουν:

- Σταδιακή ανάπτυξη ξεκινώντας με εφαρμογές διαχείρισης χαμηλού κινδύνου
- Ελεγχόμενα παιδαγωγικά πιλοτικά προγράμματα με αυστηρά πρωτόκολλα αξιολόγησης
- Δομημένη τεκμηρίωση τόσο της τεχνικής απόδοσης όσο και του ηθικού αντίκτυπου
- Μηχανισμοί τακτικής ανατροφοδότησης από τα ενδιαφερόμενα μέρη για την ενημέρωση της επαναληπτικής βελτίωσης
- Ανεξάρτητος έλεγχος αλγοριθμικών αποτελεσμάτων σε όλες τις δημογραφικές διαστάσεις.

Τεκμηριωμένα Αποτελέσματα

Το Κολλέγιο έχει αναφέρει σημαντικά οφέλη μέσω της δομημένης εφαρμογής:

- Μείωση 27% του διοικητικού φόρτου εργασίας μέσω ηθικού αυτοματισμού τεχνητής νοημοσύνης
- Βελτιωμένη εξατομίκευση της εκπαίδευσης χωρίς συμβιβασμούς στην ποιότητα της διδασκαλίας
- Βελτιωμένο διαφάνεια στην αξιολόγηση πρακτικές
- Αυξημένη ικανοποίηση των φοιτητών με την τεχνολογική ενσωμάτωση
- Επιτυχής συμμόρφωση με τις αναδυόμενες κανονιστικές απαιτήσεις



3.4.3 Ινστιτούτο Τεχνολογίας του Μπορντό: Διαφανείς Πολιτικές Αξιολόγησης Τεχνητής Νοημοσύνης

Το Ινστιτούτο Μπορντό έχει αναπτύξει πρότυπες πολιτικές που αντιμετωπίζουν μία από τις πιο ηθικά πολύπλοκες πτυχές της εκπαιδευτικής Τεχνητής Νοημοσύνης: την αξιολόγηση και αξιολόγηση των μαθητών.

Πλαίσιο πολιτικής ειδικά για την αξιολόγηση

Βασικά στοιχεία πολιτικής περιλαμβάνουν:

- Απαγόρευση πλήρως αυτοματοποιημένης αξιολόγησης υψηλού διακυβεύματος
- Υποχρεωτική ανθρώπινη αναθεώρηση όλων των αποφάσεων αξιολόγησης που επηρεάζονται από την Τεχνητή Νοημοσύνη
- Διαφανής τεκμηρίωση της αλγοριθμικής συμβολής στις αξιολογήσεις
- Δομημένος μηχανισμός προσφυγών για την αμφισβήτηση αποτελεσμάτων που επηρεάζονται από την Τεχνητή Νοημοσύνη
- Τακτικοί έλεγχοι μεροληψίας που συγκρίνουν τα αποτελέσματα αξιολόγησης σε όλους τους μαθητικούς πληθυσμούς
- Εκπροσώπηση των φοιτητών στις δομές διακυβέρνησης της αξιολόγησης

Αυτή η προσέγγιση αποτελεί παράδειγμα αυτού που οι Ramirez-Montoya et al. (2023) προσδιορίζουν ως «ανθρωποκεντρική διακυβέρνηση αξιολόγησης» που χαρακτηρίζεται από «αλγοριθμική διαφάνεια, συμμετοχή ενδιαφερομένων μερών και τυποποιημένους μηχανισμούς εποπτείας» (σελ. 149). Η έρευνά τους καταδεικνύει πώς τέτοια πλαίσια ενισχύουν τόσο την εγκυρότητα της αξιολόγησης όσο και την εμπιστοσύνη των ενδιαφερομένων μερών.

Προσέγγιση Υλοποίησης

Η στρατηγική υλοποίησης δίνει έμφαση:

- Σαφής οριοθέτηση μεταξύ κατάλληλης και ακατάλληλης αυτοματοποίησης αξιολόγησης
- Σταδιακή εφαρμογή με βάση τα διακυβεύματα και τις συνέπειες της αξιολόγησης
- Ανάπτυξη διδακτικού προσωπικού σε μεθοδολογίες αξιολόγησης ηθικής τεχνητής νοημοσύνης



- Τεχνική υποδομή που διασφαλίζει την επεξηγηματικότητα των αλγοριθμικών συστάσεων
- Πρωτόκολλα τεκμηρίωσης που επιτρέπουν την επαλήθευση της ισότητας της αξιολόγησης.

Μετρήσιμος αντίκτυπος

Το Ινστιτούτο έχει καταγράψει σημαντικά οφέλη:

- Βελτιωμένη συνέπεια αξιολόγησης διατηρώντας παράλληλα την αυτονομία του εκπαιδευτή
- Βελτιωμένη εμπιστοσύνη των φοιτητών στις μεθοδολογίες αξιολόγησης
- Μειωμένος φόρτος εργασίας αξιολόγησης χωρίς συμβιβασμούς στην ποιότητα
- Καλύτερη αναγνώριση των μαθητών που χρειάζονται πρόσθετη υποστήριξη
- Επιτυχής συμμόρφωση με τις κανονιστικές απαιτήσεις διαφάνειας της Τεχνητής Νοημοσύνης.

3.4.4 Κοινοπραξία Επαγγελματικών Σχολών του Μονάχου: Συνεργατική Διακυβέρνηση Δεδομένων

Οκτώ επαγγελματικά ιδρύματα στο Μόναχο έχουν καθιερώσει μια συνεργατική προσέγγιση στη διακυβέρνηση δεδομένων τεχνητής νοημοσύνης, καταδεικνύοντας πώς η συγκέντρωση πόρων μπορεί να ενισχύσει την ηθική εφαρμογή.

Αρχιτεκτονική κοινής διακυβέρνησης

Βασικά δομικά στοιχεία περιλαμβάνουν:

- Κοινό δεοντολογικό πλαίσιο για τη συλλογή και αξιοποίηση δεδομένων
- Κοινή τεχνική υποδομή για διαχείριση δεδομένων και ασφάλεια
- Πρωτόκολλα συλλογικής αξιολόγησης προμηθευτών και σύναψης συμβάσεων
- Κοινή εφαρμογή μεθοδολογιών ομόσπονδης μάθησης
- Συγκεντρωμένη εμπειρογνωμοσύνη για την επαλήθευση της συμμόρφωσης με τους κανονισμούς.



Αυτή η συνεργατική προσέγγιση ευθυγραμμίζεται με αυτό που οι Rodriguez-Abitia και Bribiesca -Correa (2021) ονομάζουν «διακυβέρνηση βελτιστοποιημένη ως προς τους πόρους», η οποία χαρακτηρίζεται από «κοινή εμπειρογνωμοσύνη, ενοποιημένες πολιτικές και συλλογική ικανότητα εφαρμογής» (σελ. 98). Η έρευνά τους καταδεικνύει πώς τέτοιες προσεγγίσεις μπορούν να επιτρέψουν σε μικρότερα ιδρύματα να εφαρμόσουν εξελιγμένα πλαίσια διακυβέρνησης.

Μεθοδολογία Υλοποίησης

Η προσέγγιση της Κοινοπραξίας περιλαμβάνει τα εξής χαρακτηριστικά:

- Τυποποιημένα πρωτόκολλα προστασίας δεδομένων σε όλα τα συμμετέχοντα ιδρύματα
- Κεντρική εξειδίκευση σε τεχνικές ελαχιστοποίησης δεδομένων και ανωνυμοποίησης
- Κοινοί πόροι για την αξιολόγηση τεχνικής ευπάθειας
- Συνεργατική ανάπτυξη αναλυτικών μεθοδολογιών που διασφαλίζουν την ιδιωτικότητα
- Κοινή συμμετοχή των ενδιαφερόμενων μερών στην ανάπτυξη διακυβέρνησης

Αποδεδειγμένα οφέλη

Αυτή η συνεργατική προσέγγιση έχει αποφέρει:

- Πιο εξελιγμένη διακυβέρνηση δεδομένων από ό,τι θα μπορούσαν να επιτύχουν μεμονωμένα ιδρύματα ανεξάρτητα
- Βελτιωμένη διαπραγματευτική θέση με προμηθευτές τεχνολογίας
- Βελτιωμένα πρωτόκολλα ασφάλειας δεδομένων μέσω κοινής εμπειρογνωμοσύνης
- Οικονομικά αποδοτική συμμόρφωση με τους κανονισμούς σε πολλαπλά ιδρύματα
- Πιο ολοκληρωμένος μετριασμός της προκατάληψης μέσω μεγαλύτερων και πιο ποικίλων συνόλων δεδομένων.

3.4.5 Τεχνολογικό Ινστιτούτο Δουβλίνου: Πλαίσιο πολιτικής για την τεχνητή νοημοσύνη με επίκεντρο την ισότητα

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



Το Τεχνικό Ινστιτούτο του Δουβλίνου έχει αναπτύξει πολιτικές που δίνουν ιδιαίτερη προτεραιότητα στην αλγοριθμική ισότητα σε περιβάλλοντα επαγγελματικής εκπαίδευσης.

Πλαίσιο Πολιτικής με Επίκεντρο την Ισότητα

Βασικά στοιχεία πολιτικής περιλαμβάνουν:

- Υποχρεωτικές αξιολογήσεις επιπτώσεων στην ισότητα πριν από την ανάπτυξη της Τεχνητής Νοημοσύνης
- Τακτικός αλγοριθμικός έλεγχος σε όλες τις δημογραφικές διαστάσεις
- Προτεραιότητα του εξηγήσιμες μεθοδολογίες Τεχνητής Νοημοσύνης
- Απαίτηση για μη τεχνολογικές εναλλακτικές λύσεις στα συστήματα τεχνητής νοημοσύνης
- Επίσημη ένταξη των υποεκπροσωπούμενων πληθυσμών στις δομές διακυβέρνησης.

Αυτό το πλαίσιο αποτελεί παράδειγμα αυτού που οι Hodgkin και Kumar (2023) προσδιορίζουν ως «προληπτική διακυβέρνηση ισότητας» που χαρακτηρίζεται από «συστηματική αξιολόγηση προκατάληψης, συμπεριληπτική εκπροσώπηση ενδιαφερομένων μερών και πρωτόκολλα συνεχούς παρακολούθησης» (σελ. 219). Η έρευνά τους καταδεικνύει πώς η διακυβέρνηση με επίκεντρο την ισότητα ενισχύει τόσο την ηθική εφαρμογή όσο και τα εκπαιδευτικά αποτελέσματα.

Στρατηγική Υλοποίησης

Η προσέγγιση εφαρμογής δίνει έμφαση:

- Συστηματικός έλεγχος προκατάληψης σε αλγόριθμους επαγγελματικής καθοδήγησης
- Δημογραφική παρακολούθηση των εκπαιδευτικών αποτελεσμάτων που επηρεάζονται από την Τεχνητή Νοημοσύνη
- Ανάπτυξη μεθοδολογιών αποπροσανατολισμού για δεδομένα εκπαίδευσης
- Ανάπτυξη διδακτικού προσωπικού στην ενσωμάτωση της Τεχνητής Νοημοσύνης με γνώμονα την ισότητα
- Τακτική ανατροφοδότηση από ενδιαφερόμενους φορείς από ποικίλους φοιτητικούς πληθυσμούς.

Τεκμηριωμένος αντίκτυπος



Το Ινστιτούτο έχει αποδείξει:

- Εντοπισμός και διόρθωση πιθανών προκαταλήψεων στα συστήματα επαγγελματικού προσανατολισμού
- Βελτιωμένα εκπαιδευτικά αποτελέσματα για ιστορικά υποεκπροσωπούμενες ομάδες
- Αυξημένη εμπιστοσύνη των ενδιαφερόμενων μερών στην τεχνολογική ολοκλήρωση
- Επιτυχής συμμόρφωση με τις απαιτήσεις μη διακριτικής μεταχείρισης
- Ανάπτυξη μεθοδολογιών με επίκεντρο τη μεταβιβάσιμη μετοχική συμμετοχή

3.4.6 Εφαρμογή του ALTAI σε ειδικά πλαίσια επαγγελματικής εκπαίδευσης και κατάρτισης

Ο Κατάλογος Αξιολόγησης για την Αξιόπιστη Τεχνητή Νοημοσύνη (ALTAI) έχει εφαρμοστεί με επιτυχία σε πολλά πλαίσια επαγγελματικής εκπαίδευσης, αποδεικνύοντας την πρακτική του χρησιμότητα για την ηθική εφαρμογή.

Η έρευνα των Zawacki -Richter et al. (2023) προσδιορίζει την εφαρμογή της ALTAI στην επαγγελματική εκπαίδευση ως ιδιαίτερα πολύτιμη λόγω των «συγκεκριμένων επαγγελματικών πλαισίων που χαρακτηρίζουν την επαγγελματική κατάρτιση και των σαφών επαγγελματικών επιπτώσεων των αποφάσεων που επηρεάζονται από την Τεχνητή Νοημοσύνη» (σελ. 15). Η ανάλυση της μελέτης περίπτωσης τους καταδεικνύει την προσαρμοστικότητα της ALTAI σε ποικίλα περιβάλλοντα επαγγελματικής κατάρτισης.

Εφαρμογές αξιολόγησης δεξιοτήτων που βασίζονται στην τεχνητή νοημοσύνη

Τα ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης έχουν εφαρμόσει το ALTAI για να διασφαλίσουν την ηθική εφαρμογή των αυτοματοποιημένων συστημάτων αξιολόγησης:

Ενσωμάτωση Ανθρώπινης Εποπτείας : Διασφάλιση της επικύρωσης αλγοριθμικών αξιολογήσεων από τους εκπαιδευτές

Επαλήθευση Δικαιοσύνης : Έλεγχος της ισότητας αξιολόγησης σε ποικίλους μαθητικούς πληθυσμούς



Απαιτήσεις Διαφάνειας : Σαφή και κατανοητά κριτήρια αξιολόγησης

Τεχνική Ευστάθεια : Εξασφάλιση αξιόπιστης λειτουργίας σε διάφορα πρότυπα απόδοσης των μαθητών.

Τεχνητή Νοημοσύνη στον Επαγγελματικό Προσανατολισμό και την Επαγγελματική Συμβουλευτική

Η εφαρμογή του ALTAI σε συστήματα συστάσεων σταδιοδρομίας έχει αντιμετωπίσει:

- Δυνητικός ενίσχυση του επαγγελματικός στερεότυπα
- Διαφάνεια στις συστάσεις μεθοδολογίες
- Προστασία δεδομένων στην προγνωστική μοντελοποίηση σταδιοδρομίας
- Προσβασιμότητα συστημάτων καθοδήγησης σε ποικίλους πληθυσμούς.

Προσαρμοστική Μάθηση με Τεχνητή Νοημοσύνη στην Επαγγελματική Κατάρτιση

Η εφαρμογή πλατφορμών προσαρμοστικής μάθησης έχει ωφεληθεί από την αξιολόγηση ALTAI μέσω:

- Εκτίμηση του αλγοριθμικός εξατομίκευση δικαιοσύνη
- Επαλήθευση του μάθηση μονοπάτι διαποικίληση
- Αξιολόγηση πιθανών φαινομένων φουσαλίδας φίλτρου
- Επικύρωση των ικανοτήτων εποπτείας του εκπαιδευτή.

3.4.7 Κοινά Στοιχεία Αποτελεσματικών Πολιτικών

Η ανάλυση της επιτυχούς εφαρμογής σε όλα τα ευρωπαϊκά ιδρύματα επαγγελματικής εκπαίδευσης και κατάρτισης αποκαλύπτει συνεπή πρότυπα πολιτικής:

Δομές διακυβέρνησης πολλαπλών ενδιαφερομένων μερών που περιλαμβάνουν ποικίλες προοπτικές στην ανάπτυξη και την εποπτεία

Πρωτόκολλα αξιολόγησης κινδύνου σε διαβαθμισμένες κατηγορίες που εφαρμόζουν αναλογικό έλεγχο με βάση τον πιθανό αντίκτυπο

Μηχανισμοί συνεχούς παρακολούθησης αντί για διαδικασίες έγκρισης μίας φορές



Αυστηρές απαιτήσεις διαφάνειας που διασφαλίζουν την κατανόηση των δυνατοτήτων και των περιορισμών της Τεχνητής Νοημοσύνης από τα ενδιαφερόμενα μέρη

Ουσιαστική ανθρώπινη εποπτεία, ειδικά σε επακόλουθες εκπαιδευτικές αποφάσεις

Τακτικοί κύκλοι αναθεώρησης πολιτικής που προσαρμόζονται στις εξελισσόμενες τεχνολογίες και κανονισμούς

Ολοκληρωμένα εκπαιδευτικά στοιχεία που ενισχύουν την παιδεία στην Τεχνητή Νοημοσύνη σε όλα τα ενδιαφερόμενα μέρη

Σαφείς δομές λογοδοσίας που καθορίζουν την ευθύνη για τα αποτελέσματα που επηρεάζονται από την Τεχνητή Νοημοσύνη

Οι Holmes et al. (2023) προσδιορίζουν αυτά τα συνεπή στοιχεία ως αντιπροσωπευτικά μιας «αναδυόμενης συναίνεσης σχετικά με τις θεμελιώδεις αρχές διακυβέρνησης για την εφαρμογή της εκπαιδευτικής Τεχνητής Νοημοσύνης» (σελ. 21). Η συγκριτική τους ανάλυση σε διαφορετικά θεσμικά πλαίσια καταδεικνύει την προσαρμοστικότητα αυτών των βασικών αρχών σε ποικίλα εκπαιδευτικά περιβάλλοντα.

Τα ιδρύματα που επιδεικνύουν υποδειγματική πρακτική έχουν εξισορροπήσει το δυναμικό καινοτομίας με τις κατάλληλες διασφαλίσεις, αναγνωρίζοντας ότι οι υπερβολικά περιοριστικές προσεγγίσεις ενδέχεται να περιορίσουν τα εκπαιδευτικά οφέλη, ενώ η ανεπαρκής διακυβέρνηση δημιουργεί σημαντικούς κινδύνους. Τα πιο επιτυχημένα πλαίσια έχουν ενσωματώσει ηθικές παραμέτρους σε όλο τον κύκλο ζωής της υλοποίησης αντί να τις αντιμετωπίζουν ως εκ των υστέρων παραμέτρους.

4. Ηθικές Δεξιότητες

4.1 Ορισμός Ηθικών Δεξιοτήτων

Οι ηθικές δεξιότητες αναφέρονται στην ικανότητα αξιολόγησης του κατά πόσον μια Τεχνητή Νοημοσύνη χρησιμοποιείται ηθικά σε ένα συγκεκριμένο πλαίσιο. Στο εκπαιδευτικό πλαίσιο, για παράδειγμα, οι ηθικές δεξιότητες αφορούν το κατά πόσον οι πολιτικές ενός εκπαιδευτικού κέντρου είναι ηθικές, εάν διατηρείται η ασφάλεια και το απόρρητο των δεδομένων κ.λπ.



Οι ηθικές δεξιότητες αναφέρονται στις κρίσιμες ικανότητες που απαιτούνται για την αξιολόγηση και τη διασφάλιση της ηθικής χρήσης της Τεχνητής Νοημοσύνης και των δεδομένων σε εκπαιδευτικά περιβάλλοντα, με βάση αρχές όπως η ανθρώπινη δράση, η ισότητα, η ανθρωπιά και η δικαιολογημένη επιλογή. Η ανθρώπινη δράση περιλαμβάνει την παροχή δυνατότητας στα άτομα να κάνουν ενημερωμένες επιλογές και να αναλαμβάνουν την ευθύνη για τις πράξεις τους, ενισχύοντας την αυτονομία και την λογοδοσία. Η ισότητα διασφαλίζει δίκαιη μεταχείριση και ίση πρόσβαση σε ευκαιρίες για όλους, προωθώντας την ένταξη και την απαγόρευση των διακρίσεων. Η ανθρωπιά δίνει έμφαση στον σεβασμό της ανθρώπινης αξιοπρέπειας, της ταυτότητας και της ευημερίας, δίνοντας προτεραιότητα στις ουσιαστικές ανθρώπινες συνδέσεις έναντι της οργανικής χρήσης δεδομένων. Η δικαιολογημένη επιλογή απαιτεί διαφανείς διαδικασίες λήψης αποφάσεων που βασίζονται σε ηθικές αρχές και συμμετοχή των ενδιαφερόμενων μερών, διασφαλίζοντας ότι τα συστήματα Τεχνητής Νοημοσύνης στην εκπαίδευση αναπτύσσονται υπεύθυνα και ωφελούν όλα τα εμπλεκόμενα μέρη.

4.2 Σημασία ηθικών δεξιοτήτων

Αυτές οι ηθικές παραμέτρους είναι κρίσιμες για την καθοδήγηση των εκπαιδευτικών και των εκπαιδευτικών ιδρυμάτων στην αντιμετώπιση της πολυπλοκότητας της εφαρμογής της Τεχνητής Νοημοσύνης. Τηρώντας τις ηθικές απαιτήσεις, όπως αυτές που προτείνονται από τις Οδηγίες Δεοντολογίας της Τεχνητής Νοημοσύνης, οι εκπαιδευτικοί φορείς μπορούν να μετριάσουν αποτελεσματικά τους κινδύνους, να προωθήσουν τη δικαιοσύνη και να διαφυλάξουν την ευημερία των μαθητών και του προσωπικού.

4.3 Ηθικά Δεξιότητες περιοχές και παραδείγματα

Τομέας 1: Επαγγελματική Συμμετοχή. Χρήση ψηφιακών τεχνολογιών για επικοινωνία, συνεργασία και επαγγελματική ανάπτυξη.				
Αρμοδιότητα στοιχείο				Δυνητικός δείκτης
Είναι	σε	θέση	να	Συμμετέχει ενεργά στη συνεχή επαγγελματική μάθηση σχετικά με την Τεχνητή Νοημοσύνη και
περιγράφει	κριτικά	τις		



θετικές και αρνητικές επιπτώσεις της Τεχνητής Νοημοσύνης και της χρήσης δεδομένων στην εκπαίδευση.	την ανάλυση μάθησης, καθώς και την ηθική χρήση τους.
	Ικανός να δώσει παραδείγματα συστημάτων Τεχνητής Νοημοσύνης και να περιγράψει τη σημασία τους.
	Γνωρίζει πώς αξιολογείται ο ηθικός αντίκτυπος των συστημάτων Τεχνητής Νοημοσύνης στο σχολείο.
	Γνωρίζει πώς να ξεκινά και να προωθεί στρατηγικές σε ολόκληρο το σχολείο και την ευρύτερη κοινότητά του που προωθούν την ηθική και υπεύθυνη χρήση της Τεχνητής Νοημοσύνης και των δεδομένων.
Κατανοήστε τα βασικά της Τεχνητής Νοημοσύνης και της μαθησιακής ανάλυσης	Γνωρίζοντας ότι οι αλγόριθμοι Τεχνητής Νοημοσύνης λειτουργούν με τρόπους που συνήθως δεν είναι ορατοί ή εύκολα κατανοητοί από τους χρήστες.
	Ικανός να αλληλεπιδρά και να παρέχει ανατροφοδότηση στο σύστημα Τεχνητής Νοημοσύνης για να επηρεάσει τις επόμενες προτάσεις του.
	Γνωρίζοντας ότι οι αισθητήρες που χρησιμοποιούνται σε πολλές ψηφιακές τεχνολογίες και εφαρμογές παράγουν μεγάλες ποσότητες δεδομένων, συμπεριλαμβανομένων προσωπικών δεδομένων, που μπορούν να χρησιμοποιηθούν για την εκπαίδευση ενός συστήματος Τεχνητής Νοημοσύνης.



	Γνώση των κατευθυντήριων γραμμών δεοντολογίας της ΕΕ για την τεχνητή νοημοσύνη και των μέσων αυτοαξιολόγησης.
Παράδειγμα 1: Η Άννα είναι μια δασκάλα που θέλει να διδάξει ένα μάθημα μηχανικής ρευστών και θέλει να χρησιμοποιήσει την Τεχνητή Νοημοσύνη για να δημιουργήσει μερικές εικόνες και να τις συμπεριλάβει στο powerpoint του μαθήματος. Πριν συμπεριλάβει την Τεχνητή Νοημοσύνη στο μάθημα, η Άννα έχει διαβάσει τις οδηγίες δεοντολογίας της ΕΕ για την Τεχνητή Νοημοσύνη και έχει αξιολογήσει αν είναι ηθικό να συμπεριληφθεί η χρήση της ή όχι.	
Παράδειγμα 2: Ο Λούκας διδάσκει στην ΕΕΚ Περιβαλλοντικής Εκπαίδευσης και Ελέγχου. Αρκετοί εκπαιδευτικοί του εκπαιδευτικού του κέντρου χρησιμοποιούν Τεχνητή Νοημοσύνη, επομένως έχει προτείνει μια συνάντηση με ολόκληρο το εκπαιδευτικό κέντρο για να αναλυθεί η χρήση της Τεχνητής Νοημοσύνης από τους εκπαιδευτικούς. Στόχος του Λούκας είναι να διασφαλίσει ότι όλοι οι εκπαιδευτικοί κάνουν ηθική και υπεύθυνη χρήση της Τεχνητής Νοημοσύνης και των δεδομένων.	

Τομέας 2: Ψηφιακοί Πόροι: Προμήθεια, δημιουργία και κοινή χρήση ψηφιακών πόρων	
Αρμοδιότητα στοιχείο	Δυνητικός δείκτης
Διακυβέρνηση δεδομένων	Επίγνωση των διαφόρων μορφών προσωπικών δεδομένων που χρησιμοποιούνται στην εκπαίδευση και την κατάρτιση.
	Επίγνωση των ευθυνών για τη διατήρηση της ασφάλειας και του απορρήτου των δεδομένων.



	Γνωρίζει ότι η επεξεργασία προσωπικών δεδομένων υπόκειται σε εθνικούς και ενωσιακούς κανονισμούς, συμπεριλαμβανομένου του Γενικού Κανονισμού για την Προστασία Δεδομένων (GDPR).
	Γνωρίζει ποιος έχει πρόσβαση στα δεδομένα των μαθητών, πώς παρακολουθείται η πρόσβαση και για πόσο χρονικό διάστημα διατηρούνται τα δεδομένα.
	Γνωρίζει ότι όλοι οι πολίτες της ΕΕ έχουν το δικαίωμα να μην υπόκεινται σε πλήρως αυτοματοποιημένη λήψη αποφάσεων.
	Ικανός να δώσει παραδείγματα ευαίσθητων δεδομένων, συμπεριλαμβανομένων βιομετρικών δεδομένων.
	Ικανός να σταθμίζει τα οφέλη και τους κινδύνους πριν επιτρέψει σε τρίτους να επεξεργάζονται προσωπικά δεδομένα, ειδικά όταν χρησιμοποιούν συστήματα Τεχνητής Νοημοσύνης.
Διακυβέρνηση Τεχνητής Νοημοσύνης	Γνωρίζει ότι τα συστήματα Τεχνητής Νοημοσύνης υπόκεινται σε εθνικούς και ενωσιακούς κανονισμούς (ιδίως στον προς ψήφιση νόμο περί Τεχνητής Νοημοσύνης).
	Ικανός να εξηγήσει την προσέγγιση που βασίζεται στον κίνδυνο του νόμου για την τεχνητή νοημοσύνη (υπό έγκριση).
	Γνωρίζει τις περιπτώσεις χρήσης τεχνητής νοημοσύνης υψηλού κινδύνου στην εκπαίδευση και τις σχετικές απαιτήσεις του νόμου περί τεχνητής νοημοσύνης (όταν εγκριθεί).
	Γνωρίζει πώς να ενσωματώνει ψηφιακό περιεχόμενο επεξεργασμένο/χειραγωγημένο από Τεχνητή Νοημοσύνη



	στο δικό του έργο και πώς πρέπει να αναγνωρίζεται αυτό το έργο.
	Ικανός να εξηγήσει βασικές αρχές ποιότητας δεδομένων σε συστήματα Τεχνητής Νοημοσύνης.
Παράδειγμα 1: Η Μαρία είναι καθηγήτρια στην σχολή επαγγελματικής κατάρτισης Ενεργειακής Απόδοσης και Ηλιακής Θερμικής Ενέργειας. Δημιουργεί σημειώσεις για το μάθημά της, τις οποίες θα δώσει στους μαθητές, και έχει ενσωματώσει στις σημειώσεις ψηφιακό περιεχόμενο επεξεργασμένο με τεχνητή νοημοσύνη, χρησιμοποιώντας το εργαλείο Copilot. Η Μαρία γνωρίζει ότι αυτό το περιεχόμενο πρέπει να αναγνωρισθεί και φροντίζει να το αναγνωρίσει σωστά.	
Παράδειγμα 2: Ο Jaime διδάσκει στη σχολή επαγγελματικής κατάρτισης Μηχανολόγων Κατασκευών. Χρησιμοποιεί το ChatGPT μερικές φορές για να αναπτύξει ορισμένα περιεχόμενα ή δραστηριότητες του αντικειμένου του. Κάποια στιγμή αξιολογεί την πιθανότητα να συμπεριλάβει τα προσωπικά δεδομένα των μαθητών του (όνομα, επώνυμο, ημερομηνία γέννησης και διεύθυνση) για να δημιουργήσει ένα έγγραφο που περιλαμβάνει όλες αυτές τις πληροφορίες. Γνωρίζει την ευθύνη του για τη διατήρηση της ασφάλειας και του απορρήτου των δεδομένων, επομένως αποφάσισε να μην συμπεριλάβει τα προσωπικά δεδομένα.	

Τομέας 3: Χρήση ψηφιακών τεχνολογιών και στρατηγικών για την ενίσχυση της αξιολόγησης	
Αρμοδιότητα στοιχείο	Δυνητικός δείκτης
Μοντέλα του μάθησης	Γνωρίζει ότι τα συστήματα Τεχνητής Νοημοσύνης εφαρμόζουν την κατανόηση του σχεδιασμού σχετικά με το τι είναι η μάθηση και πώς μπορεί να μετρηθεί η μάθηση· μπορεί να εξηγήσει βασικές παιδαγωγικές υποθέσεις που υποστηρίζουν ένα δεδομένο ψηφιακό σύστημα μάθησης.



Στόχοι της εκπαίδευσης	Γνωρίζει πώς ένα δεδομένο ψηφιακό σύστημα αντιμετωπίζει τους διαφορετικούς κοινωνικούς στόχους της εκπαίδευσης (προσόντα, κοινωνικοποίηση, υποκειμενοποίηση).
Ανθρώπινη δράση	Ικανός να εξετάσει τον αντίκτυπο του συστήματος Τεχνητής Νοημοσύνης στην αυτονομία των εκπαιδευτικών, την επαγγελματική ανάπτυξη και την εκπαιδευτική καινοτομία.
	Εξετάστε τις πηγές απαράδεκτης μεροληψίας στην τεχνητή νοημοσύνη που βασίζεται σε δεδομένα.
Δικαιοσύνη	Λαμβάνει υπόψη τους κινδύνους που σχετίζονται με τη συναισθηματική εξάρτηση και την αυτοεικόνα των μαθητών κατά τη χρήση διαδραστικών συστημάτων Τεχνητής Νοημοσύνης και μαθησιακών αναλύσεων.
Ανθρωπότητα	Ικανός να εξετάσει τον αντίκτυπο της Τεχνητής Νοημοσύνης και της χρήσης δεδομένων στην μαθητική κοινότητα.
	Έχω αυτοπεποίθηση στη συζήτηση των ηθικών πτυχών της Τεχνητής Νοημοσύνης και του τρόπου με τον οποίο επηρεάζουν τον τρόπο χρήσης της τεχνολογίας.
Συμμετέχει στην ανάπτυξη μαθησιακών πρακτικών που χρησιμοποιούν Τεχνητή Νοημοσύνη και δεδομένα	Μπορεί να εξηγήσει πώς οι ηθικές αρχές και αξίες λαμβάνονται υπόψη και διαπραγματεύονται στον από κοινού σχεδιασμό και τη από κοινού δημιουργία μαθησιακών πρακτικών που χρησιμοποιούν Τεχνητή Νοημοσύνη και δεδομένα (που συνδέονται με τον μαθησιακό σχεδιασμό).
Παράδειγμα 1: Η Luisa, καθηγήτρια Ηλεκτρομηχανολογικής Συντήρησης σε μια εταιρεία εκπαίδευσης, εφαρμόζει λογισμικό Τεχνητής Νοημοσύνης που	



χρησιμοποιεί προσαρμοστικά μοντέλα μάθησης για την εξατομίκευση της διδασκαλίας. Το λογισμικό αξιολογεί το αρχικό επίπεδο κάθε μαθητευόμενου και προσαρμόζει το περιεχόμενο και τις εργασίες με βάση την πρόοδό του. Αυτό επιτρέπει σε κάθε μαθητευόμενο να προοδεύει με τον δικό του ρυθμό, λαμβάνοντας άμεση ανατροφοδότηση και συγκεκριμένους πόρους για να ξεπεράσει τις δυσκολίες. Επιπλέον, η Luisa μπορεί να έχει πρόσβαση σε λεπτομερείς αναφορές για να εντοπίσει τομείς όπου οι μαθητές της χρειάζονται πρόσθετη υποστήριξη, διασφαλίζοντας την συμπεριληπτική και αποτελεσματική διδασκαλία.

Παράδειγμα 2: Ο Javi, καθηγητής Γραφικών Τεχνών στο πρόγραμμα Γραφιστικής Εκτύπωσης μεσαίου επιπέδου μιας σχολής ΕΕΚ, ενσωματώνει λογισμικό Τεχνητής Νοημοσύνης που ενισχύει τις ανθρώπινες ικανότητες βοηθώντας σε σύνθετες εργασίες σχεδιασμού. Το λογισμικό βοηθά τους μαθητές αυτοματοποιώντας επαναλαμβανόμενες εργασίες, όπως η διόρθωση χρωμάτων και οι προσαρμογές διάταξης, επιτρέποντάς τους να επικεντρωθούν σε δημιουργικές και στρατηγικές πτυχές των έργων τους. Αυτό δίνει τη δυνατότητα στους μαθητές να αναπτύξουν τις καλλιτεχνικές τους δεξιότητες και την κριτική τους σκέψη, ενώ η Τεχνητή Νοημοσύνη χειρίζεται τεχνικές λεπτομέρειες.

Τομέας 4: Αξιολόγηση: Διαχείριση και ενορχήστρωση της χρήσης ψηφιακών τεχνολογιών στη διδασκαλία και τη μάθηση.

Αρμοδιότητα στοιχείο	Δυνητικός δείκτης
Προσωπικές διαφορές	Γνωρίζοντας ότι οι μαθητές αντιδρούν με διαφορετικούς τρόπους στην αυτοματοποιημένη ανατροφοδότηση.
Αλγοριθμικό προκατάληψη	Εξετάστε τις πηγές της απαράδεκτης μεροληψίας στα συστήματα Τεχνητής Νοημοσύνης και πώς μπορούν να μετριαστούν.



Γνωστική εστίαση	Επίγνωση ότι τα συστήματα Τεχνητής Νοημοσύνης αξιολογούν την πρόοδο των μαθητών με βάση προκαθορισμένα μοντέλα γνώσης που αφορούν συγκεκριμένους τομείς.
	Γνωρίζοντας ότι τα περισσότερα συστήματα Τεχνητής Νοημοσύνης δεν αξιολογούν τη συνεργασία, τις κοινωνικές ικανότητες ή τη δημιουργικότητα.
	Επίγνωση των συνηθισμένων τρόπων χειραγώγησης της αξιολόγησης που βασίζεται στην Τεχνητή Νοημοσύνη.
Νέοι κατάχρησης τεχνολογίας	Γνωρίζοντας ότι οι μαθητές αντιδρούν με διαφορετικούς τρόπους στην αυτοματοποιημένη ανατροφοδότηση.
Παράδειγμα 1: Ο Χόρχε διδάσκει ένα μάθημα στη σχολή ΕΕΚ Εγκαταστάσεων και Συντήρησης. Αφού ολοκληρώσει το 1ο ^{θέμα} του μαθήματος, χρησιμοποιεί ένα εργαλείο Τεχνητής Νοημοσύνης για να αξιολογήσει την κατανόηση του περιεχομένου από τους μαθητές του. Επιπλέον, αυτό το εργαλείο Τεχνητής Νοημοσύνης θα παρέχει αυτόματα ανατροφοδότηση και θα προσαρμόζει το επίπεδο των ερωτήσεων με βάση τις προηγούμενες απαντήσεις. Ο Χόρχε γνωρίζει πώς αντιδρούν οι μαθητές στην αυτοματοποιημένη ανατροφοδότηση, επομένως φροντίζει να τους παρέχει εξατομικευμένη ανατροφοδότηση και να ελέγχει ο ίδιος την πρόοδο των μαθητών. Επομένως, χρησιμοποιεί την Τεχνητή Νοημοσύνη μόνο ως εργαλείο για να δώσει μια αρχική αξιολόγηση, την οποία ελέγχει και αναθεωρεί ο ίδιος στη συνέχεια.	
Παράδειγμα 2: Η Ναταλία είναι εκπαιδεύτρια σε έναν πάροχο υπηρεσιών ΕΕΚ για ανανεώσιμες πηγές ενέργειας. Ενθαρρύνει τους μαθητές να χρησιμοποιούν το Windows Copilot στα μαθήματα εκπαίδευσής της για ορισμένες δραστηριότητες. Ακόμα κι αν αυτό το εργαλείο λειτουργεί ως βοηθός Τεχνητής Νοημοσύνης που μπορεί να είναι χρήσιμο για την παραγωγικότητα και τη δημιουργικότητα, γνωρίζει ότι χρειάζονται άλλα είδη δραστηριοτήτων και αξιολόγησης για τη βελτίωση της συνεργασίας, των κοινωνικών ικανοτήτων και της δημιουργικότητας, καθώς η Τεχνητή Νοημοσύνη δεν είναι σε θέση να	



αξιολογήσει αυτές τις πτυχές. Συνεχίζει να αναλαμβάνει τις σχετικές δραστηριότητες και την αξιολόγηση ως εκπαιδευτρια η ίδια και χρησιμοποιεί την Τεχνητή Νοημοσύνη ως εργαλείο κατά τη διάρκεια των μαθημάτων της και για την προετοιμασία της εκπαίδευσης.

Τομέας 5: Ενδυνάμωση των Μαθητών: Χρήση ψηφιακών τεχνολογιών για την ενίσχυση της ένταξης, της εξατομίκευσης και της ενεργού συμμετοχής των μαθητών

Αρμοδιότητα στοιχείο	Δυνητικός δείκτης
Η Τεχνητή Νοημοσύνη ανταποκρίνεται στις ποικίλες μαθησιακές ανάγκες των μαθητών	Γνωρίζει τους διαφορετικούς τρόπους με τους οποίους τα εξατομικευμένα συστήματα μάθησης μπορούν να προσαρμόσουν τη συμπεριφορά τους (περιεχόμενο, μαθησιακή πορεία, παιδαγωγική προσέγγιση).
	Ικανός να εξηγήσει πώς ένα δεδομένο σύστημα μπορεί να ωφελήσει όλους τους μαθητές, ανεξάρτητα από τις γνωστικές, πολιτισμικές, οικονομικές ή σωματικές τους διαφορές
	Γνωρίζοντας ότι τα ψηφιακά συστήματα μάθησης αντιμετωπίζουν διαφορετικά τις διαφορετικές ομάδες μαθητών
	Ικανός να λάβει υπόψη τον αντίκτυπο στην ανάπτυξη της αυτοαποτελεσματικότητας, της εικόνας του εαυτού, της νοοτροπίας και των γνωστικών και συναισθηματικών δεξιοτήτων αυτορρύθμισης των μαθητών.



Δικαιολογημένο επιλογή	Γνωρίζει ότι η Τεχνητή Νοημοσύνη και η χρήση δεδομένων μπορεί να ωφελήσουν ορισμένους μαθητές περισσότερο από άλλους.
	Ικανός να εξηγήσει ποια στοιχεία έχουν χρησιμοποιηθεί για να δικαιολογήσουν την ανάπτυξη ενός δεδομένου συστήματος Τεχνητής Νοημοσύνης στην τάξη.
	Αναγνωρίζει την ανάγκη για συνεχή παρακολούθηση των αποτελεσμάτων της χρήσης της Τεχνητής Νοημοσύνης.
Παράδειγμα 1: Η Τερέζα, διδάσκοντας Προηγμένη Επαγγελματική Κατάρτιση στη Διοίκηση και τα Χρηματοοικονομικά σε έναν πάροχο εκπαίδευσης ενηλίκων, χρησιμοποιεί Τεχνητή Νοημοσύνη για να προσαρμόσει τις μαθησιακές εμπειρίες. Το σύστημα Τεχνητής Νοημοσύνης αξιολογεί τις ανάγκες κάθε μαθητή και προσαρμόζει τους πόρους και τις δραστηριότητες ανάλογα. Για παράδειγμα, οι οπτικοί μαθητές μπορούν να λαμβάνουν διαδραστικά εργαλεία, ενώ οι αναλυτικοί μαθητές λαμβάνουν λεπτομερείς μελέτες περιπτώσεων. Αυτή η προσέγγιση διασφαλίζει εξατομικευμένη μάθηση που ανταποκρίνεται σε διαφορετικά στυλ μάθησης και μεγιστοποιεί τη συμμετοχή των μαθητών.	
Παράδειγμα 2: Ο Óscar , διδάσκοντας Προηγμένη Επαγγελματική Κατάρτιση στον Μηχανολογικό Σχεδιασμό Κατασκευών, χρησιμοποιεί λογισμικό Τεχνητής Νοημοσύνης για τον σχεδιασμό μηχανολογικών εξαρτημάτων. Το λογισμικό προσφέρει πολλαπλές λύσεις σχεδιασμού και προσομοιώνει την απόδοση σε διαφορετικές συνθήκες. Οι μαθητές αναλύουν τα αποτελέσματα για να επιλέξουν τον πιο αποτελεσματικό σχεδιασμό, καθοδηγούμενοι από τον Óscar για να αναπτύξουν κριτικές και αναλυτικές δεξιότητες κρίσιμες για την καριέρα τους.	

Τομέας 6: Διευκόλυνση της ψηφιακής ικανότητας των μαθητών, επιτρέποντάς τους να χρησιμοποιούν δημιουργικά και υπεύθυνα τις



ψηφιακές τεχνολογίες για την πληροφόρηση, την επικοινωνία, τη δημιουργία περιεχομένου, την ευημερία και την επίλυση προβλημάτων	
Αρμοδιότητα στοιχείο	Δυνητικός δείκτης
Ηθική Τεχνητής Νοημοσύνης και Αναλυτικής Μάθησης	Ικανός να χρησιμοποιεί έργα και αναπτύξεις Τεχνητής Νοημοσύνης για να βοηθήσει τους μαθητές να μάθουν για την ηθική της Τεχνητής Νοημοσύνης και τη χρήση δεδομένων στην εκπαίδευση και την κατάρτιση.
Παράδειγμα 1: Η καθηγήτρια Σάρα διδάσκει προηγμένη ανάλυση μάρκετινγκ στους φοιτητές της που διδάσκουν επαγγελματική κατάρτιση σε ένα πανεπιστήμιο. Συζητά μια μελέτη περίπτωσης όπου η Τεχνητή Νοημοσύνη αναλύει δεδομένα κοινωνικών μέσων για στοχευμένη διαφήμιση. Η Σάρα προκαλεί μια συζήτηση σχετικά με ηθικά ζητήματα όπως η ιδιωτικότητα, η συναίνεση και η αλγοριθμική προκατάληψη. Καθοδηγεί τους φοιτητές να αναπτύξουν ηθικές κατευθυντήριες γραμμές για τη χρήση της Τεχνητής Νοημοσύνης στο μάρκετινγκ, ενισχύοντας την κατανόησή τους για την υπεύθυνη ανάπτυξη της Τεχνητής Νοημοσύνης στον κλάδο.	
Παράδειγμα 2: Ο καθηγητής Jesús διδάσκει μαγειρική διαχείριση, εστιάζοντας στον ρόλο της Τεχνητής Νοημοσύνης στην εξατομίκευση των μενού εστιατορίων. Συζητά μια μελέτη περίπτωσης όπου η Τεχνητή Νοημοσύνη αναλύει τις διατροφικές προτιμήσεις για να προσφέρει εξατομικευμένες προτάσεις μενού. Ο Jesús εμπλέκει τους εκπαιδευόμενους σε μια συζήτηση σχετικά με ηθικά ζητήματα, όπως το απόρρητο των δεδομένων, η διασφάλιση της ποικιλομορφίας των μενού και ο αντίκτυπος της Τεχνητής Νοημοσύνης στην ικανοποίηση των πελατών. Καθοδηγεί τους εκπαιδευόμενους στην ανάπτυξη ηθικών πλαισίων για τη χρήση της Τεχνητής Νοημοσύνης στην εξατομίκευση των μενού, προετοιμάζοντάς τους να εξισορροπήσουν την καινοτομία με ηθικές παραμέτρους στις λειτουργίες των εστιατορίων.	

5. Συμπέρασμα



Καθώς η Τεχνητή Νοημοσύνη συνεχίζει να διαμορφώνει το εκπαιδευτικό σκηνικό, είναι απαραίτητο η υιοθέτηση τεχνολογιών Τεχνητής Νοημοσύνης να γίνεται με τρόπο που να τηρεί τις ηθικές αξίες, να προωθεί τη δικαιοσύνη και να διασφαλίζει την ιδιωτικότητα των μαθητών. Αυτό το εγχειρίδιο έχει σκιαγραφήσει τον σκοπό και το πεδίο εφαρμογής της Τεχνητής Νοημοσύνης, υπογραμμίζοντας τη σημασία των ηθικών παραμέτρων για να διασφαλιστεί ότι οι τεχνολογίες Τεχνητής Νοημοσύνης χρησιμοποιούνται υπεύθυνα σε όλα τα εκπαιδευτικά πλαίσια.

5.1 Βασικά συμπεράσματα

Η ενσωμάτωση της Τεχνητής Νοημοσύνης στην εκπαίδευση προσφέρει τεράστιες δυνατότητες, αλλά πρέπει να καθοδηγείται από ηθικές αρχές που δίνουν προτεραιότητα στην ποικιλομορφία, την ένταξη και τη διαφάνεια. Τα εργαλεία Τεχνητής Νοημοσύνης θα πρέπει να σχεδιάζονται για να εξυπηρετούν μαθητές από διαφορετικά υπόβαθρα, λαμβάνοντας υπόψη πολιτισμικές, γλωσσικές και κοινωνικοοικονομικές διαφορές, για την προώθηση της ισότητας στα εκπαιδευτικά αποτελέσματα. Η διαφάνεια είναι εξίσου ζωτικής σημασίας — οι εκπαιδευτικοί και οι μαθητές πρέπει να κατανοούν πώς τα συστήματα Τεχνητής Νοημοσύνης λαμβάνουν αποφάσεις, ειδικά σε τομείς όπως η βαθμολόγηση, η ανατροφοδότηση ή οι εξατομικευμένες μαθησιακές διαδρομές. Η διασφάλιση της ιδιωτικότητας και της προστασίας των δεδομένων είναι απαραίτητη, απαιτώντας πλήρη συμμόρφωση με κανονισμούς όπως ο ΓΚΠΔ και την τήρηση ηθικών πρακτικών διαχείρισης δεδομένων. Τα συστήματα Τεχνητής Νοημοσύνης πρέπει επίσης να είναι βιώσιμα, ελαχιστοποιώντας τις περιβαλλοντικές τους επιπτώσεις και διασφαλίζοντας τη μακροπρόθεσμη βιωσιμότητά τους στα εκπαιδευτικά οικοσυστήματα.

Εξίσου σημαντική είναι η προώθηση της παιδείας στην Τεχνητή Νοημοσύνη μεταξύ όλων των ενδιαφερόμενων μερών. Οι εκπαιδευτικοί χρειάζονται στοχευμένη εκπαίδευση για να ενσωματώσουν ηθικά και αποτελεσματικά τα εργαλεία Τεχνητής Νοημοσύνης στην παιδαγωγική τους πρακτική, ενώ οι μαθητές θα πρέπει να είναι εξοπλισμένοι με κριτική κατανόηση των δυνατοτήτων και των περιορισμών της Τεχνητής Νοημοσύνης. Οι διοικητικοί υπάλληλοι των ιδρυμάτων διαδραματίζουν βασικό ρόλο στην ευθυγράμμιση των πολιτικών των σχολείων ή των πανεπιστημίων με τις ηθικές πρακτικές Τεχνητής Νοημοσύνης, προωθώντας ένα περιβάλλον λογοδοσίας και δικαιοσύνης. Οι υπεύθυνοι χάραξης πολιτικής πρέπει να παρέμβουν για να δημιουργήσουν ισχυρά κανονιστικά πλαίσια που καθοδηγούν την υπεύθυνη υιοθέτηση της Τεχνητής Νοημοσύνης, μετριάζουν την αλγοριθμική προκατάληψη και αποτρέπουν τα



μεροληπτικά αποτελέσματα. Τελικά, μια προσέγγιση πολλαπλών ενδιαφερόμενων μερών είναι ζωτικής σημασίας για να διασφαλιστεί ότι η Τεχνητή Νοημοσύνη βελτιώνει την εκπαίδευση με τρόπο δίκαιο, χωρίς αποκλεισμούς και βιώσιμο.

Είναι επίσης πολύ σημαντικό να ληφθούν υπόψη ορισμένα βασικά νομικά και κανονιστικά πλαίσια, όπως ο νόμος της ΕΕ για την τεχνητή νοημοσύνη (2024), ο οποίος εισάγει ένα ολοκληρωμένο νομικό πλαίσιο που κατηγοριοποιεί τα συστήματα τεχνητής νοημοσύνης με βάση τον κίνδυνο που ενέχουν για τα θεμελιώδη δικαιώματα, με σημαντικές επιπτώσεις για τα εκπαιδευτικά ιδρύματα.

Με βάση αυτό, υπάρχουν εφαρμογές Τεχνητής Νοημοσύνης υψηλού κινδύνου, όπως αυτές που χρησιμοποιούνται στις αξιολογήσεις ή τις εισαγωγές φοιτητών, οι οποίες υπόκεινται σε αυστηρές απαιτήσεις συμμόρφωσης, όπως αυστηρή τεκμηρίωση, ανθρώπινη εποπτεία και ισχυρά πρωτόκολλα διαχείρισης κινδύνων.

Επιπλέον, η διαφάνεια και η λογοδοσία δεν αποτελούν επιλογή. Τα ιδρύματα πρέπει να αποκαλύπτουν με σαφήνεια τον τρόπο λειτουργίας των συστημάτων Τεχνητής Νοημοσύνης και να διασφαλίζουν ότι οι χρήστες μπορούν να κατανοούν και να αμφισβητούν τις αυτοματοποιημένες αποφάσεις. Επιπλέον, ο Νόμος επιβάλλει προληπτικά μέτρα για τον έλεγχο και τον μετριασμό της αλγοριθμικής προκατάληψης, διασφαλίζοντας τη δικαιοσύνη και αποτρέποντας τα μεροληπτικά αποτελέσματα. Για να υποστηρίξει αυτές τις προσπάθειες, ο Κατάλογος Αξιολόγησης για Αξιόπιστη Τεχνητή Νοημοσύνη (ALTAI) προσφέρει ένα πρακτικό εργαλείο που βοηθά τα ιδρύματα να αξιολογήσουν τον ηθικό αντίκτυπο των συστημάτων Τεχνητής Νοημοσύνης τους, καλλιεργώντας μια κουλτούρα υπεύθυνης και αξιόπιστης ανάπτυξης Τεχνητής Νοημοσύνης στην εκπαίδευση.

5.2 Τελικές σκέψεις

Θα πρέπει όλοι να γνωρίζουμε ότι, καθώς η Τεχνητή Νοημοσύνη συνεχίζει να αναδιαμορφώνει την πανεπιστημιακή και επαγγελματική εκπαίδευση, το παρόν εγχειρίδιο τονίζει τη σημασία της διατήρησης μιας προσεκτικής ισορροπίας μεταξύ καινοτόμων πρακτικών και ηθικής ευθύνης. Τα οφέλη της Τεχνητής Νοημοσύνης, ξεκινώντας από εξατομικευμένες μαθησιακές οδούς έως βελτιωμένη διοικητική αποτελεσματικότητα, είναι σημαντικά, ωστόσο συνοδεύονται από κινδύνους που δεν πρέπει να παραβλέπονται ή να αγνοούνται. Ζητήματα όπως η αλγοριθμική προκατάληψη, οι παραβιάσεις της ιδιωτικής ζωής των δεδομένων και ο μειωμένος ρόλος της ανθρώπινης κρίσης απαιτούν συνεχή προσοχή. Για την αντιμετώπιση αυτών των πολυπλοκοτήτων,



πρέπει να υπάρχει συνεχής συνεργασία μεταξύ εκπαιδευτικών, διοικητικών στελεχών και υπευθύνων χάραξης πολιτικής. Φυσικά, οι ηθικές αρχές πρέπει να είναι βαθιά ενσωματωμένες στην ανάπτυξη, την ανάπτυξη και την αξιολόγηση των εργαλείων Τεχνητής Νοημοσύνης, ώστε να διασφαλίζεται ότι υποστηρίζουν την ένταξη, σέβονται την αυτονομία και ενισχύουν την εμπιστοσύνη στην εκπαιδευτική διαδικασία.

Κοιτάζοντας μπροστά, η επιτυχής ενσωμάτωση της Τεχνητής Νοημοσύνης στην εκπαίδευση θα βασιστεί όχι μόνο στην τεχνολογική ετοιμότητα, αλλά και σε ένα κοινό όραμα που θα βασίζεται στη διαφάνεια, τη δικαιοσύνη και τη δια βίου μάθηση. Ο τομέας πρέπει να προωθήσει την παιδεία στην Τεχνητή Νοημοσύνη μεταξύ όλων των ενδιαφερόμενων μερών και να δημιουργήσει ισχυρά συστήματα λογοδοσίας και εποπτείας. Δεν πρόκειται για μια εφάπαξ προσαρμογή, αλλά για ένα εξελισσόμενο ταξίδι που απαιτεί συνεχή προβληματισμό και διάλογο. Υποστηρίζοντας τις ανθρωποκεντρικές προσεγγίσεις και την ηθική καινοτομία, η εκπαιδευτική κοινότητα μπορεί να μετατρέψει την Τεχνητή Νοημοσύνη σε έναν ισχυρό σύμμαχο που ενισχύει τις μαθησιακές ευκαιρίες, διατηρώντας παράλληλα τις αξίες για ποιοτική εκπαίδευση. Για ολοκληρωμένη καθοδήγηση, τα ενδιαφερόμενα μέρη ενθαρρύνονται να συμβουλευτούν τους πόρους του έργου AI Pioneers και το Evaluation Schema for AI in Education, τα οποία είναι προσβάσιμα σε πολλές γλώσσες.

Ευχαριστίες

Θα θέλαμε να ευχαριστήσουμε όλους όσους υποστήριξαν την ανάπτυξη αυτού του εγγράφου μέσω της εποικοδομητικής κριτικής και των ανεκτίμητων συνεισφορών τους. Είμαστε βαθιά ευγνώμονες σε όλους τους εταίρους του έργου μας, οι οποίοι συνέβαλαν στο περιεχόμενο, καθώς και στο Δίκτυο Αναφοράς Pioneers AI και στους συμμετέχοντες στα σεμινάρια διαβούλευσής μας. Οι κοινές ανησυχίες και οι απόψεις τους προσέφεραν πολύτιμη καθοδήγηση σε όλη αυτή τη διαδικασία.

Τέλος, αναγνωρίζουμε τη σημασία της βιβλιογραφίας που αναφέρεται, μαζί με διάφορες άλλες πηγές, όπως επιστημονικές εργασίες, εκθέσεις, σεμινάρια, οδηγίες από εκπαιδευτικά ιδρύματα και αναρτήσεις ιστολογίου από ερευνητές και εκπαιδευτικούς. Συλλογικά, αυτές οι συνεισφορές έχουν επηρεάσει σε μεγάλο βαθμό την προσέγγισή μας στην αντιμετώπιση ηθικών ζητημάτων στη χρήση της Τεχνητής Νοημοσύνης στην εκπαίδευση σε όλα τα επίπεδα.



Αναφορές

Akgun , S., & Greenhow, C. (2022). Τεχνητή νοημοσύνη στην εκπαίδευση: Αντιμετώπιση ηθικών προκλήσεων σε περιβάλλοντα Νηπιαγωγείου-Λυκείου. *AI and Ethics* , 2 (3), 431-440.

Aldosari , SAM (2020). Το μέλλον της τριτοβάθμιας εκπαίδευσης υπό το πρίσμα των μετασχηματισμών της τεχνητής νοημοσύνης. *Διεθνές Περιοδικό Ανώτατης Εκπαίδευσης* , 9 (3), 145-151.

Attard-Frost, B., Brandusescu , A., & Lyons, K. (2024). Η διακυβέρνηση της τεχνητής νοημοσύνης στον Καναδά: Ευρήματα και ευκαιρίες από μια ανασκόπηση 84 πρωτοβουλιών διακυβέρνησης της τεχνητής νοημοσύνης. *Government Information Quarterly* , 41 (2), 101929.

Attwell , G., Bekiaridis , G., Deitmer , L., Perini, M., Roppertz , S., & Tütlys , V. (2020). Τεχνητή νοημοσύνη σε πολιτικές, διαδικασίες και πρακτικές επαγγελματικής εκπαίδευσης και κατάρτισης.

Attwell , G., Deitmer , L., Tütlys , V., Roppertz , S., & Perini, M. (2020). Ψηφιοποίηση , τεχνητή νοημοσύνη και επαγγελματικά επαγγέλματα και δεξιότητες: ποιες είναι οι ανάγκες για την εκπαίδευση εκπαιδευτικών και εκπαιδευτών;. Στο *Trends in Επαγγελματική εκπαίδευση και έρευνα κατάρτισης, Τόμος III. Πρακτικά του Ευρωπαϊκού Συνεδρίου για την Εκπαιδευτική Έρευνα (ECER), Δίκτυο Επαγγελματικής Εκπαίδευσης και Κατάρτισης (VETNET)* (σελ. 30-42).

Μπεκιαρίδης , Γ., & Attwell , Γ. (2024). Ενσωμάτωση Τεχνητής Νοημοσύνης στην Επαγγελματική Εκπαίδευση και την Εκπαίδευση Ενηλίκων: Συμπλήρωμα στο Πλαίσιο DigCompEdu . *Ubiquity Proceedings* , 4 (1).

Boninger , F., Molnar, A., & Saldaña , C. (2020). Μεγάλοι ισχυρισμοί, λίγα στοιχεία, πολλά χρήματα: Η πραγματικότητα πίσω από το πρόγραμμα Summit Learning και η ώθηση για υιοθέτηση ψηφιακών εξατομικευμένων προγραμμάτων μάθησης. *Μονάδα Έρευνας για την Εμπορευματοποίηση στην Εκπαίδευση* .

Bulathwela , S., Pérez-Ortiz, M., Holloway, C., Cukurova , M., & Shawe -Taylor, J. (2024). Η τεχνητή νοημοσύνη από μόνη της δεν θα εκδημοκρατίσει την εκπαίδευση: Σχετικά με την εκπαιδευτική ανισότητα, την τεχνο-λύση και τα εργαλεία ένταξης. *Βιωσιμότητα* , 16 (781). <https://doi.org/10.3390/su16020781>

Celik, I. (2023). Διερεύνηση των καθοριστικών παραγόντων του γραμματισμού στην τεχνητή νοημοσύνη (AI): Ψηφιακό χάσμα, υπολογιστική σκέψη, γνωστική απορρόφηση. *Τηλεματική και Πληροφορική* , 83 , 102026.

Chan, CKY (2023). Ένα ολοκληρωμένο πλαίσιο εκπαίδευσης πολιτικής για την Τεχνητή Νοημοσύνη για την πανεπιστημιακή διδασκαλία και μάθηση. *Διεθνές Περιοδικό*

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



Εκπαιδευτικής Τεχνολογίας στην Ανώτατη Εκπαίδευση, 20 (1), 38.
<https://link.springer.com/article/10.1186/s41239-023-00408-3>

Chew, E., & Chua, XN (2020). Ρομποτικός καθηγητής κινεζικής γλώσσας: Εξατομίκευση της αξιολόγησης της προόδου και της ανατροφοδότησης ή ανάληψη της εργασίας σας; *On the Horizon*, 28(3), 113–124. <https://doi.org/10.1108/OTH-04-2020-0015>

Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). Η ικανότητα τεχνητής νοημοσύνης των πρωτοετών φοιτητών ως προγνωστικός παράγοντας για την προβλεπόμενη και την πραγματική χρήση εργαλείων τεχνητής νοημοσύνης για την υποστήριξη μαθησιακών διαδικασιών στην τριτοβάθμια εκπαίδευση. *Διεθνές Περιοδικό Εκπαιδευτικής Τεχνολογίας στην Τριτοβάθμια Εκπαίδευση*, 21 (1), 18.

Ευρωπαϊκή Επιτροπή, Γενική Διεύθυνση Εκπαίδευσης, Νεολαίας, Αθλητισμού και Πολιτισμού. (2022). *Δεοντολογικές κατευθυντήριες γραμμές σχετικά με τη χρήση τεχνητής νοημοσύνης (TN) και δεδομένων στη διδασκαλία και τη μάθηση για εκπαιδευτικούς*. Υπηρεσία Εκδόσεων της Ευρωπαϊκής Ένωσης. <https://data.europa.eu/doi/10.2766/153756>

Ευρωπαϊκό Κοινοβούλιο. (2024). Νόμος της ΕΕ για την Τεχνητή Νοημοσύνη: Πρώτος κανονισμός για την τεχνητή νοημοσύνη. Ανακτήθηκε από <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Ευρωπαϊκό Κοινοβούλιο και Συμβούλιο της Ευρωπαϊκής Ένωσης. (2024). *Κανονισμός (ΕΕ) 2024/1689 της 13ης Ιουνίου 2024 για τη θέσπιση εναρμονισμένων κανόνων για την τεχνητή νοημοσύνη και την τροποποίηση των κανονισμών (ΕΚ) αριθ. 300/2008, (ΕΕ) αριθ. 167/2013, (ΕΕ) αριθ. 168/2013, (ΕΕ) 2018/858, (ΕΕ) 2018/1139 και (ΕΕ) 2019/2144 και των οδηγιών 2014/90/ΕΕ, (ΕΕ) 2016/797 και (ΕΕ) 2020/1828 (Νόμος για την Τεχνητή Νοημοσύνη) (Κείμενο που παρουσιάζει ενδιαφέρον για τον ΕΟΧ)*. Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης, σειρά L, 1–144.

Emeršič, Ž., Hrastnik, G., Meh Peer, N., & Peer, P. (2025). Στρατηγικές εκπαίδευσης πανεπιστημιακού επιπέδου εμπνευσμένες από το AIM@VET για τη διδασκαλία της όρασης υπολογιστών και της βιομετρίας. *ROSUS 2025 – Računalniška Olimpijada slik in jena uporaba v Sloveniji 2025: Zbornik 19. strokovne συνέδριο*. <https://doi.org/10.18690/um.feri.2.2025.4>

Hodgkin, E., & Kumar, A. (2023). Αλγοριθμική δικαιοσύνη σε συστήματα επαγγελματικής καθοδήγησης: Μια συγκριτική ανάλυση πλαισίων εφαρμογής. Έρευνα και Ανάπτυξη Εκπαιδευτικής Τεχνολογίας, 71(3), 215-231.

Holmes, W., Bialik, M., & Fadel, C. (2023). Τεχνητή νοημοσύνη στην εκπαίδευση: Υποσχέσεις και επιπτώσεις στη διδασκαλία και τη μάθηση. Κέντρο Ανασχεδιασμού Προγραμμάτων Σπουδών

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



JISC (2022). *Τεχνητή Νοημοσύνη στην τριτοβάθμια εκπαίδευση. Σύνοψη της τρέχουσας κατάστασης*. Αποθετήριο JISC. Ανακτήθηκε από <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>

Kochmar , E., Vu , DD, Belfer , R., Gupta, V., Serban , IV, & Pineau, J. (2020). Η αυτοματοποιημένη εξατομικευμένη ανατροφοδότηση βελτιώνει τα μαθησιακά οφέλη σε ένα έξυπνο σύστημα διδασκαλίας. Στο *Τεχνητή Νοημοσύνη στην Εκπαίδευση: 21ο Διεθνές Συνέδριο, AIED 2020, Ιφράν , Μαρόκο, 6–10 Ιουλίου 2020, Πρακτικά, Μέρος II* 21 (σελ. 140-146). Springer International Publishing.

Kumar, V., & Boulanger, D. (Οκτώβριος 2020). Εξηγήσιμη αυτοματοποιημένη βαθμολόγηση δοκιμών: Η βαθιά μάθηση έχει πραγματικά παιδαγωγική αξία. Στο *Frontiers in education* (Τόμος 5, σελ. 572367). Frontiers Media SA.

Li, Q. (2021). Η χρήση τεχνητής νοημοσύνης σε συνδυασμό με το cloud computing στον σχεδιασμό πλατφόρμας διαχείρισης εκπαιδευτικών πληροφοριών. *International Journal of Emerging Technologies in Learning*, 16(5), 32–44. <https://doi.org/10.3991/ijet.v16i05.20309>

Porter, B., & Grippa , F. (2020). Μια πλατφόρμα για ανατροφοδότηση σε πραγματικό χρόνο με δυνατότητα τεχνητής νοημοσύνης για την προώθηση της ψηφιακής συνεργασίας. *Sustainability* , 12 (24), 10243.

Prinsloo, P., & Knox, J. (2023). Διερεύνηση των ηθικών αρχών για την εφαρμογή της τεχνητής νοημοσύνης στην εκπαίδευση: Προς μια μελλοντική ατζέντα. Έρευνα και Πρακτική στη Μάθηση με Τεχνολογία, 18(1), 73-94. Ανακτήθηκε από https://www.researchgate.net/publication/368740229_Exploring-the-Ethical-Principles-for-the-Application-of-Artificial-Intelligence-in-Education-Towards-a-Future-Agenda

Ramirez -Montoya, MS, Castillo-Martínez, IM, Sanabria-Z, J., & Miranda, J. (2023). Τεχνητή νοημοσύνη στην επαγγελματική εκπαίδευση: Ηθικές προκλήσεις και πλαίσια διακυβέρνησης. Διεθνές Περιοδικό Εκπαιδευτικής Τεχνολογίας στην Ανώτατη Εκπαίδευση, 20(1), 138-155. <https://doi.org/10.1186/s41239-023-00408-3>

Rodriguez-Abitia , G., & Bribiesca-Correa, G. (2021). Αξιολόγηση της ηθικής επίδρασης της τεχνητής νοημοσύνης στην εκπαίδευση. KOKKINOS. Revista de Educación a Distancia, 6(1). <https://doi.org/10.34627/redvol6iss1e202303>

ρωμαϊκός Etxebarrieta , G., Orcasitas-Vicandi, M., & Antzaka , A. (2024). *Σχήμα αξιολόγησης για την Τεχνητή Νοημοσύνη στην εκπαίδευση σχετικά με τα δεδομένα, την ιδιωτικότητα, την ηθική και τις αξίες της ΕΕ (WP5)* . AIPioneers.org. Ανακτήθηκε από: <https://aipioneers.org/evaluation-schema-for-ai-in-education-on-data-privacy-ethics-and-eu-values-wp5/>

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



Roppertz , S. (2020). Τεχνητή Νοημοσύνη και Επαγγελματική Εκπαίδευση και Κατάρτιση – Η Προοπτική των Γερμανών Εκπαιδευτικών Κτηνιατρικής. Στα *Πρακτικά του Συνεδρίου του Ευρωπαϊκού Δικτύου Εξ Αποστάσεως και Ηλεκτρονικής Μάθησης (EDEN)* (Τεύχος 2, σελ. 207-216). Ευρωπαϊκό Δίκτυο Εξ Αποστάσεως και Ηλεκτρονικής Μάθησης.

Rott , KJ, Lao, L., Petridou , E., & Schmidt-Hertha, B. (2022). Ανάγκες και απαιτήσεις για πρόσθετο προσόν Τεχνητής Νοημοσύνης κατά τη διάρκεια διττής επαγγελματικής κατάρτισης: Αποτελέσματα από μελέτες μαθητευόμενων και εκπαιδευτικών. *Υπολογιστές και εκπαίδευση: Τεχνητή νοημοσύνη* , 3 , 100102.

Şenocak , D., Bozkurt, A., & Koçdar , S. (2024). Διερεύνηση των ηθικών αρχών για την εφαρμογή της τεχνητής νοημοσύνης στην εκπαίδευση: Προς μια μελλοντική ατζέντα. Στο *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (σελ. 200-213). IGI Global.

Seufert, S. (2024). Τεχνητή Νοημοσύνη στην Επαγγελματική Εκπαίδευση και Κατάρτιση (EEK): Αξιολόγηση της αποδοχής της Τεχνητής Νοημοσύνης από τους ηγέτες της EEK στην Ελβετία.

Tang, J., & Hai, L. (2021). Κατασκευή και διερεύνηση ενός έξυπνου συστήματος αξιολόγησης για εκπαιδευτικές εφαρμογές μέσω τεχνολογίας τεχνητής νοημοσύνης. *Διεθνές Περιοδικό Αναδυόμενων Τεχνολογιών στη Μάθηση*, 16(5), 17–31. <https://doi.org/10.3991/ijet.v16i05.20293>

Tuomi , I., Cachia , R., & Villar- Onrubia , D. (2023). Σχετικά με το μέλλον της τεχνολογίας στην εκπαίδευση: Αναδυόμενες τάσεις και πολιτικές επιπτώσεις. *Υπηρεσία Εκδόσεων της Ευρωπαϊκής Ένωσης, Λουξεμβούργο* .

UNESCO (2023a). *ChatGPT και τεχνητή νοημοσύνη στην τριτοβάθμια εκπαίδευση*. Ανακτήθηκε από: https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

UNESCO (2023b). *Οδηγίες για την παραγωγική τεχνητή νοημοσύνη στην εκπαίδευση και την έρευνα* . Ανακτήθηκε από: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

Vázquez- Cano , E., Mengual -Andrés, S., & López-Meneses, E. (2021). Chatbot για τη βελτίωση της στίξης στην ισπανική γλώσσα και την ενίσχυση ανοιχτών και ευέλικτων μαθησιακών περιβαλλόντων. *Διεθνές Περιοδικό Εκπαιδευτικής Τεχνολογίας στην Ανώτατη Εκπαίδευση* , 18 , 1-20.

Villegas-Ch, W., García-Ortiz, J., Mullo -Ca, K., Sánchez- Viteri , S., & Roman- Cañizares , M. (2021). Υλοποίηση ενός εικονικού βοηθού για την ακαδημαϊκή διαχείριση ενός πανεπιστημίου με τη χρήση τεχνητής νοημοσύνης. *Future Internet* , 13 (4), 97.

Williamson, B., & Eynon , R. (2020). Ιστορικά θέματα, ελλείποντες κρίκοι και μελλοντικές κατευθύνσεις στην Τεχνητή Νοημοσύνη στην εκπαίδευση. *Μάθηση, Μέσα και Τεχνολογία* , 45 (3), 223-235.

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι, ωστόσο, μόνο του/των συγγραφέα/ων και δεν αντικατοπτρίζουν απαραίτητα εκείνες της Ευρωπαϊκής Ένωσης ή του Εκτελεστικού Οργανισμού Ευρωπαϊκής Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA φέρουν ευθύνη γι' αυτές.



Williamson, B., Macgilchrist, F., & Potter, J. (2023). Επανεξέταση της Τεχνητής Νοημοσύνης, του αυτοματισμού και της επεξεργασίας δεδομένων στην εκπαίδευση. *Μάθηση, Μέσα και Τεχνολογία*, 48 (1), 1–5. <https://doi.org/10.1080/17439884.2023.2167830>

Wongvorachan, T., Lai, KW, Bulut, O., Tsai, YS, & Chen, G. (2022). Τεχνητή νοημοσύνη: Μετασχηματίζοντας το μέλλον της ανατροφοδότησης στην εκπαίδευση. *Journal of Applied Testing Technology*, 95-116.

Zawacki -Richter, O., Marín, VI, Bond, M., & Gouverneur, F. (2023). Συστηματική ανασκόπηση της έρευνας σχετικά με τις εφαρμογές τεχνητής νοημοσύνης στην τριτοβάθμια εκπαίδευση – πού είναι οι εκπαιδευτικοί; Διεθνές Περιοδικό Εκπαιδευτικής Τεχνολογίας στην Τριτοβάθμια Εκπαίδευση, 20(1). <https://doi.org/10.1186/s41239-023-00408-3>

ΕΠΙΚΟΙΝΩΝΙΑ

www.aipioneers.org

gorka.roman@ehu.es

maria.orcasitas@ehu.es

natalia.louleli@ehu.eus



AI PIONEERS

Artificial Intelligence in education & training

Manual de política y ética en la enseñanza de la inteligencia artificial

Junio de 2025

Beneficiario Principal del entregable: Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

Paquete de trabajo 5

Entregable D5.2 "Manual sobre política y ética en la educación en IA"

Número de proyecto: 101087261

Nombre del proyecto: La IA y el futuro de la

Educación Acrónimo del proyecto: AI

Pioneers

Coordinador del proyecto: Universidad de Bremen, Instituto de Tecnología y Educación (ITB) Convocatoria: ERASMUS-EDU-2022-PI-FORWARD

Tema: ERASMUS-EDU-2022-PI-FORWARD-LOT1

Tipo de acción: Subvenciones ERASMUS a tanto alzado

Autoridad que concede la subvención: Agencia Ejecutiva en el ámbito Educativo y Cultural Fecha de inicio del proyecto:

01/01/2023

Fecha de finalización del

proyecto: 31/12/2025

Duración del proyecto: 36

meses



Co-funded by
the European Union



AI PIONEERS
Artificial intelligence in education & training



AI PIONEERS
Artificial Intelligence in education & training

Versión inglesa del documento

Junio de 2025



Manual de política y ética en la enseñanza de la inteligencia artificial

Gorka Roman Etxebarrieta (Autor)

Maria Orcasitas-Vicandi (Autor)

Natalia Louleli (Autor)

Asier León Nuñez (Autor)

Itsaso Biota Piñeiro (Autor)

Ana Boyano Murillo (Autor)

Raquel Justo Blanco (Autor)

Josune Rodriguez Negro (Autor)



Índice

Manual de política y ética en la enseñanza de la inteligencia artificial	3
1. Introducción.....	5
1.1 Objeto y ámbito de aplicación del manual	6
1.2 La importancia de las consideraciones éticas en la adopción de la IA	6
1.3 Manual para educadores, administradores y responsables políticos.....	8
2. Comprender la IA en la EFP y la educación de adultos.....	10
2.1. El papel de la IA en la educación	11
2.2. La IA en la educación y formación profesional y de adultos	12
3. Desarrollo de directrices éticas	13
3.1 Importancia de las políticas éticas en la enseñanza de la IA.....	13
3.2 Requisitos clave de los sistemas de IA fiables	16
3.3 Consideración de los aspectos jurídicos de la aplicación de la IA en los centros de FP	18
3.4 Estudios de casos y ejemplos de políticas eficaces de IA en centros de FP ..	25
4. Habilidades éticas	34
4.1 Definición de competencias éticas	34
4.2 Importancia de las competencias éticas.....	34
4.3 Ámbitos y ejemplos de competencias éticas	34
5. Conclusión.....	43
5.1 Principales conclusiones.....	43
5.2 Reflexiones finales.....	45
Agradecimiento	45
Referencias.....	46



1. Introducción

El proyecto AI Pioneers se centra en avanzar en la integración de la Inteligencia Artificial (IA) en la educación, especialmente en la educación de adultos y la educación y formación profesional (EFP). Este proyecto ERASMUS+ Forward Looking comenzó en enero de 2023 y está financiado hasta diciembre de 2025.

Con la creciente influencia de la IA en la educación y el aumento de la demanda de apoyo profesional en el uso de la IA y el desarrollo de habilidades, el proyecto tiene como objetivo mantener una amplia red mundial de pioneros de la IA en todos los países socios y más allá. El establecimiento de la Red de Referencia de Pioneros de la IA es un elemento clave, dirigido a educadores, responsables políticos, planificadores educativos, desarrolladores de IA y otras partes interesadas. Al implicar a los más interesados en el uso de la IA en la educación, el proyecto pretende crear un eje central que ayude a la integración de la IA en la educación de adultos y la EFP, fomentando futuras iniciativas y proyectos educativos de IA. Más información sobre la red y cómo participar en ella en el [sitio web del proyecto](#).

Otro objetivo importante era desarrollar un Suplemento al Marco DigCompEDU (Bekiaridis, 2024). Este documento describe las habilidades y competencias necesarias que los educadores necesitan para comprender, evaluar y utilizar las tecnologías de la IA en la educación, con un enfoque específico en la educación de adultos y la EFP. Este suplemento alinea las competencias de IA con las seis áreas centrales del Marco DigCompEdu (Redecker, 2017): Compromiso profesional, Recursos digitales, Enseñanza y aprendizaje, Evaluación, Empoderamiento de los alumnos y Facilitación de la competencia digital del alumno. Para cada área, el suplemento:

1. Analiza las aplicaciones de la IA en el sector
2. Propone actividades para desarrollar las competencias pertinentes
3. Sugiere niveles de progresión para el desarrollo de competencias
4. Identifica los retos y proporciona estrategias para superarlos.

El proyecto sigue desarrollando y compartiendo una serie de recursos a través del AI Pioneers Toolkit, disponible en el sitio web del proyecto. Este conjunto de herramientas ofrece recomendaciones curriculares para la integración de la IA en la educación, destaca los beneficios y los riesgos del uso de la IA en entornos educativos y proporciona directrices de buenas prácticas para la IA en la educación (Tommasi & Perini, 2024). También incluye breves descripciones de herramientas para diferentes fines pedagógicos, como asistentes de



aprendizaje, herramientas para crear materiales educativos y herramientas de apoyo al aprendizaje entre iguales.

El proyecto está realizando una importante labor a la hora de abordar los retos éticos que plantea la IA en los entornos educativos. Este documento acompaña a un esquema de evaluación y a . Ambos están diseñados para orientar a instituciones y educadores en la aplicación responsable de la IA, teniendo en cuenta al mismo tiempo la privacidad y las implicaciones éticas.

El Esquema de evaluación de la IA en la educación se centra en directrices prácticas para la adopción de herramientas de IA, garantizando que se ajustan a los principios éticos y los valores de la UE. Es un gran recurso para que las instituciones educativas tomen decisiones informadas sobre la integración de la IA.

Y éste, el *Handbook on Policy and Ethics in AI Education* es una guía más completa, que ofrece políticas y marcos éticos para abordar las preocupaciones actuales.

1.1 Objeto y ámbito de aplicación del manual

El Handbook on Policy and Ethics in AI Education (Manual de política y ética en la enseñanza de la IA) es un recurso fundamental para educadores, responsables políticos e instituciones que deseen abordar las complejidades de la IA en entornos educativos. Ayuda a crear un marco que garantice que las consideraciones éticas ocupan un lugar central, protegiendo la privacidad de los estudiantes, promoviendo la equidad y fomentando un uso responsable.

Una ventaja fundamental es la orientación práctica que ofrece a los educadores, que les permite utilizar la IA de forma responsable y ayudar a los estudiantes a comprender sus capacidades, limitaciones e implicaciones éticas. Esto será fundamental para fomentar una generación de estudiantes que no sólo dominen la tecnología, sino que también sean éticamente conscientes de su uso. Para los responsables políticos, el marco coherente que ofrece el manual orienta a los usuarios hacia enfoques más normalizados en todas las instituciones, lo que es esencial para mitigar los riesgos de la IA, como los sesgos en los algoritmos o las violaciones de la privacidad. También puede ayudar a garantizar que la IA no se integre por sí misma, sino de forma que mejore la experiencia educativa sin comprometer las normas éticas.

1.2 La importancia de las consideraciones éticas en la adopción de la IA



Las consideraciones éticas son cruciales en la adopción de la IA, especialmente en campos como la educación de adultos y la formación profesional (EFP), donde el impacto de la IA en alumnos, educadores e instituciones puede ser significativo. Estas consideraciones ayudan a garantizar que la IA se utilice de forma responsable, inclusiva y transparente, y que apoye los objetivos más amplios de la educación sin comprometer los valores fundamentales.

Algunas de las consideraciones clave que cubre el esquema son;

Diversidad e inclusión: Los sistemas y herramientas de IA deben diseñarse para atender a una amplia gama de alumnos, incluidos los que tienen diferentes antecedentes, capacidades y necesidades de aprendizaje. En la educación de adultos y la EFP, los alumnos suelen tener experiencias vitales diversas, como distintos niveles de educación previa, grupos de edad y experiencia laboral. La IA debe ser flexible e integradora, garantizando que su uso apoye y no excluya a ningún grupo en particular. Quienes adopten éticamente la IA deben ser conscientes de las formas en que la tecnología podría perpetuar los prejuicios o crear brechas digitales para mitigarlos.

Transparencia: Para que la IA se adopte éticamente en los entornos educativos, la transparencia es esencial. Las instituciones deben explicar claramente cómo funcionan los sistemas de IA, los datos que recogen y los procesos de toma de decisiones que los sustentan. Los alumnos y los educadores deben entender cómo se utilizan las herramientas de IA en el aula o en el entorno de formación, y cómo los sistemas pueden afectar a los resultados del aprendizaje. Esta transparencia fomenta la confianza en las tecnologías de IA y garantiza la responsabilidad en su uso.

Privacidad y protección de datos: La educación de adultos y las instituciones de EFP a menudo manejan datos personales sensibles, como perfiles de alumnos, resultados de evaluaciones e historial profesional. El uso ético de la IA en este contexto debe dar prioridad a la privacidad y la protección de datos. Las instituciones deben garantizar el cumplimiento de las leyes de protección de datos (por ejemplo, GDPR) y garantizar que los datos de los alumnos se recopilen, almacenen y utilicen de manera responsable. El consentimiento para la recopilación de datos debe ser informado y voluntario, y los alumnos deben tener control sobre su información personal.

Sostenibilidad: La IA debe integrarse de forma que se ajuste a la sostenibilidad a largo plazo de los sistemas educativos. Esto significa considerar tanto el impacto medioambiental de las tecnologías de IA como su papel en el fomento



de prácticas sostenibles dentro de los programas de educación y formación. En el contexto de la EFP, la IA también puede apoyar la sostenibilidad en las industrias preparando a los alumnos con competencias que promuevan las tecnologías verdes y las prácticas ecológicas.

Alfabetización en IA: A medida que la Inteligencia Artificial se integra más en la educación, es fundamental que tanto los educadores como los alumnos adquieran los conocimientos necesarios. En la educación de adultos y la EFP, esto implica dotar a los alumnos de los conocimientos y habilidades necesarios para comprender las tecnologías de la IA y utilizarlas eficazmente en sus carreras. Esta capacitación permite a los alumnos desenvolverse en un mercado laboral en constante evolución y garantizar que no se queden atrás. Para los educadores, la alfabetización en IA les ayuda a evaluar críticamente las herramientas de IA e incorporarlas de forma que mejoren el aprendizaje sin comprometer las normas éticas.

Justicia y equidad: Los sistemas de IA deben estar diseñados para tratar a todos los alumnos de forma justa, garantizando que ningún grupo se vea desfavorecido por decisiones o algoritmos automatizados. En la educación de adultos y la EFP, esto es especialmente importante, ya que estos campos a menudo atienden a alumnos que ya pueden enfrentarse a retos sociales o económicos. Es fundamental que la IA no refuerce las desigualdades existentes ni cree nuevas barreras de acceso, sino que apoye las oportunidades equitativas para todos los alumnos.

1.3 Manual para educadores, administradores y responsables políticos

Este manual está destinado a educadores, administradores y responsables políticos en el contexto de la adopción de la IA en la educación de adultos y la formación profesional (EFP). Se sabe que la adopción de la IA y sus consideraciones éticas afectan de manera diferente a cada uno de los grupos mencionados:

Educadores:

Para los educadores, la integración ética de la IA en la educación de adultos y la EFP consiste en mejorar la enseñanza y el aprendizaje sin socavar los principios de equidad, privacidad y transparencia. Las herramientas de IA pueden ofrecer experiencias de aprendizaje personalizadas, agilizar las tareas administrativas y proporcionar información sobre el progreso de los estudiantes. Sin embargo, los



educadores deben asegurarse de que la IA apoye los diversos estilos y necesidades de aprendizaje, fomentando un entorno de aprendizaje inclusivo.

Hay muchas consideraciones clave para los educadores en relación con la alfabetización en IA, el uso ético y la transparencia con los alumnos.

Alfabetización en IA: Los educadores deben estar equipados con las habilidades necesarias para comprender, evaluar e integrar eficazmente la IA en sus métodos de enseñanza. Esto incluye el conocimiento del funcionamiento de las herramientas de IA y sus implicaciones éticas en el aula.

Uso ético: Los educadores deben asegurarse de que la IA se utiliza para mejorar el aprendizaje, no para sustituir la interacción humana. Deben estar preparados para abordar cualquier problema de parcialidad o discriminación que pueda surgir de los sistemas de IA.

Transparencia con los alumnos: Los educadores deben ser transparentes con los alumnos sobre cómo se utiliza la IA en el proceso de aprendizaje y los datos que recopila. Esto ayuda a generar confianza y garantiza que los estudiantes sientan que se respeta su privacidad.

Administradores:

Para los administradores de centros de educación de adultos y de EFP, la adopción de la IA implica supervisar la integración de estas tecnologías en los sistemas de la institución y garantizar que se dispone de la infraestructura necesaria para soportar las herramientas de IA. También deben dar prioridad a las políticas que garanticen el uso ético de la IA, manteniendo al mismo tiempo la integridad institucional.

Los administradores deben tener en cuenta muchos aspectos clave, como la seguridad y la privacidad de los datos, la equidad y el acceso, y la formación y el apoyo continuos.

Seguridad y privacidad de los datos: Los administradores deben garantizar la aplicación de medidas sólidas de protección de datos, ya que la IA se basa en la recopilación de grandes cantidades de datos de los alumnos. El cumplimiento de la normativa de protección de datos (como el GDPR) es esencial.

Equidad y acceso: La IA debe desplegarse de forma que garantice un acceso equitativo a todos los alumnos, independientemente de su origen



socioeconómico. Los administradores deben abogar por políticas que impidan que las herramientas de IA exacerben las desigualdades existentes.

Formación y apoyo continuos: Los administradores deben asignar recursos al desarrollo profesional continuo para que los educadores puedan mantenerse al día sobre los avances de la IA y las mejores prácticas éticas.

Responsables políticos:

Los responsables políticos desempeñan un papel crucial en la creación de los marcos y normativas que rigen el uso de la IA en la educación. Deben velar por que la adopción de la IA se ajuste a objetivos educativos más amplios, respete los derechos humanos y mejore las oportunidades de aprendizaje permanente.

Hay una serie de consideraciones clave para los responsables políticos en relación con las directrices éticas y la regulación, la promoción de la alfabetización en IA y la garantía de la inclusión y la accesibilidad.

Directrices éticas y reglamentación: Los responsables políticos deben establecer normas y reglamentos éticos que guíen el uso de la IA en la educación. Esto incluye abordar cuestiones como la transparencia, la privacidad de los datos y la equidad en los sistemas educativos impulsados por la IA.

Promover la alfabetización en IA: Los responsables políticos deben promover la alfabetización en IA en todo el sistema educativo, garantizando que tanto los educadores como los alumnos tengan las herramientas necesarias y la comprensión para navegar por las tecnologías de IA.

Garantizar la inclusión y la accesibilidad: Es vital que las políticas de IA den prioridad a la educación inclusiva, asegurándose de que las tecnologías de IA no dejen atrás a ciertos grupos. Esto puede implicar incentivar la investigación y el desarrollo de soluciones de IA adaptadas a alumnos diversos.

2. Comprender la IA en la EFP y la educación de adultos

La integración de la Inteligencia Artificial (IA) en la educación presenta oportunidades de transformación junto con importantes retos. La investigación subraya el potencial de la IA para mejorar la enseñanza mediante tareas automatizadas, aprendizaje personalizado e innovaciones pedagógicas basadas en datos, mientras que las herramientas de IA generativa amplían las posibilidades de una educación interactiva y reflexiva. Sin embargo, las preocupaciones éticas en torno a la privacidad de los datos, la influencia



comercial y la brecha digital exigen una gobernanza sólida. En la educación profesional y de adultos, la IA no sólo informa el diseño curricular y la predicción de habilidades para la preparación de la mano de obra, sino que también necesita fomentar la alfabetización digital y el compromiso crítico con las implicaciones sociales de la IA. Equilibrar los beneficios de la IA con políticas equitativas y éticas sigue siendo crucial para configurar ecosistemas educativos inclusivos y orientados al futuro.

2.1. El papel de la IA en la educación

Un informe político recientemente publicado por el Centro Común de Investigación (CCI) de la Comisión Europea, titulado [On the Futures of Technology in Education: Tendencias emergentes e implicaciones políticas](#), (Tuomi, Cachia & Villar-Onrubia, 2023), ofrece una visión global de cómo los avances tecnológicos, especialmente en Inteligencia Artificial (IA), están transformando las prácticas, los procesos y las instituciones educativas. El documento destaca las importantes y diversas repercusiones de la IA en la educación, que afectan a los procesos de enseñanza y aprendizaje, los métodos de evaluación y las estructuras organizativas.

La IA se considera una solución potencial para reducir la carga de trabajo de los profesores, especialmente en tareas como corregir los deberes de los alumnos. Al automatizar estas tareas repetitivas y lentas, la IA puede hacer más atractiva la enseñanza y permitir que los profesores se concentren más en las actividades docentes que en las tareas administrativas (Li, 2021; Tang & Hai, 2021; Villegas-Ch et al., 2021). Esta automatización podría revolucionar la evaluación formativa (Kumar & Boulanger, 2020), permitiendo una retroalimentación continua para los estudiantes y facilitando la evaluación en tiempo real (Porter & Grippa, 2020) y el aprendizaje personalizado o adaptativo (Kochmar et al., 2020; Wongvorachan et al., 2022).

Además, se espera que la IA generativa y los modelos fundacionales tengan un impacto sustancial en la educación. Tecnologías como el subtítulo automático de vídeos, la traducción y el resumen de vídeos ofrecen oportunidades innovadoras para mejorar la pedagogía (Chew y Chua, 2020; Vázquez-Cano et al., 2021). Estas tecnologías pueden ayudar a desarrollar "compañeros inteligentes" o "compañeros de aprendizaje", desplazando la atención del aprendizaje memorístico a la metacognición y la reflexión. Esto requiere una cuidadosa redistribución de responsabilidades entre humanos y máquinas, determinando qué actividades delegar en la IA.



También debemos destacar el papel de la IA en la "dataficación" de la educación, donde los datos a gran escala sobre las actividades educativas se utilizan para reforzar y fundamentar nuevos enfoques pedagógicos, tecnologías de aprendizaje y políticas educativas (Williamson et al., 2023). El objetivo es registrar los procesos de aprendizaje, no solo los resultados, para mejorar el compromiso y el rendimiento de los estudiantes (Porter y Grippa, 2020).

Sin embargo, las preocupaciones éticas sobre el tratamiento de datos personales, el posible uso indebido, la dependencia de infraestructuras controladas en gran medida por entidades comerciales y algunos otros retos son cuestiones fundamentales que hay que abordar (Seufert, 2024; Wongvorachan et al., 2022). La integración de la IA en la educación plantea cuestiones importantes sobre los modelos de gobernanza y la formulación de políticas, lo que exige una reevaluación de las políticas educativas para incorporar las capacidades y limitaciones de la IA (Attard-Frost et al., 2024; Boninger et al., 2020; Williamson & Eynon, 2020). También debemos abordar la importancia de evaluar críticamente los riesgos de exacerbar la brecha digital y el dominio de los intereses comerciales en el sector (Bulathwela et al., 2024; Celik, 2023). Regular la innovación tecnológica para garantizar visiones educativas diversas y alinear la política industrial con las teorías del aprendizaje son aspectos cruciales que debemos considerar.

El papel de la IA en la educación es una herramienta transformadora y polifacética que ofrece oportunidades para mejorar la eficiencia, la personalización y la innovación en la enseñanza y el aprendizaje. Sin embargo, también plantea importantes retos y consideraciones éticas que deben abordarse para garantizar un uso equitativo, eficaz y responsable de la IA en los entornos educativos.

2.2. La IA en la educación y formación profesional y de adultos

Integrar la enseñanza de la IA en los planes de estudios implica no solo enseñar los fundamentos de la IA y sus aplicaciones, sino también explorar sus repercusiones sociales. Esto incluye fomentar el pensamiento crítico sobre las capacidades, limitaciones y dimensiones éticas de la IA, como las preocupaciones en torno a la parcialidad y la privacidad (Akgun & Greenhow, 2022; Aldosari, 2020; Attwell et al., 2020; Delcker, et al., 2024; Seufert, 2024; Wongvorachan et al., 2022). Mejorar la alfabetización digital también es esencial, ya que comprender cómo interactuar con los sistemas de IA e interpretarlos se está convirtiendo en una habilidad clave en la mano de obra moderna, incluidos los profesores (Bekiaridis y Attwell, 2024; Delcker, et al., 2024; Roppertz, 2020; Rott et al., 2022; Seufert, 2024). A medida que la IA sigue evolucionando, la



atención se está desplazando no solo hacia el uso de herramientas de IA, sino también hacia su comprensión y evaluación crítica, preparando a los alumnos para un futuro en el que la IA forme parte integral de la vida cotidiana y el trabajo (Rott et al., 2022).

En consecuencia, preparar a la mano de obra para adquirir nuevas competencias en la era de la IA es otro aspecto crucial de la educación de adultos y la EFP (Attwell et al., 2020; Roppertz, 2020; Rott et al., 2022). El papel de la IA en la identificación de tendencias emergentes y la predicción de futuras necesidades de cualificación es inestimable. Esta información permite a los programas educativos adaptar sus planes de estudios para satisfacer las necesidades cambiantes del mercado laboral (Seufert, 2024). La mejora y el reciclaje de las competencias se convierten en temas centrales, y los programas guiados por la IA permiten a adultos y profesionales adquirir nuevas competencias o actualizar las existentes, manteniéndose así al día de los avances tecnológicos (Attwell et al., 2020; Emeršič et al., 2024; Delcker, et al., 2024). Además, la formación orientada a la IA garantiza que los alumnos no sólo sean técnicamente expertos, sino que también estén preparados para trabajar en entornos integrados de IA. Esta preparación es clave para garantizar que la mano de obra esté preparada para los retos y oportunidades que presenta el rápido avance del campo de la IA, haciendo que la educación de adultos y la EFP sean más relevantes, eficaces y alineadas con el futuro panorama laboral.

3. Desarrollo de directrices éticas

3.1 Importancia de las políticas éticas en la enseñanza de la IA

Las políticas éticas en la educación sobre IA son esenciales para garantizar que las tecnologías de IA se utilicen de manera que promuevan la equidad, la transparencia y la responsabilidad, protegiendo al mismo tiempo los derechos y el bienestar de los alumnos y educadores. A medida que la IA se integra cada vez más en los entornos educativos, contar con directrices éticas claras es crucial para gestionar sus riesgos potenciales y aprovechar sus beneficios. A continuación se exponen las principales razones por las que las políticas éticas son vitales en la educación sobre IA:

Protección de los derechos y la intimidad de los alumnos

Los sistemas de IA en la educación a menudo requieren la recopilación y el tratamiento de datos personales, incluidos los expedientes académicos, los



comportamientos de aprendizaje y otra información sensible. Las políticas éticas garantizan que los datos se traten con respeto y de conformidad con la normativa sobre privacidad (por ejemplo, el GDPR). Estas políticas ayudan a proteger la privacidad de los alumnos garantizando la transparencia en la recopilación de datos, dando a los alumnos el control sobre sus datos y garantizando que se utilicen de forma responsable y segura.

Garantizar la justicia y la equidad

Uno de los problemas éticos más importantes de la IA en la educación es la posibilidad de que los algoritmos sesgados perjudiquen a determinados grupos de alumnos. Los sistemas de IA pueden perpetuar inadvertidamente las desigualdades existentes por motivos de raza, género, estatus socioeconómico u otros factores. Las políticas éticas son necesarias para guiar el desarrollo y el despliegue de herramientas de IA que sean inclusivas y equitativas, garantizando que todos los alumnos tengan el mismo acceso a las oportunidades, independientemente de sus antecedentes o características personales.

Fomento de la transparencia y la rendición de cuentas

Las tecnologías de IA pueden ser complejas y opacas, lo que dificulta que educadores, alumnos y administradores comprendan plenamente cómo se toman las decisiones. Las políticas éticas ayudan a garantizar que los sistemas de IA sean transparentes, es decir, que las partes interesadas sepan cómo funcionan las herramientas de IA, qué datos se utilizan y cómo se toman las decisiones. Esta transparencia genera confianza y responsabilidad, lo que permite a educadores y alumnos comprender mejor las tecnologías de IA e interactuar con ellas.

Fomentar la alfabetización y el uso responsable de la IA

A medida que la IA se integra más en el proceso de aprendizaje, es importante que tanto los educadores como los alumnos comprendan cómo funciona y sus posibles implicaciones. Las políticas éticas en la educación sobre IA deben fomentar la alfabetización en IA, garantizando que tanto educadores como alumnos estén equipados con los conocimientos y habilidades para utilizar la IA de forma responsable y crítica. Esta comprensión permite a los usuarios maximizar los beneficios de la IA siendo conscientes de sus limitaciones y daños potenciales.

Fomentar el desarrollo ético y la innovación



Las tecnologías de IA en la educación deben desarrollarse con una sólida base ética para garantizar que se utilicen en beneficio de todos los alumnos. Las políticas éticas proporcionan un marco para que los desarrolladores, investigadores e instituciones educativas diseñen sistemas de IA que se ajusten a valores educativos como la equidad, la capacitación y el respeto a la diversidad. Esto también garantiza que las tecnologías de IA se diseñen y utilicen de forma que apoyen el aprendizaje permanente y los objetivos más amplios de la educación.

Hacer frente al riesgo de desplazamiento

A medida que los sistemas de IA automatizan determinados aspectos de la educación (por ejemplo, la calificación, la tutoría o las tareas administrativas), existe el riesgo de que desplacen a los educadores o limiten la interacción humana. Las políticas éticas pueden ayudar a orientar el uso responsable de la IA garantizando que la automatización complemente a los educadores humanos, en lugar de sustituirlos. Las políticas pueden fomentar el uso de la IA para mejorar la eficacia docente y la eficiencia administrativa, manteniendo al mismo tiempo el valor de la presencia humana en la experiencia educativa.

Crear confianza en los sistemas de IA

Para que la IA se adopte con éxito en entornos educativos, educadores, estudiantes y administradores deben confiar en la tecnología. Las políticas éticas desempeñan un papel importante a la hora de establecer y mantener esta confianza, ya que proporcionan directrices claras sobre cómo debe utilizarse la IA y qué normas éticas deben seguirse. Cuando los sistemas de IA se desarrollan y despliegan en consonancia con estas políticas, es más probable que las partes interesadas los adopten y utilicen con eficacia.

Orientar el futuro de la IA en la educación

A medida que la IA sigue evolucionando, es crucial que las políticas éticas proporcionen una orientación a largo plazo para su integración en los sistemas educativos. Estas políticas ayudan a crear un marco para el futuro desarrollo de las tecnologías de IA que esté en consonancia con los valores de la sociedad, garantizando que la IA se utilice para mejorar los resultados del aprendizaje, fomentar la innovación y apoyar el aprendizaje permanente de maneras que sean beneficiosas y éticas.

En general, las políticas éticas en la educación en IA son necesarias para guiar el desarrollo y el uso responsables de las tecnologías de IA. Estas políticas



garantizan que la IA se utilice de forma que respete la privacidad del alumno, promueva la equidad, fomente la transparencia y genere confianza. Además, ayudan a educadores, alumnos e instituciones a afrontar los retos éticos que surgen a medida que la IA sigue transformando los entornos educativos. Al establecer estas directrices éticas, los sistemas educativos pueden aprovechar el poder de la IA al tiempo que salvaguardan los valores que son fundamentales para una educación eficaz e inclusiva.

3.2 Requisitos clave para sistemas de IA fiables

La integración de la Inteligencia Artificial (IA) en la educación ofrece inmensas oportunidades para mejorar la enseñanza, el aprendizaje y la administración. Sin embargo, para garantizar que la IA se utilice de forma ética y responsable, las instituciones educativas deben desarrollar directrices claras para la implementación ética de los sistemas de IA. Numerosos marcos, directrices y auditorías publicados (por ejemplo, AI Audit 2023a de la UNESCO y Guidance for Generative AI in Education and Research 2023b) proporcionan información valiosa sobre lo que constituye una IA digna de confianza. Estos marcos hacen hincapié en principios clave como la **transparencia, la equidad, la responsabilidad y el respeto de los derechos humanos**. Sin embargo, la tarea de desarrollar y aplicar estos principios puede resultar desalentadora para las instituciones, especialmente en un panorama tecnológico en rápida evolución.

Como parte de AI Pioneers, hemos desarrollado un esquema de evaluación (Roman Etxebarrieta et al., 2024) diseñado para simplificar y contextualizar estos principios para entornos educativos. Nuestro esquema proporciona un enfoque estructurado para ayudar a las instituciones a navegar por las complejidades de la implementación de la IA, garantizando al mismo tiempo el cumplimiento de las normas éticas. Mientras haya varias directrices disponibles, nuestro esquema pretende ofrecer una herramienta práctica y fácil de usar que guiará a las instituciones educativas en el desarrollo de sus propias directrices éticas de IA, teniendo en cuenta sus contextos y necesidades únicos.

Una de las ideas centrales de la Ley de IA de la UE (Parlamento Europeo y Consejo de la Unión Europea, 2024) es que la responsabilidad principal de garantizar que los sistemas de IA se desarrollen de forma ética recae en los desarrolladores y empresas que los crean, y que las directrices a seguir dependerán del nivel de riesgo asociado a los sistemas de IA. Sin embargo, los centros educativos son responsables de garantizar que estos sistemas se implanten y utilicen de forma ética dentro de sus entornos educativos. Los centros deben ser conscientes de las cuestiones éticas clave para comprender



tanto las limitaciones como el potencial de los sistemas de IA que adopten. Esto implica no sólo cumplir con las políticas y los requisitos legales existentes, sino también anticipar y abordar los retos que puedan surgir a medida que evolucionen estas tecnologías.

En concreto, nuestro esquema de evaluación se centra en la identificación de las preocupaciones éticas y en el desarrollo de una política o estrategia que pueda apoyar y promover el uso de la IA en los centros educativos respetando las directrices éticas clave. Estas directrices incluyen: **diversidad y equidad, transparencia y responsabilidad, privacidad y protección de datos, seguridad y protección, sostenibilidad y bienestar social, empoderamiento de profesores y alumnos, participación democrática en la planificación de políticas educativas y prácticas de IA, autonomía, diseño ético (relacionado con la sostenibilidad y la accesibilidad), comercialización** (Şenocak et al., 2024), **agencia de profesores y alumnos, y alfabetización en IA.**

El esquema de evaluación consta de varias secciones, basadas en las dimensiones de Chan (2023), ajustadas con la revisión de Şenocak et al. (2024) y las directrices éticas de la Comisión Europea (2022), cada una centrada en diferentes aspectos de la implementación de la IA en entornos educativos. Estas secciones guían a las instituciones en el desarrollo de marcos éticos y garantizan que la IA se utilice de forma responsable. La primera sección se centra en la **Evaluación del uso actual de la IA y las preocupaciones éticas**, y ayuda a las instituciones a evaluar su madurez actual en materia de IA, lo que les permite estimar el grado de preparación del centro educativo para implantar la IA o ampliar su uso (JISC, 2022).

La segunda sección se centra en la **gobernanza y la supervisión de la IA** y, por tanto, aborda cuestiones relacionadas con el cumplimiento de las políticas regionales e internacionales, que abarcan la privacidad, la protección de datos y la equidad. Aborda cuestiones éticas complejas como la no discriminación y la transparencia, que también son consideraciones clave para una IA ética. La tercera sección, **Consideraciones operativas**, ofrece orientación sobre la formación de profesores, alumnos y personal, la promoción de la alfabetización en IA, la garantía de la supervisión humana y el apoyo a la participación democrática en la toma de decisiones relacionadas con la IA. Garantizar que todos los usuarios estén adecuadamente formados y sean conscientes de las capacidades y limitaciones de la IA es clave para el éxito de la implantación de la IA y garantiza la igualdad de acceso para todos.

La cuarta sección se centra en **las consideraciones pedagógicas**, teniendo en cuenta cuestiones importantes como los dilemas éticos en las evaluaciones



basadas en la IA, la agencia del estudiante y el desarrollo de competencias, especialmente en la educación y formación profesional y de adultos (Attwell et al., 2021; UNESCO, 2019). La quinta y última sección se centra en **Otras consideraciones: Sostenibilidad y diseño ético, que** ofrece orientación para abordar los impactos a largo plazo y garantizar un uso accesible de la IA.

El esquema de evaluación proporciona una herramienta práctica para que las instituciones educativas se orienten en la implantación de la IA, identifiquen los retos éticos y desarrollen políticas acordes con las normas internacionales. Para obtener información más detallada, consulte el esquema de evaluación completo [Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)] y la presentación de Genially que se pueden encontrar [en el sitio web AI Pioneers](#).

3.3 Consideración de los aspectos jurídicos de la aplicación de la IA en los centros de FP

3.3.1 Marco normativo de la Ley de Inteligencia Artificial de la UE

La Ley de IA de la Unión Europea establece un marco normativo exhaustivo que clasifica los sistemas de inteligencia artificial en función de sus riesgos potenciales, proporcionando una orientación esencial para los centros de educación y formación profesional (EFP) que implementan estas tecnologías. Este enfoque basado en el riesgo garantiza una supervisión reglamentaria proporcionada al tiempo que fomenta la innovación en contextos educativos (Parlamento Europeo, 2024).

Como Zawacki-Richter et al. (2023) articulan en su revisión sistemática de las consideraciones éticas en la educación mejorada por IA, los marcos reguladores deben equilibrar la innovación tecnológica con la protección de los derechos fundamentales. Su análisis identifica la Ley de IA de la UE como un intento pionero de establecer "parámetros claros para el despliegue responsable de la IA en entornos de aprendizaje, reconociendo al mismo tiempo el potencial transformador de estas tecnologías" (p. 8).

Clasificación de riesgos para la IA educativa

Los sistemas de IA desplegados en entornos de EFP entran predominantemente en la categoría de **alto riesgo** debido a su importante impacto en las experiencias educativas de los estudiantes y en sus futuras oportunidades



profesionales. La Ley de IA de la UE clasifica los sistemas de IA del siguiente modo:

Riesgo inaceptable: se prohíben rotundamente los sistemas que supongan una amenaza clara para los derechos fundamentales o la seguridad.

Alto riesgo: aplicaciones en sectores críticos, incluido el educativo, que exigen un riguroso cumplimiento de la normativa.

Riesgo limitado: Sistemas con obligaciones de transparencia que exigen la divulgación de la interacción de la IA.

Riesgo mínimo: aplicaciones con requisitos normativos limitados más allá de los marcos existentes.

Implicaciones de la clasificación de alto riesgo para los centros de FP:

- Protocolos obligatorios de evaluación y mitigación de riesgos
- Requisitos rigurosos en materia de documentación
- Mecanismos de supervisión humana
- Obligaciones de transparencia para las partes interesadas
- Evaluaciones periódicas de la conformidad.

La investigación de Ramírez-Montoya et al. (2023) subraya que la clasificación de alto riesgo de los sistemas educativos de IA requiere "estructuras de gobernanza rigurosas que trasciendan el cumplimiento técnico para abordar preocupaciones sociales más amplias" (p. 142). Su análisis de los retos de implantación en los centros de formación profesional europeos pone de relieve la necesidad de una interpretación contextualizada de los requisitos normativos.

3.3.2 Consideraciones jurídicas clave para la implantación de la EFP

Requisitos de transparencia y rendición de cuentas

La Ley de Inteligencia Artificial de la UE exige una amplia transparencia en el despliegue de la IA en los centros educativos:

- Los centros de FP deben informar a los estudiantes, al personal y a las partes interesadas cuando se utilicen sistemas de IA en la evaluación, las recomendaciones de aprendizaje o la toma de decisiones administrativas.



- La documentación debe articular el propósito, la funcionalidad y las limitaciones de las aplicaciones de IA
- Las decisiones basadas en la IA que afecten a la progresión o certificación de los estudiantes deben ser explicables y justificables.
- Los procesos algorítmicos deben ser auditables para permitir la validación de los resultados.

Prinsloo y Knox (2023) identifican la transparencia como un principio ético fundamental en la implementación de la IA educativa, señalando que "la opacidad algorítmica presenta desafíos únicos en contextos de aprendizaje donde la confianza entre educadores y estudiantes constituye la base de una pedagogía eficaz" (p. 76). Su marco para la transparencia de la IA educativa va más allá de la divulgación técnica para abarcar la comprensión contextual accesible a diversas partes interesadas.

Protección de datos y respeto de la intimidad

La implantación de la IA en contextos de EFP exige un cumplimiento meticuloso de los marcos de protección de datos:

- Todos los sistemas de IA deben cumplir las disposiciones del Reglamento General de Protección de Datos (RGPD)
- Deben establecerse mecanismos claros de consentimiento para el tratamiento de datos personales
- Los principios de minimización de datos deben regir las prácticas de recogida y conservación
- La limitación de la finalidad garantiza que los datos se utilicen únicamente para objetivos educativos específicos.
- Las salvaguardias técnicas y organizativas deben proteger la integridad y confidencialidad de los datos de los estudiantes.

Holmes et al. (2023) observan que las consideraciones de protección de datos en contextos educativos van más allá del cumplimiento legal para abarcar "cuestiones éticas más amplias relativas a la agencia del estudiante, las dinámicas de poder y la posible vigilancia educativa" (p. 12). Su trabajo subraya la importancia de contextualizar los marcos de protección de datos dentro de las consideraciones éticas específicas de la educación.

Prevención de prejuicios y garantía de equidad



La Ley de IA de la UE establece requisitos explícitos para mitigar el sesgo algorítmico en contextos educativos:

- Los centros de FP deben realizar evaluaciones periódicas de los sistemas de IA para prevenir la discriminación
- Las metodologías de evaluación deben valorar los posibles impactos dispares entre grupos demográficos
- Los datos de entrenamiento algorítmico deben ser representativos de la diversidad de los estudiantes
- Los protocolos de supervisión continua deben identificar y corregir los sesgos emergentes en los sistemas operativos
- Se requiere una auditoría periódica de los resultados entre las poblaciones estudiantiles para verificar un impacto equitativo.

La investigación de Hodgkin y Kumar (2023) que examina el sesgo algorítmico en los sistemas de orientación profesional identifica "metodologías sistemáticas para evaluar el impacto dispar en diversas poblaciones de estudiantes" (p. 217) como esencial para la implementación equitativa de la IA. Su marco para la evaluación exhaustiva de los sesgos proporciona una guía práctica para las instituciones de EFP que navegan por los requisitos de equidad.

Herramientas de IA en el aula (riesgo limitado)

Numerosas aplicaciones de IA, como: Asistentes de tutoría potenciados por IA, sistemas de detección de plagio, motores de recomendación de recursos de aprendizaje, herramientas de automatización administrativa y simulaciones virtuales de laboratorio en entornos de EFP pueden entrar dentro de la clasificación de riesgo limitado:

Estas aplicaciones deben informar claramente a los usuarios sobre:

- La naturaleza IA del sistema
- Limitaciones de las capacidades algorítmicas
- Disposiciones sobre supervisión humana
- Mecanismos de impugnación de los resultados.

3.3.3 Calendario de aplicación y planificación de la conformidad

Los centros de EFP deben elaborar planes estratégicos de aplicación en consonancia con el calendario de aplicación gradual de la Ley de IA de la UE:



2^(nd) de febrero de 2025: Prohibición de los sistemas de IA que planteen riesgos inaceptables

2^(nd) de agosto de 2025: Aplicación de códigos de buenas prácticas y designación de autoridades reguladoras nacionales

2^(nd) de agosto de 2026: Aplicación general de la Ley de IA, incluidas las normas para los sistemas de IA de alto riesgo.

2^(nd) de agosto de 2027: pleno cumplimiento de las obligaciones específicas para los sistemas de IA de alto riesgo.

Una planificación prudente del cumplimiento requiere:

- Inventario exhaustivo de las aplicaciones de IA existentes y previstas
- Evaluación de la clasificación de riesgos de la cartera de tecnología educativa
- Análisis de los desfases con respecto a los requisitos reglamentarios
- Hoja de ruta estructurada con objetivos por etapas
- Asignación de recursos para la documentación y el control del cumplimiento.

3.3.4 Responsabilidades legales y responsabilidad civil

Los centros de FP que implantan sistemas de IA asumen importantes responsabilidades jurídicas:

Las estructuras de rendición de cuentas deben designar responsabilidades claras para la gobernanza del sistema de IA

Los marcos de responsabilidad deben abordar los daños potenciales derivados de errores o sesgos algorítmicos

Los acuerdos contractuales con los proveedores de tecnología deben delimitar explícitamente la asignación de responsabilidades

Los protocolos de indemnización deben abordar las posibles reclamaciones derivadas de decisiones influidas por la IA.

Las consideraciones en materia de seguros deben reflejar las nuevas exposiciones a la responsabilidad civil.



Las instituciones educativas deben reconocer que el despliegue de sistemas de IA, incluso los adquiridos a terceros, no las exime de la responsabilidad legal por los resultados que afecten a los estudiantes.

Rodríguez-Abitia y Bribiesca-Correa (2021) identifican las consideraciones de responsabilidad como una dimensión de la gobernanza de la IA educativa que a menudo se pasa por alto, señalando que "las instituciones educativas se enfrentan a riesgos legales únicos cuando delegan funciones tradicionalmente humanas en sistemas algorítmicos" (p. 94). Su análisis de la jurisprudencia emergente proporciona un precedente valioso para las instituciones de EFP que desarrollan marcos integrales de responsabilidad.

3.3.5 Lista de evaluación para una IA fiable (ALTAI) en la aplicación de la EFP

La Lista de Evaluación de la Inteligencia Artificial Confiable (ALTAI) proporciona a los centros de FP un marco estructurado para la aplicación ética y el cumplimiento de la normativa. Esta herramienta práctica hace operativas las Directrices Éticas para una Inteligencia Artificial Confiable desarrolladas por el Grupo de Expertos de Alto Nivel de la UE sobre Inteligencia Artificial.

En su análisis comparativo de los marcos de evaluación ética, Zawacki-Richter et al. (2023) identifican ALTAI como particularmente valioso en contextos educativos debido a su "cobertura integral de las dimensiones técnicas y socio-éticas relevantes para los entornos de aprendizaje" (p. 14). Su investigación demuestra la adaptabilidad de ALTAI a diversos contextos educativos, incluida la formación profesional.

Componentes clave de ALTAI relevantes para la implantación de la EFP

ALTAI estructura la evaluación en siete dimensiones fundamentales especialmente relevantes en contextos de formación profesional:

Agencia humana y supervisión

- Las aplicaciones de EFP deben apoyar la autonomía de los instructores en lugar de suplantar su criterio profesional
- La autonomía de los estudiantes debe preservarse mediante el consentimiento informado y la transparencia en las decisiones
- Debe mantenerse una supervisión humana significativa en las decisiones importantes que afectan a la progresión de los estudiantes.



Robustez técnica y seguridad

- Los sistemas de IA deben funcionar con fiabilidad en diversos entornos de formación profesional
- Los protocolos de seguridad deben proteger contra posibles manipulaciones o accesos no autorizados.
- Los procedimientos alternativos deben garantizar la continuidad de la enseñanza en caso de fallo del sistema.

Privacidad y gobernanza de datos

- Los datos de los estudiantes utilizados en la formación y el funcionamiento de la IA deben protegerse mediante una gobernanza global.
- Los principios de minimización de datos deben guiar las prácticas de recogida y conservación
- Unas sólidas medidas de seguridad deben salvaguardar la información sensible de accesos no autorizados.

Transparencia

- Las capacidades y limitaciones de la IA deben comunicarse claramente a las partes interesadas en la educación
- Los procesos de decisión deben poder explicarse en un lenguaje no técnico
- La documentación debe permitir verificar el rendimiento y la conformidad del sistema.

Diversidad, no discriminación y equidad

- La aplicación específica de la EFP debe tener en cuenta la diversidad de orígenes y necesidades de aprendizaje de los estudiantes.
- Las metodologías de evaluación deben valorar el impacto de la equidad en todas las dimensiones demográficas
- Las estrategias para mitigar los prejuicios deben tener en cuenta las consideraciones de diversidad propias de la formación profesional.

Bienestar social y medioambiental

- Deben evaluarse sistemáticamente las amplias repercusiones sociales de la IA en la formación profesional
- Hay que evaluar las consecuencias medioambientales de las infraestructuras informáticas



- Deben tenerse en cuenta las posibles implicaciones de la certificación de competencias en IA para el mercado laboral.

Rendición de cuentas

- Deben establecerse estructuras de gobernanza claras que determinen la responsabilidad de los sistemas de IA
- Los mecanismos de auditoría deben permitir verificar el cumplimiento de la normativa
- Los procedimientos de reparación deben abordar los resultados perjudiciales o las consecuencias imprevistas.

Aplicación de ALTAI en contextos de EFP

Los centros de FP pueden hacer operativa la ALTAI mediante protocolos de evaluación sistemática:

La evaluación inicial de la aplicación compara las aplicaciones de IA propuestas con criterios de fiabilidad.

La verificación previa a la implantación garantiza el cumplimiento de los requisitos técnicos y éticos

La auditoría posterior a la implantación valida el rendimiento en diversas poblaciones de estudiantes.

La reevaluación periódica identifica los problemas éticos emergentes a medida que evolucionan los sistemas.

Al integrar ALTAI en los marcos de gobernanza, las instituciones de EFP pueden demostrar la diligencia debida en el cumplimiento de la normativa y, al mismo tiempo, aumentar la confianza de las partes interesadas en la aplicación de la IA.

3.4 Estudios de casos y ejemplos de políticas eficaces de IA en centros de FP

3.4.1 Marcos políticos europeos fundamentales

Las políticas institucionales eficaces se basan en marcos fundacionales establecidos a escala europea:



Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial (AI HLEG)

El HLEG de IA de la Comisión Europea ha elaborado orientaciones esenciales a través de:

Directrices éticas para una IA digna de confianza: establecimiento de siete requisitos para la implantación ética de la IA

Recomendaciones políticas y de inversión: Orientación estratégica para el desarrollo sostenible de la IA

Assessment List for Trustworthy AI: Ofrece herramientas prácticas de aplicación para las organizaciones.

Estos marcos fundacionales han servido de base a las políticas institucionales concretas de los proveedores europeos de EFP.

Como observan Holmes et al. (2023) en su análisis de los marcos de gobernanza de la IA, las directrices HLEG de IA representan "un puente crítico entre los principios éticos abstractos y las prácticas de aplicación concretas" (p. 18). Su investigación rastrea la influencia de estas directrices en diversos contextos educativos, demostrando su adaptabilidad a los entornos de formación profesional.

3.4.2 Escuela Técnica Superior de Copenhague: Marco global de gobernanza de la IA

La Escuela Técnica Superior de Copenhague ha establecido una estructura de gobernanza a varios niveles que demuestra una práctica ejemplar en la aplicación de la IA en la EFP.

Arquitectura política y estructura de gobierno

El planteamiento de la institución se caracteriza por:

- Comité de Ética de la Inteligencia Artificial compuesto por administradores, educadores, estudiantes, representantes de la industria y expertos técnicos.
- Protocolos de aprobación escalonados basados en la clasificación algorítmica del riesgo
- Formación obligatoria en AI para todo el personal docente y administrativo



- Formalización de los derechos de los estudiantes en materia de transparencia algorítmica e impugnabilidad de las decisiones
- Protocolos de evaluación estructurados que miden las dimensiones educativa y ética.

Este enfoque está en consonancia con el marco de participación de las partes interesadas de Prinsloo y Knox (2023), que hace hincapié en la importancia de la "representación estructurada a través de las jerarquías institucionales y las comunidades afectadas" (p. 82). Su investigación demuestra cómo la gobernanza multilateral mejora tanto el rigor ético como la eficacia de la aplicación.

Metodología de aplicación

Entre los elementos clave de la aplicación figuran:

- Despliegue por fases, empezando por las aplicaciones administrativas de bajo riesgo.
- Proyectos pedagógicos piloto controlados con protocolos de evaluación rigurosos
- Documentación estructurada tanto de los resultados técnicos como del impacto ético.
- Mecanismos regulares de retroalimentación de las partes interesadas para informar sobre la mejora iterativa.
- Auditoría independiente de los resultados algorítmicos a través de dimensiones demográficas.

Resultados documentados

El Colegio ha informado de importantes beneficios gracias a su aplicación estructurada:

- Reducción del 27% de la carga de trabajo administrativo mediante la automatización ética de la IA
- Mayor personalización educativa sin comprometer la calidad de la enseñanza
- Mayor transparencia en las prácticas de evaluación
- Mayor satisfacción de los estudiantes con la integración tecnológica
- Cumplimiento satisfactorio de los nuevos requisitos normativos



3.4.3 Instituto Tecnológico de Burdeos: Políticas transparentes de evaluación de la IA

El Instituto de Burdeos ha elaborado políticas modelo que abordan uno de los aspectos éticamente más complejos de la IA educativa: la valoración y evaluación de los alumnos.

Marco político específico para la evaluación

Entre los elementos políticos clave figuran:

- Prohibición de la evaluación totalmente automatizada de altas apuestas
- Revisión humana obligatoria de todas las decisiones de evaluación influidas por la IA
- Documentación transparente de la contribución algorítmica a las evaluaciones
- Mecanismo de recurso estructurado para impugnar los resultados influidos por la IA
- Auditorías periódicas de sesgos en las que se comparan los resultados de las evaluaciones entre poblaciones de estudiantes.
- Representación de los estudiantes en las estructuras de gobernanza de la evaluación

Este enfoque ejemplifica lo que Ramírez-Montoya et al. (2023) identifican como "gobernanza de la evaluación centrada en el ser humano", caracterizada por "transparencia algorítmica, participación de las partes interesadas y mecanismos de supervisión formalizados" (p. 149). Su investigación demuestra cómo estos marcos mejoran tanto la validez de la evaluación como la confianza de las partes interesadas.

Enfoque de aplicación

La estrategia de aplicación hace hincapié en:

- Delimitación clara entre automatización de la evaluación apropiada e inapropiada
- Aplicación gradual en función de los intereses y las consecuencias de la evaluación
- Desarrollo del profesorado en metodologías éticas de evaluación de la IA



- Infraestructura técnica que garantiza la explicabilidad de las recomendaciones algorítmicas
- Protocolos de documentación que permitan verificar la equidad de la evaluación.

Impacto mensurable

El Instituto ha documentado importantes beneficios:

- Mayor coherencia en la evaluación, al tiempo que se preserva la autonomía del instructor
- Mayor confianza de los estudiantes en las metodologías de evaluación
- Reducción de la carga de trabajo de evaluación sin comprometer la calidad
- Mejor identificación de los estudiantes que necesitan apoyo adicional
- Cumplimiento normativo satisfactorio de los requisitos de transparencia de la IA.

3.4.4 Consorcio de Escuelas de Formación Profesional de Múnich: Gobierno colaborativo de datos

Ocho centros de formación profesional de Múnich han establecido un enfoque colaborativo para la gobernanza de los datos de IA, demostrando cómo la puesta en común de recursos puede mejorar la aplicación ética.

Arquitectura de gobernanza compartida

Los elementos estructurales clave incluyen:

- Marco ético común para la recogida y utilización de datos
- Infraestructura técnica compartida para la gestión de datos y la seguridad
- Protocolos de colaboración para la evaluación y contratación de proveedores
- Aplicación conjunta de metodologías de aprendizaje federado
- Peritaje conjunto para la verificación del cumplimiento de la normativa.

Este enfoque colaborativo se ajusta a lo que Rodríguez-Abitia y Bribiesca-Correa (2021) denominan "gobernanza con optimización de recursos", caracterizada por "conocimientos compartidos, políticas unificadas y capacidad de ejecución colectiva" (p. 98). Su investigación demuestra cómo estos enfoques pueden



permitir a las instituciones más pequeñas aplicar marcos de gobernanza sofisticados.

Metodología de aplicación

El planteamiento del Consorcio se caracteriza por:

- Protocolos normalizados de protección de datos en todas las instituciones participantes
- Experiencia centralizada en técnicas de minimización y anonimización de datos
- Recursos compartidos para la evaluación técnica de la vulnerabilidad
- Desarrollo colaborativo de metodologías analíticas que preserven la privacidad
- Participación conjunta de las partes interesadas en el desarrollo de la gobernanza

Beneficios demostrados

Este planteamiento de colaboración ha dado sus frutos:

- Gobernanza de datos más sofisticada de lo que las instituciones individuales podrían lograr de forma independiente.
- Mejora de la posición negociadora con los proveedores de tecnología
- Mejora de los protocolos de seguridad de datos gracias a la experiencia compartida
- Cumplimiento rentable de la normativa en múltiples instituciones
- Mitigación más exhaustiva de los sesgos mediante conjuntos de datos más amplios y diversos.

3.4.5 Instituto Técnico de Dublín: Marco político de la IA centrado en la equidad

El Instituto Técnico de Dublín ha desarrollado políticas que priorizan específicamente la equidad algorítmica en contextos de formación profesional.

Marco político centrado en la equidad

Entre los elementos políticos clave figuran:

- Evaluaciones obligatorias del impacto sobre la equidad antes del despliegue de la IA



- Auditoría algorítmica periódica de las dimensiones demográficas
- Priorización de metodologías de IA explicables
- Necesidad de alternativas no tecnológicas a los sistemas de IA
- Inclusión formalizada de las poblaciones infrarrepresentadas en las estructuras de gobierno.

Este marco ejemplifica lo que Hodgkin y Kumar (2023) identifican como "gobernanza proactiva de la equidad" caracterizada por "una evaluación sistemática de los sesgos, una representación inclusiva de las partes interesadas y protocolos de supervisión continua" (p. 219). Su investigación demuestra cómo la gobernanza centrada en la equidad mejora tanto la aplicación ética como los resultados educativos.

Estrategia de aplicación

El enfoque de aplicación hace hincapié en:

- Pruebas sistemáticas de sesgo en algoritmos de orientación profesional
- Seguimiento demográfico de los resultados educativos influidos por la IA
- Desarrollo de metodologías de depuración de datos de formación
- Perfeccionamiento del profesorado en la integración de la Inteligencia Artificial consciente de la equidad
- Opiniones periódicas de las partes interesadas procedentes de diversas poblaciones estudiantiles.

Impacto documentado

El Instituto lo ha demostrado:

- Identificación y corrección de posibles sesgos en los sistemas de orientación profesional
- Mejores resultados educativos para los grupos históricamente infrarrepresentados
- Mayor confianza de las partes interesadas en la integración tecnológica
- Cumplimiento satisfactorio de los requisitos de no discriminación
- Desarrollo de metodologías transferibles centradas en la equidad

3.4.6 Aplicación de ALTAI en contextos específicos de EFP



La Lista de Evaluación de la Inteligencia Artificial de Confianza (ALTAI) se ha aplicado con éxito en numerosos contextos de formación profesional, lo que demuestra su utilidad práctica para la aplicación ética.

La investigación de Zawacki-Richter et al. (2023) identifica la implementación de ALTAI en la formación profesional como particularmente valiosa debido a los "contextos profesionales concretos que caracterizan la formación profesional y las implicaciones profesionales explícitas de las decisiones influenciadas por la IA" (p. 15). El análisis de su estudio de caso demuestra la adaptabilidad de ALTAI a diversos entornos de formación profesional.

Aplicaciones de evaluación de competencias basadas en IA

Los centros de FP han aplicado ALTAI para garantizar una implantación ética de los sistemas de evaluación automatizada:

Integración de la supervisión humana: Garantizar la validación de las evaluaciones algorítmicas por parte de los instructores

Verificación de la equidad: Comprobación de la equidad de la evaluación en poblaciones estudiantiles diversas

Requisitos de transparencia: Criterios de evaluación explícitos y comprensibles

Solidez técnica: Garantizar un funcionamiento fiable con distintos patrones de rendimiento de los alumnos.

IA en orientación profesional y asesoramiento vocacional

Se ha abordado la aplicación de ALTAI en los sistemas de recomendación de carreras profesionales:

- Posible refuerzo de los estereotipos profesionales
- Transparencia en las metodologías de recomendación
- Protección de datos en la modelización predictiva de la carrera profesional
- Accesibilidad de los sistemas de orientación en poblaciones diversas.

Aprendizaje adaptativo basado en IA para la formación profesional

La implantación de plataformas de aprendizaje adaptativo se ha beneficiado de la evaluación de ALTAI a través de:



- Evaluación de la equidad de la personalización algorítmica
- Verificación de la diversificación de las vías de aprendizaje
- Evaluación de los posibles efectos de las burbujas de filtración
- Validación de las capacidades de supervisión del instructor.

3.4.7 Elementos comunes de las políticas eficaces

El análisis de la aplicación con éxito en todas las instituciones europeas de EFP revela pautas políticas coherentes:

Estructuras de gobernanza multipartitas que impliquen diversas perspectivas en el desarrollo y la supervisión.

Protocolos de evaluación de riesgos por niveles que aplican un escrutinio proporcionado basado en el impacto potencial.

Mecanismos de supervisión continua en lugar de procesos de aprobación puntuales.

Sólidos requisitos de transparencia que garanticen a las partes interesadas la comprensión de las capacidades y limitaciones de la IA.

Supervisión humana significativa, especialmente en las decisiones educativas importantes

Ciclos periódicos de revisión de las políticas para adaptarlas a la evolución de las tecnologías y la normativa

Componentes educativos integrales para la alfabetización en IA de todas las partes interesadas.

Estructuras claras de rendición de cuentas que establezcan la responsabilidad de los resultados influidos por la IA.

Holmes et al. (2023) identifican estos elementos coherentes como representativos de un "consenso emergente en relación con los principios fundacionales de gobernanza para la implementación de la IA educativa" (p. 21). Su análisis comparativo entre diversos contextos institucionales demuestra la adaptabilidad de estos principios básicos a diversos entornos educativos.

Las instituciones que demuestran una práctica ejemplar han equilibrado el potencial de innovación con las salvaguardias adecuadas, reconociendo que los



enfoques demasiado restrictivos pueden limitar los beneficios educativos, mientras que una gobernanza insuficiente crea riesgos significativos. Los marcos más eficaces han integrado las consideraciones éticas en todo el ciclo de vida de la aplicación, en lugar de tratarlas como consideraciones a posteriori.

4. Habilidades éticas

4.1 Definición de competencias éticas

Las competencias éticas se refieren a la capacidad de evaluar si una IA se utiliza de forma ética en un contexto específico. En el contexto educativo, por ejemplo, las competencias éticas consisten en saber si las políticas de un centro educativo son éticas, si se mantiene la seguridad y privacidad de los datos, etc.

Las competencias éticas se refieren a las capacidades críticas necesarias para evaluar y garantizar el uso ético de la IA y los datos en entornos educativos, basándose en principios como la agencia humana, la equidad, la humanidad y la elección justificada. La agencia humana implica permitir a las personas tomar decisiones informadas y asumir la responsabilidad de sus actos, fomentando la autonomía y la responsabilidad. La equidad garantiza un trato justo y la igualdad de acceso a las oportunidades para todos, promoviendo la inclusión y la no discriminación. La humanidad hace hincapié en el respeto de la dignidad humana, la identidad y el bienestar, dando prioridad a las conexiones humanas significativas sobre el uso instrumental de los datos. La elección justificada requiere procesos transparentes de toma de decisiones basados en principios éticos y en la participación de las partes interesadas, garantizando que los sistemas de IA en la educación se desplieguen de forma responsable y beneficien a todas las partes implicadas.

4.2 Competencias éticas Importancia

Estas consideraciones éticas son cruciales para guiar a los educadores y las instituciones educativas en la navegación por las complejidades de la implementación de la IA. Al adherirse a requisitos éticos como los propuestos por las Directrices Éticas de la IA, las partes interesadas en la educación pueden mitigar eficazmente los riesgos, promover la equidad y defender el bienestar de los estudiantes y el personal por igual.

4.3 Ámbitos y ejemplos de competencias éticas



Área 1: Compromiso profesional. Uso de las tecnologías digitales para la comunicación, la colaboración y el desarrollo profesional.

Elemento de competencia	Indicadores potenciales
Es capaz de describir de forma crítica las repercusiones positivas y negativas de la IA y el uso de datos en la educación.	Participa activamente en el aprendizaje profesional continuo sobre IA y analítica del aprendizaje y su uso ético.
	Capaz de dar ejemplos de sistemas de IA y describir su relevancia.
	Conoce cómo se evalúa en la escuela el impacto ético de los sistemas de IA.
	Sabe cómo iniciar y promover estrategias en toda la escuela y su comunidad en general que fomenten el uso ético y responsable de la IA y los datos.
Comprender los fundamentos de la IA y el análisis del aprendizaje	Consciente de que los algoritmos de IA funcionan de formas que los usuarios no suelen ver ni comprender fácilmente.
	Puede interactuar con el sistema de inteligencia artificial y darle información para influir en sus recomendaciones.
	Consciente de que los sensores utilizados en muchas tecnologías y aplicaciones digitales generan grandes cantidades de datos, incluidos datos personales, que pueden utilizarse para entrenar un sistema de IA.
	Conocimiento de las directrices éticas sobre IA de la UE y de los instrumentos de autoevaluación.



Ejemplo 1: Ana es una profesora que quiere impartir una lección de mecánica de fluidos, y quiere utilizar la IA para generar algunas imágenes e incluirlas en el powerpoint de la asignatura. Antes de incluir la IA en la lección, Ana ha leído las directrices éticas sobre IA de la UE y ha evaluado si es ético incluir su uso o no.

Ejemplo 2: Lucas imparte clases en la FP de Educación y Control Ambiental. Varios profesores de su centro educativo utilizan la IA, por lo que ha propuesto una reunión con todo el centro educativo para analizar el uso de la IA por parte de los profesores. El objetivo de Lucas es asegurarse de que todos los profesores hacen un uso ético y responsable de la IA y de los datos.

Área 2: Recursos digitales: Buscar, crear y compartir recursos digitales

Elemento de competencia	Indicadores potenciales
Gobernanza de datos	Conocimiento de las distintas formas de datos personales utilizados en la educación y la formación.
	Consciente de las responsabilidades en el mantenimiento de la seguridad y la privacidad de los datos.
	Sabe que el tratamiento de datos personales está sujeto a la normativa nacional y de la UE, incluido el Reglamento general de protección de datos (RGPD).
	Conoce quién tiene acceso a los datos de los alumnos, cómo se controla el acceso y cuánto tiempo se conservan los datos.



	Sabe que todos los ciudadanos de la UE tienen derecho a no ser objeto de una toma de decisiones totalmente automatizada.
	Capaz de dar ejemplos de datos sensibles, incluidos los datos biométricos.
	Capaz de sopesar los beneficios y los riesgos antes de permitir a terceros el tratamiento de datos personales, especialmente cuando se utilizan sistemas de IA.
Gobernanza de la IA	Sabe que los sistemas de IA están sujetos a la normativa nacional y de la UE (en particular, la Ley de IA que debe aprobarse).
	Capaz de explicar el enfoque basado en el riesgo de la Ley de AI (que debe adoptarse).
	Conoce los casos de uso de IA de alto riesgo en la educación y los requisitos asociados en virtud de la Ley de IA (cuando se adopte).
	Sabe cómo incorporar contenidos digitales editados/manipulados con IA en su propio trabajo y cómo debe acreditarse dicho trabajo.
	Capaz de explicar los principios clave de la calidad de los datos en los sistemas de IA.
Ejemplo 1: María es profesora en el centro de FP de Eficiencia Energética y Energía Solar Térmica. Está creando apuntes para su curso y los entrega a los alumnos, y ha incorporado algunos contenidos digitales editados por AI a los apuntes utilizando la herramienta Copilot. María sabe que este contenido debe ser acreditado y se asegura de acreditarlo correctamente.	
Ejemplo 2: Jaime imparte clases en el centro de FP de Fabricación Mecánica. En ocasiones utiliza ChatGPT para desarrollar algunos contenidos o	



actividades de su asignatura. En algún momento valora la posibilidad de incluir los datos personales de sus alumnos (nombre, apellidos, fecha de nacimiento y dirección) para crear un documento que incluya toda esta información. Consciente de su responsabilidad de mantener la seguridad y privacidad de los datos, decide no incluir los datos personales.

Ámbito 3: Utilización de tecnologías y estrategias digitales para mejorar la evaluación

Elemento de competencia	Indicadores potenciales
Modelos de aprendizaje	Sabe que los sistemas de IA ponen en práctica la concepción del diseñador sobre qué es el aprendizaje y cómo puede medirse; puede explicar los supuestos pedagógicos clave en los que se basa un determinado sistema de aprendizaje digital.
Objetivos de la educación	Conoce cómo un determinado sistema digital aborda los diferentes objetivos sociales de la educación (cualificación, socialización, subjetivación).
Agencia humana	Capaz de considerar el impacto del sistema de IA en la autonomía del profesorado, el desarrollo profesional y la innovación educativa.
	Considere las fuentes de sesgo inaceptables en la IA basada en datos.
Equidad	Considera los riesgos relacionados con la dependencia emocional y la autoimagen de los estudiantes cuando se utilizan sistemas interactivos de IA y analítica del aprendizaje.
Humanidad	Capaz de considerar el impacto de la IA y el uso de datos en la comunidad estudiantil.



	Confianza para debatir los aspectos éticos de la IA y cómo influyen en la forma de utilizar la tecnología.
Participa en el desarrollo de prácticas de aprendizaje que utilizan IA y datos	Puede explicar cómo se tienen en cuenta y se negocian los principios y valores éticos en el codiseño y la cocreación de prácticas de aprendizaje que utilizan IA y datos (vinculado al diseño del aprendizaje).
Ejemplo 1: Luisa, profesora de Mantenimiento Electromecánico en una empresa de formación, implanta un software de IA que utiliza modelos de aprendizaje adaptativo para personalizar la enseñanza. El software evalúa el nivel inicial de cada aprendiz y ajusta el contenido y las tareas en función de su progreso. Esto permite a cada aprendiz de avanzar a su propio ritmo, recibiendo información inmediata y recursos específicos para superar las dificultades. Además, Luisa puede acceder a informes detallados para identificar las áreas en las que sus alumnos necesitan apoyo adicional, lo que garantiza una enseñanza integradora y eficaz.	
Ejemplo 2: Javi, profesor de Artes Gráficas en el programa de Impresión Gráfica de nivel medio de un centro de FP, integra un software de IA que mejora la capacidad humana ayudándole en tareas de diseño complejas. El software ayuda a los estudiantes automatizando tareas repetitivas, como la corrección del color y los ajustes de maquetación, lo que les permite centrarse en los aspectos creativos y estratégicos de sus proyectos. Esto permite a los estudiantes desarrollar sus habilidades artísticas y su pensamiento crítico mientras la IA se ocupa de los detalles técnicos.	

Área 4: Evaluación: Gestionar y orquestar el uso de las tecnologías digitales en la enseñanza y el aprendizaje.	
Elemento de competencia	Indicadores potenciales



Diferencias personales	Consciente de que los estudiantes reaccionan de distintas maneras a los comentarios automatizados.
Sesgo algorítmico	Considere las fuentes de sesgos inaceptables en los sistemas de IA y cómo pueden mitigarse.
Enfoque cognitivo	Consciente de que los sistemas de IA evalúan el progreso de los alumnos basándose en modelos predefinidos de conocimientos específicos del ámbito.
	Consciente de que la mayoría de los sistemas de IA no evalúan la colaboración, las competencias sociales ni la creatividad.
	Conocimiento de las formas habituales de manipular la evaluación basada en la IA.
Nuevas formas de abusar de la tecnología	Consciente de que los profesores pueden confiar demasiado en los comentarios generados por la IA, que a veces pueden parecer impersonales y/o no abordar las necesidades específicas de los alumnos o sus malentendidos.
Ejemplo 1: Jorge imparte una asignatura en el centro de FP de Instalaciones y Mantenimiento. Tras finalizar el ^{primer} tema de la asignatura, utiliza una herramienta de IA para evaluar la adquisición de los contenidos por parte de sus alumnos. Además, esta herramienta de IA proporciona automáticamente feedback y adapta el nivel de las preguntas en función de las respuestas anteriores. Jorge es consciente de cómo reaccionan los alumnos a los comentarios automáticos, por lo que se asegura de proporcionarles comentarios personalizados y de comprobar él mismo los progresos de los estudiantes. Por lo tanto, utiliza la IA sólo como herramienta para dar una evaluación inicial, que él mismo comprueba y revisa posteriormente.	
Ejemplo 2: Natalia es formadora en un proveedor de servicios de EFP para energías renovables. Anima a los alumnos a utilizar Windows Copilot en sus clases de formación para determinadas actividades. Aunque esta herramienta funciona como un asistente de IA que puede ser útil para la productividad y la	



creatividad, es consciente de que se necesitan otros tipos de actividades y evaluaciones para mejorar la colaboración, las competencias sociales y la creatividad, ya que la IA no es capaz de evaluar estos aspectos. Ella misma sigue encargándose de las actividades y evaluaciones pertinentes como formadora y utiliza la IA como herramienta durante sus clases y para la preparación de la formación.

Área 5: Capacitación de los alumnos: Uso de las tecnologías digitales para mejorar la inclusión, la personalización y la participación activa de los alumnos

Elemento de competencia	Indicadores potenciales
La IA responde a las diversas necesidades de aprendizaje de los alumnos	Conoce las distintas formas en que los sistemas de aprendizaje personalizado pueden adaptar su comportamiento (contenido, trayectoria de aprendizaje, enfoque pedagógico).
	Capaz de explicar cómo un sistema determinado puede beneficiar a todos los alumnos, independientemente de sus diferencias cognitivas, culturales, económicas o físicas.
	Consciente de que los sistemas de aprendizaje digital tratan de forma diferente a los distintos grupos de estudiantes
	Capaz de considerar el impacto en el desarrollo de la autoeficacia, la autoimagen, la mentalidad y las habilidades de autorregulación cognitiva y afectiva del alumno.



Elección justificada	Sabe que la IA y el uso de datos pueden beneficiar a unos alumnos más que a otros.
	Ser capaz de explicar qué pruebas se han utilizado para justificar la implantación de un determinado sistema de IA en el aula.
	Reconoce la necesidad de un seguimiento constante de los resultados del uso de la IA.

Ejemplo 1: Teresa, que imparte Formación Profesional de Grado Superior en Administración y Finanzas en un proveedor de educación para adultos, utiliza la IA para adaptar las experiencias de aprendizaje. El sistema de IA evalúa las necesidades de cada alumno y ajusta los recursos y las actividades en consecuencia. Por ejemplo, los alumnos visuales pueden recibir herramientas interactivas, mientras que los analíticos reciben estudios de casos detallados. Este enfoque garantiza un aprendizaje personalizado que satisface los diversos estilos de aprendizaje y maximiza el compromiso de los estudiantes.

Ejemplo 2: Óscar, que imparte Formación Profesional de Grado Superior en Diseño de Fabricación Mecánica, emplea un software de IA para diseñar piezas mecánicas. El software ofrece múltiples soluciones de diseño y simula el rendimiento en diferentes condiciones. Los estudiantes analizan los resultados para seleccionar el diseño más eficiente, guiados por Óscar para desarrollar habilidades críticas y analíticas cruciales para sus carreras.

Ámbito 6: Facilitar la competencia digital de los alumnos permitiéndoles utilizar de forma creativa y responsable las tecnologías digitales para la información, la comunicación, la creación de contenidos, el bienestar y la resolución de problemas.

Elemento de competencia	Indicadores potenciales
-------------------------	-------------------------



Ética de la IA y la analítica del aprendizaje	Capaz de utilizar proyectos y despliegues de IA para ayudar a los estudiantes a aprender sobre la ética de la IA y el uso de datos en la educación y la formación.
Ejemplo 1: La profesora Sara enseña análisis de marketing avanzado a sus alumnos de formación profesional en una universidad. Comenta un caso práctico en el que la IA analiza los datos de las redes sociales para la publicidad dirigida. Sara promueve un debate sobre cuestiones éticas como la privacidad, el consentimiento y el sesgo algorítmico. Guía a los estudiantes en el desarrollo de directrices éticas para el uso de la IA en el marketing, fomentando su comprensión del despliegue responsable de la IA en el sector.	
Ejemplo 2: El profesor Jesús enseña gestión culinaria, centrándose en el papel de la IA en la personalización de los menús de los restaurantes. Comenta un caso práctico en el que la IA analiza las preferencias dietéticas para ofrecer recomendaciones de menú personalizadas. Jesús invita a los alumnos a debatir cuestiones éticas como la privacidad de los datos, la diversidad de los menús y el impacto de la IA en la satisfacción del cliente. Guía a los alumnos en el desarrollo de marcos éticos para el uso de la IA en la personalización de menús, preparándoles para equilibrar la innovación con las consideraciones éticas en las operaciones de restauración.	

5. Conclusión

A medida que la IA sigue configurando el panorama educativo, es esencial que la adopción de las tecnologías de la IA se haga de una manera que defienda los valores éticos, promueva la equidad y salvaguarde la privacidad de los estudiantes. Este manual ha esbozado el propósito y el alcance de la implementación de la IA, subrayando la importancia de las consideraciones éticas para garantizar que las tecnologías de IA se utilicen de forma responsable en todos los contextos educativos.

5.1 Principales conclusiones

La integración de la IA en la educación encierra un inmenso potencial, pero debe guiarse por principios éticos que den prioridad a la diversidad, la inclusión y la transparencia. Las herramientas de IA deben estar diseñadas para dar cabida a



alumnos de diversos orígenes, teniendo en cuenta las diferencias culturales, lingüísticas y socioeconómicas para promover la equidad en los resultados educativos. La transparencia es igualmente vital: los educadores y los estudiantes deben entender cómo toman decisiones los sistemas de IA, especialmente en áreas como la calificación, la retroalimentación o las rutas de aprendizaje personalizadas. Garantizar la privacidad y la protección de datos es esencial, lo que requiere el pleno cumplimiento de normativas como el GDPR y la adhesión a prácticas éticas de gestión de datos. Los sistemas de IA también deben ser sostenibles, minimizando su impacto medioambiental y garantizando su viabilidad a largo plazo en los ecosistemas educativos.

Igualmente importante es promover la alfabetización en IA entre todas las partes interesadas. Los educadores necesitan formación específica para integrar de forma ética y eficaz las herramientas de IA en su pedagogía, mientras que los estudiantes deben estar equipados con una comprensión crítica de las capacidades y limitaciones de la IA. Los administradores institucionales desempeñan un papel clave a la hora de alinear las políticas escolares o universitarias con las prácticas éticas de la IA, fomentando un entorno de responsabilidad y equidad. Los responsables políticos deben intervenir para crear marcos normativos sólidos que guíen la adopción responsable de la IA, mitiguen el sesgo algorítmico y eviten resultados discriminatorios. En última instancia, un enfoque de múltiples partes interesadas es crucial para garantizar que la IA mejore la educación de una manera que sea justa, inclusiva y sostenible.

También es muy importante tener en cuenta algunos marcos jurídicos y normativos clave, como la Ley de IA de la UE (2024), que introduce un marco jurídico global que clasifica los sistemas de IA en función de su riesgo para los derechos fundamentales, con importantes implicaciones para los centros educativos.

A partir de ahí están las aplicaciones de IA de alto riesgo, como las que se utilizan en las evaluaciones o admisiones de estudiantes, que están sujetas a estrictos requisitos de cumplimiento, incluida una documentación rigurosa, supervisión humana y sólidos protocolos de gestión de riesgos.

Además, la transparencia y la rendición de cuentas no son una opción. Las entidades deben revelar claramente cómo funcionan los sistemas de IA y garantizar que los usuarios puedan comprender y cuestionar las decisiones automatizadas. Además, la Ley exige medidas proactivas para auditar y mitigar el sesgo algorítmico, garantizando la equidad y evitando resultados



discriminatorios. Para apoyar estos esfuerzos, la Lista de Evaluación de la Inteligencia Artificial Confiable (ALTAI) ofrece una herramienta práctica que ayuda a las instituciones a evaluar el impacto ético de sus sistemas de IA, fomentando una cultura de despliegue de IA responsable y confiable en la educación.

5.2 Reflexiones finales

Todos deberíamos ser conscientes del hecho de que, a medida que la IA sigue remodelando la educación universitaria y profesional, este manual hace hincapié en la importancia de mantener un cuidadoso equilibrio entre las prácticas innovadoras y la responsabilidad ética. Los beneficios de la IA, desde los itinerarios de aprendizaje personalizados hasta la mejora de la eficiencia administrativa, son significativos, pero conllevan riesgos que no deben pasarse por alto ni ignorarse. Cuestiones como el sesgo algorítmico, las violaciones de la privacidad de los datos y la disminución del papel del juicio humano exigen una atención constante. Para sortear estas complejidades, debe existir una colaboración constante entre educadores, administradores y responsables políticos. Por supuesto, los principios éticos deben estar profundamente arraigados en el desarrollo, despliegue y evaluación de las herramientas de IA para garantizar que apoyan la inclusión, respetan la autonomía y refuerzan la confianza en el proceso educativo.

De cara al futuro, el éxito de la integración de la IA en la educación dependerá no sólo de la preparación tecnológica, sino también de una visión compartida basada en la transparencia, la equidad y el aprendizaje permanente. El sector debe fomentar la alfabetización en IA entre todas las partes interesadas y crear sistemas sólidos de rendición de cuentas y supervisión. No se trata de un ajuste puntual, sino de un viaje evolutivo que exige reflexión y diálogo continuos. Al defender los enfoques centrados en el ser humano y la innovación ética, la comunidad educativa puede transformar la IA en un poderoso aliado que mejore las oportunidades de aprendizaje, preservando al mismo tiempo los valores de una educación de calidad. Para una orientación completa, se anima a las partes interesadas a consultar los recursos del proyecto AI Pioneers y el Esquema de evaluación de la IA en la educación, accesibles en varios idiomas.

Acuse de recibo s

Queremos dar las gracias a todos los que han contribuido a la elaboración de este documento con sus críticas constructivas y sus inestimables aportaciones. Estamos profundamente agradecidos a todos los socios de nuestro proyecto, que han contribuido al contenido, así como a la Red de Referencia de Pioneros



de la IA y a los participantes en nuestros seminarios de consulta. Sus preocupaciones y opiniones compartidas ofrecieron una valiosa orientación a lo largo de este proceso.

Por último, reconocemos la importancia de la bibliografía de referencia, junto con otros recursos como artículos científicos, informes, seminarios, directrices de instituciones educativas y entradas de blog de investigadores y educadores. Colectivamente, estas contribuciones han informado en gran medida nuestro enfoque para abordar las consideraciones éticas en el uso de la IA en la educación a todos los niveles.

Referencias

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.

Aldosari, S. A. M. (2020). El futuro de la educación superior a la luz de las transformaciones de la inteligencia artificial. *Revista Internacional de Educación Superior*, 9(3), 145-151.

Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). La gobernanza de la inteligencia artificial en Canadá: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, 41(2), 101929.

Attwell, G., Bekiaridis, G., Deitmer, L., Perini, M., Roppertz, S., & Tütlys, V. (2020). Inteligencia artificial en políticas, procesos y prácticas de educación y formación profesional.

Attwell, G., Deitmer, L., Tütlys, V., Roppertz, S., & Perini, M. (2020). Digitalización, inteligencia artificial y ocupaciones y competencias profesionales: ¿cuáles son las necesidades de formación de profesores y formadores?. En *Trends in vocational education and training research, Vol. III. Proceedings of the European Conference on Educational Research (ECER), Vocational Education and Training Network (VETNET)* (pp. 30-42).

Bekiaridis, G., & Attwell, G. (2024). Integración de la Inteligencia Artificial en la Formación Profesional y la Educación de Adultos: A Supplement to the DigCompEdu Framework. *Actas de Ubiquity*, 4(1).

Boninger, F., Molnar, A., & Saldaña, C. (2020). Grandes afirmaciones, pocas pruebas, mucho dinero: The Reality behind the Summit Learning Program and the Push to Adopt Digital Personalized Learning Programs. *Commercialism in Education Research Unit*.



Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). La inteligencia artificial por sí sola no democratizará la educación: On educational inequality, techno-solutionism and inclusive tools. *Sustainability*, 16(781). <https://doi.org/10.3390/su16020781>

Celik, I. (2023). Exploring the determinants of artificial intelligence (Ai) literacy: Brecha digital, pensamiento computacional, absorción cognitiva. *Telematics and Informatics*, 83, 102026.

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *Revista Internacional de Tecnología Educativa in Higher Education*, 20(1), 38. <https://link.springer.com/article/10.1186/s41239-023-00408-3>

Chew, E., & Chua, X. N. (2020). Tutor robótico de chino: ¿Personalizar la evaluación de los progresos y la retroalimentación o hacerse cargo de tu trabajo? *On the Horizon*, 28(3), 113-124. <https://doi.org/10.1108/OTH-04-2020-0015>

Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). First-year students AI-competence as a predictor for intended and de facto use of AI-tools for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 18.

Comisión Europea, Dirección General de Educación, Juventud, Deporte y Cultura. (2022). *Directrices éticas sobre el uso de la inteligencia artificial (IA) y los datos en la enseñanza y el aprendizaje para educadores*. Oficina de Publicaciones de la Unión Europea. <https://data.europa.eu/doi/10.2766/153756>

Parlamento Europeo. (2024). Ley de IA de la UE: Primer reglamento sobre inteligencia artificial. Obtenido de <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Parlamento Europeo y Consejo de la Unión Europea. (2024). *Reglamento (UE) 2024/1689, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre inteligencia artificial y se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Ley de Inteligencia Artificial) (Texto pertinente a efectos del EEE)*. Diario Oficial de la Unión Europea, serie L, 1-144.

Emeršič, Ž., Hrastnik, G., Meh Peer, N., & Peer, P. (2025). Estrategias educativas de nivel universitario inspiradas en AIM@VET para la enseñanza de la visión por



ordenador y la biometría. *ROSUS 2025 - Računalniška obdelava slik in njena uporaba v Sloveniji 2025: Zbornik 19. strokovne konference*. <https://doi.org/10.18690/um.feri.2.2025.4>

Hodgkin, E., & Kumar, A. (2023). Equidad algorítmica en los sistemas de orientación profesional: A comparative analysis of implementation frameworks. *Educational Technology Research and Development*, 71(3), 215-231.

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education: Promises and implications for teaching and learning. Centro de Rediseño Curricular

JISC (2022). *AI in tertiary Education. A summary of the current state of play*. Repositorio JISC. Obtenido de <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020). Automated personalized feedback improves learning gains in an intelligent tutoring system. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6-10, 2020, Proceedings, Part II 21* (pp. 140-146). Springer International Publishing.

Kumar, V., & Boulanger, D. (2020, octubre). Explainable automated essay scoring: El aprendizaje profundo realmente tiene valor pedagógico. En *Frontiers in education* (Vol. 5, p. 572367). Frontiers Media SA.

Li, Q. (2021). The use of artificial intelligence combined with cloud computing in the design of education information management platform. *International Journal of Emerging Technologies in Learning*, 16(5), 32-44. <https://doi.org/10.3991/ijet.v16i05.20309>

Porter, B., y Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sostenibilidad*, 12(24), 10243.

Prinsloo, P., y Knox, J. (2023). Exploración de los principios éticos para la aplicación de la inteligencia artificial en la educación: Towards a future agenda. *Research and Practice in Technology Enhanced Learning*, 18(1), 73-94. Obtenido de https://www.researchgate.net/publication/368740229_Exploring-the-Ethical-Principles-for-the-Implementation-of-Artificial-Intelligence-in-Education-Towards-a-Future-Agenda



Ramírez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., & Miranda, J. (2023). Inteligencia artificial en la formación profesional: Retos éticos y marcos de gobernanza. *International Journal of Educational Technology in Higher Education*, 20(1), 138-155. <https://doi.org/10.1186/s41239-023-00408-3>

Rodríguez-Abitia, G., y Bribiesca-Correa, G. (2021). Evaluando el impacto ético de la inteligencia artificial en la educación. *RED. Revista de Educación a Distancia*, 6(1). <https://doi.org/10.34627/redvol6iss1e202303>

Roman Etxebarrieta, G., Orcasitas-Vicandi, M., & Antzaka, A. (2024). *Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)*. AIPioneers.org. Obtenido de: <https://aipioneers.org/evaluation-schema-for-ai-in-education-on-data-privacy-ethics-and-eu-values-wp5/>

Roppertz, S. (2020). Artificial Intelligence and Vocational Education and Training- Perspective of German Vet Teachers. En *European Distance and E-Learning Network (EDEN) Conference Proceedings* (No. 2, pp. 207-216). European Distance and E-Learning Network.

Rott, K. J., Lao, L., Petridou, E., & Schmidt-Hertha, B. (2022). Necesidades y requisitos de una cualificación adicional de IA durante la formación profesional dual: Results from studies of apprentices and teachers. *Computers and education: Artificial intelligence*, 3, 100102.

Şenocak, D., Bozkurt, A., & Koçdar, S. (2024). Exploring the Ethical Principles for the Implementation of Artificial Intelligence in Education: Towards a Future Agenda. En *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 200-213). IGI Global.

Seufert, S. (2024). Artificial Intelligence in Vocational Education and Training (VET): Evaluating VET Leaders' Acceptance of AI in Switzerland.

Tang, J., & Hai, L. (2021). Construction and exploration of an intelligent evaluation system for educational app through artificial intelligence technology. *International Journal of Emerging Technologies in Learning*, 16(5), 17-31. <https://doi.org/10.3991/ijet.v16i05.20293>

Tuomi, I., Cachia, R., & Villar-Onrubia, D. (2023). Sobre el futuro de la tecnología en la educación: Tendencias emergentes e implicaciones políticas. *Oficina de Publicaciones de la Unión Europea, Luxemburgo*.

UNESCO (2023a). *ChatGPT e inteligencia artificial en la enseñanza superior*. Obtenido de: <https://www.iesalc.unesco.org/wp->



[content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf](#)

UNESCO (2023b). *Orientaciones para la IA generativa en la educación y la investigación*. Obtenido de:

<https://unesdoc.unesco.org/ark:/48223/pf0000386693>

Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot para mejorar el aprendizaje de la puntuación en español y potenciar entornos de aprendizaje abiertos y flexibles. *Revista Internacional de Tecnología Educativa en la Educación Superior*, 18, 1-20.

Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S., & Roman-Cañizares, M. (2021). Implementación de un asistente virtual para la gestión académica de una universidad con el uso de inteligencia artificial. *Internet del Futuro*, 13(4), 97.

Williamson, B., y Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235.

Williamson, B., Macgilchrist, F., y Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1), 1-5. <https://doi.org/10.1080/17439884.2023.2167830>

Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology*, 95-116.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2023). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00408-3>

CONTACTO

www.aipioneers.org

gorka.roman@ehu.es

maria.orcasitas@ehu.es

natalia.louleli@ehu.eus



AI PIONEERS

Artificial Intelligence in education & training

Manuale sulle Politiche e sull'Etica dell'IA nell'Istruzione

Giugno 2025

Curatore principale del deliverable: Università dei Paesi Baschi/Euskal
Herriko Unibertsitatea (UPV/EHU)

Work Package 5

Deliverable D5.1 "Manuale sulle Politiche e sull'Etica dell'IA nell'Istruzione"

Numero del progetto: 101087261

Nome del progetto: AI and the future of

Education Acronimo del progetto: AI

Pioneers

Coordinatore del progetto: Università di Brema, Istituto Tecnologia ed
Educazione (ITB)

Call: ERASMUS-EDU-2022-PI-FORWARD

Tema: ERASMUS-EDU-2022-PI-FORWARD-LOT1

Tipo di azione: ERASMUS Lump Sum Grants

Autorità concedente: Agenzia Esecutiva Europea per
l'Istruzione e la Cultura

Data di inizio del progetto: 01/01/2023

Data di fine del progetto: 31/12/2025

Durata del progetto: 36 mesi



Co-funded by
the European Union



AI PIONEERS
Artificial Intelligence in education & training



AI PIONEERS
Artificial Intelligence in education & training

Versione Italiana del Deliverable

Giugno 2025



Manuale sulle Politiche e sull'Etica dell'IA nell'Istruzione

Gorka Roman Etxebarrieta (Autore)

Maria Orcasitas-Vicandi (Autore)

Natalia Louleli (Autore)

Asier León Nuñez (Autore)

Itsaso Biota Piñeiro (Autore)

Ana Boyano Murillo (Autore)

Raquel Justo Blanco (Autore)

Josune Rodriguez Negro (Autore)



Indice

Manuale sulle Politiche e sull'Etica dell'IA nell'Istruzione	3
1. Introduzione.....	5
1.2 L'importanza delle considerazioni etiche nell'adozione dell'IA	7
1.3 Un manuale per educatori, amministratori e responsabili politici	8
2. Comprendere l'IA nell'Istruzione e Formazione Professionale (VET) e nell'Educazione degli Adulti	10
2.1. Il ruolo dell'IA nell'istruzione.....	11
2.2. L'IA nell'istruzione degli Adulti e nella Formazione Professionale	12
3. Sviluppo di linee guida etiche	13
3.1 L'importanza delle politiche etiche dell'IA nell'istruzione	13
3.2 Requisiti chiave per sistemi di IA affidabili	15
3.3 Considerare gli aspetti legali dell'implementazione dell'IA negli istituti VET	18
3.4 Studi di caso ed esempi di politiche efficaci sull'IA negli istituti di istruzione e formazione professionale	25
4. Competenze Etiche	34
4.1 Definizione di Competenze Etiche	34
4.2 Importanza delle Competenze Etiche.....	35
4.3 Aree delle Competenze Etiche ed Esempi	35
5. Conclusione.....	43
5.1 Punti chiave	44
5.2 Considerazioni finali	45
Ringraziamenti	Errore. Il segnalibro non è definito.



1. Introduzione

Il progetto AI Pioneers si concentra sul potenziamento dell'integrazione dell'Intelligenza Artificiale (IA) nell'istruzione, in particolare nell'Istruzione degli Adulti e nella Formazione Professionale (VET). Questo progetto ERASMUS+ Forward Looking è iniziato a gennaio 2023 e sarà finanziato fino a dicembre 2025.

Con l'influenza crescente dell'IA nell'istruzione e l'aumento della domanda di supporto professionale per il suo utilizzo e per lo sviluppo di competenze, il progetto mira a sostenere una rete globale e strutturata di "AI Pioneers" nei paesi partner e oltre. La creazione della Rete di Riferimento degli AI Pioneers rappresenta un elemento chiave, rivolto a educatori, policy maker, pianificatori dell'istruzione, sviluppatori di IA e altri stakeholder rilevanti. Coinvolgendo coloro che sono maggiormente interessati all'uso dell'IA nell'istruzione, il progetto intende costruire un hub centrale per supportare l'integrazione dell'IA nell'Istruzione degli Adulti e nella VET, promuovendo iniziative future e progetti educativi sull'IA. Maggiori informazioni sulla rete e su come partecipare sono disponibili sul [sito web del progetto](#).

Un altro obiettivo importante è stato lo sviluppo di un Supplemento al Quadro di Riferimento DigCompEDU (Bekiaridis, 2024). Questo documento definisce le competenze e le abilità necessarie affinché gli educatori possano comprendere, valutare e utilizzare le tecnologie di IA in ambito educativo, con un focus specifico sull'Istruzione degli Adulti e sulla VET. Il supplemento allinea le competenze in IA con le sei aree fondamentali del Quadro DigCompEdu (Redecker, 2017): Coinvolgimento Professionale, Risorse digitali, Insegnamento e apprendimento, Valutazione, Potenziamento degli studenti e Facilitazione della competenza digitale degli studenti. Per ciascuna area, il supplemento:

1. Analizza le applicazioni dell'IA all'interno dell'area
2. Propone attività per sviluppare le competenze pertinenti
3. Suggerisce livelli di progressione per la costruzione delle competenze
4. Identifica le sfide e fornisce strategie per superarle

Il progetto continua a sviluppare e diffondere una serie di risorse tramite l'AI Pioneers Toolkit, disponibile sul sito web del progetto. Questo strumento offre raccomandazioni curriculari per l'integrazione dell'IA nell'istruzione, mette in evidenza i benefici e i rischi dell'uso dell'IA nei contesti educativi, e fornisce linee



guida di buone pratiche per l'utilizzo dell'IA nell'istruzione (Tommasi & Perini, 2024). Include inoltre brevi panoramiche di strumenti destinati a diversi scopi didattici, come assistenti all'apprendimento, strumenti per la creazione di materiali educativi e soluzioni a supporto dell'apprendimento tra pari.

Il progetto sta svolgendo un lavoro importante per affrontare le sfide etiche che l'IA pone in ambito educativo. Questo documento accompagna uno schema di valutazione ed entrambi sono pensati per guidare le istituzioni e gli educatori nell'implementazione dell'IA in modo responsabile, tenendo conto della privacy e delle implicazioni etiche.

Lo *Schema di Valutazione per l'IA nell'Istruzione* si concentra su linee guida pratiche per l'adozione degli strumenti di IA, garantendo che siano in linea con i principi etici e i valori dell'UE. Rappresenta una risorsa preziosa per le istituzioni educative che desiderano prendere decisioni consapevoli sull'integrazione dell'IA.

Infine, questo *Manuale sulle Politiche e sull'Etica dell'IA nell'Istruzione* costituisce una guida più approfondita, offrendo politiche e quadri etici per affrontare le sfide attuali.

1.1 Scopo e ambito del manuale

Il *Manuale sulle Politiche e sull'Etica dell'IA nell'Istruzione* è una risorsa fondamentale per gli educatori, i responsabili politici e le istituzioni che intendono affrontare le complessità dell'IA in ambito educativo. Contribuisce a creare un quadro di riferimento per garantire che le considerazioni etiche siano in primo piano, proteggendo la privacy degli studenti, promuovendo l'equità e incoraggiando un uso responsabile.

Uno dei principali vantaggi del manuale è l'orientamento pratico che offre agli educatori, consentendo loro di utilizzare l'IA in modo consapevole e responsabile, e di aiutare gli studenti a comprenderne capacità, limiti e implicazioni etiche. Questo sarà fondamentale per formare una generazione di apprendenti non solo competenti nell'uso della tecnologia, ma anche consapevoli sul piano etico. Per i responsabili politici, il quadro coerente proposto dal manuale guida gli utenti verso approcci più standardizzati tra le istituzioni, aspetto essenziale per mitigare i rischi legati all'IA, come i bias negli algoritmi o le violazioni della privacy. Il manuale contribuisce inoltre a garantire che l'integrazione dell'IA non avvenga fine a sé stessa, ma sia orientata a valorizzare l'esperienza educativa senza compromettere i principi etici.



1.2 L'importanza delle considerazioni etiche nell'adozione dell'IA

Le considerazioni etiche sono fondamentali nell'adozione dell'IA, soprattutto in ambiti come l'educazione degli adulti e la formazione professionale (VET), dove l'impatto dell'IA su discenti, educatori e istituzioni può essere significativo. Tali considerazioni aiutano a garantire che l'IA venga utilizzata in modo responsabile, inclusivo e trasparente, e che supporti gli obiettivi più ampi dell'istruzione senza compromettere i valori fondamentali.

Tra le principali considerazioni affrontate dallo schema troviamo:

Diversità e Inclusione: i sistemi e gli strumenti di IA devono essere progettati per rispondere alle esigenze di una vasta gamma di studenti, inclusi quelli con background, abilità e bisogni formativi differenti. Nell'educazione degli adulti e nella VET, i partecipanti provengono spesso da esperienze di vita eterogenee, come diversi livelli di istruzione pregressa, fasce d'età ed esperienze lavorative. L'IA deve essere flessibile e inclusiva, assicurando che il suo utilizzo sostenga, e non escluda, alcun gruppo. Chi adotta l'IA in modo etico deve essere consapevole dei modi in cui la tecnologia può perpetuare bias o creare divari digitali, così da poterli prevenire e contrastare.

Trasparenza: affinché l'IA sia adottata in modo etico nei contesti educativi, la trasparenza è essenziale. Le istituzioni devono essere chiare su come funzionano i sistemi di IA, quali dati raccolgono e quali sono i processi decisionali sottostanti. Gli studenti ed educatori devono comprendere come vengono utilizzati gli strumenti di IA in aula o negli ambienti formativi, e in che modo questi possono influenzare i risultati dell'apprendimento. La trasparenza favorisce la fiducia nelle tecnologie di IA e garantisce che vi sia responsabilità nel loro utilizzo.

Privacy and protezione dei dati: le istituzioni che si occupano di educazione degli adulti e VET trattano spesso dati personali sensibili, come profili degli studenti, risultati delle valutazioni e carriera lavorativa. L'uso etico dell'IA in questo contesto deve dare priorità alla privacy e alla protezione dei dati. Le istituzioni devono garantire la conformità alle normative sulla protezione dei dati (come il GDPR) e assicurarsi che le informazioni degli studenti siano raccolte, archiviate e utilizzate in modo responsabile. Il consenso alla raccolta dei dati deve essere informato e volontario, e i discenti devono poter avere il controllo sui propri dati personali.

Sostenibilità: l'IA deve essere integrata in modo coerente con la sostenibilità a lungo termine dei sistemi educativi. Questo significa considerare sia l'impatto



ambientale delle tecnologie di IA, sia il loro ruolo nel promuovere pratiche sostenibili all'interno dei programmi di istruzione e formazione. Nel contesto della VET, l'IA può anche supportare la sostenibilità nei settori industriali, preparando gli studenti con competenze che promuovono tecnologie verdi e pratiche eco-sostenibili.

Competenze di alfabetizzazione all'IA: man mano che l'IA viene integrata nell'istruzione, è fondamentale sviluppare competenze di alfabetizzazione all'IA sia tra gli educatori che tra gli studenti. Nell'educazione degli adulti e nella VET, ciò significa fornire agli studenti le conoscenze e le abilità per comprendere le tecnologie di IA e utilizzarle efficacemente nel proprio percorso professionale. Questo tipo di empowerment permette agli studenti di orientarsi nel mercato del lavoro in evoluzione, evitando che vengano lasciati indietro. Per gli educatori, l'alfabetizzazione all'IA consente di valutare criticamente gli strumenti basati su IA e integrarli in modo da arricchire l'apprendimento, senza compromettere gli standard etici.

Correttezza ed equità: i sistemi di IA devono essere progettati per trattare tutti gli studenti in modo equo, assicurando che nessun gruppo venga svantaggiato da decisioni automatizzate o algoritmi. Nell'educazione degli adulti e nella VET, questo aspetto è particolarmente importante in quanto questi settori spesso si rivolgono a studenti che possono già affrontare sfide sociali o economiche. È fondamentale che l'IA non rafforzi disuguaglianze esistenti né crei nuove barriere all'accesso, ma al contrario promuova opportunità eque per tutti gli studenti.

1.3 Un manuale per educatori, amministratori e responsabili politici

Questo manuale è pensato per educatori, amministratori e policy maker nel contesto dell'adozione dell'IA nell'educazione degli adulti e nella formazione professionale (VET). È noto che l'adozione dell'IA e le relative considerazioni etiche impattano ciascuno di questi gruppi in modi differenti:

Educatori:

Per gli educatori, l'integrazione etica dell'IA nell'educazione degli adulti e nell'IFP consiste nel migliorare l'insegnamento e l'apprendimento senza compromettere i principi di equità, privacy e trasparenza. Gli strumenti basati su IA possono offrire esperienze di apprendimento personalizzate, semplificare attività amministrative e fornire informazioni utili sui progressi degli studenti. Tuttavia, gli educatori



devono assicurarsi che l'IA supporti diversi stili ed esigenze di apprendimento, promuovendo un ambiente didattico inclusivo.

Ci sono molte considerazioni chiave per gli educatori riguardo l'alfabetizzazione all'IA, l'uso etico e la trasparenza con gli studenti:

Alfabetizzazione all'IA: gli educatori devono essere dotati delle competenze necessarie per comprendere, valutare e integrare efficacemente l'IA nelle proprie metodologie didattiche. Questo include la conoscenza del funzionamento degli strumenti di IA e delle loro implicazioni etiche all'interno della classe.

Uso etico: gli educatori devono garantire che l'IA venga utilizzata per potenziare l'apprendimento e non per sostituire l'interazione umana. Devono inoltre essere pronti ad affrontare eventuali problemi legati a bias o discriminazioni derivanti dai sistemi di IA.

Trasparenza con gli studenti: gli educatori dovrebbero essere trasparenti con gli studenti sull'uso dell'IA nel processo di apprendimento e sui dati raccolti. Questo contribuisce a costruire fiducia e garantisce che gli studenti si sentano rispettati nella loro privacy.

Amministratori:

Per gli amministratori delle istituzioni che operano nell'ambito dell'educazione degli adulti e della formazione professionale (VET), l'adozione dell'IA comporta la supervisione dell'integrazione di queste tecnologie nei sistemi dell'istituzione e l'assicurarsi che l'infrastruttura sia adeguata a supportare gli strumenti basati sull'IA. Devono inoltre dare priorità a politiche che garantiscano un uso etico dell'IA, salvaguardando al contempo l'integrità istituzionale.

Esistono numerose considerazioni chiave per gli amministratori che riguardano la sicurezza e la privacy dei dati, l'equità e l'accesso, e la formazione e il supporto continui:

Sicurezza e privacy dei dati: gli amministratori devono garantire l'adozione di solide misure di protezione dei dati, poiché l'IA si basa sulla raccolta di grandi quantità di dati degli studenti. È fondamentale rispettare le normative sulla protezione dei dati, come il GDPR.

Equità e accesso: l'IA dovrebbe essere implementata in modo da garantire pari opportunità a tutti gli studenti, indipendentemente dal loro background socio-economico. Gli amministratori dovrebbero promuovere politiche che impediscano agli strumenti di IA di accentuare disuguaglianze già esistenti.



Formazione e supporto continui: gli amministratori devono destinare risorse alla formazione professionale continua, affinché gli educatori possano rimanere aggiornati sui progressi dell'IA e sulle migliori pratiche etiche.

Policy maker:

I policy maker svolgono un ruolo cruciale nella creazione dei quadri normativi e regolamentari che disciplinano l'uso dell'IA nell'istruzione. Devono assicurarsi che l'adozione dell'IA sia coerente con gli obiettivi educativi più ampi, rispetti i diritti umani e promuova opportunità di apprendimento permanente.

Esistono numerose considerazioni fondamentali per i policy maker, in particolare riguardo alle linee guida e alle normative etiche, alla promozione dell'alfabetizzazione all'IA, e alla garanzia di inclusività e accessibilità:

Linee guida e regolamenti etici: i policy maker devono stabilire standard etici e regolamenti che guidino l'uso dell'IA nell'istruzione. Ciò include la necessità di affrontare tematiche come la trasparenza, la privacy dei dati e l'equità nei sistemi educativi basati sull'IA.

Promozione dell'alfabetizzazione all'IA: i policy maker dovrebbero favorire la diffusione dell'alfabetizzazione all'IA in tutto il sistema educativo, assicurandosi che sia gli educatori che gli studenti dispongano degli strumenti e delle conoscenze necessari per orientarsi tra le tecnologie di IA.

Garanzia di inclusività e accessibilità: è fondamentale che le politiche sull'IA promuovano un'istruzione inclusiva, garantendo che le tecnologie di IA non escludano nessun gruppo. Ciò può includere l'incentivazione alla ricerca e allo sviluppo di soluzioni di IA pensate per rispondere alle esigenze di diversi studenti.

2. Comprendere l'IA nell'Istruzione e Formazione Professionale (VET) e nell'Educazione degli Adulti

L'integrazione dell'Intelligenza Artificiale (IA) nell'istruzione presenta opportunità trasformative ma anche sfide significative. La ricerca evidenzia il potenziale dell'IA nel migliorare l'insegnamento attraverso l'automazione di compiti, l'apprendimento personalizzato e l'innovazione didattica basata sui dati, mentre gli strumenti di IA generativa ampliano le possibilità di un'educazione più interattiva e riflessiva. Tuttavia, le preoccupazioni etiche relative alla privacy dei dati, all'influenza commerciale e al divario digitale richiedono una solida



governance. Nell'ambito dell'educazione degli adulti e della formazione professionale, l'IA non solo influisce sulla progettazione dei curriculum e sulla previsione delle competenze per l'inserimento nel mondo del lavoro, ma richiede anche la promozione della competenza digitale e di un coinvolgimento critico rispetto alle implicazioni sociali dell'IA. Rimane cruciale bilanciare i benefici dell'IA con politiche eque ed etiche, per plasmare ecosistemi educativi inclusivi e orientati al futuro.

2.1. Il ruolo dell'IA nell'istruzione

Un recente rapporto politico pubblicato dal Joint Research Centre (JRC) della Commissione Europea, intitolato [On the Futures of Technology in Education: Emerging Trends and Policy Implications](#), (Tuomi, Cachia & Villar-Onrubia, 2023), offre una panoramica completa su come i progressi tecnologici, in particolare nel campo dell'Intelligenza Artificiale (IA) stiano trasformando le pratiche educative, i processi e le istituzioni. Il documento evidenzia gli impatti significativi e diversificati dell'IA sull'istruzione, influenzando i processi di insegnamento e apprendimento, i metodi di valutazione e le strutture organizzative.

L'IA è vista come una potenziale soluzione per ridurre il carico di lavoro degli insegnanti, in particolare per attività come la correzione dei compiti. Automatizzando questi compiti ripetitivi e dispendiosi in termini di tempo, l'IA può rendere l'insegnamento più attrattivo e consentire agli insegnanti di concentrarsi maggiormente sulle attività didattiche piuttosto che su quelle amministrative (Li, 2021; Tang & Hai, 2021; Villegas-Ch et al., 2021). Questa automazione potrebbe rivoluzionare la valutazione formativa (Kumar & Boulanger, 2020), consentendo di dare feedback continui agli studenti e facilitando la valutazione in tempo reale (Porter & Grippa, 2020) e l'apprendimento personalizzato o adattivo (Kochmar et al., 2020; Wongvorachan et al., 2022).

Inoltre, l'IA generativa e i modelli fondativi sono destinati ad avere un impatto sostanziale sull'istruzione. Tecnologie come la sottotitolazione automatica dei video, la traduzione automatica e il riassunto video offrono opportunità innovative per potenziare la pedagogia (Chew & Chua, 2020; Vazquez-Cano et al., 2021). Queste tecnologie possono contribuire allo sviluppo di "compagni intelligenti" o "partner di apprendimento", spostando l'attenzione dalla memorizzazione meccanica alla metacognizione e alla riflessione. Questo comporta la necessità di una redistribuzione attenta delle responsabilità tra esseri umani e macchine, per stabilire quali attività delegare all'IA.



Va inoltre sottolineato il ruolo dell'IA nella datificazione dell'istruzione, ovvero l'utilizzo su larga scala dei dati relativi alle attività educative per informare e rafforzare nuovi approcci pedagogici, tecnologie didattiche e politiche educative (Williamson et al., 2023). L'obiettivo è registrare i processi di apprendimento, non solo i risultati, per migliorare l'impegno e il successo degli studenti (Porter & Grippa, 2020).

Tuttavia, le preoccupazioni etiche relative al trattamento dei dati personali, al possibile uso improprio, alla dipendenza da infrastrutture in gran parte controllate da entità commerciali e ad altre sfide rappresentano questioni critiche da affrontare (Seufert, 2024; Wongvorachan et al., 2022). L'integrazione dell'IA nell'istruzione solleva interrogativi rilevanti sui modelli di governance e sulla legislazione, rendendo necessaria una revisione delle politiche educative che tenga conto delle potenzialità e dei limiti dell'IA (Attard-Frost et al., 2024; Boninger et al., 2020; Williamson & Eynon, 2020). Dobbiamo anche affrontare l'importanza di valutare criticamente i rischi di esacerbare il divario digitale e il dominio degli interessi commerciali nel settore (Bulathwela et al., 2024; Celik, 2023). Regolare l'innovazione tecnologica per garantire visioni educative diverse e allineare la politica industriale alle teorie dell'apprendimento sono aspetti cruciali da considerare.

Il ruolo dell'IA nell'istruzione è quello di uno strumento trasformativo e multifunzionale che offre opportunità per migliorare l'efficienza, la personalizzazione e l'innovazione nell'insegnamento e nell'apprendimento. Tuttavia, presenta anche sfide significative e considerazioni etiche che devono essere affrontate per garantire un uso equo, efficace e responsabile dell'IA nei contesti educativi.

2.2. L'IA nell'istruzione degli Adulti e nella Formazione Professionale

Integrare l'educazione all'Intelligenza Artificiale (IA) nei curricula non significa solo insegnare i fondamenti dell'IA e le sue applicazioni, ma anche esplorarne gli impatti sociali. Questo include lo sviluppo del pensiero critico riguardo alle capacità, ai limiti e alle dimensioni etiche dell'IA, come ad esempio le preoccupazioni legate ai bias (pregiudizi) e alla privacy (Akgun & Greenhow, 2022; Aldosari, 2020; Attwell et al., 2020; Delcker et al., 2024; Seufert, 2024; Wongvorachan et al., 2022).

Il potenziamento dell'alfabetizzazione digitale è anch'esso essenziale, poiché saper interagire con i sistemi di IA e interpretarli sta diventando una abilità chiave nel mondo del lavoro moderno, anche per gli insegnanti (Bekiaridis & Attwell, 2024; Delcker et al., 2024; Roppertz, 2020; Rott et al., 2022; Seufert, 2024). Con l'evoluzione continua dell'IA, l'attenzione si sta spostando non solo sull'utilizzo



degli strumenti di IA, ma anche sulla loro comprensione e valutazione critica, preparando gli studenti a un futuro in cui l'IA è parte integrante della vita e del lavoro quotidiani (Rott et al., 2022).

Di conseguenza, la preparazione della forza lavoro alle nuove competenze nell'era dell'IA è un altro aspetto cruciale dell'educazione degli adulti e dell'istruzione e formazione professionale (Attwell et al., 2020; Roppertz, 2020; Rott et al., 2022). Il ruolo dell'IA nell'identificare le tendenze emergenti e nel prevedere le competenze richieste in futuro è di grande valore. Queste informazioni permettono ai programmi educativi di adattare i curricula per soddisfare le esigenze in continua evoluzione del mercato del lavoro (Seufert, 2024). La riqualificazione (reskilling) e l'aggiornamento delle competenze (upskilling) diventano temi centrali, con programmi guidati dall'IA che consentono ad adulti e professionisti di acquisire nuove competenze o aggiornare quelle esistenti, mantenendosi così al passo con l'innovazione tecnologica (Attwell et al., 2020; Emeršič et al., 2024; Delcker et al., 2024). Inoltre, la formazione supportata dall'IA garantisce che gli studenti non siano solo tecnicamente preparati, ma anche pronti a lavorare in ambienti integrati con l'IA. Questa preparazione è fondamentale per garantire che la forza lavoro sia pronta ad affrontare le sfide e le opportunità offerte dal campo dell'IA in rapida evoluzione, rendendo l'istruzione degli adulti e la formazione professionale più rilevanti, efficaci e allineate ai futuri scenari occupazionali.

3. Sviluppo di linee guida etiche

3.1 L'importanza delle politiche etiche dell'IA nell'istruzione

Le politiche etiche nell'ambito dell'educazione all'intelligenza artificiale (IA) sono fondamentali per garantire che le tecnologie IA vengano utilizzate in modo equo, trasparente e responsabile, tutelando i diritti e il benessere di studenti e insegnanti. Con l'integrazione crescente dell'IA nei contesti educativi, è cruciale disporre di linee guida etiche chiare, che permettano di gestire i potenziali rischi e di sfruttarne i benefici. Di seguito sono riportati alcuni motivi chiave per cui le politiche etiche sono essenziali nell'uso dell'IA nell'istruzione:

Tutela dei diritti e della privacy degli studenti

I sistemi di IA utilizzati in ambito educativo richiedono spesso la raccolta e l'elaborazione di dati personali, come registri scolastici, comportamenti di apprendimento e altre informazioni sensibili. Le politiche etiche garantiscono che questi dati vengano trattati con rispetto e in conformità con le normative sulla



privacy (ad es. il GDPR). Tali politiche aiutano a proteggere la privacy degli studenti assicurando trasparenza nella raccolta dei dati, controllo da parte degli studenti sulle proprie informazioni e utilizzo responsabile e sicuro dei dati raccolti.

Garantire la correttezza e l'equità

Una delle preoccupazioni etiche più critiche legate all'uso dell'IA nell'istruzione è la possibilità che gli algoritmi siano influenzati da bias, svantaggiando determinati gruppi di studenti. I sistemi di IA possono involontariamente perpetuare disuguaglianze esistenti basate su razza, genere, status socio-economico o altri fattori. Le politiche etiche sono quindi necessarie per guidare lo sviluppo e l'implementazione di strumenti di IA che siano inclusivi ed equi, garantendo pari opportunità a tutti gli studenti, indipendentemente dal loro background o dalle loro caratteristiche personali.

Promuovere trasparenza e responsabilità

Le tecnologie basate sull'IA possono essere complesse e opache, rendendo difficile per insegnanti, studenti e amministratori comprendere appieno come vengono prese le decisioni. Le politiche etiche aiutano a garantire che i sistemi di IA siano trasparenti, ovvero che gli stakeholder sappiano come funzionano gli strumenti, quali dati vengono utilizzati e come vengono prese le decisioni. Questa trasparenza favorisce la fiducia e la responsabilità, permettendo a educatori e studenti di comprendere e interagire meglio con le tecnologie IA.

Sostenere l'alfabetizzazione all'IA e l'uso responsabile

Con l'integrazione crescente dell'IA nei processi di apprendimento, è fondamentale che educatori e studenti comprendano il funzionamento di tali tecnologie e le loro implicazioni. Le politiche etiche dovrebbero promuovere **l'alfabetizzazione all'IA**, assicurando che tutti siano dotati delle competenze necessarie per utilizzare l'IA in modo critico e responsabile. Questa comprensione consente di sfruttare al meglio i benefici dell'IA, essendo al contempo consapevoli delle sue limitazioni e dei potenziali rischi.

Promuovere uno sviluppo e un'innovazione etica

Le tecnologie IA nell'educazione devono essere sviluppate su solide basi etiche per garantire che siano utilizzate a beneficio di tutti gli studenti. Le politiche etiche offrono un quadro di riferimento per sviluppatori, ricercatori e istituzioni educative affinché progettino sistemi di IA coerenti con valori educativi quali equità, responsabilizzazione e rispetto della diversità. Questo garantisce anche che le



tecnologie di IA siano ideate e utilizzate in modi che supportino l'apprendimento permanente e gli obiettivi più ampi dell'istruzione.

Affrontare il rischio di sostituzione

Poiché i sistemi di IA automatizzano alcuni aspetti dell'educazione (come la valutazione, il tutoraggio, i compiti amministrativi), esiste il rischio che possano sostituire gli educatori o limitare l'interazione umana. Le politiche etiche possono guidare un uso responsabile dell'IA, garantendo che l'automazione integri il lavoro umano, anziché sostituirlo. Le politiche possono incoraggiare l'uso dell'IA per migliorare l'efficacia didattica e l'efficienza amministrativa, mantenendo comunque il valore della presenza umana nell'esperienza educativa.

Costruire fiducia nei sistemi di IA

Affinché l'IA venga adottata con successo nei contesti educativi, è necessario che educatori, studenti e amministratori si fidino della tecnologia. Le politiche etiche svolgono un ruolo fondamentale nel creare e mantenere questa fiducia, fornendo linee guida chiare sull'uso dell'IA e sugli standard etici da rispettare. Quando i sistemi di IA sono sviluppati e implementati in conformità con tali politiche, è più probabile che gli stakeholder li accettino e li utilizzino efficacemente.

Guidare il futuro dell'IA nell'istruzione

Poiché l'IA continua a evolversi, è essenziale che le politiche etiche offrano una guida a lungo termine per la sua integrazione nei sistemi educativi. Queste politiche contribuiscono a creare un quadro per lo sviluppo futuro delle tecnologie di IA che sia allineato ai valori della società, assicurando che l'IA venga utilizzata per migliorare i risultati dell'apprendimento, promuovere l'innovazione e sostenere l'apprendimento permanente in modo vantaggioso ed etico.

In generale, le politiche etiche nell'istruzione che fa uso dell'IA sono necessarie per guidare lo sviluppo e l'uso responsabile delle tecnologie dell'intelligenza artificiale. Tali politiche garantiscono che l'IA venga impiegata nel rispetto della privacy degli studenti, promuovendo l'equità, incoraggiando la trasparenza e costruendo fiducia. Inoltre, aiutano educatori, studenti e istituzioni a gestire le sfide etiche che emergono man mano che l'IA trasforma gli ambienti educativi. Attraverso l'adozione di queste linee guida etiche, i sistemi educativi possono sfruttare il potere dell'IA salvaguardando i valori fondamentali di un'educazione efficace e inclusiva.

3.2 Requisiti chiave per sistemi di IA affidabili



L'integrazione dell'Intelligenza Artificiale (IA) nell'istruzione offre enormi opportunità per migliorare l'insegnamento, l'apprendimento e l'amministrazione. Tuttavia, per garantire un uso etico e responsabile dell'IA, le istituzioni educative devono sviluppare linee guida chiare per l'implementazione etica dei sistemi di IA. Numerosi quadri di riferimento, linee guida e strumenti di audit pubblicati (ad esempio l'AI Audit dell'UNESCO 2023a e la Guidance for Generative AI in Education and Research 2023b) offrono preziosi spunti su cosa renda un sistema di IA affidabile. Tali quadri sottolineano principi fondamentali **come trasparenza, equità, responsabilità e rispetto dei diritti umani**. Tuttavia, per le istituzioni, sviluppare e applicare questi principi può risultare impegnativo, soprattutto in un contesto tecnologico in rapida evoluzione.

Nel contesto del progetto AI Pioneers, è stato sviluppato uno schema di valutazione (Roman Etxebarrieta et al., 2024) pensato per semplificare e contestualizzare tali principi nell'ambito educativo. Il nostro schema fornisce un approccio strutturato per aiutare le istituzioni a orientarsi nella complessità dell'implementazione dell'IA, garantendo al contempo il rispetto degli standard etici. Pur esistendo diverse linee guida, lo schema si propone come uno strumento pratico e accessibile per supportare le istituzioni educative nello sviluppo di proprie linee guida etiche sull'IA, tenendo conto del proprio contesto e delle proprie necessità specifiche.

Uno degli elementi centrali dell'AI Act dell'UE (Parlamento Europeo e Consiglio dell'Unione Europea, 2024) è che la responsabilità primaria per lo sviluppo etico dei sistemi di IA ricade su sviluppatori e aziende produttrici, e che le linee guida da seguire variano in base al livello di rischio associato al sistema di IA. Tuttavia, i centri educativi sono responsabili della corretta implementazione e dell'uso etico di tali sistemi nel proprio ambiente. È fondamentale che i centri siano consapevoli delle principali questioni etiche per comprendere sia i limiti sia le potenzialità dei sistemi che adottano. Questo implica non solo il rispetto delle normative vigenti, ma anche la capacità di anticipare e affrontare le sfide che possono emergere con l'evoluzione delle tecnologie.

In particolare, il nostro schema di valutazione si concentra sull'identificazione delle problematiche etiche e sulla costruzione di politiche o strategie in grado di supportare e promuovere l'uso dell'IA nei centri educativi nel rispetto delle linee guida etiche fondamentali. Queste linee guida includono: **diversità ed equità, trasparenza e responsabilità, protezione dei dati e privacy, sicurezza e protezione, sostenibilità e benessere sociale, valorizzazione di docenti e studenti, partecipazione democratica nella pianificazione educativa e nelle pratiche di IA, autonomia, progettazione etica (relativa a sostenibilità e**



accessibilità), commercializzazione (Şenocak et al., 2024), **agency di docenti e studenti e alfabetizzazione all'IA.**

Lo schema di valutazione si articola in diverse sezioni, basate sulle dimensioni proposte da Chan (2023), adattate con la revisione di Şenocak et al. (2024) e le linee guida etiche della Commissione Europea (2022), ciascuna focalizzata su un diverso aspetto dell'implementazione dell'IA nei contesti educativi. Queste sezioni guidano le istituzioni nello sviluppo di quadri etici e nell'uso responsabile dell'IA. La prima sezione si concentra sulla **Valutazione dell'uso attuale dell'IA e delle preoccupazioni etiche**, aiutando le istituzioni a valutare il proprio livello di maturità in materia di IA, per stimare il grado di preparazione del centro educativo nell'implementare o ampliare l'uso dell'IA (JISC, 2022).

La seconda sezione riguarda la **Governance e monitoraggio dell'IA**, e affronta questioni legate alla conformità alle normative regionali e internazionali, in particolare in materia di privacy, protezione dei dati ed equità. Affronta inoltre questioni etiche complesse come la non discriminazione e la trasparenza, anch'esse fondamentali per un'IA etica. La terza sezione, **Considerazioni operative**, fornisce indicazioni sulla formazione di insegnanti, studenti e personale, promuove l'alfabetizzazione all'IA, garantisce la supervisione umana e sostiene la partecipazione democratica alle decisioni relative all'IA. Assicurarsi che tutti gli utenti siano adeguatamente formati e consapevoli delle capacità e dei limiti dell'IA è essenziale per una buona implementazione e garantisce parità di accesso.

La quarta sezione è dedicata alle **Considerazioni pedagogiche**, e prende in esame questioni rilevanti come i dilemmi etici nella valutazione tramite IA, l'agency degli studenti e lo sviluppo delle competenze, con particolare riferimento all'Istruzione e Formazione Professionale (VET) e all'educazione degli adulti (Attwell et al., 2021; UNESCO, 2019). La quinta e ultima sezione riguarda **Altre considerazioni: sostenibilità e progettazione etica**, offrendo orientamenti per affrontare gli impatti a lungo termine e garantire un uso accessibile dell'IA.

Lo schema di valutazione fornisce uno strumento pratico alle istituzioni educative per orientarsi nell'implementazione dell'IA, identificare le sfide etiche e sviluppare politiche allineate agli standard internazionali. Per informazioni più dettagliate, consultare lo schema di valutazione completo [Schema di valutazione per l'IA nell'istruzione su dati, privacy, etica e valori dell'UE (WP5)] e la presentazione di Genially, disponibili sul [sito web di AI Pioneers](#).



3.3 Considerare gli aspetti legali dell'implementazione dell'IA negli istituti VET

3.3.1 Il quadro normativo dell'AI Act dell'UE

L'AI Act dell'Unione Europea stabilisce un quadro normativo completo che classifica i sistemi di intelligenza artificiale in base ai potenziali rischi, fornendo indicazioni fondamentali per le istituzioni dell'istruzione e formazione professionale (VET) che intendono implementare queste tecnologie. Questo approccio basato sul rischio garantisce un controllo normativo proporzionato, promuovendo al contempo l'innovazione nei contesti educativi (Parlamento Europeo, 2024).

Come illustrato da Zawacki-Richter et al. (2023) nella loro revisione sistematica sulle considerazioni etiche nell'educazione potenziata dall'IA, i quadri regolatori devono bilanciare l'innovazione tecnologica con la tutela dei diritti fondamentali. La loro analisi identifica l'AI Act dell'UE come un tentativo pionieristico di stabilire "parametri chiari per un'implementazione responsabile dell'IA negli ambienti di apprendimento, riconoscendone al contempo il potenziale trasformativo" (p. 8).

Classificazione del rischio per l'IA educativa

I sistemi di IA utilizzati negli ambienti VET ricadono prevalentemente nella categoria ad alto rischio, a causa del loro impatto significativo sull'esperienza educativa degli studenti e sulle loro opportunità professionali future. L'AI Act dell'UE suddivide i sistemi di IA come segue:

Rischio inaccettabile: i sistemi che pongono chiare minacce ai diritti fondamentali o alla sicurezza sono proibiti in modo assoluto

Alto rischio: applicazioni in settori critici, tra cui l'istruzione, che richiedono una rigorosa conformità normativa

Rischio limitato: sistemi con obblighi di trasparenza che richiedono la divulgazione dell'interazione con l'IA

Rischio minimo: applicazioni con requisiti normativi limitati, regolati da quadri giuridici già esistenti

Implicazioni della classificazione ad alto rischio per le istituzioni VET:

- Obbligo di valutazione del rischio e adozione di protocolli di mitigazione



- Requisiti rigorosi di documentazione
- Presenza di meccanismi di supervisione umana
- Obblighi di trasparenza verso gli stakeholder coinvolti
- Valutazioni regolari di conformità

La ricerca di Ramirez-Montoya et al. (2023) sottolinea che la classificazione ad alto rischio dei sistemi di IA nell'istruzione richiede "strutture di governance rigorose che vadano oltre la semplice conformità tecnica, affrontando questioni sociali più ampie" (p. 142). La loro analisi delle sfide di implementazione in vari istituti VET europei evidenzia la necessità di un'interpretazione contestualizzata dei requisiti normativi.

3.3.2 Considerazioni legali chiave per l'implementazione dell'IA nella VET

Requisiti di trasparenza e responsabilità

L'AI Act dell'Unione Europea impone requisiti rigorosi di trasparenza e responsabilità per l'uso dell'intelligenza artificiale nei contesti educativi, inclusi quelli della formazione professionale (VET). In particolare:

- Le istituzioni VET devono informare chiaramente studenti, docenti e altri stakeholder quando vengono impiegati sistemi di IA in attività come valutazioni, raccomandazioni per l'apprendimento o decisioni amministrative
- La documentazione relativa all'IA deve descrivere in modo dettagliato lo scopo del sistema, il suo funzionamento e i suoi limiti
- Qualsiasi decisione automatizzata che abbia impatto sulla progressione dello studente o sul conseguimento di certificazioni deve essere comprensibile e giustificabile
- I processi algoritmici devono essere verificabili per consentire la convalida dei risultati

Prinsloo e Knox (2023) identificano la trasparenza come un principio etico fondamentale nell'implementazione dell'IA, osservando che "l'opacità algoritmica presenta sfide uniche nei contesti di apprendimento dove la fiducia tra educatori e studenti costituisce la base di una pedagogia efficace" (p. 76). Il loro quadro di riferimento per la trasparenza dell'IA educativa va oltre la mera divulgazione tecnica, includendo una comprensione contestuale accessibile a una varietà di soggetti interessati.



Protezione dei Dati e Conformità alla Privacy

L'implementazione dell'IA nei contesti VET richiede una rigorosa aderenza ai quadri normativi in materia di protezione dei dati:

- Tutti i sistemi di IA devono rispettare le disposizioni del Regolamento Generale sulla Protezione dei Dati (GDPR)
- Devono essere stabiliti meccanismi chiari di consenso per il trattamento dei dati personali
- I principi di minimizzazione dei dati devono governare le pratiche di raccolta e conservazione dei dati
- Il principio di limitazione delle finalità garantisce che i dati vengano utilizzati esclusivamente per obiettivi educativi specifici
- Misure di sicurezza tecniche e organizzative devono tutelare l'integrità e la riservatezza dei dati degli studenti.

Holmes et al. (2023) osservano che le considerazioni sulla protezione dei dati nei contesti educativi vanno oltre la semplice conformità normativa, includendo "questioni etiche più ampie riguardanti l'autonomia degli studenti, le dinamiche di potere e il potenziale per la sorveglianza educativa" (p. 12). Il loro lavoro sottolinea l'importanza di contestualizzare i quadri normativi sulla protezione dei dati all'interno delle specificità etiche del settore educativo.

Prevenzione dei bias e garanzia di equità

Il Regolamento sull'Intelligenza Artificiale dell'UE stabilisce requisiti espliciti per mitigare i bias algoritmici nei contesti educativi:

- Le istituzioni di VET devono condurre valutazioni regolari dei bias nei sistemi di IA per prevenire discriminazioni
- **Assess** Le metodologie di valutazione devono analizzare i potenziali impatti differenziati tra gruppi demografici
- I dati di addestramento degli algoritmi devono essere rappresentativi della diversità studentesca
- Devono essere attivati protocolli di monitoraggio continuo per identificare e correggere i bias emergenti nei sistemi operativi
- È richiesta una revisione regolare dei risultati tra le popolazioni studentesche per verificare l'equità degli effetti.



La ricerca di Hodgkin e Kumar (2023) sull'analisi dei bias nei sistemi di orientamento professionale evidenzia la necessità di “metodologie sistematiche per valutare l'impatto differenziato tra popolazioni studentesche eterogenee” (p. 217) come elemento essenziale per una implementazione equa dell'IA. Il loro quadro di riferimento per una valutazione completa dei pregiudizi fornisce una guida pratica per gli istituti di formazione professionale che si trovano ad affrontare i requisiti di equità.

L'IA negli strumenti utilizzati in classe (rischio limitato)

Numerose applicazioni di AI come: assistenti al tutoraggio dotati di intelligenza artificiale, sistemi di rilevamento del plagio, motori di raccomandazione per le risorse didattiche, strumenti di automazione amministrativa e simulazioni di laboratori virtuali nell'ambito VET possono rientrare nella classificazione di rischio limitato:

Queste applicazioni devono fornire agli utenti informazioni chiare su:

- La natura dell'IA del sistema
- Limitazioni delle capacità algoritmiche
- Disposizioni di supervisione umana
- Meccanismi di contestazione dei risultati

3.3.3 Timeline di attuazione e pianificazione della conformità

Gli istituti VET devono sviluppare piani di attuazione strategici allineati con le tempistiche di applicazione graduale dell'AI Act dell'UE:

2 febbraio 2025: Divieto dei sistemi di IA che comportano rischi inaccettabili

2 agosto 2025: Applicazione dei codici di buona condotta e designazione delle autorità nazionali di regolamentazione

2 agosto 2026: Applicazione generale dell'AI Act, comprese le norme per i sistemi di IA ad alto rischio

2 agosto 2027: Piena applicazione degli obblighi specifici per i sistemi di IA ad alto rischio

Una pianificazione prudente della conformità richiede:

- Inventario completo delle applicazioni di IA esistenti e pianificate



- Valutazione della classificazione del rischio del portfolio tecnologico educativo
- Analisi delle lacune rispetto ai requisiti normativi
- Tabella di marcia strutturata per l'attuazione con obiettivi intermedi
- Allocazione delle risorse per la documentazione e il monitoraggio della conformità

3.3.4 Responsabilità legali e considerazioni sulle responsabilità

Gli istituti VET che implementano sistemi di IA si assumono importanti responsabilità legali:

Le **strutture di responsabilità** devono assegnare in modo chiaro le responsabilità per la governance dei sistemi di IA

I **quadri di responsabilità** devono affrontare i potenziali danni derivanti da errori o bias algoritmici

Gli **accordi contrattuali** con i fornitori tecnologici devono delineare esplicitamente l'allocazione delle responsabilità

I **protocolli di indennizzo** devono affrontare i potenziali reclami derivanti da decisioni influenzate dall'IA

Le **considerazioni assicurative** dovrebbero riflettere le nuove esposizioni al rischio di responsabilità

Le istituzioni educative devono essere consapevoli che l'adozione di sistemi di IA, anche se acquistati da terze parti, non le esonera dalla responsabilità legale per gli esiti che incidono sugli studenti.

Rodriguez-Abitia e Bribiesca-Correa (2021) identificano le considerazioni sulla responsabilità come una dimensione spesso trascurata nella governance dell'IA educativa, osservando che “le istituzioni educative affrontano esposizioni legali specifiche quando delegano a sistemi algoritmici funzioni tradizionalmente svolte da esseri umani” (p. 94). La loro analisi della giurisprudenza emergente fornisce precedenti utili per le istituzioni VET che intendono sviluppare quadri di responsabilità completi.

3.3.5 Lista di valutazione per un'IA affidabile (ALTAI) nell'implementazione dell'istruzione e della formazione professionale



La Lista di valutazione per un'Intelligenza Artificiale affidabile (ALTAI) fornisce agli istituti di istruzione e formazione professionale un quadro strutturato per l'implementazione etica e la conformità normativa. Questo strumento pratico rende operativi i Principi Etici per un'IA Affidabile sviluppati dal Gruppo dell'UE di esperti di alto livello sull'intelligenza artificiale.

Nella loro analisi comparativa dei quadri di valutazione etica, Zawacki-Richter et al. (2023) individuano ALTAI come particolarmente utile nei contesti educativi, grazie alla sua "copertura completa degli aspetti sia tecnici che socio-etici rilevanti per gli ambienti di apprendimento" (p. 14). La loro ricerca dimostra l'adattabilità di ALTAI a contesti educativi diversi, inclusa la formazione professionale.

Componenti chiave di ALTAI rilevanti per l'implementazione dell'istruzione e della formazione professionale

ALTAI struttura la valutazione su sette dimensioni fondamentali, particolarmente rilevanti nei contesti di formazione professionale:

Agenzia umana e supervisione

- Le applicazioni VET devono supportare l'autonomia degli insegnanti, senza sostituire il giudizio professionale
- L'autonomia degli studenti deve essere tutelata attraverso il consenso informato e la trasparenza nelle decisioni
- Deve essere garantita una supervisione umana significativa nelle decisioni che incidono sul percorso formativo degli studenti

Robustezza Tecnica e Sicurezza

- I sistemi di intelligenza artificiale devono funzionare in modo affidabile in diversi ambienti di formazione professionale
- I protocolli di sicurezza devono proteggere da potenziali manipolazioni o accessi non autorizzati
- Procedure di emergenza devono assicurare la continuità didattica in caso di malfunzionamenti del sistema

Privacy e Governance dei Dati



- I dati degli studenti utilizzati per l'addestramento e il funzionamento dei sistemi di IA devono essere protetti attraverso una governance solida
- I principi di minimizzazione dei dati devono guidare le pratiche di raccolta e conservazione
- Misure di sicurezza robuste devono salvaguardare le informazioni sensibili da accessi non autorizzati

Trasparenza

- Le capacità e i limiti dell'IA devono essere comunicati in modo chiaro agli stakeholder del settore educativo
- I processi decisionali devono essere spiegabili in un linguaggio non tecnico
- La documentazione deve consentire la verifica delle prestazioni e della conformità del sistema

Diversità, Non Discriminazione ed Equità

- Le implementazioni nelle VET devono tenere conto della diversità degli studenti e delle loro esigenze formative.
- Le metodologie di valutazione devono esaminare l'impatto equitativo attraverso le dimensioni demografiche
- Le strategie di mitigazione dei bias devono affrontare le specificità della formazione professionale

Benessere Sociale e Ambientale

- Devono essere valutati sistematicamente gli impatti sociali ampi dell'IA nella formazione professionale
- Devono essere considerati gli effetti ambientali dell'infrastruttura computazionale
- Le implicazioni sul mercato del lavoro legate alla certificazione di competenze tramite IA devono essere analizzate

Responsabilità



- Devono essere stabilite strutture di governance chiare che identifichino chi è responsabile dei sistemi di IA
- I meccanismi di audit devono consentire la verifica della conformità alle normative
- Devono esistere procedure per affrontare risultati dannosi o conseguenze impreviste

Application of ALTAI in VET Contexts

Le istituzioni dell'Istruzione e Formazione Professionale (IfFP) possono mettere in pratica ALTAI attraverso protocolli di valutazione sistematici:

La valutazione iniziale dell'implementazione valuta le applicazioni di IA proposte rispetto ai criteri di affidabilità

La verifica prima dell'implementazione assicura la conformità ai requisiti tecnici ed etici

La verifica successiva all'implementazione convalida le prestazioni di diverse popolazioni di studenti

La rivalutazione periodica consente di individuare nuove criticità etiche man mano che i sistemi si evolvono

Integrando ALTAI nei quadri di governance, gli istituti di istruzione e formazione professionale possono dimostrare la dovuta diligenza nella conformità alle normative, rafforzando al contempo la fiducia degli stakeholder nell'implementazione dell'IA.

3.4 Studi di caso ed esempi di politiche efficaci sull'IA negli istituti di istruzione e formazione professionale

3.4.1 Quadri di riferimento delle politiche europee

Le politiche istituzionali efficaci si fondano su quadri di riferimento stabiliti a livello europeo:

Gruppo di esperti ad alto livello sull'intelligenza artificiale (AI HLEG)



L'AI HLEG della Commissione europea ha sviluppato linee guida essenziali tra cui:

Linee guida etiche per un'IA affidabile: definizione di sette requisiti per un'implementazione etica dell'IA

Raccomandazioni sulle politiche e sugli investimenti: Fornisce una direzione strategica per lo sviluppo sostenibile dell'IA

Assessment List for Trustworthy AI: offre strumenti pratici di attuazione per le organizzazioni.

Questi quadri fondamentali hanno orientato politiche istituzionali concrete all'interno degli enti europei di istruzione e formazione professionale.

Come osservano Holmes et al. (2023) nella loro analisi dei quadri di governance dell'IA, le linee guida dell'AI HLEG rappresentano "un ponte cruciale tra principi etici astratti e pratiche concrete di implementazione" (p. 18). La loro ricerca evidenzia l'influenza di tali linee guida in diversi contesti educativi, dimostrandone l'adattabilità anche nell'ambito della formazione professionale.

3.4.2 Copenhagen Technical College: Quadro completo di governance dell'IA

Il Copenhagen Technical College ha istituito una struttura di governance multilivello che rappresenta una pratica esemplare nell'implementazione dell'IA nell'Istruzione e Formazione Professionale (VET).

Architettura delle Politiche e Struttura di Governance

L'approccio dell'istituto prevede:

- Un Comitato Etico sull'IA composto da più stakeholder, tra cui amministratori, docenti, studenti, rappresentanti dell'industria ed esperti tecnici
- Protocolli di approvazione a livelli, basati sulla classificazione del rischio algoritmico
- Formazione obbligatoria sull'IA per tutto personale docente e amministrativo
- Diritti formalizzati degli studenti in merito alla trasparenza degli algoritmi e alla contestabilità delle decisioni



Protocolli di valutazione strutturati che misurano sia le dimensioni educative sia quelle etiche

Questo approccio è in linea con il quadro di riferimento di Prinsloo e Knox (2023) per il coinvolgimento degli stakeholder, che sottolinea l'importanza di una "rappresentanza strutturata attraverso le gerarchie istituzionali e le comunità interessate" (p. 82). La loro ricerca dimostra come la governance multi-stakeholder rafforzi sia il rigore etico che l'efficacia dell'implementazione.

Metodologia di Implementazione

Elementi chiave dell'implementazione includono:

- Introduzione graduale a partire da applicazioni amministrative a basso rischio
- Piloti pedagogici controllati con protocolli di valutazione rigorosi
- Documentazione strutturata sia delle prestazioni tecniche che dell'impatto etico
- Meccanismi regolari di feedback degli stakeholder per informare i miglioramenti iterativi

Verifica indipendente dei risultati algoritmici attraverso le diverse dimensioni demografiche

Risultati documentati

Il College ha riportato benefici significativi attraverso un'implementazione strutturata:

- Riduzione del 27% del carico di lavoro amministrativo grazie all'automazione etica dell'IA
- Miglioramento della personalizzazione educativa senza compromettere la qualità dell'insegnamento
- Maggiore trasparenza nelle pratiche di valutazione
- Aumento della soddisfazione degli studenti per l'integrazione tecnologica
- Conformità efficace ai requisiti normativi emergenti



3.4.3 Bordeaux Institute of Technology: Politiche Trasparenti per la Valutazione con l'IA

L'Istituto di Bordeaux ha sviluppato politiche modello che affrontano uno degli aspetti eticamente più complessi dell'IA educativa: la valutazione degli studenti.

Quadro normativo specifico per la valutazione

Elementi chiave delle politiche adottate:

- Proibizione dell'uso esclusivo di sistemi automatizzati per valutazioni ad alto impatto
- Revisione obbligatoria da parte di un essere umano per ogni decisione valutativa influenzata dall'IA
- Documentazione trasparente del contributo algoritmico alle valutazioni
- Meccanismi strutturati per la contestazione dei risultati influenzati dall'IA
- Verifiche regolari dei bias tramite il confronto dei risultati delle valutazioni tra diverse popolazioni studentesche
- Rappresentanza studentesca all'interno delle strutture di governance delle valutazioni

Questo approccio rappresenta un esempio di ciò che Ramirez-Montoya et al. (2023) definiscono “governance della valutazione centrata sull'essere umano”, caratterizzata da “trasparenza algoritmica, partecipazione degli stakeholder e meccanismi formali di supervisione” (p. 149). La loro ricerca mostra come tali strutture migliorino sia la validità delle valutazioni che la fiducia degli stakeholder.

Approccio di implementazione

La strategia di implementazione enfatizza:

- Chiara delimitazione tra automatismi di valutazione appropriati e inappropriati
- Implementazione graduale in base all'importanza e alle conseguenze della valutazione



- Formazione del corpo docente sulle metodologie etiche di valutazione con IA
- Infrastruttura tecnica che garantisca la spiegabilità delle raccomandazioni algoritmiche
- Protocolli di documentazione che permettano la verifica dell'equità nella valutazione

Impatto misurabile

L'Istituto ha documentato benefici significativi:

- Maggiore coerenza nelle valutazioni, pur mantenendo l'autonomia dei docenti
- Maggiore fiducia degli studenti nei metodi di valutazione
- Riduzione del carico di lavoro legato alla valutazione senza compromettere la qualità
- Migliore identificazione degli studenti che necessitano di ulteriore supporto
- Successo della conformità normativa con i requisiti di trasparenza dell'IA

3.4.4 Consorzio delle Scuole Professionali di Monaco: Governance Collaborativa dei Dati

Otto istituti professionali di Monaco hanno adottato un approccio collaborativo alla governance dei dati sull'IA, dimostrando come la condivisione delle risorse possa migliorare l'implementazione etica.

Architettura di governance condivisa

Elementi strutturali chiave:

- Quadro etico comune per la raccolta e l'utilizzo dei dati
- Infrastruttura tecnica condivisa per la gestione e la sicurezza dei dati
- Protocolli di valutazione e contrattazione collaborativa dei fornitori
- Implementazione congiunta di metodologie di apprendimento federato



- Competenze comuni per la verifica della conformità normativa

Questo approccio collaborativo è in linea con quanto Rodriguez-Abitia e Bribiesca-Correa (2021) definiscono “governance ottimizzata per le risorse”, caratterizzata da “competenze condivise, politiche unificate e capacità collettiva di attuazione” (p. 98). La loro ricerca dimostra come tali approcci possano consentire anche alle istituzioni più piccole di implementare quadri di governance complessi ed efficaci.

Metodologia di implementazione

L'approccio del Consorzio prevede:

- Protocolli standardizzati per la protezione dei dati tra le istituzioni partecipanti
- Competenze centralizzate nelle tecniche di minimizzazione e anonimizzazione dei dati
- Risorse condivise per la valutazione della vulnerabilità tecnica
- Sviluppo collaborativo di metodologie analitiche che tutelino la privacy
- Coinvolgimento congiunto degli stakeholder nello sviluppo della governance

Benefici dimostrati

Questo approccio collaborativo ha prodotto:

- Una governance dei dati più sofisticata di quella che le singole istituzioni potrebbero ottenere autonomamente
- Maggiore potere negoziale nei confronti dei fornitori di tecnologie
- Miglioramento dei protocolli di sicurezza dei dati grazie alla condivisione delle competenze
- Cost-effective regulatory compliance across multiple institutions
- Conformità normativa economicamente vantaggiosa tra più istituzioni
- Mitigazione dei bias più efficace grazie all'uso di dataset più ampi e diversificati



3.4.5 Dublin Technical Institute: Quadro Politico sull'IA Incentrato sull'Equità

Il Dublin Technical Institute ha sviluppato politiche che danno priorità specifica all'equità algoritmica nei contesti dell'istruzione e formazione professionale

Quadro politico centrato sull'equità

Elementi chiave della politica:

- Valutazioni obbligatorie dell'impatto sull'equità prima del dispiegamento di sistemi IA
- Audit algoritmico regolare su tutte le dimensioni demografiche
- Priorità all'uso di metodologie IA spiegabili
- Obbligo di prevedere alternative non tecnologiche ai sistemi IA
- Inclusione formalizzata delle popolazioni sottorappresentate nelle strutture di governance

Questo quadro rappresenta un esempio di ciò che Hodgkin e Kumar (2023) definiscono “governance proattiva dell'equità”, caratterizzata da “valutazione sistematica dei bias, rappresentanza inclusiva degli stakeholder e protocolli di monitoraggio continuo” (p. 219). La loro ricerca dimostra come una governance centrata sull'incentrata migliori sia l'implementazione etica che i risultati educativi.

Strategia di implementazione

L'approccio adottato pone l'accento su:

- Test sistematici dei bias negli algoritmi di orientamento professionale
- Monitoraggio demografico degli esiti educativi influenzati dall'IA
- Sviluppo di metodologie di debiasing per i dati di formazione
- Formazione del corpo docente per un'integrazione dell'IA attenta all'equità
- Feedback regolare degli stakeholder da parte di popolazioni studentesche diverse

Impatto documentato



L'Istituto ha evidenziato:

- Individuazione e correzione di bias nei sistemi di orientamento professionale
- Miglioramento degli esiti educativi per i gruppi storicamente sottorappresentati
- Maggiore fiducia degli stakeholder nell'integrazione tecnologica
- Conformità con i requisiti di non discriminazione
- Sviluppo di metodologie incentrate sull'equità, trasferibili ad altri contesti

3.4.6 Applicazione di ALTAI nei Contesti VET specifici

Lista di valutazione per un'IA affidabile (ALTAI) è stata applicata con successo in numerosi contesti di istruzione e formazione professionale (VET), dimostrando la sua utilità pratica per un'implementazione etica dell'IA.

La ricerca di Zawacki-Richter et al. (2023) identifica l'implementazione dell'ALTAI nella formazione professionale sia particolarmente preziosa a causa dei "contesti professionali concreti che caratterizzano la formazione professionale e delle implicazioni esplicite di carriera delle decisioni influenzate dall'IA" (p. 15). L'analisi di casi studio dimostra l'adattabilità di ALTAI a diversi ambienti di formazione professionale.

Applicazioni di valutazione delle competenze basate sull'IA

Gli istituti di formazione professionale hanno applicato ALTAI per garantire un'implementazione etica dei sistemi di valutazione automatizzati:

Integrazione della supervisione umana: assicurare la convalida delle valutazioni algoritmiche da parte dei docenti

Verifica dell'equità: testare l'equità delle valutazioni tra diverse popolazioni studentesche

Requisiti di trasparenza: rendere i criteri di valutazione espliciti e comprensibili

Robustezza tecnica: Garantire un funzionamento affidabile tra diversi modelli di rendimento degli studenti



L'IA per l'Orientamento Professionale e la Consulenza Formativa

L'applicazione di ALTAI nei sistemi di raccomandazione di carriera ha affrontato:

- Potenziale rafforzamento degli stereotipi professionali
- Trasparenza nelle metodologie di raccomandazione
- La protezione dei dati nei modelli predittivi di carriera
- Accessibilità dei sistemi di orientamento a diverse popolazioni

Apprendimento Adattivo Basato su IA nella Formazione Professionale

L'implementazione di piattaforme di apprendimento adattivo ha tratto beneficio dalla valutazione secondo ALTAI grazie a:

- Valutazione dell'equità della personalizzazione algoritmica
- Verifica della diversificazione dei percorsi di apprendimento
- Analisi dei potenziali effetti "bolla di filtraggio"
- Validazione della capacità di supervisione da parte dei docenti

3.4.7 Elementi comuni delle politiche efficaci

L'analisi delle implementazioni di successo nelle istituzioni VET europee rivela modelli ricorrenti nelle politiche adottate:

Strutture di governance multi-stakeholder che includono prospettive diverse nello sviluppo e nella supervisione

Protocolli di valutazione del rischio a livelli che applicano un controllo proporzionato all'impatto potenziale

Meccanismi di monitoraggio continuo piuttosto che processi di approvazione una tantum

Requisiti di trasparenza solidi, per garantire la comprensione da parte degli stakeholder delle capacità e dei limiti dell'IA

Supervisione umana significativa, soprattutto nelle decisioni educative importanti



Cicli regolari di revisione delle politiche, in linea con l'evoluzione tecnologica e normativa

Componenti formative complete, volte a sviluppare la alfabetizzazione all'IA tra tutti gli stakeholder

Strutture di responsabilità chiare, che definiscono chi è responsabile degli esiti influenzati dall'IA

Holmes et al. (2023) identificano questi elementi costanti come rappresentativi di un "consenso emergente sui principi fondamentali di governance per l'implementazione dell'IA in ambito educativo" (p. 21). La loro analisi comparativa, condotta in contesti istituzionali diversi, dimostra l'adattabilità di questi principi fondamentali a diversi ambienti educativi.

Le istituzioni che mostrano pratiche esemplari sono riuscite a bilanciare il potenziale innovativo con adeguate misure di salvaguardia, riconoscendo che approcci eccessivamente restrittivi possono limitare i benefici educativi, mentre una governance insufficiente comporta rischi significativi. I framework più efficaci hanno integrato le considerazioni etiche in tutte le fasi del ciclo di implementazione, anziché trattarle come un aspetto secondario.

4. Competenze Etiche

4.1 Definizione di Competenze Etiche

Le competenze etiche si riferiscono alla capacità di valutare se l'IA viene utilizzata in modo etico in uno specifico contesto. In ambito educativo, ad esempio, questo significa saper determinare se le politiche di un centro educativo sono eticamente corrette, se la sicurezza dei dati e la privacy degli utenti sono rispettate, e così via.

Le competenze etiche si riferiscono alle capacità critiche necessarie per valutare e garantire l'uso etico dell'IA e dei dati all'interno dei contesti educativi, sulla base di principi quali l'autonomia umana, l'equità, l'umanità e la scelta giustificata. La capacità di agire consapevolmente implica mettere le persone di compiere scelte consapevoli e di assumersi la responsabilità delle proprie azioni, promuovendo così autodeterminazione e responsabilizzazione. L'equità garantisce un trattamento imparziale e un accesso equo alle opportunità per tutti, favorendo l'inclusione e contrastando ogni forma di discriminazione. L'umanità sottolinea l'importanza di rispettare la dignità, l'identità e il benessere delle persone, dando priorità alle relazioni umane significative rispetto all'uso strumentale dei dati. La scelta giustificata richiede processi decisionali



trasparenti, fondati su principi etici e sulla partecipazione degli stakeholder, per garantire che i sistemi di IA nell'istruzione vengano utilizzati in modo responsabile e portino benefici a tutte le parti coinvolte.

4.2 Importanza delle Competenze Etiche

Queste considerazioni etiche sono fondamentali per orientare educatori e istituzioni educative nella gestione delle complessità legate all'implementazione dell'IA. Seguendo i principi etici delineati, ad esempio, dalle Linee guida sull'etica dell'IA, i soggetti coinvolti nell'istruzione possono mitigare efficacemente i rischi, promuovere l'equità e tutelare il benessere di studenti e personale.

4.3 Aree delle Competenze Etiche ed Esempi

Area 1: Impegno Professionale. Uso delle tecnologie digitali per la comunicazione, la collaborazione e lo sviluppo professionale.	
Elemento di competenza	Indicatori potenziali
È in grado di descrivere criticamente gli impatti positivi e negativi dell'uso dell'IA e dei dati nell'istruzione	Partecipa attivamente ad attività di formazione continua sull'intelligenza artificiale, sull'analisi dei dati nell'apprendimento e sul loro uso etico.
	Comprende le nozioni di base relative all'intelligenza artificiale e all'analisi dei dati per l'apprendimento.
	Conoscere le modalità di valutazione dell'impatto etico dei sistemi di IA all'interno della scuola.
	Saper come avviare e promuovere strategie, nella scuola e nella comunità, per un uso etico e responsabile dell'IA e dei dati.
Comprende le nozioni di base relative all'IA e all'analisi dei dati per l'apprendimento	Avere consapevolezza del fatto che gli algoritmi di IA operano spesso in modi non visibili o difficilmente comprensibili dagli utenti.



	Saper interagire con un sistema di IA e fornire feedback per influenzare le raccomandazioni generate.
	Avere consapevolezza che i sensori integrati in molte tecnologie e applicazioni digitali generano grandi quantità di dati, anche personali, che possono essere utilizzati per addestrare sistemi di IA.
	Conoscere le linee guida etiche sull'IA dell'Unione Europea e gli strumenti di autovalutazione disponibili.
Esempio 1: Ana è un'insegnante che vuole preparare una lezione di meccanica dei fluidi e desidera utilizzare un sistema di intelligenza artificiale per generare alcune immagini da inserire nella presentazione PowerPoint della lezione. Prima di integrare l'IA nel percorso didattico, Ana ha consultato le linee guida etiche dell'Unione Europea sull'intelligenza artificiale e ha valutato se l'uso dello strumento sia eticamente appropriato.	
Esempio 2: Lucas insegna nel percorso VET di Educazione e Controllo Ambientale. Poiché diversi docenti del suo istituto utilizzano l'intelligenza artificiale, ha proposto un incontro con tutto il personale educativo per analizzare le modalità di utilizzo dell'IA da parte degli insegnanti. L'obiettivo di Lucas è garantire che tutti i docenti facciano un uso etico e responsabile dell'IA e dei dati.	

Area 2: Risorse Digitali: Ricercare, creare e condividere risorse digitali	
Elemento di competenza	Indicatori potenziali



Governance dei dati	Avere consapevolezza delle diverse tipologie di dati personali utilizzate nei contesti educativi e formativi.
	Riconoscere le responsabilità legate alla sicurezza e alla protezione della privacy dei dati.
	Conoscere il fatto che il trattamento dei dati personali è regolato da normative nazionali e dell'UE, inclusa la General Data Protection Regulation (GDPR).
	Sapere chi ha accesso ai dati degli studenti, come viene monitorato tale accesso e per quanto tempo i dati vengono conservati.
	Conoscere il diritto di ogni cittadino dell'UE a non essere soggetto a processi decisionali completamente automatizzati.
	Saper fornire esempi di dati sensibili, inclusi quelli biometrici.
	Essere in grado di valutare i benefici e i rischi prima di consentire a terze parti di trattare dati personali, in particolare quando si utilizzano sistemi di intelligenza artificiale.
Governance dell'IA	Essere a conoscenza del fatto che i sistemi di intelligenza artificiale sono soggetti a normative nazionali e dell'Unione Europea (in particolare al futuro AI Act).
	Saper descrivere l'approccio basato sul rischio previsto dall'AI Act (in fase di adozione).
	Conosce i casi di utilizzo dell'IA ad alto rischio nell'istruzione e i relativi requisiti ai sensi della legge sull'IA (se adottata).



	Sapere come incorporare i contenuti digitali modificati/manipolati dall'IA nel proprio lavoro e come tale lavoro dovrebbe essere accreditato.
	Saper spiegare i principi fondamentali legati alla qualità dei dati nei sistemi di intelligenza artificiale.
Esempio 1: Maria insegna presso la scuola VET di Efficienza Energetica ed Energia Solare Termica. Sta preparando dei materiali didattici da distribuire agli studenti e ha inserito in questi appunti alcuni contenuti digitali modificati con strumenti di IA, utilizzando Copilot. Maria è consapevole che tali contenuti devono essere correttamente attribuiti e si assicura di indicarne la fonte in modo appropriato.	
Esempio 2: Jaime insegna presso la scuola VET di Produzione Meccanica. Utilizza occasionalmente ChatGPT per sviluppare contenuti o attività legati alla propria materia. A un certo punto valuta la possibilità di includere i dati personali degli studenti (nome, cognome, data di nascita e indirizzo) in un documento. Consapevole della propria responsabilità nella protezione dei dati e del rispetto della privacy, decide di non inserire tali informazioni.	

Area 3: Utilizzo delle tecnologie digitali e di strategie per migliorare la valutazione		
Elemento competenza	di	Indicatori potenziali
Modelli apprendimento	di	Si riconosce che i sistemi di IA riflettono la concezione dell'apprendimento adottata dai progettisti e il modo in cui questo viene misurato; si è in grado di spiegare le ipotesi pedagogiche su cui si basa un dato sistema di apprendimento digitale.
Obiettivi dell'educazione		Si è in grado di analizzare come un sistema digitale affronta i diversi obiettivi sociali dell'educazione (qualificazione, socializzazione, formazione dell'identità personale).



Autonomia responsabilità individuale	e	Si è in grado di riflettere sull'impatto del sistema di IA sull'autonomia dell'insegnante, sullo sviluppo professionale e sull'innovazione educativa.
		Si considerano le fonti di bias inaccettabili nei sistemi di IA basati sui dati.
Equità		Si prendono in considerazione i rischi legati alla dipendenza emotiva e all'autopercezione degli studenti nell'uso di sistemi di IA interattiva e dell'analisi dei dati di apprendimento.
Umanità		Si valuta l'impatto dell'utilizzo di IA e dati sulla comunità studentesca.
		Si è in grado di discutere con consapevolezza gli aspetti etici dell'IA e come questi influenzano l'uso delle tecnologie.
Partecipazione allo sviluppo di pratiche di apprendimento che utilizzano IA e dati		Si è in grado di spiegare come i principi e i valori etici siano presi in considerazione e negoziati nel co-design e nella co-creazione di pratiche didattiche che utilizzano IA e dati (in relazione al design dell'apprendimento).
Esempio 1: Luisa, docente di Manutenzione Elettromeccanica presso un ente di formazione, adotta un software basato su IA che utilizza modelli di apprendimento adattivo per personalizzare l'insegnamento. Il software valuta il livello iniziale di ciascun apprendista e adatta contenuti e compiti in base ai progressi individuali. In questo modo, ogni apprendista può procedere con il proprio ritmo, ricevendo feedback immediati e risorse specifiche per superare eventuali difficoltà. Inoltre, Luisa può consultare report dettagliati per identificare le aree in cui gli studenti necessitano di ulteriore supporto, garantendo così un insegnamento inclusivo ed efficace.		
Esempio 2: Javi, docente di Arti Grafiche nel percorso di formazione professionale di livello medio in Stampa Grafica, integra un software di IA che		



supporta la capacità umana assistendo nelle attività di progettazione più complesse. Il software automatizza compiti ripetitivi, come la correzione dei colori e l'adattamento dell'impaginazione, permettendo agli studenti di concentrarsi sugli aspetti creativi e strategici dei loro progetti. In questo modo, gli studenti possono sviluppare le proprie competenze artistiche e il pensiero critico, mentre l'IA gestisce le componenti tecniche.

Area 4: Valutazione: Gestire e organizzare l'uso delle tecnologie digitali nella didattica

Elemento di competenza	Indicatori potenziali
Differenze individuali	Consapevolezza del fatto che gli studenti reagiscono in modi diversi ai feedback automatizzati.
Bias algoritmico	Capacità di considerare le fonti di bias inaccettabile nei sistemi di IA e i possibili metodi per mitigarli.
Focalizzazione cognitiva	Consapevolezza che i sistemi di IA valutano i progressi degli studenti sulla base di modelli di conoscenza specifici del dominio.
	Consapevolezza che la maggior parte dei sistemi di IA non valuta competenze collaborative, sociali o creative.
	Conoscenza dei modi più comuni per manipolare le valutazioni basate su IA.
Nuove modalità di uso improprio della tecnologia	Consapevolezza del fatto che gli studenti reagiscono in modi diversi ai feedback automatizzati.



Esempio 1: Jorge teaches a subject in the Installations and Maintenance VET school. After finishing the 1st topic of the subject, he uses an AI tool to assess the acquisition of the contents by his students. Furthermore, this AI tool will automatically provide feedback and adapt the level of the questions based on previous answers. Jorge is aware of how students react to automated feedback, so he makes sure to provide them personalised feedback and to check the students' progress himself. Therefore, he uses the AI only as a tool to give an initial assessment, which is checked and revised by him subsequently.

Esempio 2: Natalia è formatrice presso un ente di formazione professionale nel settore delle energie rinnovabili. Durante le lezioni incoraggia gli studenti a utilizzare Windows Copilot per svolgere alcune attività. Sebbene lo strumento possa favorire produttività e creatività, Natalia è consapevole che sono necessarie altre tipologie di attività e valutazione per sviluppare competenze collaborative, sociali e creative, che l'IA non è in grado di valutare. Continua quindi a progettare e gestire personalmente tali attività, utilizzando l'IA solo come strumento di supporto alla didattica e alla preparazione delle lezioni.

Area 5: Potenziare gli studenti: Utilizzare le tecnologie digitali per promuovere l'inclusione, la personalizzazione e il coinvolgimento attivo degli studenti

Elemento di competenza	Indicatori potenziali
Uso dell'IA per rispondere ai diversi bisogni educativi degli studenti	Conoscenza delle diverse modalità con cui i sistemi di apprendimento personalizzato possono adattare il proprio comportamento (contenuti, percorsi di apprendimento, approccio pedagogico).
	Capacità di spiegare in che modo un determinato sistema può risultare vantaggioso per tutti gli studenti, indipendentemente dalle differenze cognitive, culturali, economiche o fisiche.



	Consapevolezza del fatto che i sistemi di apprendimento digitale trattano in modo diverso i vari gruppi di studenti.
	Capacità di considerare l'impatto sullo sviluppo dell'autoefficacia, dell'autopercezione, della mentalità di apprendimento e delle capacità di autoregolazione cognitiva e affettiva degli studenti.
Scelta giustificata	Consapevolezza del fatto che l'uso dell'IA e dei dati può avvantaggiare alcuni studenti più di altri.
	Capacità di spiegare quali evidenze sono state utilizzate per giustificare l'adozione di un determinato sistema di IA in aula.
	Riconoscimento della necessità di monitorare costantemente i risultati derivanti dall'uso dell'IA.
Esempio 1: Teresa, docente di Formazione Professionale Superiore in Amministrazione e Finanza presso un ente per l'istruzione degli adulti, utilizza l'IA per personalizzare le esperienze di apprendimento. Il sistema valuta i bisogni di ciascun studente e adatta risorse e attività di conseguenza. Ad esempio, agli studenti visivi vengono proposti strumenti interattivi, mentre a quelli con un approccio più analitico vengono offerti casi studio dettagliati. Questo approccio consente un apprendimento personalizzato che valorizza i diversi stili cognitivi e ottimizza il coinvolgimento.	
Esempio 2: Óscar, docente di Formazione Professionale Superiore in Progettazione Meccanica, utilizza un software di IA per la progettazione di componenti meccanici. Il programma propone diverse soluzioni progettuali e ne simula il funzionamento in condizioni differenti. Gli studenti analizzano i risultati per scegliere il progetto più efficiente, sviluppando così – con la guida del docente – competenze critiche e analitiche fondamentali per il loro futuro professionale.	



Area 6: Facilitare la competenza digitale degli studenti consentendo loro di utilizzare in modo creativo e responsabile le tecnologie digitali per l'informazione, la comunicazione, la creazione di contenuti, il benessere e la risoluzione di problemi

Elemento di competenza	Indicatori potenziali
Etica dell'IA e dell'analisi dei dati nell'apprendimento	Capacità di utilizzare progetti e applicazioni di IA per aiutare gli studenti a comprendere l'etica dell'IA e dell'uso dei dati nell'istruzione e nella formazione.

Esempio 1: La professoressa Sara insegna analisi di marketing avanzato a futuri docenti della formazione professionale in ambito universitario. Durante il corso, presenta un caso studio in cui l'IA analizza dati provenienti dai social media per finalità pubblicitarie mirate. Sara stimola un dibattito su temi etici quali la privacy, il consenso e i bias algoritmici. Guida gli studenti nello sviluppo di linee guida etiche per l'utilizzo dell'IA nel marketing, promuovendo una comprensione critica e responsabile dell'impiego dell'IA nel settore.

Esempio 2: Il professor Jesús insegna gestione culinaria, con particolare attenzione al ruolo dell'IA nella personalizzazione dei menu nei ristoranti. Presenta un caso studio in cui l'IA analizza le preferenze alimentari per offrire raccomandazioni personalizzate. Jesús coinvolge gli studenti in un dibattito su questioni etiche come la privacy dei dati, la diversità nei menu e l'impatto dell'IA sulla soddisfazione del cliente. Li guida nello sviluppo di quadri etici per l'utilizzo dell'IA nella personalizzazione dei menu, preparandoli a coniugare innovazione e responsabilità etica nella gestione della ristorazione.

5. Conclusione

Con l'evolversi dell'Intelligenza Artificiale nei contesti educativi, è fondamentale che l'adozione di queste tecnologie avvenga nel rispetto dei valori etici, promuovendo l'equità e tutelando la privacy degli studenti. Questo manuale ha illustrato lo scopo e l'ambito dell'implementazione dell'IA, evidenziando l'importanza di considerazioni etiche per garantire un uso responsabile delle tecnologie dell'IA in tutti i contesti educativi.



5.1 Punti chiave

L'integrazione dell'Intelligenza Artificiale nell'istruzione offre un enorme potenziale, ma deve essere guidata da principi etici che mettano al centro la diversità, l'inclusione e la trasparenza. Gli strumenti basati su IA dovrebbero essere progettati per rispondere ai bisogni di studenti provenienti da contesti culturali, linguistici e socio-economici differenti, promuovendo equità nei risultati educativi. La trasparenza è altrettanto essenziale: docenti e studenti devono poter comprendere come i sistemi di IA prendano decisioni, specialmente in ambiti delicati come la valutazione, il feedback o la personalizzazione dei percorsi di apprendimento. La tutela della privacy e la protezione dei dati sono imprescindibili, e richiedono piena conformità a regolamenti come il GDPR e l'adozione di pratiche etiche nella gestione dei dati. Inoltre, i sistemi di IA devono essere sostenibili, riducendo il loro impatto ambientale e garantendo una presenza duratura negli ecosistemi educativi.

È altrettanto importante promuovere l'alfabetizzazione all'IA tra tutti gli attori coinvolti. I docenti hanno bisogno di una formazione mirata per integrare in modo etico ed efficace gli strumenti di IA nella propria didattica, mentre gli studenti devono sviluppare una comprensione critica delle potenzialità e dei limiti di tali tecnologie. Gli amministratori delle istituzioni educative svolgono un ruolo cruciale nell'allineare le politiche interne a pratiche etiche nell'uso dell'IA, promuovendo un ambiente di responsabilità ed equità. I policymaker devono contribuire alla creazione di quadri normativi solidi che guidino un'adozione responsabile dell'IA, attenuino i bias algoritmici e prevengano risultati discriminatori. In definitiva, un approccio multi-stakeholder è fondamentale per garantire che l'IA migliori l'educazione in modo equo, inclusivo e sostenibile.

È inoltre essenziale tenere conto dei principali quadri normativi e regolatori, come l'AI Act dell'UE (2024), che introduce una cornice giuridica articolata, classificando i sistemi di IA in base al rischio che rappresentano per i diritti fondamentali, con implicazioni significative per le istituzioni educative.

In base a tale classificazione, le applicazioni di IA ad alto rischio, come quelle impiegate per la valutazione degli studenti o per le ammissioni, sono soggette a rigorosi requisiti di conformità, tra cui una documentazione dettagliata, la supervisione umana e protocolli solidi di gestione del rischio.

Inoltre, trasparenza e responsabilità non sono opzionali: le istituzioni devono rendere esplicito il funzionamento dei sistemi di IA e garantire che gli utenti possano comprenderne e contestarne le decisioni automatizzate. L'AI Act



prevede anche misure proattive per audit e mitigazione dei bias algoritmici, al fine di assicurare equità e prevenire discriminazioni. A sostegno di questi obiettivi, la Assessment List for Trustworthy AI (ALTAI) offre uno strumento pratico per aiutare le istituzioni a valutare l'impatto etico dei loro sistemi di IA, promuovendo una cultura dell'uso responsabile e affidabile dell'IA nel settore educativo.

5.2 Considerazioni finali

Dobbiamo tutti essere consapevoli che, mentre l'Intelligenza Artificiale continua a rimodellare l'istruzione universitaria e la formazione professionale, questo manuale sottolinea l'importanza di mantenere un attento equilibrio tra pratiche innovative e responsabilità etica. I benefici dell'IA, dai percorsi di apprendimento personalizzati all'efficienza amministrativa, sono significativi ma accompagnati da rischi che non devono essere trascurati. Questioni come i bias algoritmici, le violazioni della privacy dei dati e la progressiva riduzione del ruolo del giudizio umano richiedono un'attenzione costante. Per affrontare queste complessità, è fondamentale una collaborazione continua tra docenti, dirigenti scolastici e decisori politici. I principi etici devono essere profondamente integrati nello sviluppo, nell'adozione e nella valutazione degli strumenti di IA, affinché tali tecnologie promuovano l'inclusione, rispettino l'autonomia e rafforzino la fiducia nei processi educativi.

Guardando al futuro, il successo dell'integrazione dell'IA nell'istruzione non dipenderà soltanto dalla prontezza tecnologica, ma anche da una visione condivisa fondata sulla trasparenza, sull'equità e sull'apprendimento permanente. Il settore dovrà promuovere l'alfabetizzazione all'IA tra tutti gli attori coinvolti e costruire sistemi solidi di responsabilità e supervisione. Non si tratta di un adeguamento una tantum, bensì di un percorso in continua evoluzione, che richiede riflessione costante e dialogo aperto. Adottando approcci centrati sulla persona e un'innovazione etica, la comunità educativa può trasformare l'IA in un alleato prezioso, capace di ampliare le opportunità di apprendimento salvaguardando al contempo i valori di un'educazione di qualità. Per un orientamento più approfondito, si invitano tutti gli interessati a consultare le risorse del progetto AI Pioneers e lo Schema di Valutazione per l'IA nell'Istruzione, disponibili in più lingue.

Ringraziamenti

Desideriamo ringraziare tutte le persone che hanno sostenuto lo sviluppo di questo documento attraverso critiche costruttive e contributi preziosi. Siamo profondamente grati a tutti i partner del progetto, che hanno collaborato attivamente ai contenuti, così come alla Rete di Riferimento di AI Pioneers e ai partecipanti ai nostri seminari di consultazione. Le loro preoccupazioni condivise



e i loro punti di vista hanno rappresentato una guida fondamentale durante tutto il processo.

Infine, riconosciamo l'importanza della letteratura di riferimento, insieme a numerose altre risorse come articoli scientifici, rapporti, seminari, linee guida di istituzioni educative e post di blog redatti da ricercatori e docenti. Nel loro insieme, questi contributi hanno arricchito significativamente il nostro approccio alle considerazioni etiche sull'uso dell'intelligenza artificiale nell'istruzione, a tutti i livelli.

Riferimenti bibliografici

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440.

Aldosari, S. A. M. (2020). The future of higher education in the light of artificial intelligence transformations. *International Journal of Higher Education*, 9(3), 145-151.

Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, 41(2), 101929.

Attwell, G., Bekiaridis, G., Deitmer, L., Perini, M., Roppertz, S., & Tütlys, V. (2020). Artificial intelligence in policies, processes and practices of vocational education and training.

Attwell, G., Deitmer, L., Tütlys, V., Roppertz, S., & Perini, M. (2020). Digitalisation, artificial intelligence and vocational occupations and skills: what are the needs for training teachers and trainers?. In *Trends in vocational education and training research, Vol. III. Proceedings of the European Conference on Educational Research (ECER), Vocational Education and Training Network (VETNET)* (pp. 30-42).

Bekiaridis, G., & Attwell, G. (2024). Integrating Artificial Intelligence in Vocational and Adult Education: A Supplement to the DigCompEdu Framework. *Ubiquity Proceedings*, 4(1).

Boninger, F., Molnar, A., & Saldaña, C. (2020). Big Claims, Little Evidence, Lots of Money: The Reality behind the Summit Learning Program and the Push to Adopt Digital Personalized Learning Programs. *Commercialism in Education Research Unit*.

Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational



inequality, techno-solutionism and inclusive tools. *Sustainability*, 16(781). <https://doi.org/10.3390/su16020781>

Celik, I. (2023). Exploring the determinants of artificial intelligence (Ai) literacy: Digital divide, computational thinking, cognitive absorption. *Telematics and Informatics*, 83, 102026.

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://link.springer.com/article/10.1186/s41239-023-00408-3>

Chew, E., & Chua, X. N. (2020). Robotic Chinese language tutor: Personalising progress assessment and feedback or taking over your job? *On the Horizon*, 28(3), 113–124. <https://doi.org/10.1108/OTH-04-2020-0015>

Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). First-year students AI-competence as a predictor for intended and de facto use of AI-tools for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 18.

European Commission, Directorate-General for Education, Youth, Sport, and Culture. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/153756>

European Parliament. (2024). EU AI Act: First regulation on artificial intelligence. Retrieved from <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. Official Journal of the European Union, L series, 1–144.

Emeršič, Ž., Hrastnik, G., Meh Peer, N., & Peer, P. (2025). AIM@VET-inspired university level education strategies for teaching computer vision and biometrics. *ROSUS 2025 – Računalniška obdelava slik in njena uporaba v Sloveniji 2025: Zbornik 19. strokovne konference*. <https://doi.org/10.18690/um.feri.2.2025.4>



Hodgkin, E., & Kumar, A. (2023). Algorithmic fairness in vocational guidance systems: A comparative analysis of implementation frameworks. *Educational Technology Research and Development*, 71(3), 215-231.

Holmes, W., Bialik, M., & Fadel, C. (2023). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign

JISC (2022). *AI in tertiary Education. A summary of the current state of play*. JISC Repository. Retrieved from <https://repository.jisc.ac.uk/8783/1/ai-in-tertiary-education-report-june-2022.pdf>

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2020). Automated personalized feedback improves learning gains in an intelligent tutoring system. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II* 21 (pp. 140-146). Springer International Publishing.

Kumar, V., & Boulanger, D. (2020, October). Explainable automated essay scoring: Deep learning really has pedagogical value. In *Frontiers in education* (Vol. 5, p. 572367). Frontiers Media SA.

Li, Q. (2021). The use of artificial intelligence combined with cloud computing in the design of education information management platform. *International Journal of Emerging Technologies in Learning*, 16(5), 32–44. <https://doi.org/10.3991/ijet.v16i05.20309>

Porter, B., & Grippa, F. (2020). A platform for AI-enabled real-time feedback to promote digital collaboration. *Sustainability*, 12(24), 10243.

Prinsloo, P., & Knox, J. (2023). Exploring the ethical principles for the implementation of artificial intelligence in education: Towards a future agenda. *Research and Practice in Technology Enhanced Learning*, 18(1), 73-94. Retrieved from https://www.researchgate.net/publication/368740229_Exploring-the-Ethical-Principles-for-the-Implementation-of-Artificial-Intelligence-in-Education-Towards-a-Future-Agenda

Ramirez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., & Miranda, J. (2023). Artificial intelligence in vocational education: Ethical challenges and governance frameworks. *International Journal of Educational Technology in Higher Education*, 20(1), 138-155. <https://doi.org/10.1186/s41239-023-00408-3>



Rodriguez-Abitia, G., & Bribiesca-Correa, G. (2021). Assessing the ethical impact of artificial intelligence in education. RED. Revista de Educación a Distancia, 6(1). <https://doi.org/10.34627/redvol6iss1e202303>

Roman Etxebarrieta, G., Orcasitas-Vicandi, M., & Antzaka, A. (2024). *Evaluation schema for AI in education on data, privacy, ethics, and EU values (WP5)*. AIPioneers.org. Retrieved from: <https://aipioneers.org/evaluation-schema-for-ai-in-education-on-data-privacy-ethics-and-eu-values-wp5/>

Roppertz, S. (2020). Artificial Intelligence and Vocational Education and Training—Perspective of German Vet Teachers. In *European Distance and E-Learning Network (EDEN) Conference Proceedings* (No. 2, pp. 207-216). European Distance and E-Learning Network.

Rott, K. J., Lao, L., Petridou, E., & Schmidt-Hertha, B. (2022). Needs and requirements for an additional AI qualification during dual vocational training: Results from studies of apprentices and teachers. *Computers and education: Artificial intelligence*, 3, 100102.

Şenocak, D., Bozkurt, A., & Koçdar, S. (2024). Exploring the Ethical Principles for the Implementation of Artificial Intelligence in Education: Towards a Future Agenda. In *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 200-213). IGI Global.

Seufert, S. (2024). Artificial Intelligence in Vocational Education and Training (VET): Evaluating VET Leaders' Acceptance of AI in Switzerland.

Tang, J., & Hai, L. (2021). Construction and exploration of an intelligent evaluation system for educational app through artificial intelligence technology. *International Journal of Emerging Technologies in Learning*, 16(5), 17–31. <https://doi.org/10.3991/ijet.v16i05.20293>

Tuomi, I., Cachia, R., & Villar-Onrubia, D. (2023). On the futures of technology in education: Emerging trends and policy implications. *Publications Office of the European Union, Luxembourg*.

UNESCO (2023a). *ChatGPT and artificial intelligence in higher education*. Retrieved from: https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

UNESCO (2023b). *Guidance for generative AI in education and research*. Retrieved from: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>



Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18, 1-20.

Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S., & Roman-Cañizares, M. (2021). Implementation of a virtual assistant for the academic management of a university with the use of artificial intelligence. *Future Internet*, 13(4), 97.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223-235.

Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1), 1–5. <https://doi.org/10.1080/17439884.2023.2167830>

Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology*, 95-116.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2023). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00408-3>

CONTACT

www.aipioneers.org

gorka.roman@ehu.es

maria.orcasitas@ehu.es

natalia.louleli@ehu.eus