PREVENTING AND FIGHTING EARLY LEAVING FROM EDUCATION & TRAINING WITH AI

PREVENTING AND FIGHTING EARLY LEAVING
FROM EDUCATION & TRAINING WITH AI

stAI

Act before it happens: Preventing and fighting Early Leaving from Education and Training with Artificial Intelligence

T3.2 Data-driven decision support in Education: guidelines outlining opportunities, recommendations and concerns

Work Package 3

Co-funded by
the European Union

Author: STAI Consortium

PROJECT INFORMATION

Project acronym:

stAI

Project title: Act before it happens: Preventing and fighting Early Leaving from Education and Training with Artificial Intelligence

Project N°:2023-1-IS01-KA220-SCH-000157308

Sub-programme or KA:

KA220-SCH - Cooperation partnerships in school Education

Website: www.stai-edu.eu

Consortium:

Table of Contents

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Introduction

Education is entering a new era where data and Artificial Intelligence (AI) are becoming powerful tools to support students and educators. This document focuses on creating practical and accessible guidelines to help schools and education systems use AI-driven tools effectively to tackle the challenges of Early Leaving from Education and Training (ELET). These tools can help identify students at risk, recommend support strategies, and make decision-making processes more informed and proactive.

This document builds on the insights gathered from earlier work, aiming to explore how AI-powered systems can be integrated into real educational settings. The goal is to ensure that these technologies are not only effective but also ethical, transparent, and inclusive. By examining their potential benefits, such as better monitoring of students and personalised interventions, alongside the concerns they raise, including privacy issues, data biases, and the fairness of AI decisions, we aim to create a balanced approach that schools can trust.

Ultimately, this document is about helping educators and policymakers navigate this complex but exciting landscape. The guidelines we develop will provide clear, actionable advice for using AI in ways that support students and foster equality, inclusion, and transparency. Through collaboration with project partners and input from real-world educational contexts, these guidelines will pave the way for safer, smarter, and more effective tools to reduce dropout rates and give every student a better chance to succeed.

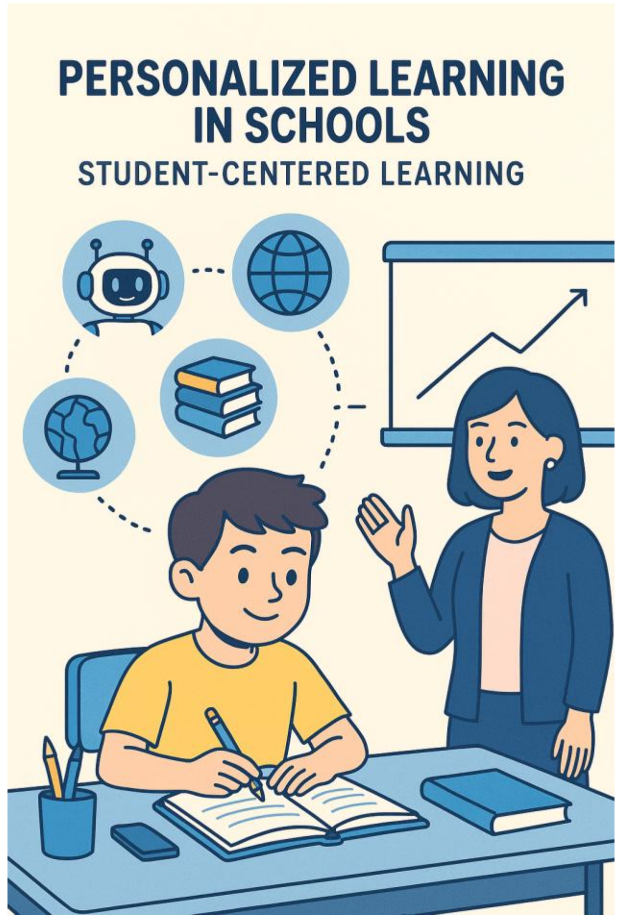
Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Opportunities

1. Personalised Learning

Student-Centred Learning Approaches



Personalised learning represents a transformative shift in education, moving away from standardised curricula towards student-centred approaches that adapt to each learner's unique journey. Instead of grouping students by age or predefined levels, personalised systems adapt continuously to learners' needs, preferences and pace. This shift in focus empowers students to take an active role in their learning journey, fostering greater engagement and ownership over their educational path.

Data-Driven Adaptation of Learning Content

Artificial intelligence plays a central role in enabling personalised learning. By analysing structured data, such as assessment scores and attendance patterns, as well as unstructured inputs, such as time spent on tasks and interaction patterns, AI can build detailed learning profiles. These profiles inform the dynamic adaptation of materials, formats and levels of complexity, ensuring that each learner is presented with tasks that are neither too simple nor too challenging. The system's real-time adaptability creates an evolving learning environment that responds to students' progress and needs, thereby helping to maintain motivation and reduce frustration.

Tailored pathways reduce disengagement.

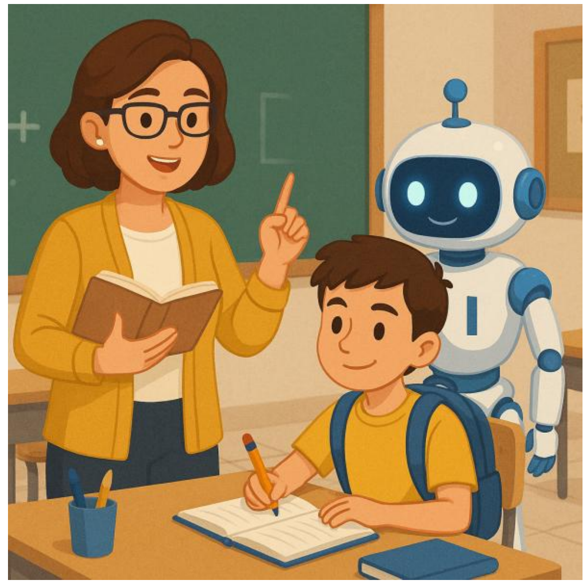
A key advantage of personalised learning is its ability to address the underlying causes of disengagement from school before they escalate. When students cannot relate to the content, feel they are consistently behind their peers or perceive school as irrelevant, their motivation suffers. However, by offering multiple learning paths that are aligned with individual interests and capabilities, personalised learning can restore relevance and purpose

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

to students' educational experience. This is particularly important for learners who do not fit traditional educational models or have experienced setbacks.

Supporting differentiated goals and definitions of success



Traditional systems often rely on narrow definitions of achievement. In contrast, personalised learning takes a broader view of success, including not only academic proficiency, but also progress in skills such as collaboration, problem solving and emotional self-regulation. AI tools can track these dimensions and provide feedback that recognises incremental improvement. This encourages persistence and self-confidence - two critical elements in preventing early school leaving.

Promoting flexibility and learning continuity

Personalised learning allows for greater flexibility in terms of when, where and how students learn. This is particularly beneficial for learners who are facing external challenges, such as health issues, caring responsibilities or socio-economic barriers. By decoupling learning from rigid schedules and physical locations, AI-enabled systems help to ensure continuity and access, thereby reducing the likelihood of temporary setbacks leading to permanent disengagement.

Supporting Self-Awareness in Learning

Another key feature of personalised learning is its ability to encourage metacognition. Real-time dashboards, progress trackers and intelligent feedback guide students through content and encourage them to reflect on their learning process. The ability to understand how they learn, identify effective strategies, and adapt accordingly empowers students to become more autonomous and resilient.

Conditions for successful implementation

To realise the full benefits of personalised learning, it must be thoughtfully integrated into the broader educational ecosystem. This involves ensuring equitable access to technology, digital infrastructure and user-friendly, inclusive learning platforms. It is essential that teachers are trained not only in using AI tools, but also in understanding how to interpret their outputs, maintain pedagogical autonomy, and manage the ethical implications of data use.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

2. Early Intervention

Introduction

Early school leaving is a major challenge for all European Union member states. It affects not only the individual path of young people but also social cohesion, economic productivity, and long-term competitiveness of European societies. A young person who leaves school before acquiring basic skills is at high risk of unemployment, social exclusion, and limited personal development opportunities. The European Union has recognised the seriousness of this phenomenon and has set ambitious targets for its reduction. Early intervention is considered the most effective strategy to prevent dropout, as it addresses risk factors in time and provides personalised support to vulnerable students.

European context and definitions

Early school leaving (ELET) is defined as the percentage of young people aged 18–24 who have completed at most lower secondary education and who are not engaged in education or training. In 2023, the EU average rate was 9.5%, close to the 2030 target of below 9%. However, differences between states are significant. In some Nordic and Central European countries, the rate is under 5%, while in Southern and Eastern states it often exceeds 12–14%.

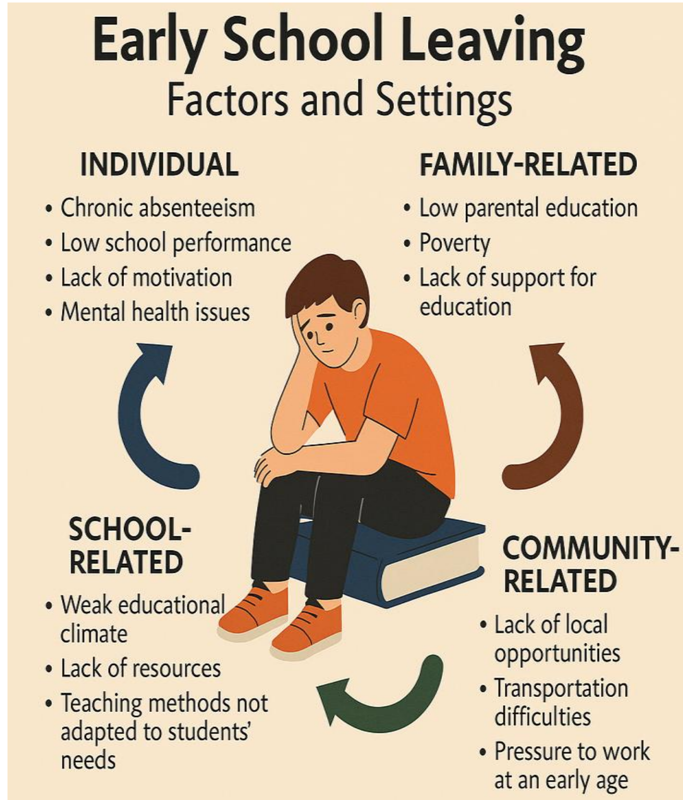
The reasons young people leave education prematurely are not singular. They reflect an interwoven set of influences: personal struggles, family circumstances, institutional shortcomings, and the wider community environment. No single factor alone explains dropout; rather, it is the interaction of these conditions that leads to disengagement. The following sections examine these dimensions in detail.

Main risk factors are multiple and interconnected:

Individual Factors

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**



The experiences and capacities of the learner are central to understanding early leaving.

Chronic absenteeism is one of the most powerful predictors. When students are frequently absent, they fall behind academically, lose social ties, and begin to see school as peripheral to their daily life.

Low academic performance deepens this risk. Repeated failures erode confidence, making students feel incapable of meeting expectations. In systems where grade repetition is common, students may become discouraged when their progress appears blocked.

Lack of motivation often arises when students do not see the relevance of education to their future. If curricula feel abstract, disconnected from real life or employment prospects, engagement declines.

Mental health challenges such as anxiety, depression, or behavioural difficulties add another layer of vulnerability. Without adequate support, these issues can undermine attendance, concentration, and performance.

Policy responses: Early-warning systems that track attendance and performance can identify risks before they become entrenched. Schools need to provide targeted tutoring, personalised guidance, and access to school-based counselling. Curriculum reforms that highlight relevance—through project-based learning, vocational options, and career guidance—can strengthen motivation.

Family-Related Factors

Family background exerts a powerful influence on educational outcomes.

Parental education is a key determinant. When parents themselves have limited schooling, they may lack the knowledge or confidence to support their children's studies or to engage actively with teachers and schools.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Poverty and material deprivation create practical obstacles: the inability to afford school supplies, reliable internet, or safe transport can reduce participation. Economic pressures may also push young people toward early work, particularly in families facing unemployment or precarious jobs.

Limited family support - whether due to long working hours, language barriers, or social marginalisation—can weaken the encouragement students need to persevere. In some households, family instability or conflict further erodes the conditions for learning.

Policy responses: Reducing the influence of socio-economic background requires targeted support. This includes financial measures such as free school meals, transport subsidies, and income support for disadvantaged families. Programmes that build parental engagement—through workshops, home visits, or community mediators—help strengthen the link between home and school. Broader social policies that alleviate child poverty also directly support educational attainment.

School-Related Factors

The school environment can either protect against or contribute to early leaving.

A **weak educational climate**, marked by low expectations, limited discipline, or negative peer relations, diminishes the sense of belonging. Bullying and exclusion, if unaddressed, often drive vulnerable students away.

Insufficient resources prevent schools from offering the targeted support that struggling students need. Overcrowded classrooms, lack of specialised staff, and limited access to extracurricular activities reduce opportunities for engagement.

Teaching methods not adapted to diverse needs create further risks. Traditional, lecture-based approaches may alienate students with different learning styles, while limited differentiation leaves some behind. Curricula that are too rigid or exam-driven may fail to capture the interest of those seeking practical or applied learning.

Policy responses: Strengthening school climate requires whole-school strategies to promote inclusion, anti-bullying programmes, and student participation in decision-making. Investment in teacher training can equip educators with inclusive pedagogies and differentiated instruction skills. Adequate funding for disadvantaged schools ensures they can offer tutoring, extracurricular activities, and smaller class sizes. Expanding vocational and project-based learning options makes school more relevant and appealing to a wider range of students.

Co-funded by the European Union
The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Community-Related Factors

The broader community context shapes whether education is accessible and worthwhile.

Lack of local opportunities - whether in training, apprenticeships, or employment—can weaken the perceived value of completing education. If pathways from school to work are unclear, students may lose motivation.

Transportation difficulties create concrete barriers, especially in rural or peripheral regions. Long, costly, or unreliable commutes discourage regular attendance.

Pressure to work at an early age arises in communities where low-skilled jobs are available or where cultural norms emphasise early economic contribution. For some families, immediate income outweighs the promise of long-term benefits from further schooling.

Policy responses: Investing in reliable and affordable transport networks is crucial to ensure access for rural students. Partnerships between schools, employers, and training providers can help align education with local labour markets and make pathways visible. Community-based youth centres, mentoring programmes, and extracurricular initiatives provide alternative spaces for engagement. Measures to regulate child labour and promote youth employment opportunities tied to training can also reduce premature school-to-work transitions.

Conclusion

Early school leaving is not the result of a single cause, but of a web of interconnected factors spanning the individual, family, school, and community. Tackling it requires a holistic policy approach. Efforts must start with prevention: monitoring attendance, supporting students at risk, and making education relevant and engaging. Families need to be empowered as partners, with financial and social supports to overcome material barriers. Schools must be equipped with the resources, staff, and pedagogies to include all learners. Communities must ensure that education is accessible and clearly linked to meaningful opportunities.

By addressing these dimensions in an integrated manner, the EU and its Member States can move closer to the 2030 target and, more importantly, ensure that every young person has the chance to complete their education and to enter adulthood with confidence, opportunity, and hope.

Strategic framework and European policies

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

The European Union has developed several instruments to support member states:
- Council Recommendation "Pathways to School Success" (2022) – sets an integrated framework based on prevention, early intervention, and compensatory measures.
- European Education Area 2030 – sets clear targets: reducing early school leaving to below 9%, ensuring at least 96% participation in early childhood education, and reducing the share of low achievers in reading, mathematics, and science to below 15%.
- Digital Education Action Plan 2021–2027 – supports digital inclusion, provides resources for teachers and students, and combats inequalities in access to online education.
- European Child Guarantee (2021) – ensures access to quality education, healthcare, healthy meals, and decent living conditions.
- EU funds: through ESF+, Erasmus+, and RRF, member states receive resources for school infrastructure, teacher support, and implementation of programs to reduce dropout.

Early intervention – concept and mechanisms

Early intervention is essential to prevent early school leaving. It involves timely and coordinated actions when a student begins to show signs of risk. The most important mechanisms include:
- Early warning systems (EWS): monitor absenteeism, declining grades, and problematic behaviour, enabling schools to act quickly.
- Individualised support plans: tailored to each student's needs, including mentoring, tutoring, and remedial activities.
- Psychological counselling and emotional support: essential for students facing anxiety, depression, or family issues.
- Parental and community involvement: through home visits, regular meetings, workshops, and family support programs.
- Assisted transitions: from primary to lower secondary or from lower secondary to upper secondary/VET, to prevent disengagement and difficult adaptation.

Impact and benefits of early intervention

Early intervention brings multiple benefits:
- Increases chances of completing education and integrating into the labor market.
- Reduces the risk of poverty and social exclusion.
- Improves the well-being of students and families.
- Adds economic value through a better-qualified workforce.

Studies show that investments in prevention and early intervention are more effective and less costly than compensatory measures taken after a student has already left school.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Conclusions

Early school leaving remains a complex problem, but early intervention is one of the most effective solutions. The European Union provides the strategic framework and resources, but success depends on how each member state, including Romania, manages to implement measures adapted to its own context. Through cooperation between schools, families, communities, and European institutions, the goal of reducing early school leaving to below 9% by 2030 can be achieved.

3. Improved Teaching Strategies

Improved Teaching Strategies

Improving teaching strategies is central to creating inclusive, engaging, and equitable learning environments, particularly for students at risk of Early Leaving from Education and Training (ELET). While personalised learning and early intervention focus on the learner, improved teaching strategies address the methods and systems educators use to deliver effective instruction. The integration of data-driven decision support tools enables a shift from reactive to proactive teaching, supporting more responsive, informed, and innovative pedagogical practices.

From Intuition to Evidence-Informed Pedagogy

Traditional teaching often relies on educator's intuition and experience. While valuable, this approach may not always identify or address early signs of disengagement. Data analytics and AI technologies offer tools to systematically analyse student engagement, behaviour, and outcomes, supporting educators in making timely instructional decisions. For example, AI-based systems can assist teachers in adapting their strategies by offering real-time dashboards and smart alerts that prompt timely pedagogical adjustments. The RADAR system (Embarak & Hawarna, 2024) and the PDAR/LIS tools (Baneres, Rodriguez-Gonzalez, & Guerrero-Roldan, 2023) exemplify how real-time data feedback loops allow educators to spot and respond to learning problems early, often before traditional assessments would flag a concern. Though these systems were developed for tertiary education, the underlying principles of predictive analytics and continuous monitoring are equally relevant in school settings.

Blended and Active Learning as Strategic Pedagogies

Teaching strategies that combine digital tools with collaborative, project-based methods can be highly effective in preventing ELET. The Icelandic *Austur-Vestur Sköpunarsmiðjur* (Austur-Vestur Sköpunarsmiðjur, n.d.) and *Mixtúra* (Mixtúra, n.d.) initiatives illustrate how creative technologies and makerspaces can re-engage students by emphasising hands-on learning, problem-solving, and agency. These environments give students new ways to demonstrate learning and foster skills such as persistence, teamwork, and creativity. Likewise, Italy's *Classe AI* (Classe AI, 2025) integrates AI education into classroom

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

activities, encouraging both teachers and students to explore AI critically and creatively. This supports the development of digital literacy and demystifies technologies that are often seen as distant or opaque.

Professional Empowerment and Reflective Practice

Improved teaching strategies are only sustainable when educators are supported in their professional development. As seen in Reykjavík's *Mixtúra* centre, embedding continuous professional learning into local education policy creates a culture where innovation is a systemic practice, not a one-off initiative. Portugal's National School Success Programme emphasises teacher-led, localised action planning, where schools develop context-specific strategies based on self-assessment data. This reinforces teacher agency and fosters ownership over change by encouraging educators to reflect on their practice and monitor their impact, creating a cycle of iterative improvement. AI can further support this by creating highly personalised professional development pathways for teachers, suggesting relevant research, training modules, or even providing AI-driven coaching based on their specific classroom challenges.

The Evolving Role of AI in Pedagogy

Looking forward, the role of AI can expand beyond current applications to become a true "cognitive partner" for both teachers and students. Advanced AI teaching assistants could act as a sophisticated co-pilot for teachers, not just flagging at-risk students but also suggesting pedagogical approaches for difficult topics, automating the creation of diverse assessment tools, and handling routine administrative tasks. This would free up teachers to focus on direct student interaction and complex pedagogical planning. For students, AI tutors could engage them in Socratic dialogues to prompt critical thinking or help them visualise their own learning journeys to develop metacognitive skills and self-regulation.

Strategic Implications and a Human-Centred Future

These enhanced teaching strategies directly contribute to reducing dropout risk by increasing student engagement, offering diverse pathways to success, and creating supportive learning environments where students feel seen and motivated. However, to harness these future possibilities, several strategic considerations are critical. As AI becomes more integrated, robust ethical frameworks, strong privacy protections (like GDPR), and active mitigation of algorithmic bias are non-negotiable. Technology should always be positioned as a tool to augment and empower teachers, not replace their professional judgment or human connection. For any critical decisions, a human must remain in the loop. Finally, ensuring digital equity—where all students and teachers have access to the necessary technology and skills—will be paramount to prevent AI from widening existing inequalities

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Concerns

1. Data Privacy and Security

As educational institutions increasingly rely on artificial intelligence (AI) and data-driven decision-making to personalise learning experiences, optimise student performance, and enhance administrative efficiency, concerns surrounding privacy, security, and compliance with European Union (EU) regulations have become more pressing. The implementation of AI in education must be carefully evaluated to ensure ethical use, legal compliance, and the protection of students' fundamental rights.

One of the foremost concerns in AI-driven educational tools is the accessibility of these systems to minors. Under **GDPR Article 8**, processing the personal data of children under 16 (or 13 in some EU countries) requires verifiable parental consent. If the AI system is designed for vulnerable groups such as neurodivergent students or individuals with disabilities, additional safeguards must be in place to avoid bias and ensure transparency.

Many AI-powered educational applications require access to user devices, often collecting data from contact lists, media files, microphones, cameras, geolocation, and even biometric sensors. Such extensive data access raises red flags regarding data minimisation principles under **GDPR Article 5**. Institutions and technology providers must assess whether the requested permissions are strictly necessary for the tool's functionality and whether users can provide informed consent before granting access.

Personal Data Collection and Processing

The type and extent of personal data collected by AI systems must be justified under **GDPR Article 9**, which imposes strict conditions on processing special categories of data, such as biometric information, health records, or behavioural analytics. AI systems used in education often collect vast amounts of data, including students' academic records, interaction patterns, and even emotional responses.

If data collection is disproportionate to the intended purpose, it violates the principle of **data proportionality** and could lead to unnecessary exposure of sensitive information.

Legal Basis and Parental Consent

Educational institutions and AI providers must establish a lawful basis for processing personal

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

data. Depending on the use case, this could be **explicit consent, contractual necessity, or legitimate interest**. However, if the AI system involves profiling, automated decision-making, or behavioural tracking, **GDPR Article 22** grants individuals the right to contest and opt out of such processing. Additionally, institutions must ensure that parental consent mechanisms for minors are robust, preventing unauthorised data collection from children without adequate safeguards.

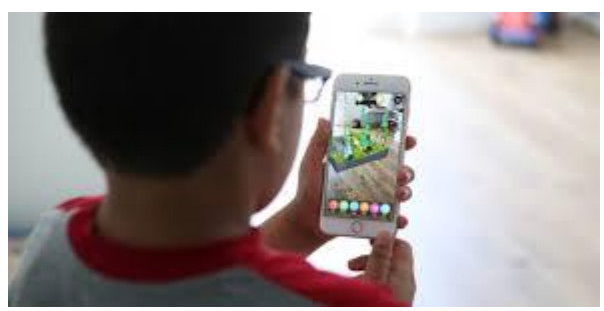
Data Storage, Security, and Retention

A crucial aspect of AI governance in education is the **storage location and duration** of collected data. If student data is stored outside the EU, **Standard Contractual Clauses (SCCs)** must be in place to ensure compliance with GDPR. Furthermore, the principle of **storage limitation (GDPR Article 5)** mandates that data should only be retained for as long as necessary to fulfil its intended purpose.

Cybersecurity risks also pose a significant challenge, as unauthorised access to educational data can lead to identity theft, targeted advertising, or even manipulation of student records. Institutions must implement **encryption, strict access controls, and anonymisation techniques** to mitigate these risks and prevent breaches.

AI-Specific Concerns in Educational Tools

Beyond standard privacy concerns, AI-driven education platforms introduce **unique risks** that require specialised oversight:



- **Adaptive Learning and Behavioural Tracking:** AI systems that analyse student interactions and adapt content accordingly must be transparent about data usage.
- **Emotion Recognition and Biometrics:** If AI tools use **facial recognition, voice analysis, or body sensors**, they must comply with **GDPR Article 9** and justify the necessity of such data collection.
- **Augmented Reality (AR) and Environment Scanning:** AI-based **AR features** that scan the user's surroundings may inadvertently capture **third-party data** without consent.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

- **Generative AI and Virtual Assistants:** If AI chatbots or tutoring agents interact dynamically with students, mechanisms must be in place to prevent **misinformation, biases, or unauthorised data storage**.

Third-Party Data Sharing and Compliance

The involvement of third-party vendors in AI-powered education raises additional concerns about **data sharing**. If student data is transferred to third parties for **analytics, marketing, or AI model training**, **Data Processing Agreements (DPAs) under GDPR Article 28** are mandatory. Institutions must ensure **full transparency** regarding which entities have access to data and for what purposes.

Risk Assessment and Transparency

To mitigate the risks associated with AI in education, institutions should conduct **Data Protection Impact Assessments (DPIAs) as required by GDPR Article 35**. DPIAs help identify **potential risks, evaluate mitigation strategies, and ensure compliance with legal and ethical standards**. Additionally, educational institutions must provide clear and age-appropriate privacy policies to students and parents, ensuring they understand their rights and the implications of AI-based decisions.

 Project N°:2023-1-IS01-KA220-SCH-000157308

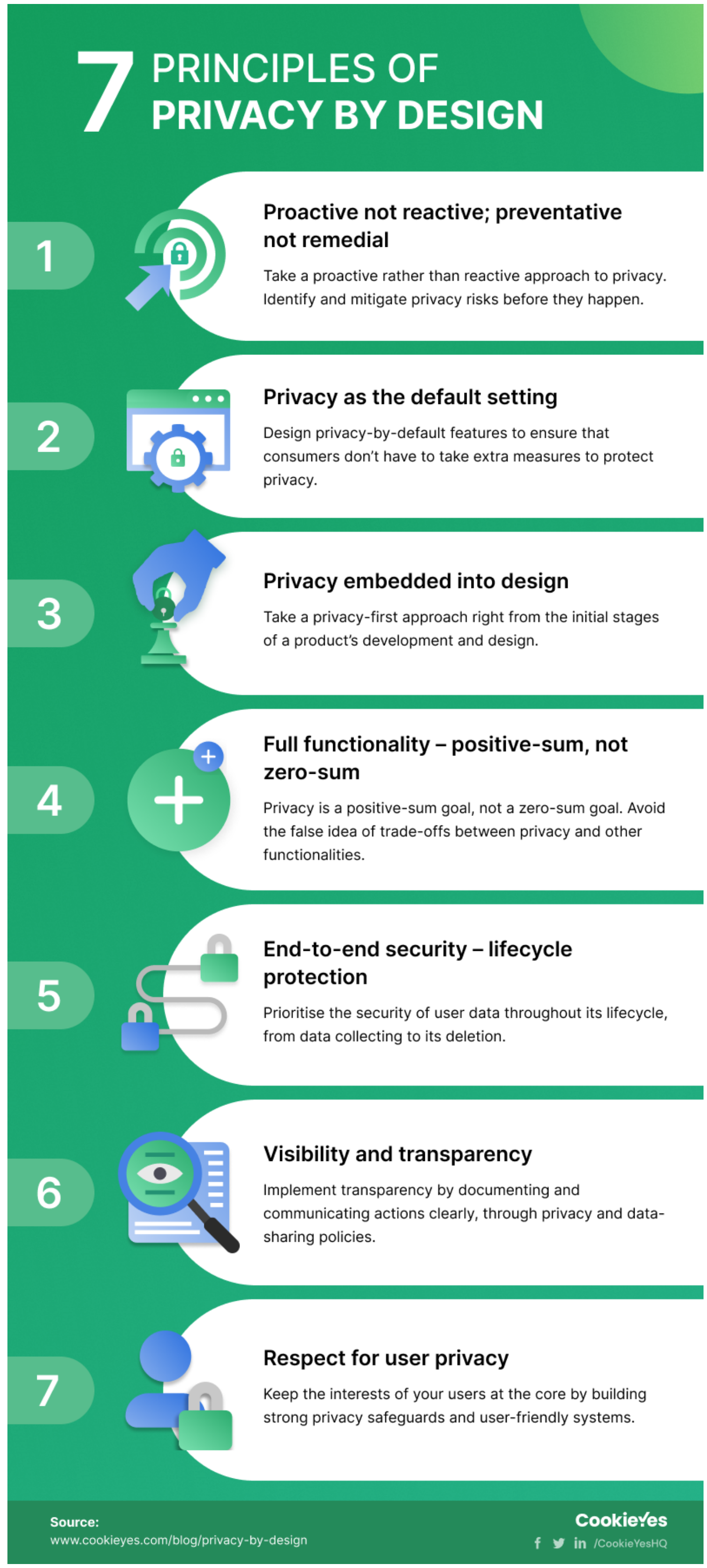
Data Protection by Design and User Control

Following the **Privacy by Design (GDPR Article 25)** principle, AI tools should be configured to collect **the minimum necessary data** by default. Students and parents must have easy access to **privacy settings, opt-out mechanisms, and data deletion options**. Moreover, institutions must ensure compliance with **data subject rights** (Articles 15-22), enabling users to **access, correct, delete, or port their data** when necessary.

Age Assurance and Safeguards for Children

With increasing concerns about **age verification**, AI-based education platforms must balance security with privacy. While age assurance mechanisms can help protect minors, poorly implemented solutions may introduce **new risks**, such as excessive data collection or flawed identification methods. Institutions should prioritise solutions that provide **effective safeguards without violating privacy rights**.

Data-driven decision support in education holds immense potential to revolutionise learning experiences, offering **personalised education, improved accessibility, and efficient administrative processes**.



However, these benefits come with significant **privacy, security, and ethical challenges**. By adhering to EU regulations, implementing **strong data protection measures, conducting risk assessments, and**

ensuring transparency, educational institutions can harness AI's potential while safeguarding the rights and interests of students.

As AI in education continues to evolve, the **balancing act between innovation and responsibility** will remain critical. Policymakers, educators, and AI developers must work collaboratively to create **trustworthy, fair, and privacy-compliant AI systems** that truly enhance education without compromising fundamental rights.

2. Overreliance on Data

The vast amount of data in society has led to the use of algorithm-driven technologies, such as artificial intelligence (AI) and machine learning (ML), to organise and interpret information. However, overreliance on data can be problematic if these systems are not properly designed. AI depends on quantifiable information, raising concerns about whether complex realities can truly be reduced to data. Ultimately, human biases, perceptions, and worldviews shape the design and assessment of these systems, imposing greater limitations than the technology itself. Even as machines learn, humans define their starting point and learning framework (Kallberg, 2022).

Overreliance on data refers to an excessive dependence on data-driven models for decision-making without considering context, human judgment, or external factors especially when the data used is flawed, limited, or lacks diversity.

Balancing data quantity and quality

Machine learning models, especially deep learning algorithms, often demand substantial amounts of data to identify intricate patterns. However, quality should never be sacrificed for quantity (Pagadala, 2023). In the intricate realm of machine learning, finding the perfect equilibrium between data quantity and quality is akin to balancing on a tightrope. On one hand, more data provides the model with a broader perspective, allowing it to generalize better to unseen data. On the other hand, high-quality data ensures the model learns from reliable patterns, resulting in accurate predictions. While quantity ensures comprehensiveness, quality guarantees accuracy.

- Quantity over quality: If data are abundant, but much of it is of low quality, the model might learn incorrect patterns, leading to unreliable predictions.
- Quality over quantity: Conversely, having high-quality but insufficient data can limit the model's ability to generalize to diverse scenarios.

Enhancing data quality

Data, as defined by Andrea Jones-Rooy, is an imperfect approximation of some aspect of the world at a certain time and place (Jones-Rooy, 2019). As long as data is considered cold, hard, infallible truth, we run the risk of generating and reinforcing a lot of inaccurate understandings of the world around us. For that reason, data quality directly impacts AI

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

models' performance, accuracy, and reliability (Redman, 2018). High-quality data enables models to make better predictions and produce more reliable outcomes, fostering trust and confidence among users (Ataman, 2024).

Ensuring data quality also means addressing biases present in the data, which is essential to avoid perpetuating and amplifying these biases in AI-generated outputs. Furthermore, a diverse and representative dataset enhances an AI model's ability to generalise well across different situations and inputs, ensuring its performance and relevance across various contexts and user groups. Ultimately, maintaining data quality is key to realising the full potential of AI systems in delivering value, driving innovation, and ensuring ethical outcomes.

The key components of quality data in AI are accuracy, consistency, completeness, timeliness, relevance, integrity, diversity, and representativeness.

- Accuracy: Precise and error-free data ensures that AI models learn from correct information, reducing the risk of misleading predictions and poor decision-making. Conversely, errors in data input can lead to incorrect decisions or misguided insights, potentially causing significant harm to both organizations and individuals.
- Completeness: Comprehensive datasets that capture all relevant variables prevent information gaps, enabling the model to identify meaningful patterns and relationships. On the contrary, data incompleteness refers to the situation where the dataset used for training a machine learning model lacks certain information or has missing values for some of its features. Incomplete datasets can hinder AI algorithms from detecting critical patterns and correlations, compromising the model's ability to generalise effectively to new, unseen data.
- Consistency: Uniform data across sources and timeframes avoids contradictions and ensures the model receives coherent input, improving stability and reliability, whereas inconsistent data can lead to confusion and misinterpretation, impairing the performance of AI systems.
- Relevance: Data must be directly related to the problem being addressed. Relevant data contributes directly to the problem at hand, helping AI systems to focus on the most important variables and relationships. Irrelevant data can clutter models and introduce noise, reducing the model's effectiveness.
- Timeliness: Up-to-date data ensures that the model reflects current trends and behaviours, which is critical for adapting to changing environments, whereas outdated data results in irrelevant or misleading outputs. Hence, the importance of training systems with data freshness to ensure high AI performance.
- Diversity and Representativeness: Diverse and unbiased datasets help models generalise well across different populations and scenarios, minimising the risk of biased outcomes. If the data collected is not representative of the entire population, it can lead to biased conclusions. The model's predictions will be a direct reflection of that data.
- Integrity: Secure and unaltered data protects against corruption or manipulation, ensuring trustworthy results.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

To illustrate the impact of data quality issues, particularly data incompleteness, consider a dropout prediction system that relies solely on historical student data such as grades, attendance, and engagement metrics. However, the model overlooks non-quantifiable factors like mental health issues or personal challenges. Consequently, students who are emotionally struggling but maintaining strong academic performance may go unnoticed and not receive the necessary support, ultimately reducing the model's effectiveness in identifying at-risk students.

Enabling models to learn and adapt over time

The environment in which machine learning operates may evolve or differ from what the algorithms were developed to face. While this can happen in many ways, two of the most frequent are concept drift and covariate shift. With the former, the relationship between the inputs the system uses and its outputs isn't stable over time or may be misspecified (Babic, Cohen, Evgeniou, & Gerke, 2021).

Suppose a university uses a machine learning model trained on pre-pandemic data to predict student dropout risk. The model relies on factors such as class attendance, participation in campus activities, and in-person academic support usage. However, during and after the COVID-19 pandemic, the university has shifted to online and hybrid learning. Student behaviours and engagement patterns change—attendance is now virtual, campus activities are limited, and students rely more on digital resources.

The model, trained on in-person learning data, fails to account for these behavioural shifts. It inaccurately flags students as high-risk because traditional indicators (like physical attendance) no longer apply. Conversely, it may overlook at-risk students struggling with online learning due to poor internet access, isolation, or mental health challenges.

The model's predictive accuracy declines because the relationship between input data and dropout risk has changed. This shift in data patterns reflects concept drift, highlighting the need to retrain the model with updated data to adapt to the new learning environment.

It is crucial to empower models to continuously learn and adapt over time without requiring complete retraining. Implementing continual and transfer learning allows models to effectively leverage knowledge from related tasks, enhancing adaptability and performance.

Covariate shifts occur when the data fed into an algorithm during its use differs from the data that trained it. This can happen even if the patterns the algorithm learned are stable and there's no concept drift (Babic et al., 2021). Covariate shift is a specific type of dataset shift often encountered in machine learning. It is when the distribution of input data shifts between the training environment and the live environment.

A university develops a machine learning model to predict student dropout risk based on historical data. The model uses features like standardised test scores, attendance records, and participation in extracurricular activities to assess risk.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

The university later adopted a test-optional admissions policy, significantly reducing the number of students submitting standardised test scores. Additionally, online learning has become more prevalent, changing attendance tracking methods.

Although the relationship between predictors and dropout risk (the conditional distribution) remains the same, the distribution of input features (covariates) has shifted. For example, the model now receives fewer standardised test scores, making this feature less informative. Besides, attendance data collected from online platforms differs from traditional in-person records. As a result, the model's performance deteriorates because it was trained on data where standardised test scores and in-person attendance were reliable predictors. This change in the input data distribution without altering the outcome relationship exemplifies a covariate shift. To address this, the model must be retrained or adjusted to account for the new feature distributions.

To mitigate covariate shifts—changes in input data distribution between training and testing—various strategies can be employed. Domain adaptation aligns training and testing data distributions through techniques like feature normalisation and adversarial learning. Importance weighting adjusts training sample influence based on distribution differences, while robust feature selection focuses on stable features to minimise sensitivity to shifts. Regularisation methods enhance generalisation, and data augmentation introduces diverse scenarios to the model. Ensemble methods combine multiple models to balance distributional variances, and online learning allows continuous model updates with new data. Additionally, adapting batch normalisation statistics to the target domain further improves robustness. Collectively, these approaches help maintain consistent model performance across varying data distributions.

Integrating data with prior experience, human expertise, and domain knowledge

As Andrea Jones-Rooy writes in Quartz (Jones-Rooy, 2019), data only exists in the first place because humans chose to collect it, and they collected it using human-made tools. Therefore, data is as fallible as people. Data is essential for discovery, but it requires human judgment to select, refine, and transform it into meaningful insights. Ultimately, the value of data depends on its quality and the expertise and skills of the individual interpreting and wielding it (Jones-Rooy, 2019; Taylor, 2020). In addition to correct data handling, taking into account domain knowledge, understanding the process behind the data, and considering contextual information are crucial for the effective performance of machine learning models. These elements contribute to more accurate, reliable, and meaningful outcomes.

When data is used alongside other decision-making factors, such as expert judgment and contextual understanding, the impact of flawed analyses is minimised, as these additional factors can compensate for potential inaccuracies. However, when data serves as the primary basis for decision-making, flawed analyses can lead to significant errors and poor outcomes. The key takeaway is not to dismiss the value of data but to emphasise the importance of integrating data and analytical insights with complementary decision-making tools, such as prior experience and domain expertise. Relying solely on data as the ultimate source of truth can undermine the decision-making process and increase the risk of error (Nolis, 2018).

For example, in predicting the risk of early school leaving, it may be valuable to incorporate potential causal relationships identified by teachers to complement insights derived from data analysis. Teachers' observations and firsthand experiences can provide context-specific factors that data alone may overlook, leading to a more comprehensive and accurate risk assessment.

Integrating domain knowledge is essential for enhancing model performance, such as embedding expert insights into feature selection, model constraints, and decision logic. Additionally, combining data-driven models with rule-based or symbolic reasoning systems can improve interpretability and reliability. Designing human-in-the-loop systems is equally important, enabling human oversight in critical decision-making processes and allowing for timely intervention and feedback to ensure more accurate and responsible outcomes. Besides, causal inference and reasoning enhance predictive models by identifying true cause-and-effect relationships rather than relying solely on correlations. This leads to more robust and generalizable predictions, as models can better understand how changes in input variables influence outcomes. As a result, models become more interpretable and reliable, particularly in complex or dynamic environments where understanding causal mechanisms is crucial for accurate and actionable predictions.

Enhancing transparency, guiding decision-making, and building trust and understanding

People strategically choose whether or not to engage with an AI explanation, demonstrating empirically that there are scenarios where AI explanations reduce overreliance (Vasconcelos et al., 2023). The use of interpretable models or explainability tools makes AI decisions transparent and understandable to stakeholders. Explanations clarify decision-making, helping users identify when models depend too heavily on data or biased patterns. Explanations also empower domain experts to intervene, ensuring data-driven insights are balanced with real-world expertise.

3. Ethical Considerations

While AI and data-driven tools have great potential to improve education and reduce Early Leaving from Education and Training (ELET), their use also raises important ethical questions. To ensure these technologies are used responsibly, it's essential to address concerns around fairness, transparency, privacy, and their overall impact on students and educators. Below are some key ethical considerations to keep in mind:

1. Avoiding Bias and Ensuring Fairness

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

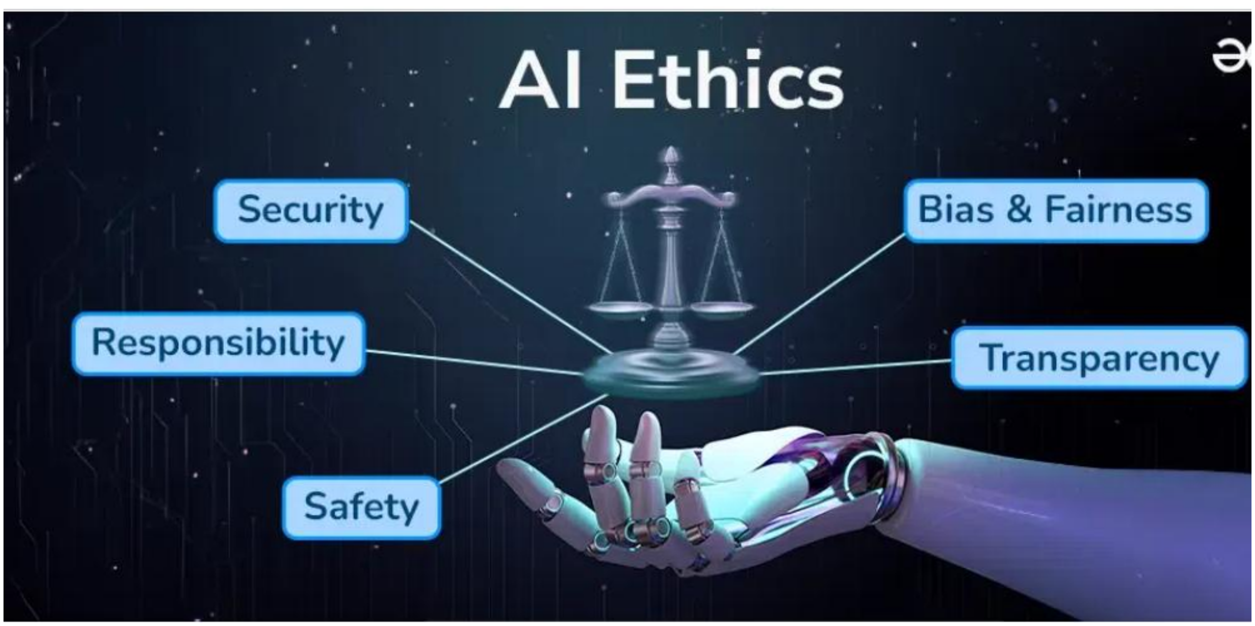
AI systems rely on data to make decisions, but if that data reflects existing inequalities, it could lead to unfair outcomes. For example, students from disadvantaged backgrounds might be incorrectly flagged as "at-risk" simply because of patterns in historical data. This can reinforce stereotypes and worsen inequalities. To prevent this, we must carefully examine the data used and ensure it represents all groups fairly. AI tools should be designed to promote inclusion, not discrimination, and to treat every student as an individual.

2. Making Decisions Clear and Understandable

AI systems often work like black boxes, making decisions that are difficult to explain or understand. For teachers, parents, and students, this lack of clarity can be frustrating and may reduce trust in the system. To address this, AI tools need to provide clear, easy-to-understand explanations for their recommendations. Educators should be able to understand why a student is flagged as "at-risk" and how the system reached that conclusion.

3. Supporting Students Without Stigmatising Them

Labelling students as "at-risk" can unintentionally harm their confidence or lead to lower expectations from teachers. It's important to use AI insights to support students in a positive way, focusing on their potential and offering opportunities for growth. These tools should help empower students and strengthen the relationship between educators and learners, not define or limit their futures.



4. Protecting Privacy and Gaining Consent

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Students' personal data must be handled with care and respect. Schools and organisations need to follow strict privacy regulations, like GDPR, ensuring that all data is securely stored and used only for educational purposes. Students and their families should understand how their data will be used and have the right to give or deny consent.

5. Maintaining Accountability and Oversight

AI tools should complement, not replace, human judgment. If a system makes an incorrect prediction, there must be a way to review and fix it. Educators should remain in control of how AI recommendations are applied, and clear accountability processes should be in place to address any mistakes.

6. Involving the Right People in the Design

AI tools should be designed with input from the people they affect—teachers, students, parents, and policymakers. Including these perspectives helps ensure the tools address real needs and are ethical and practical to use in everyday educational settings.

7. Balancing Technology with Human Connection

Finally, while AI can be a helpful tool, education is fundamentally about human relationships. Teachers, not algorithms, understand the unique needs and emotions of their students. Any AI system should be used to enhance, not replace, the personal connections that are so vital in education.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

Conclusion

The integration of Artificial Intelligence (AI) and data-driven tools into the field of education represents a paradigm shift in how schools understand, support, and respond to students' needs. As explored throughout this document, the use of AI technologies offers promising opportunities to personalize learning, anticipate risks such as Early Leaving from Education and Training (ELET), and enhance teaching and administrative processes. These technologies are not futuristic novelties—they are becoming embedded in the present-day educational landscape.

One of the most powerful contributions of AI is its ability to support personalized and student-centered learning. By dynamically adjusting content, pace, and feedback based on real-time data, AI systems empower students to follow individualized learning paths. This responsiveness helps increase engagement, reduce frustration, and prevent students from falling behind. Such adaptive systems are especially valuable in diverse classrooms, where learners may have varying backgrounds, learning styles, or support needs.

Furthermore, AI enhances the capacity of educators and school systems to detect early warning signs of disengagement. Through predictive analytics and behavioral pattern analysis, it becomes possible to flag students who may be at risk of leaving education prematurely. Importantly, these insights enable schools to take proactive and targeted measures—ranging from academic support to socio-emotional interventions—thereby shifting the focus from reaction to prevention.

However, the adoption of AI in education cannot proceed without a strong foundation in ethical awareness and regulatory compliance. The concerns related to data privacy, especially when dealing with minors, are serious and multifaceted. Institutions must ensure that data collection, storage, and processing practices are fully aligned with the General Data Protection Regulation (GDPR), providing transparency and meaningful consent to learners and their families. Furthermore, safeguards must be in place to prevent misuse of sensitive information and to uphold the rights and dignity of all users.

Equally important is the issue of overreliance on data. While quantitative data can provide valuable insights, it must be contextualised within the broader realities of students' lives and the judgment of educators. Models trained solely on past performance or narrow indicators risk overlooking the human, social, and emotional dimensions of learning. Integrating data with teacher expertise, school knowledge, and students' lived experiences creates a more balanced and accurate decision-making process.

As AI technologies evolve, models must be able to learn and adapt to changing educational environments. The concept of concept drift and covariate shift reminds us that student behaviour, learning contexts, and social realities are dynamic. Systems designed for static environments risk becoming obsolete or producing misleading outputs

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

if not regularly retrained with current data. Thus, continuous evaluation, human oversight, and flexible adaptation are essential features of trustworthy educational AI systems.

Another critical consideration is fairness and inclusivity. AI systems must be carefully designed to avoid perpetuating biases embedded in historical data. For instance, if previous school data underrepresents disadvantaged groups, algorithms may unintentionally penalise these same students. Ethical AI in education should actively promote equity by correcting for systemic imbalances, ensuring that all learners are treated fairly regardless of background, ability, or context.

Ultimately, successful AI integration in education depends on the active participation of all stakeholders. Teachers, students, parents, school leaders, and policymakers must be engaged in the design, implementation, and evaluation of AI tools. Their feedback ensures that technologies remain aligned with educational values, classroom realities, and students' best interests. Bottom-up involvement is particularly important to create systems that are adaptable, inclusive, and responsive to real-world challenges.

In conclusion, AI is not a silver bullet, but a powerful tool—one that must be used wisely, ethically, and collaboratively. When properly designed and implemented, AI can support more responsive, equitable, and engaging learning environments. Through shared responsibility, ongoing reflection, and a commitment to student wellbeing, we can ensure that technological innovation serves the goal of inclusive and quality education for all.

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

References

1. Austur-Vestur sköpunarsmiðjur. (n.d). https://sites.google.com/rvkskolar.is/austurvestur/
2. Ataman, A. (2024). Data Quality in AI: Challenges, Importance & Best Practices. Retrieved from AIMultiple RESEARCH website: https://research.aimultiple.com/data-quality-ai/
3. Babic, B., Cohen, I. G., Evgeniou, T., & Gerke, S. (2021). When Machine Learning Goes Off the Rails: A guide to managing the risks. Retrieved from https://hbr.org/2021/01/when-machine-learning-goes-off-the-rails
4. Bañeres, D., Rodríguez-González, M. E., Guerrero-Roldán, A. E., & Cortadas, P. (2023). An early warning system to identify and intervene online dropout learners. *International Journal of Educational Technology in Higher Education*, 20(1). https://doi.org/10.1186/s41239-022-00371-5
5. Embarak, O. H., & Hawarna, S. (2024). Enhancing Student Success with XAI-Powered RADAR: An Automated AI-driven System for Early Detection of At-risk Students. *Procedia Computer Science*, 231(2023), 151–160. https://doi.org/10.1016/j.procs.2023.12.187
6. Jones-Rooy, A. (2019). I'm a data scientist who is skeptical about data. Retrieved from https://qz.com/1664575/is-data-science-legit
7. Kallberg, J. (2022). The risks of over-reliance on quantifiable data. Retrieved from https://www.linkedin.com/pulse/risks-over-reliance-quantifiable-data-kallberg-ph-d-cism-cissp/
8. Mixtúra. (n.d.). https://reykjavik.is/en/residents-children-youth/mixtura
9. Nolis, J. (2018). You're relying on data too much: Making decisions worse, not better. Retrieved from https://towardsdatascience.com/youre-relying-on-data-too-much-250d4edc70c3
10. Pagadala, S. (2023). Maximizing Machine Learning Success: The Data Dilemma Unraveled. Retrieved from https://pagadala-saiteja.medium.com/maximizing-machine-learning-success-the-data-dilemma-unraveled-87b941236cf8
11. Redman, T. C. (2018). If Your Data Is Bad, Your Machine Learning Tools Are Useless. Retrieved from https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless
12. Taylor, P. (2020). The Problem With An Over-Reliance On Data. Retrieved from https://paulitaylor.com/2020/11/08/the-problem-with-an-over-reliance-on-data/
13. Vasconcelos, H., Jörke, M., Grunde-Mclaughlin, M., Gerstenberg, T., Bernstein, M. S., & Krishna, R. (2023). Explanations Can Reduce Overreliance on AI Systems During Decision-Making. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), 1–38. https://doi.org/10.1145/3579605
14. European Commission. (2016). General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679). Official Journal of the European Union. Retrieved from https://eur-lex.europa.eu
15. European Data Protection Board (EDPB). (2020). Guidelines on Data Protection by Design and by Default. Retrieved from https://edpb.europa.eu
16. Council of Europe. (2021). Artificial Intelligence, Education, and Human Rights: A Policy Brief. Strasbourg, France: Council of Europe.
17. Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1
18. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson Education.
19. Selwyn, N. (2022). Should Robots Replace Teachers? AI and the Future of Education. Polity Press.
20. UNESCO. (2023). AI and Education: Guidance for Policy-Makers. United Nations Educational, Scientific and Cultural Organization. Retrieved from https://unesdoc.unesco.org

Co-funded by the European Union

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. **Project N°:2023-1-IS01-KA220-SCH-000157308**

21. West, S. M., Whittaker, M., & Crawford, K. (2019). Discriminating Systems: Gender, Race, and Power in AI. AI Now Institute. Retrieved from https://ainowinstitute.org