COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

DGII/EDU/AIED(2025)02rev
Original : English
Strasbourg, 19 February 2025

Feasibility study for a European reference framework for the evaluation of educational technologies

Short version

prepared by Council of Europe consultants:

Irene-Angelica CHOUNTA, Beth HAVINGA, Maroun JNED, Dora KATSAMORI, Xenia ZIOUVELOU

CONTENTS

1 KEY DEFINITIONS

Adaptive Tutoring Systems or Intelligent Tutoring Systems (ITS) or intelligent interactive learning environments or personalised learning systems (NB Some of these terms are contested): AI-driven tools that might provide step-by-step tutorials, practice exercises, scaffolding mechanisms (e.g. recommendations, feedback, suggestions, and prompts), and assessments, individualised for each learner, usually through topics in well-defined structured subjects such as mathematics or physics.

AI Literacy: Having competencies in both the human and technological dimensions of Artificial Intelligence, at a level appropriate for the individual (i.e. according to their age and interests).

AI systems: Shorthand term encompassing AI-driven tools, applications, software, networks, etc. ***In this report, we use the term to refer to all such systems that are used within learning or educational contexts, and/or for learning or educational purposes.***

Artificial Intelligence (AI): Artificial Intelligence is notoriously challenging to define and understand. Accordingly, we offer two complementary definitions:

> A set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being. Current developments aim, for instance, to be able to entrust a machine with complex tasks previously delegated to a human. (Council of Europe, 2021)
>
> Machine-based systems that can, given a set of human-defined objectives, make predictions, recommendations, or decisions that influence real or virtual environments. AI systems interact with us and act on our environment, either directly or indirectly. Often, they appear to operate autonomously and can adapt their behaviour by learning about the context. (UNICEF, 2021)

Artificial Intelligence and education (AI&ED): The various connections between AI and education that include what might be called 'learning *with* AI', 'learning *about* AI', and 'preparing *for* AI'. Learning *with* AI has also been called 'Artificial Intelligence for education'.

Artificial Intelligence in education (AIED): An academic field of enquiry, established in the 1980s, that primarily researches AI systems to support learning (i.e. 'learning *with* AI').

Chatbots: systems designed to respond automatically to messages through the interpretation of natural language. Typically, these are used to provide support in response to queries (e.g. "Where is my next class?", "Where can I find information about my assessment?").

e-Proctoring/Proctoring: The use of AI-driven systems to monitor learners taking examinations with the purpose of detecting fraud and cheating.

Educators: Shorthand term encompassing teachers and other professionals in formal education and early childhood care, including school psychologists, pedagogues, librarians, teaching assistants and tutors.

Learners: Shorthand term to encompass children and young people in formal education (i.e. pupils and students) and people of all ages engaged in formal, informal or non-formal education (in accordance with the principle of lifelong learning) (Committee of Ministers, 2019)

Learning Analytics: Gathering, analysing, and visualising big data, especially as generated by digital devices, about learners and learning processes, with the aim of supporting or enhancing teaching and learning.

Robotics: Movable machines that perform tasks either automatically or with a degree of autonomy.

Machine Learning (ML): A type of AI, the type that is currently dominant, that uses algorithms and statistical models to analyse big data, identify data patterns, draw inferences and adapt, without specific step-by-step instructions.

Natural Language Processing (NLP) or Speech to text and Natural Language Generation: Systems that use AI to transcribe, interpret, translate and create text and spoken language.

Personalised learning systems: See Adaptive Tutoring Systems

Profiling: The automated processing of personal data to analyse or predict aspects of that person's performance, economic situation, health, personal preferences, interests, reliability, behaviour, location, or movements.

2 EXECUTIVE SUMMARY

2.1 Purpose

This feasibility study examines current practices in the evaluation and governance of Artificial Intelligence in Education (AIED) systems (hereafter referred to as AI Systems) in education, and their alignment with the Council of Europe's core values of democracy, human rights, and the rule of law. It maps key trends, risks, and stakeholder needs while analysing existing evaluation frameworks and identifying structural gaps. The study highlights areas requiring safeguarding, such as cognitive development, human oversight, digital citizenship, and democratic participation. Ultimately, it explores potential models for a European reference framework, proposing harmonised standards and governance mechanisms to ensure ethical, transparent, and effective AI integration in education.

2.2 Background

The Council of Europe has taken significant steps to address the intersection of Artificial Intelligence (AI) and education, with a focus on safeguarding its core values of human rights, democracy, and the rule of law. In 2022, the Council published the report "Artificial Intelligence and Education – A Critical View through the Lens of Human Rights, Democracy, and the Rule of Law," (AI&ED report), which examined the potential impact of AI on education. This report identified both opportunities and challenges posed by AI, emphasizing the need for ethical AI systems that respect data and intellectual property rights while addressing genuine educational challenges rather than perpetuating existing inequalities. The report highlighted the importance of evidence-based evaluations and called for the involvement of multiple stakeholders, such as teachers, policymakers, parents and administrators, in the development and deployment of AI systems. The report's findings were validated by a survey conducted among Council of Europe member states, reinforcing the necessity for regulatory measures specific to AI in education. As a result, and as agreed by the Education Ministers in 2024, the Education Department of the Council published a preparatory study advocating for a legal instrument to regulate the use of AI in educational settings.

In May 2024, the Council of Europe member states signed the Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law, reinforcing the commitment to uphold legal standards in AI deployment. While this convention addresses AI's impact broadly, both the AI&ED report and the Preparatory Study for the Development of a Legal Instrument on Regulating the use of AI in education underscore the specific relevance of AI in education, arguing that educational environments require targeted governance at a regulatory, ethical, pedagogical and human rights level. Concerns include, for example, the risks of AI reinforcing existing biases and limiting democratic participation by shaping educational experiences through biased algorithms. Although AI Systems are touted as providing potential for personalization and adaptive learning, challenges also exist related to bias and unequal access to quality education as well as undesired or unintended cognitive and pedagogical impacts.

Educational institutions and policymakers must ensure that AI applications comply with legal standards related to data protection, transparency, and fairness. Current regulatory frameworks, such as the EU's General Data Protection Regulation (GDPR), or the AI Act provide a foundation, but additional, sector-specific measures are needed to address the application of AI in education as a special case. This is due, for example, to the potential societal impact e.g., considering the unique vulnerabilities of children in educational settings or equity regarding access of learners to technology. There is a need to assess the legal, ethical, pedagogical and cognitive implications of AI in education, particularly concerning issues such as algorithmic bias, data usage, and the long-term impact on children's development.

Many countries are actively developing policies aimed at integrating AI into formal and informal educational settings, often as part of broader national digital transformation strategies. These policies frequently recognize education as a key sector for fostering inclusive digital access and democratic engagement. However, a significant gap remains in providing concrete guidance for evaluating and certifying AI Systems in education. Furthermore, although there has been a push for regulatory oversight, including age restrictions and data privacy safeguards, progress has been slow, with only limited policy action.

Recognizing this, there has been a growing call for supranational guidance on risk mitigation. At the September 2024 pre-conference workshops, representatives from 31 member states proposed a two-tier regulatory model—a common core of fundamental requirements for all member states, complemented by flexible, localized regulations. This aligns with repeated recommendations from the *Second Working Conference on Regulating the Use of AI Systems in Education*, where participants emphasized the need for a framework that balances harmonisation with national adaptability to reflect diverse educational priorities and ensure acceptance of the outcomes.

Given the rapid evolution of AI technologies, it is essential to establish a comprehensive legal, ethical framework evaluating the responsible and safe integration of AI Systems within education in member states. This framework must uphold democratic values and provide practical, flexible guidance that ensures that, if in use, AI Systems serve educational purposes while safeguarding the rights of learners and educators.

2.3 Areas of potential impact that require safeguarding

The rapid growth of AI in education presents a range of significant challenges that require careful management to ensure ethical and responsible use, which protects human rights and societal values. One primary concern is the protection of sensitive student data, which necessitates rights impact assessments (i.e. HUDERIA), robust cybersecurity measures and transparent data handling practices. Ensuring the quality, accuracy, and pedagogical soundness of AI Systems is also critical, requiring rigorous evaluation processes to prevent the dissemination of ineffective or harmful educational tools and practices. Educators must be adequately trained to identify, evaluate and, where applicable, integrate AI effectively into their teaching while maintaining pedagogical integrity. Additionally, equitable access to AI Systems, but most importantly human-mediated education, is essential to prevent exacerbating existing inequalities and the digital divide, particularly across different socio-economic and geographic regions. AI algorithms, if not carefully designed, risk perpetuating biases present in training data, leading to unequal educational opportunities and reinforcing social prejudices. The implementation of AI Systems also raises concerns about their impact on education, potentially undermining the autonomy and agency of teachers and students. Furthermore, adaptive learning technologies, while potentially beneficial for some learners, may inadvertently create disparities and have unintended consequences for other learners. These risks highlight the need for ongoing scrutiny and evaluation to ensure that AI systems in education support inclusivity, fairness, and the best interests of learners and educators alike (see chapters 4 and 6 for further details).

There is a critical need to safeguard key areas of educational and societal impact as AI Systems are integrated into education. Being able to effectively measure **the impact of AI Systems on cognitive development,** for example, is a pressing concern, and cognitive challenges through human-AI interaction must be adequately addressed. AI technologies, such as chatbots and social robots, may potentially promote problem-solving and engagement, but there is also concern that an **over-reliance on AI can undermine motivation, memory retention, and cognitive performance.** Excessive screen time and digital interruptions also pose risks, potentially leading to cognitive overload and conditions such as "digital dementia." This calls for careful monitoring and the implementation of cognitive-forcing mechanisms to mitigate the negative effects of AI-supported learning and decision-making.

Human involvement and control are equally crucial in AI deployment within education. Human-Centred Artificial Intelligence (HCAI) emphasizes the need to maintain human oversight and control to ensure that AI systems enhance rather than replace human capacities. This approach values human expertise, ethical considerations, and flexible thinking, which are essential to maintaining trust and preventing biases and unintended consequences. Despite the promises of AI System developers to provide personalized learning experiences, the importance of human educators remains central to fostering emotional and social intelligence. **The lack of clear evidence regarding AI's effectiveness**, compared to traditional educational methods, reinforces the need for a balanced, human-centred approach that sees AI Systems as potential supportive tools rather than a replacement for human educators.

Digital citizenship represents another area requiring focused attention, as AI significantly influences how individuals engage in the digital world. The Council of Europe emphasizes the need for digital literacy, ethical technology use, and active participation in digital society. **A robust review framework should, therefore, take into account competencies such as media literacy, privacy awareness, and ethical decision-making to empower students and educators in navigating AI Systems responsibly.** The European Year of Digital Citizenship Education in 2025 and recent policy resolutions highlight the importance of aligning AI education with human rights, democracy, and the rule of law. Additionally, competency provision is vital, as AI demands new skills in critical thinking, data literacy, and ethical decision-making. Various supranational frameworks, such as UNESCO's AI Competency Framework for Teachers, provide guidance on developing AI-related competencies, ensuring educators are equipped to assess and implement AI solutions effectively.

There is also an urgent need to safeguard democratic participation, as AI has the potential to be used to distort public opinion and spread misinformation. Developing media and AI literacy (covering socio-technical, ethical, and human rights aspects) curricula and promoting responsible digital engagement are essential to counter AI-driven manipulation and reinforce democratic values within educational systems.

2.4 Existing evaluation frameworks and gaps

One of the key methods for ensuring the quality of technology throughout procurement and implementation is independent evaluation which requires developers and service providers to submit various types of evidence about their products. While in some countries, such as Germany and Australia, the public sector has developed rigorous evaluation tools, these primarily address security, data protection, and technical aspects of AI Systems for education environments. There is little uniformity in how evaluation frameworks in general address the unique challenges posed by AI across multiple regions.

There are differing motivations and needs driving the current evaluation and certification landscape. This presents a challenge as there is a lack of commonly accepted benchmarks for assessing the quality of AI Systems. The various stakeholder groups prioritize different outcomes, ranging from academic rigour to market viability, leading to inconsistencies in how AI Systems are assessed and implemented. While public sector frameworks, for example, tend to focus on operational safety and compliance, they often don't directly address practical classroom needs. Conversely, private sector initiatives, although commercially validating and responsive to market demands, may lack impartiality or uptake by procuring bodies. Research-led evaluations, while academically rigorous, often fail to translate into practical implementation suggestions for educators and policymakers. This diverse ecosystem of evaluation efforts can be supported through the harmonisation of diverse perspectives and reference standards for an adaptable approach to AI evaluation in education.

Several structural gaps currently hinder the establishment of robust evaluation systems. The absence of widely accessible testing or co-creation environments and regulatory sandboxes limits the ability of developers to refine AI Systems based on real-world educational contexts,

while the lack of standardized, comparable information about AI Systems complicates decision-making for educators and procurement officials. This is compounded by inconsistencies in terminology and definitions of AI Systems in education across different frameworks, making it difficult to determine which technologies should undergo formal evaluation and certification. Addressing these gaps requires the development of clear guidance that can support stakeholders throughout the product life cycle: from development and procurement to implementation and continuous improvement. Establishing a European reference framework would help bridge these gaps, offering guidance regarding key evaluation criteria, whilst providing local flexibility.

2.5 Possible constructions and components of a reference framework

This feasibility study explores multiple approaches to developing a European Reference Framework for evaluating educational technologies, incorporating feedback from the *pre-conference workshops* and the *Second Working Conference on AI Regulation in Education*. It assesses the technical and practical feasibility of various models in addressing key challenges, for example, ethical compliance, pedagogical alignment, data privacy, transparency, and adaptability across diverse educational contexts.

2.5.1 Key feasibility findings

Benefits for stakeholders

A reference framework would provide guidance on common key issues for all member states whilst ensuring regulatory flexibility for areas that can be addressed differently locally or nationally. Additionally, a framework can provide clarity and benchmarks for evidence-informed decision making as well as AI Systems development for stakeholders.

Challenges to implementation

Despite the potential benefits, multiple barriers must be addressed including:

- Regulatory alignment: Member states have diverse education policies, making it difficult to establish a single, enforceable framework.
- Industry resistance: A formal certification mechanism may be perceived as an obstacle to innovation.
- Sustainability and funding: Maintaining a high-quality evaluation system requires long-term investment and coordination.
- Technical disparities: Varying levels of digital infrastructure and AI readiness across Europe could create implementation gaps.

Evaluation of proposed models

Several approaches were analysed for their feasibility:

- Self-regulation or simple standards development: Encourages industry responsibility but lacks enforceability, leading to inconsistencies in implementation.
- A formal Council of Europe certification mechanism: Ensures compliance and trust but requires significant financial and administrative resources.
- A harmonized framework for all member states: Would provide consistency but is difficult to implement due to legal and structural differences between countries.
- A European reference framework of evaluation (recommended model): A hybrid approach that balances consistency with flexibility, incorporating harmonized standards, and practical guidance to address the diverse needs of member states flexibly (see chapter 8).

2.6 Next steps and implementation actions

Building on these findings, the feasibility study suggests next steps for designing, implementing, and sustaining a European reference framework.

Designing and developing the reference framework

- Establishing a common core set of evaluation standards with a key focus on human rights, democracy and the rule of the law and establishing guidance on further criteria which can be adapted to local educational contexts.
- Developing a system of evaluation that accommodates varying levels of technological readiness.
- Providing guidance and support mechanisms, including non-binding guidance documents, toolkits, and best practices to assist educators, policymakers, and developers.
- Exploring the use of testing environments and sandboxes to allow controlled experimentation with AI Systems in education before widespread deployment.

Establishing the governance structure

- Creating a central governance body under the Council of Europe to oversee the framework's implementation, coordinate with national agencies, and ensure compliance with human rights and democratic principles.
- Defining roles and responsibilities of stakeholders, ensuring structured participation from educators, industry representatives, policymakers, and civil society.
- Implementing mechanisms for continuous stakeholder dialogue, allowing the framework to evolve based on technological advancements and emerging challenges.

Securing sustainability from the outset

- Developing multi-stakeholder feedback possibilities so that the framework upholds the ideas of fostering innovation whilst ensuring safety and the safeguarding of human rights, democracy and the rule of the law
- Ensuring dedicated budget and team capacity to enable regular reviews and changes, as well as to facilitate engagement with Member States and other key stakeholders.
- Developing support mechanisms to aid member states in their development of local review and evaluation mechanisms.

Additionally**,** the Council of Europe could develop and maintain a database covering national evaluation outcomes for AI Systems as well as key information that could support evidence-based decision making practices. This database could serve as a key reference tool for stakeholders, providing a centralized repository of vetted AI Systems

2.7 Conclusion

The development of a European reference framework for the evaluation of AI Systems is an important step in ensuring that AI Systems that are intended for deployment in education environments provide relevant evidence and align with European values, protecting the rights of learners and educators. An evaluation system will ensure that evidence-based approaches are used to support decision-making and policy development. The complexity of the current evaluation landscape calls for a structured yet adaptable approach. A framework which provides centralised guidance around key areas aligned with Council of Europe values and flexible implementation methods for member states regarding other criteria offers a balanced solution based on what Member States have requested and enables Member States to safeguard key potential areas of impact. This feasibility study offers suggestions for the successful implementation of such a framework. The extended executive summary outlines

the main options open to the Council of Europe for a reference framework and makes recommendations for a possible solution

3 POSSIBLE CONSTRUCTIONS AND COMPONENTS OF A REFERENCE FRAMEWORK

A European reference framework should aim to create a robust, fair, and transparent system for reviewing AI Systems, fostering innovation while maintaining high standards of education and ethical integrity. It should support the development of AI Systems that are not only effective but also equitable and responsible, ultimately contributing to a more inclusive and high-quality education system across Europe. Moreover, the framework should support the responsible use of AI Systems in education by addressing considerations such as data privacy, algorithmic transparency, and bias mitigation. By setting clear guidelines and benchmarks, the framework can help prevent the misuse of AI Systems and ensure that they contribute positively to the educational environment by providing necessary evidence to decision makers. Additionally, a reference framework can promote equity by ensuring that any AI Systems in use are accessible to all students, including those with diverse learning needs and backgrounds.

The primary objective of establishing a European reference framework for reviewing AI Systems is to ensure that these technologies are effectively evaluated and implemented according to multiple criteria. There are different methods that could be used to facilitate the creation and implementation of a reference framework for AI Systems. The pros and cons to each of these types will be explored here.

3.1 Self-regulation

As explored in the preparatory study for the development of a legal instrument on regulating the use of AI Systems in education (CoE, 2024c), self-regulation and voluntary guidelines or frameworks are often employed by AI stakeholders as a strategy to convince legislators that formal legal measures are unnecessary (Calo, 2017). This approach allows AI providers, along with their customers and users, to assert that they are addressing ethical concerns and questions surrounding AI, all without creating binding obligations or real-world impacts on AI applications and practices. However, this strategy brings up significant concerns, such as the absence of enforcement mechanisms, a lack of accountability, and unclear responsibilities to society at large. For instance, the public sector often views commercially driven practices as insufficient because they focus more on legitimising the product and are often driven by the companies themselves (Lindroos Cermakova et. al, 2024). An analysis of 22 major AI ethics guidelines highlighted that these principles or principled practice, while often well-intentioned, "are rather weak and pose no imminent threat" to AI stakeholders, offering little more than symbolic gestures (Hagendorff, 2020).

UNESCO has called on governments and ministries of education to build their own capacities for reviewing and validating AI systems, reducing the reliance on industry self-regulation (UNESCO, 2021). The Council of Europe's Office of HR Commissioner has called for the establishment of Independent Public Oversight Offices[1]. The vast majority of the European EdTech community embrace the idea of regulation with 87% of respondents to the European EdTech Map Survey[2] stating that they believe that evidence-based testing and regulation would help them (Havinga & Clary, 2024).

Whilst a code of conduct is also regularly used as a tool to support the implementation of certain regulatory measures, for example as part of the implementation of the GDPR by CNIL in France, or as necessary components for public private partnerships, these are also subject to the same criticisms as other self-regulation practices, as there are rarely standards for these or bodies that can uphold or enforce them.

1 https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64
2This is a yearly survey sent to the 1580 mapped EdTech Organisations be the European EdTech Alliance (www.edtechmap.eu)

3.2 Certification (CoE)

The creation of a certification program for AI Systems by the Council of Europe constitutes another possibility for safeguarding important values and principles while promoting quality and responsible AI innovation in education. Such a certification could build trust in AI Systems that demonstrate compliance with the fundamental values promoted by the Council of Europe such as democracy, human rights and rule of law, as well as developing criteria for the responsible development of AI Systems for education.

This certification would be the outcome of a rigorous evaluation process and present as a quality stamp/trust label associated with trust assurance and opportunities for the certificate holders. This could significantly accelerate the adoption and integration of AI Systems as it would empower technology developers, educators, institutions, and parents by alleviating their concerns around specific criteria.

Potentially, an open, online database of all the AI systems that have been awarded certificates could be created providing awareness and an ecosystem of trusted AI Systems. This could lead to multiple stakeholders feeling more empowered in their decision-making about AI Systems from the educator, parent and school to the EdTech organisations, ministries and national governments.

A certification process, providing a quality label to Youth Centres has been developed by the Council of Europe's Youth Department[3] and provides an example of certification practices. It would be important to clarify in this regard, however, how a Council of Europe certification for AI Systems can enhance or build upon local and other existing certification practices, align with both local and supranational regulatory measures, and ensure a harmonised approach. Additionally, the full scope of the certification must be determined, e.g., whether this measure should initially only support technologies entering foundational or mandatory schooling or also include segments like higher education or life-long learning. A Council of Europe certification would require personnel investments and a team with expertise in the evaluation criteria.

3.3 The European Pharmacopoeia

The European Pharmacopoeia[4] (Ph. Eur.) is an integral part of the European Directorate for the Quality of Medicines & HealthCare (EDQM[5]) that is the single reference work for the quality control of medicines in the signatory states of the Convention[6] on its elaboration. The purpose of the European Pharmacopoeia is to promote public health by providing recognised common standards for the quality of medicines and their components. Thus, forming a basis for the safe use of medicines by patients.

The European Pharmacopoeia publishes official quality standards for medicines and their ingredients in Europe, which provide a legal and scientific basis for quality control during the development, production and marketing processes. These global quality standard for medicines[7] stipulate the qualitative and quantitative composition and the tests that are to be carried out on medicines, on the raw materials used in production of medicines and on the intermediates of synthesis. All producers of medicines and/or substances for pharmaceutical

3 https://www.coe.int/en/web/youth/quality-label-for-youth-centres
4. https://www.edqm.eu/en/european-pharmacopoeia-ph.-eur.-11th-edition
5. https://www.edqm.eu/en/
6. The legal framework that make the European Pharmacopoeia mandatory, are: (1) the Convention developed by the Council of Europe on the Elaboration of a European Pharmacopoeia,(2) a Protocol adopted in 1994 and amending the Convention to prepare for the accession of the European Union and defining the respective powers of the European Union and its member states within the European Pharmacopoeia Commission, (3) European Union Directive 2001/83/EC as amended, on medicines for human use and and Regulation (EU) 2019/6 on veterinary medicinal products. These maintain the mandatory character of European Pharmacopoeia monographs when requesting marketing authorisation (MA).
7. These global quality standards for medicines address the needs of the regulatory authorities, those engaged in the quality control of medical products and their constituents and manufacturers of medical products and their individual components.

use must, therefore, apply these quality standards[8] in order to market their products in the signatory states of the Convention.

These standards are legally binding as laid down in the Council of Europe Convention on the Elaboration of a European Pharmacopoeia (CoE, 1974) and in European Union and national pharmaceutical legislation. These common standards become mandatory on the same date in all states, which are parties to the convention, thus ensuring consistent quality across Europe and facilitating the free movement of medicinal products within the region.

3.3.1 European AIEDutopeia

Following the example of the “European Pharmacopoeia" the creation of a European AIEdutopeia, could become the single reference work for the quality, including the safety, regulation conformity, promotion of human rights, democracy and the rule of the law, of AI Systems. Such an initiative would ensure that official quality standards (aligned with the Councils values) e.g. safeguarding ethical, pedagogical, and cognitive development criteria for AI in education in Europe are ensured.

This set of standards could be legally binding as in the case of Ph. Eur. or could adopt a more flexible scheme, offering an assessment framework without binding regulation. In the latter case an assessment framework could establish the benchmarks for quality, ethics, and pedagogical effectiveness among others acting as a trust guidance for the developers of AI Systems (for the design, development, deployment) as well as to educators and institutions (as a set of criteria for assessing AI Systems that they wish to adopt).

This approach, however, leads to absolutes of acceptable practice, e.g., either the product has the required ingredients and testing, or it does not, which may not effectively capture the development lifecycles of AI Systems or the desire for iterative co-creation through testing environments and sandbox approaches. This can lead to a system of products being described for either ‘on-label’ and ‘off-label’ use. Whilst an AIEDutopia approach would be extremely beneficial in regards to required testing and evidence before implementing a product in education environments or for education purposes, helping users better understand which implementation scenarios have been tested and approved by the developer and regulator, it may not only be prohibitive to smaller scale innovation practices, but also not reflect the requests of the member states for the Council of Europe to provide supranational guidance whilst enabling member state control and flexibility over evaluation.

3.4 A single, harmonised framework for all member states

Standards, evaluation mechanisms and certification practices all provide evidence to the procurement agencies and decision makers. The existing frameworks for the evaluation of AI Systems tend to address different needs, incentives, and stakeholders (Lindroos Cermakova et. al., 2024). Existing frameworks and standards for evaluation “vary greatly in detail and specificity both in the number of criteria being addressed and the manner in which perceived impact or changes in learning are measured” (Kucirkova et. al, 2023) and rarely accommodate multiple focus areas or stakeholders in their conceptualisation and implementation. Evaluation frameworks are also highly culturally specific, are trying to directly answer local needs (Lindroos Cermakova et. al, 2024), and are tied into local policy or societal requirements for the use of technologies in education (see chapter 3).

Whilst there is a need to have a more harmonised, interoperable and unified approach to dealing with the different focus areas and the evaluation of AI Systems used in education environments, the vast differences between existing and developing standards and evaluation systems would make the task of defining one, single and comprehensive framework extremely

8. For example, the 11th Edition is legally binding in 39 European countries as of 1 January 2023 and applied in more than 130 countries worldwide and is continually updated and modernised to meet users' needs.

difficult. The fact that these existing standards may be built on hyperlocal regulations or requirements also make them difficult to adapt to new environments: For some criteria, what is important to one region or nation may not have any place in the evaluation of technology in another. Additionally, this model would see the Council of Europe as an enforcement body and the manager of this framework with input from the member states. This, however, goes against the requests of the member states as gathered in the pre-conference workshops in October 2024, to provide a two-tier model of guidance and support[9].

3.5 A European standards and enforcement committee (CoE Standard setting bodies)

Standards will play a vital role in supporting compliance to any regulations by defining concrete technical requirements to adhere to. However, as noted in the preparatory study for the development of a legal instrument on regulating the use of AI Systems in education (CoE 2024c), general AI standards rarely address education-specific issues. This can increase issues of access and inclusion in educational settings and the devaluation of regional or minority languages and their protection and promotion contributing to the building of a Europe based on democracy and cultural diversity[10]. As mentioned in the European Charter for Regional or Minority Languages, regional or minority languages are part of Europe's cultural heritage and their protection and promotion contribute to the building of a Europe based on democracy and cultural diversity. Talking about AI Systems in education the values of accessibility and equity are of great importance, therefore the lack of equal access to advanced technologies, including AI, could lead to the widen of the existing educational disparities. One obstacle that could lead to educational disparities is the issue of language, as AI models and generative AI solutions are primarily based on (American)English, making it difficult to be tailored to the specific needs of each education system. Design and technology standards are also not obliged to uphold SDG (Sustainable Development Goals) 4 or to ensure that EdTech is inclusive, equitable and open to all, which could unintentionally exacerbate issues of unfairness, exclusion, and poor pedagogical practice: issues that have been raised in relation to the implementation of AI Systems. Whilst there are some AI standards under development[11], even with standards in place, mechanisms must still be developed to help all stakeholders ensure they are compliant, can be fairly evaluated, and the participation in these systems must be ensured by those most interested in protecting fundamental human rights and the public interest (McFadden et., al, 2021). **It is important to note that the standards developed by regular standards bodies are voluntary, non-binding instruments and often exist behind a paywall.** The current European standards committee focussed on Learning Technologies (CEN/TC-353), for example, does not currently cover any AI related topics, and has not suggested any new standards in the past two years.

3.6 A reference framework as guidance for member states to develop robust evaluation systems

The existing evaluation efforts of different stakeholder groups within member states and internationally range from research institutions and policy makers, through to private organisations or NGOs and educators These evaluation systems comprise of diverse and often incomparable criteria as the various stakeholders have different focal points, and motivations for the relative evaluation systems and their outcomes. A number of member states are already developing their own national review systems for EdTech or AI Systems but are also seeing

9 These workshops were carried out with 31 Member State representatives seeking to understand current regulatory practice and main areas of combined need.

10 The European Charter for Regional or Minority Languages https://www.coe.int/en/web/european-charter-regional-orminority-languages/about-the-charter Together with the Framework Convention for the Protection of National Minorities it constitutes the Council of Europe's commitment to the protection of national minorities

11For example, ISO (the International Standardization Organization) in sub-committee JTC1/SC42 https://www.iso.org/committee/6794475.html

the challenges in their own work to date[12]. Additionally, there is a call for supranational guidance and support in identifying key areas of risk and mitigation with respondents from the 31 member states represented at the pre-conference workshops in September 2024 expressing a need, for example, for primary and secondary regulations: a common core set and flexible localised sets. A solution to these challenges would be the creation of a reference framework for member states, which outlines core areas of review, responsibilities and suggestions for implementation based on examples thus offering supranational guidance whilst allowing for a secondary level of more flexible and localised implementation and alignment. This would enable member states to align their own evaluation systems with local requirements whilst maintaining internationally agreed standards for core criteria. It would also mean that the work of the Council of Europe would support the current developments in member states. **In this way, the member states themselves would be enforcing bodies of any evaluation systems, and the Council of Europe would provide clear guidance and regulation where appropriate and regarding a core set of agreed upon criteria.**

The reference framework should detail relevant areas that need to be reviewed, addressing the perspectives of multiple stakeholders to avoid the disconnect currently experienced, e.g. by educators trying to interpret reviews with a research focus, or by developers aiming to align their AI Systems with the needs of technical review processes, but also requiring these to be applicable to educational practice. **The framework must be developed with an understanding that not all AI Systems can or should be compliant with all criteria depending on their own scope and intended purpose and that certain types of assessment or evidence may look different across the development lifecycle of an AI System.** Furthermore, this framework should outline the relevance of existing legal requirements that must be taken into consideration.

It is the group's recommendation that the Council follows this approach.

3.6.1 Components of a reference framework

In May 2024, the Council of Europe adopted **a principle-based approach** framework convention on AI. This has the goal of,

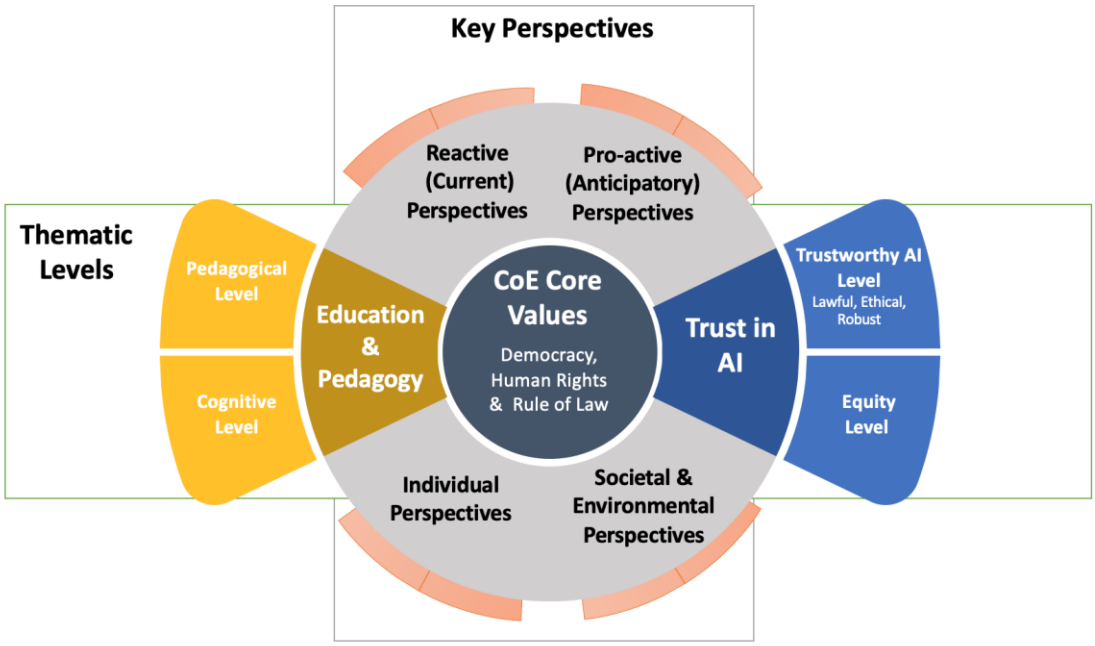
- Ensuring activities within the **lifecycle of AI Systems** are fully consistent with **human rights, democracy and the rule of law**, and
- Setting **inherent principles** related to **activities within the lifecycle of AI systems.**

When exploring the necessary components of a reference framework and ensuring this is aligned with the framework convention on AI, it is important to start from the **core values of the Council of Europe,** that is Democracy, Human Rights and the Rule of Law. From there, different **thematic levels**, which are important to the implementation of trustworthy AI in education (aligned with the European Commission, High Level Expert Group (HLEG, 2019[13]) and the safeguarding of users and education systems can be introduced. As well as different perspectives: that cover different time horizons - including current perspectives (reactive), anticipatory perspectives (proactive), and different ecosystem stakeholder perspectives, that is the perspectives of individuals, society and of the environment (see figure 2 as an illustrative example of such a framework).

12. A representative at the October conference on AI Regulation from a public sector evaluation mechanism stated, “it’s important to develop certification that has meaning but we need to communicate meaning on the level that is understood by teachers and in schools and there are current issues of communication, especially regarding who the recipient of the outcome is”.

13 According to the Guidelines of the HLEG, (2019) trustworthy AI should be (1) lawful, that is respecting all applicable laws and regulations, (2) ethical, that is respecting ethical principles and value and finally (3) robust both from a technical perspective while taking into account its social environment. According to the Guidelines, the 7 key requirements that AI systems should meet in order to be deemed trustworthy are: Human agency & oversight, Technical Robustness & Safety (by design), Privacy & Data Governance, Transparency, Diversity, Non-discrimination, & Fairness, Societal & Environmental Well-being, and Accountability (see EC, HLEG 2019, Ethics Guidelines for Trustworthy AI).

Figure 2 – An example of core values, thematic levels and key perspectives to be aligned in the development of a review framework



This example illustrates how the core Council of Europe values could be broken down across various dimensions and aligned with relevant principles as well as measures that can be put in place and evaluated against to determine adherence to these principles. A method like this can assist in defining the areas for evaluation, their connection to relevant principles and ways to ensure adherence to these.

3.6.2 Reflecting a common core of values, standards, and needs

In order to adequately outline the core areas that should be covered by a reference framework as well as the needs and requirements of the stakeholders that any evaluation is addressing, it will be necessary to conduct a comprehensive mapping of these ensuring that all of the main motivations for evaluation according to the various stakeholder groups are explored, and that the outputs from the evaluation mechanisms as well as alignments with existing regulations and review mechanisms are adequately addressing these. In this way, the reference framework can achieve a harmonisation of existing frameworks by supporting their local and contextualised development with supranational guidance. Additionally, it will be possible to ensure alignment with other, existing standardisation and regulatory bodies within local and European contexts.

3.6.3 Adaptability to local needs

As was suggested by Member States, there is a need to ensure that any measure offered by the Council of Europe could identify a common core set of requirements that must be addressed by all member states and a secondary, flexible set of requirements that should be addressed but may look different in various local contexts.

In this way, a reference framework answers the needs expressed by the member states and it is the recommendation of this group that the reference framework also list and outline all the key areas that should be addressed in a local evaluation mechanism of AI Systems, whilst leaving it to the member state as to how they answer these issues and address these areas. For example, the reference framework could adopt an approach of outlining areas and related needs from the relevant stakeholder groups that should be addressed in a local context with some core elements, being mandatorily standardised with others more flexibly described.

Additionally, it will be necessary to understand the type of evaluation and evidence requirements needed at different times in development lifecycles of products and what is possible to evaluate across those cycles[14].

14. For example, the structures found in the evaluation mechanism of AI for Education https://ai-for-education.org/qa-for-ai/

3.7 SWOT Analysis

Table 10 – SWOT analysis of possible evaluation solutions for the Council of Europe

	Strengths	Weaknesses	Opportunities	Threats
Self-Regulation	• **Speed** of creation and implementation can be accelerated as these do not undergo regular regulatory processes	• **non-enforceability** of the regulation • **potential to actually increase mistrust** due to a reliance on statements of the industry with no checks and balances. • **No mechanisms for redress** if there are any fundamental breaches of these regulations	• **potential to partner with the industry** to create a multi-tiered evaluation approach, which includes self-regulatory practices for certain, clearly identified components.	• **The ability for certain industry players to distort or warp regulatory practice to their own commercial needs and hide this behind a statement of trust.**
Certification (CoE)	• **Standardisation and quality** based on a set of clear criteria • **Increased trust and transparency** in AI Systems • **Enhanced adoption** of AI Systems • **Fostering Innovation** across the AIED developer ecosystem	• **Cost and financial implications** for maintaining this initiative • Need for **the continuous update of this standard** • **Certification scope** to be decided (limited scope or broad)	• **Collaboration potential w**ith key relevant stakeholders (educational institutions, AI expert groups, etc) • Potential to create a **Global Standard** • **Revenue generation** could balance costs and ensure sustainability (if the certification is fee-based)	• **Competition** (with existing certification bodies) • **Bureaucratic processes & Certification process bureaucracy**
European Pharmacopoeia	• **Harmonised** set of recognised common standards for the quality of AI Systems • **Globally accepted standardised quality control mechanism** aligned with the CoE values for AI Systems • **Increased public trust** in AI Systems from an Internationally accepted standards • **Legally binding standards**	• **Formation timeline for such an initiative** (long-process)	• **Knowledge transfer** from the European Pharmacopoeia • **Fostering the CoE core values for AI Systems** • **Lack of harmonised common standards** *(External factor)*	• **Industrial resistance from educational technology companies**

	Strengths	Weaknesses	Opportunities	Threats
A single European Framework	• **A common framework for all member states** • **The ability to add local context** 'on top' of standards	• **the different cultural, regulatory and motivational reasoning behind the various existing standards and evaluation mechanisms make it almost impossibl**e to develop a comprehensive common framework • **The Council of Europe would be an enforcing and certifying body** and would need to develop capacity for this	• **an analysis of existing regulatory measures** can increase the understanding of gaps, which could be filled by a Council-led approach	• **Member states will need to move faster** than the development of this minimum standard could take in its own development • **Member state cultural and motivational reasoning may not be able to be harmonised** • There are a number of commercial players currently exploring this opportunity
European Committee (CoE Standard Setting body)	• **Common standards** to align against • Developed from **Council Values**	• **Current standards committees are in existence** and working on aligned issues • **Standards bodies have limited success** as standards are often behind paywalls or only developed by certain communities • The Council may not have the capacity or mandate to manage such standards committees	• **A better alignment with and development of European standardisation work.**	• **Standards are only successful when there is wide scale adoption and a reason for the adoption** • The need for adoption would still have to be led through efforts like a guidance framework.
A reference Framework	• **Member states would be enforcing and certifying bodies** following a core set of agreed supranational measures and a flexible set of national measures • **The council would align and develop guidance for both the core and national measures**	• **The national measures may not be adopted in a uniform way** however this can be counteracted through the regulation and other framework conventions	• **To better align the work of the member states** • **To provide targeted support to member states through international agreement and guidance** • **Provide flexibility to the member states**	• **An inability to define the two sets of measurements** and a lack of infrastructure in member states to allow that to occur.

4 CONCLUSIONS / RECOMMENDATIONS AND NEXT STEPS

It is clear that significant gaps persist, challenging the alignment of AI Systems with the diverse and democratic goals of education. The divide between rigorous academic research and the practical application of EdTech tools highlights a pressing need for frameworks that not only provide robust evidence but also translate findings into actionable, context-sensitive guidance for educators, learners, and policymakers. Furthermore, the tension between commercial priorities and the ethical imperatives of equity and sustainability underscores the importance of democratic accountability in evaluation systems, which must go beyond short-term efficacy. Equally critical is the absence of universally adaptable metrics that respect regional and cultural diversity, as education systems demand flexible approaches tailored to their unique contexts. Finally, the lack of systematic mechanisms to integrate evaluation findings into large-scale AI Systems development processes hampers the ability to refine tools in real-world environments, limiting their potential to meet the evolving needs of learners globally.

A review of the different approaches to developing a reference framework for a review system of AI Systems has led to the recommendation of a reference framework as guidance for member states to develop robust evaluation systems. This would consist of both a core set of supranational evaluation criteria and a set of recommendations to be addressed flexibly in national practice. The alignment of these with core Council of Europe values, across clearly defined principles and thematically grouped should assist member states in the implementation of the framework. Furthermore, addressing requirements for evidence across the development lifecycle of AI Systems means ensuring the creation of evidence and adherence to evaluation throughout the entire development process and not only at certain stages.

4.1 Developmental measures and necessary steps

In order to develop a reference framework, it will be necessary to address the sustainability and governance structures of any evaluation mechanism to ensure their adoption and integration into the wider education ecosystem.

4.1.1 Sustainability

The framework must be considered as a living document, which requires an ongoing process for reconsidering its contents and structure at regular intervals with updates where appropriate. In order to ensure the sustainability of the reference framework, there must be a dedicated budget and capacity to manage at a minimum a yearly check and review of the reference framework e.g., through a pool of certified experts, the adoption and communication of any changes made to the reference framework, and the work together with the member states to address how the reference framework is working in implementation and what other support mechanisms are necessary to aid member states in their development of local review and evaluation mechanisms. Additionally, it will be essential to have multi-stakeholder feedback to ensure that the framework upholds the ideas of fostering innovation whilst ensuring safety and the safeguarding of human rights, democracy and the rule of the law.

4.1.2 Development of the reference framework

In a next step, the main structures and the key components of the core elements of both the supranational and national criteria for the reference framework will need to be identified and validated together in a multi-stakeholder process. Examples of implementation, alignment with existing regulatory frameworks, and requirements across the relevant milestones of the development lifecycle will also need to be addressed. This can be done with a core group of experts, a governance committee representing multi stakeholder perspectives and through regular (e.g. quarterly) validation sessions with the member states. How regularly evaluations should be repeated and the period of validity of an evaluation considering the rapid pace of AI Systems development must be defined. Likewise, it will be important to determine at what intervals the evaluation criteria need to be reassessed.

4.1.3 Governance

With the suggestion from member states for the Council of Europe to develop core evaluation and regulation criteria whilst providing guidance on elements that can be addressed more flexibly locally, it will be important to ensure governance structures for the development of the reference framework, which not only reflect the needs of these two distinct levels of responsibility, but also take a broader multistakeholder approach. In this way, it is recommended that a steering committee be developed, which can consist of representatives from research, pedagogy, policy, industry, and practical implementation (e.g. educators and learners), for example, to address feasibility and practicality, as well as determining relevance and aligning the framework with the needs of these stakeholders guided by the values of the Council of Europe.

During the phase of development of the reference framework, it will be imperative to conduct stakeholder group meetings, both individually and in multistakeholder groups to verify needs and assess feasibility of the relevant elements of the evaluation as well as to adequately identify the core, Council of Europe led elements and the flexible, member state led components of the reference framework. These stakeholder group meetings can take place online and require pre and post involvement to guide discussions and ensure feedback is appropriately captured.

4.1.4 Risk management

A key risk to the implementation of a reference framework is the extended timeline and potential for delay in the adoption of the proposed legal framework regulating AI and education. Should this become a barrier to implementation, it would be possible to refer to current legal frameworks and existing conventions whilst providing further guidance on gaps that have been identified in the preparatory study for the development of a legal instrument on regulating the use of AI Systems in education (CoE, 2024c). Additionally, an inability to reach consensus on which issues should be addressed as core, supranational criteria and which should be addressed more flexibly and at national or regional level could present a risk to the development of the guidance framework.

4.2 Support mechanisms for immediate implementation

While the reference framework is under development, there are steps that can be taken to support the ecosystem, some of which can be implemented quickly and provide immediate help.

4.2.1 Guidance documents and guidelines

Understanding the need for immediate support as member states and local policy and decision makers are in the process of developing their own systems for AI regulation, it is suggested that the Council of Europe prepare guidance documents in the form, for example, of guidelines (see 5.3.5. for some examples). Complementary to the Council of Europe's policy toolbox[15], these can address different stakeholder groups and can detail key areas of information that should be taken into consideration by member states in relation to these stakeholder groups and in order to ensure relevant areas of evaluation are covered and key issues are addressed.

These guidelines should not be limited to, but could outline the following topics:

- Developing and implementing AI Systems according to ethical considerations
- Understanding how to assess data privacy and security
- Risk assessment and mitigation policies

15. In development

- How to create regular AI auditing systems following supranational and national regulations
- Establishing continuous monitoring and feedback processes to ensure efficacy over time
- Understanding what information should be required and provided to maintain transparency
- Establishing the role of the educator, decision maker and user of AI Systems in human oversight requirements in practice
- How to assess technologies for practice (e.g. a guide to the right questions) based on the Council values and principles

Guidelines like these, explicitly guiding the development, procurement and implementation of new technologies exist in other sectors, such as the health sector, with examples such as the German guide for data protection measures in health data (BMWi, 2018) providing support in understanding the legal and data protection requirements, case-studies for comparison, and technical requirements and solutions for development. Documents like these could play an important role in ensuring appropriate and safe education settings whilst defining key contacts and case studies for a comparison of systems and opportunities.

4.2.2 AI literacy support

AI literacy is essential for navigating the assessment, use and potential integration of AI Systems into education. Building on existing ICT and digital competencies, AI literacy bridges the technological and human dimensions, equipping learners with the knowledge to understand AI Systems and their societal impacts. The Council of Europe emphasizes the importance of addressing AI's influence on human rights and privacy, while the European Parliament's AI Act highlights the need for interdisciplinary, inclusive approaches to AI literacy. UNESCO's AI Competency Framework advocates embedding AI learning into curricula, emphasizing human-centered interaction, ethical practices, and sustainability.

AI literacy must encompass not just technical knowledge but also an understanding of AI's impact on human and societal dimensions. Digital competence, as outlined by the Council of Europe, underpins these efforts, with added emphasis on preparing citizens to manage the complex challenges of AI in a rapidly evolving digital landscape. Raising awareness among all stakeholders is vital to fostering well-being and informed engagement with AI. It will be essential to provide guidance on this topic aligned with key aspects of the reference framework or to align with existing measures which cover relevant topics.

4.2.3 Structured Dialogue

Developing a European reference framework for reviewing AI Systems requires structured dialogue among stakeholders such as educators, policymakers, developers, and students. The multifaceted nature of AI Systems, from learning platforms to personalized tools, demands a balanced, inclusive approach to ensure the framework addresses ethical concerns, pedagogical impacts, and data privacy. By promoting transparency and accountability, a structured dialogue fosters trust, identifies shared values and challenges, and ensures the framework remains adaptive to technological and educational advancements. The active involvement of teachers, students, and parents is critical to successfully assessing AI Systems for use , enabling informed adoption and understanding risks.

4.2.4 Testing environments

Testing environments provide essential support for the evaluation and refinement of AI Systems, ensuring they meet the varied needs of European education systems. A model for developing integrated testing environments within Europe such as the Helsinki Education Hub (referenced in 5.3.1.) could enhance development standards by promoting innovation while

upholding safety and regulatory requirements and connecting into local research and policy networks. By establishing clearly defined framework architectures, these testing environments could strengthen public-private partnerships and encourage collaboration and resource sharing. Instead of a supranational testing environment, supporting local development could provide a sustainable way of ensuring ongoing development support and adherence to local regulations and cultural practices.

4.2.5 Sandboxes

By integrating Sandboxes as experimentation facilities for AI Systems as well as from a regulatory perspective (such as those of CNIL referenced in 5.3.2), it is possible to adopt a participatory, co-innovation approach in designing impactful AI Systems. This could be done both at a supranational level for core components of a Council of Europe led reference framework and also providing guidance for the development at local and inter-regional levels. Local sandboxes could also play a key role in providing bottom-up, evidence-based input for the updates of the reference framework. The support of existing regulatory sandboxes and the development of new possibilities to create sandboxes could be an essential support mechanism for the implementation of a reference framework.

4.2.6 Common database of evaluated AI Systems

There is currently no unified database or website of AI Systems or EdTech systems which clearly shows their evaluation certifications, areas of evidence that can be attested, or which outlines comparable data about AI Systems in a meaningful way for key stakeholders (as discussed in chapter 5.3.3). There would be an opportunity for the Council of Europe to develop and maintain a database which encompasses national evaluation outcomes and key areas of information regarding the AI Systems which can lead to evidence-based decision-making practices and provide a trusted source of information to all key stakeholders.