

# PROMOTING INFERENTIAL PROCESSES IN EDUCATIONAL CONTEXTS IN THE AGE OF ARTIFICIAL INTELLIGENCE

## PROMUOVERE I PROCESSI INFERENZIALI NEI CONTESTI EDUCATIVI NELL'ERA DELL'INTELLIGENZA ARTIFICIALE



### Double Blind Peer Review

#### Citation

Nuzzaci, A. (2025). Promoting inferential processes in educational contexts in the age of artificial intelligence. *Giornale italiano di educazione alla salute, sport e didattica inclusiva*, 9(Suppl.1).

#### Doi:

[https://doi.org/10.32043/gsd.v9i2\\_Sup.1557](https://doi.org/10.32043/gsd.v9i2_Sup.1557)

#### Copyright notice:

© 2024 this is an open access, peer-reviewed article published by Open Journal System and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

[gsdjournal.it](http://gsdjournal.it)

ISSN: 2532-3296

ISBN: 978-88-6022-515-3

Antonella Nuzzaci  
University of Messina (ITALY)  
[antonella.nuzzaci@unicme.it](mailto:antonella.nuzzaci@unicme.it)



### ABSTRACT

In the age of artificial intelligence, fostering inferential processes presents a crucial educational challenge. This paper explores how different forms of inference—deduction, induction, and abduction—are embedded in both human reasoning and computational models and examines how they can be effectively promoted in school and university settings. By comparing natural cognition with AI system architectures (symbolic, sub-symbolic, and neuro-symbolic), the article proposes instructional strategies, digital tools, and learning environments designed to support critical thinking, intuitive reasoning, and epistemic responsibility. Special attention is given to Explainable AI (XAI) as a pedagogical lever, to the cultivation of metacognitive skills, and to the role of artificial intelligence as an epistemic partner. The aim is to outline an educational model that integrates logical rigor, inferential creativity, and responsible digital citizenship.

Nell'era dell'intelligenza artificiale, promuovere i processi inferenziali rappresenta una sfida educativa cruciale. Questo articolo esplora come le diverse forme di inferenza — deduzione, induzione e abduzione — siano presenti sia nel ragionamento umano sia nei modelli computazionali, e analizza come possano essere efficacemente sviluppate nei contesti scolastici e universitari. Mettendo a confronto la cognizione naturale con le architetture dei sistemi di IA (simbolica, sub-simbolica e neuro-simbolica), l'articolo propone strategie didattiche, strumenti digitali e ambienti di apprendimento finalizzati a sostenere il pensiero critico, il ragionamento intuitivo e la responsabilità epistemica. Particolare attenzione è rivolta all'Intelligenza Artificiale Spiegabile (XAI) come leva pedagogica, allo sviluppo delle competenze metacognitive e al ruolo dell'IA come partner epistemico. L'obiettivo è delineare un modello educativo che integri rigore logico, creatività inferenziale e cittadinanza digitale responsabile.

Received 03/07/2025

Accepted 22/07/2025

Published 30/07/2025

## Introduction

In the contemporary era, where artificial intelligence (AI) is becoming increasingly prevalent in various areas of knowledge, training, and daily life, it is becoming increasingly urgent to question how it is redefining inferential processes in cognitive and educational contexts. AI systems do not merely execute technical tasks; they make inferences—deductive, inductive, or abductive—that inform decisions, classifications, recommendations, and user interactions.

By 'inference', we mean the process by which a conclusion is reached from given or perceived information. In education, this process is fundamental: any activity involving understanding, critical judgment, problem-solving, or conceptual creation is based on some form of inference. The ability to think inferentially is therefore one of the fundamental pillars of human intelligence and deep learning (Johnson-Laird & Byrne, 1991).

In the context of AI, symbolic models, which are based on formal logic, and subsymbolic models, which are based on neural networks and pattern recognition, incorporate inferences in different ways. For example, the former do so explicitly and the latter do so in an opaque and often uninterpretable way. This gives rise to the contemporary paradox whereby AI can outperform humans in specific tasks without explaining its reasoning, resulting in powerful yet opaque systems (Pedwell, 2023).

This inferential opacity is particularly problematic in educational settings where AI is integrated into adaptive learning environments, automated tutoring systems, and predictive assessment and knowledge representation tools (Ane & Nepa, 2024).

In such environments, it is crucial that students and faculty can not only use AI outputs but also understand the underlying logic. The ability to interrogate AI inferences, assess their coherence, recognize their limitations, and engage in critical dialogue with them has become a new literacy necessary for contemporary education (Holmes, Bialik, & Fadel, 2019; UNESCO, 2019).

According to Kahneman (2011), human thought operates at two levels: quick and intuitive 'System 1', and reflective and deductive 'System 2'. Education systems must therefore foster environments in which both systems are cultivated, and AI is used to stimulate richer inferential processes, rather than to replace them. Recent research by Chen et al. (2023) demonstrates the fundamental role of human intuition in evaluating the reliability of AI-generated responses. Three forms of intuition emerge in this sense: intuition about the result (does the answer make sense?), intuition about salient characteristics (what matters?), and intuition about the limits of AI (when can AI make mistakes?). Teaching these skills involves fostering critical and metacognitive thinking, as well as higher-order skills.

Indeed, the introduction of AI into educational practices cannot be viewed as merely a technological improvement; it necessitates a redefinition of training paradigms and brings new epistemological, ethical, and didactic questions to the forefront. As Polanyi (1966) demonstrated, intuition is not an irrational residue, but rather a form of

inference based on tacit knowledge and the recognition of meaningful patterns. In the age of AI, educating for inference therefore also means educating for critical intuition, contextual evaluation, and epistemic responsibility, which is understood as the duty to critically justify information, especially when produced by automatic systems (Floridi, 2023). In a school setting, this involves teaching students to be skeptical of the responses of machines.

This study aims to analyze these issues by providing a theoretical reconstruction of the main inferential models (deduction, induction, abduction), paying particular attention to their implementation in intelligent systems and the differences in comparison with human thought. It also examines emerging computational models that attempt to integrate intuition and logic, focusing on the formative value of neurosymbolic intelligence and the challenge of explainability. Finally, it reflects on strategies to promote inferential education in the age of AI through curricula, adaptive environments, and teaching practices inspired by the 'epistemology of uncertainty'.

Ultimately, as Umberto Eco (1979) teaches us, every interpretative process is an 'inferential walk', a heuristic path guided by hypotheses, clues, and changing contexts in which meaning is constructed through the interplay of logic and intuition. The future school's role will be to teach students to consciously navigate these processes alongside others and intelligent technologies. Educating people in inference, critical reflection, and collaboration with AI means training them to be readers of the world who can navigate uncertain information, and automatic explanations and make conscious choices.

## **Inferential models between human cognition and computational architecture**

### *The nature of inferential processes*

Inferential thinking is a fundamental function of the human mind. It enables us to form judgments and draw conclusions based on incomplete, ambiguous, or uncertain information, as well as develop hypotheses and make decisions. From a logical and cognitive point of view, inferential processes are traditionally divided into three main categories: deduction, induction, and abduction. Each represents a different way of proceeding from the known to the unknown.

Deduction is the process by which one arrives at conclusions that follow logically from certain premises. It is a form of closed reasoning, typically employed in axiomatic systems, mathematical logic, and symbolic artificial intelligence. A deductive system is valid if the conclusions are necessarily true given true premises. However, precisely because of its rigidity, the deduction can be inflexible in real-world contexts where information may be incomplete or contradictory (Johnson-Laird & Byrne, 1991).

**Box: 1****Deduction – Didactic example (Science: states of matter)****Background: Scientific rules**

The teacher introduces two rules on the behavior of substances:

1. **Rule 1:** If a substance changes from liquid to gas at room temperature, then it is a **volatile substance**.
2. **Rule 2:** All volatile substances **evaporate easily**.

Case: Ethyl alcohol evaporates at room temperature (i.e. changes from liquid to gas).

**Step-by-step deduction**

1. **Premise 1:** Ethyl alcohol changes from liquid to gas at room temperature.  
→ **Conclusion 1:** Ethyl alcohol is a **volatile substance**.
2. **Premise 2:** Ethyl alcohol is a volatile substance. → **Conclusion 2:** Ethyl alcohol **evaporates easily**.

**Final deductive conclusion:**

Ethyl alcohol evaporates easily.

**What happens if information is missing?**

If a student only knows that alcohol evaporates, **but does not know the definition of "volatile,"** then **he cannot deduce** whether he belongs to that category or apply the second rule.

**Result:** The deductive process is interrupted, as happens in expert systems if data is missing.

**What students learn from this example**

- Scientific **rules** can be used to make **deductive logical reasoning**, not just observations.
- Deduction helps to **link cause and effect** in a rigorous way.
- It teaches how **rule-based automated systems** reason: you need all the information to reach a conclusion.
- It is useful for **assessing the consistency of scientific explanations** and intelligent models.

**Instructional Tip:** This type of activity can be used as an introductory exercise in a logic or science module. It is useful to have students build their deductive chains on real cases. Possible assessment: Ask students to identify missing premises or check the consistency of inferences.

**Useful digital tools** to develop deductive thinking include platforms such as ASSISTments, GeoGebra (with rules), which offer structured and traceable exercises.

**Useful reference:**

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Table 1. Didactic example of deduction: states of matter

Induction, on the other hand, involves generalizing from specific examples. It forms the basis of most scientific methods and human learning: we observe patterns and use them to infer general laws. However, while induction is a powerful tool for building knowledge, it carries a margin of uncertainty; new observations can challenge or invalidate previous generalizations (Bruner, 1961). In education, promoting inductive thinking means developing the ability to observe, abstract, formulate hypotheses, and recognize patterns in new situations, thus facilitating flexible and critical learning.

## Box: 2

### Induction – Teaching Example (Science: Heat Conduction)

#### Background: Laboratory observation

During a classroom experiment, students observe that:

1. A **metal** spoon heats up quickly when submerged in hot water.
2. A **copper** bar heats up quickly when touched with a heat source.
3. An **aluminum tube** heats up quickly when exposed to the sun.

#### Inductive reasoning

From these **observations**, students notice a regularity:

→ All **observed metals** heat up easily.

#### General (inductive) conclusion:

**Metals conduct heat well.**

#### Limits of induction

This **generalization** is **plausible**, but **not certain**:

- If in the future a metal is found that **does not** conduct heat well, the rule should be **revised**.

Induction is always subject to **verification and revision** based on new observations.

#### What students learn from this example

- That inductive thinking arises **from the observation and comparison of real cases**.
- Which is fundamental in **the construction of scientific knowledge**.
- That inductive conclusions are **open, flexible**, and must be **put to the test**.
- That recognizing **patterns and regularities** is essential for critical learning.

**Didactic tip:** In the classroom, this inferential form can be stimulated by proposing open-ended problems, experiments with unexpected outcomes, or guided simulations. Students can compare different hypotheses to explain an event and assess its plausibility. An evaluation rubric can include originality of the hypothesis, consistency with the data, clarity of exposition.

**Digital tools** to enhance inductive reasoning, it is possible to use interactive virtual laboratories such as those of the PhET platform.

#### Useful reference:

Bruner, J. S. (1961). *The Act of Discovery*. Harvard Educational Review, 31(1), 21–32.

Table 2. *Induction Teaching Example: Heat Conduction*

Abduction, as theorized by Charles S. Peirce (1931–1958), is the process of inference by which the most probable explanation for an observed phenomenon is sought. Unlike deduction, which starts with general rules to arrive at conclusions, and induction, which generalizes from observations, abduction starts with a surprising or unexpected fact and looks for a plausible hypothesis to explain it. While it does not guarantee certainty, it opens up space for hypothesis, intuition, and discovery. It is particularly relevant in scientific research, as well as in learning and problem-solving processes.

### Box: 3

#### **Abduction – Didactic example (Science: an unexpected phenomenon)**

##### **Background: Laboratory observation**

During a laboratory activity, students notice a surprising phenomenon:

In a plant that is kept in the dark for several days, the leaves begin to **turn yellow**.

It was not expected: the other conditions were normal (water, temperature, air).

##### **Abductive reasoning**

At this point, students ask themselves a question:

"Why are the leaves turning yellow, if the plant has adequate water and temperature?"

Based on their knowledge, they formulate a **possible explanation**:

→ "Maybe the lack of light prevents the plant from doing photosynthesis, and this causes the leaves to turn yellow."

##### **Plausible hypothesis (abduction):**

The **lack of light** could explain the observed phenomenon.

##### **Characteristics of abduction**

**It starts from an unexpected or anomalous fact** (yellowing leaves).

**Look for a plausible**, but not certain, hypothesis.

**It can guide the design of an experiment** to test the hypothesis (e.g. comparing two plants: one in the dark, one in the light).

##### **What students learn from this example**

That **abduction is useful for formulating initial explanations**, when you do not yet have all the information.

Which is a form of reasoning **close to intuition and scientific discovery**.

That the hypothesis formulated must then be **verified with subsequent experiments or observations**.

That it is an approach widely used in **diagnosis, problem solving, research and creativity**.

**Didactic tip:** In the classroom, this inferential form can be stimulated by proposing open-ended problems, experiments with unexpected outcomes, or guided simulations. Students can compare different hypotheses to explain an event and assess its plausibility. An evaluation rubric can include: originality of the hypothesis, consistency with the data, clarity of exposition.

**Digital tools** to stimulate abduction in educational contexts, chatbots or generative systems such as ChatGPT can be used.

**Useful reference:**

Abduction is a form of creative, hypothetical inference that plays a fundamental role in forming hypotheses and intuitive understanding. In a school setting, activities that stimulate abductive reasoning include reconstructing a historical event or solving an open scientific problem. For instance, when presented with unusual data that contradicts the accepted version of a historical fact, students are encouraged to propose and compare plausible hypotheses to explain them and evaluate their consistency. By doing this, they learn to construct explanations based on clues, context, and interpretive logic, rather than simply looking for 'the right answer'. Abduction is also the form that most closely resembles how modern AI systems, especially those based on neural networks, generate 'plausible' results in the absence of explicit rules. For instance, an AI

system that generates image captions might look at a photo of a wet dog next to a puddle and produce the phrase "The dog played in the rain". There are no explicit rules linking each visual configuration to a specific sentence; instead, the system formulates an educated guess based on learned data and context through abduction. As with human abduction, this is a 'reasonable' conjecture, but not necessarily true. Such dynamics are also evident when interacting with conversational interfaces such as ChatGPT. These interfaces demonstrate that abduction is operational in language models (Hassani & Silva, 2023).

In an educational context, the three types of inference are dynamically intertwined. Teaching cannot be limited to transmitting deductive knowledge; it must also encourage inductive exploration and abductive flexibility. As Kahneman (2011) points out, humans unconsciously alternate between fast, intuitive processes and slow, reflective processes, corresponding to the dynamics of System 1 and System 2. Educating individuals to be aware of these processes is essential to train them to think autonomously and critically, even when interacting with intelligent systems.

Therefore, inferential models are not only logical abstractions, but also cognitive and didactic matrices that structure the relationship between knowledge, understanding, and decision-making. It is therefore essential to understand how these models are implemented — or replaced — in AI systems, and the risks and opportunities they entail in a training context. For instance, in an environment where learning is supported by an AI-powered intelligent tutor, the system can observe how a student responds to a variety of mathematical problems. Based on this data, the AI can infer (by abduction or induction) that the student has misunderstood a specific concept, such as the distributive property. The AI does not merely apply fixed rules, but rather builds a hypothesis based on the observed behavior and proposes targeted activities to address the identified gaps. A similar approach is described by Ane and Nepa (2024), who propose a predictive model representing students' knowledge and adapting educational interventions through inferential metrics based on Bloom's taxonomy (Ane & Nepa, 2024; Di Tore, 2023).

This demonstrates that the inferential models adopted by AI are not merely abstract logical schemes; they have a concrete influence on teaching strategies, formative assessment, and the type of educational relationship established in these contexts. Understanding these mechanisms is crucial for critically assessing the impact of AI on education.

Type of inference	Theoretical definition	Didactic example	Classroom application	Learning objective
Deduction	From general rules to firm conclusions	Case: Ethyl alcohol	Construction of deductive rules with disciplinary materials	Logical rigorosity, consistency
Induction	From observed cases to general rules	Heat conduction	Guided observation, discussion of generalizations	Critical thinking, flexibility
Abduction	Best Plausible Explanation	Yellowing leaves	Search for multiple hypotheses, comparison of explanations	Creativity, hypothesis, exploration

Table 4. Summary table "practical theory  $\leftrightarrow$ " for each inferential form

### Automatic deduction in symbolic systems

In the early days of AI development, deduction formed the theoretical and operational basis of "symbolic AI" systems. These systems were based on first-order logic and formalized inference rules, as well as explicit knowledge representation (McCarthy & Hayes, 1969). These systems operated in a manner analogous to formal human reasoning: given certain premises and logical rules, they could infer valid conclusions deterministically. This approach was fundamental to the development of expert systems in fields such as medicine, law, and technology, where traceable and verifiable decision-making was essential (Davis & Lenat, 1982; Feigenbaum & Lester, 1977). However, while automatic deduction is powerful in closed and well-structured environments, it shows important limitations in open, complex, and dynamic contexts, such as educational ones. Its logical rigidity renders it unsuitable for managing the ambiguity, incompleteness, and variability of human knowledge. In education, for instance, authentic learning frequently involves dealing with unforeseen situations, ill-defined problems, and evolving concepts, which cannot be resolved using purely deductive logic.

Furthermore, automatic deduction suffers from a kind of 'cognitive weakness': it lacks intentionality and the ability to construct meaning and learn autonomously from experience. This clearly distinguishes it from human deduction, which is intertwined with intuition, context, and the ability to select relevant information. Therefore, integrating tools based on formal deduction into educational practices alone risks producing a 'sterile logic', devoid of conceptual depth and disconnected from students' real educational needs.

In the field of education, however, deduction can play an important role when used as the subject of metacognitive reflection. Teaching students the principles of formal logic, such as syllogisms and conditional inferences, can help them develop argumentative skills and critical thinking abilities. However, for these tools to be effective, they must be integrated into a broader framework that also encompasses



open, intuitive, and contextual inferential processes, such as induction and abduction. Only in this way does deduction fully assume its didactic value, favoring integrated learning that meets the real needs of students.

This need for integration also reflects the ongoing transformations in the field of artificial intelligence: sub-symbolic AI is based on neural networks, whereas symbolic AI is based on logical rules, and neuro-symbolic AI integrates the two. In shifting from symbolic AI to sub-symbolic AI, the scientific community has recognized that to address real-world complexity it is necessary to move beyond automatic deduction toward models that incorporate learning, intuition, and adaptation. This shift has profound implications for education: it is no longer enough to teach what to know; we must also teach how to infer, when to trust inferences, and when to question them.

### **Abduction and intuition in contemporary computational models**

Unlike deduction, which guarantees the logical validity of conclusions, and induction, which identifies patterns in data, abduction is a form of exploratory and often incomplete inference. It involves formulating the most plausible explanation for a set of observations, even when certain rules or complete data are lacking (Peirce, 1931–1958). Abduction is therefore particularly suited to contexts of uncertainty, ambiguity, and discovery, such as those that characterize creative thinking and authentic learning. During a science lesson, the teacher proposes a simple experiment: leave a glass of water out for a day. The next day, the pupils observe that the water level has decreased. Confronted with this unexpected outcome, the children formulate various hypotheses: perhaps someone drank the water, maybe the glass has a hole, or perhaps the water evaporated. Through observation and comparing hypotheses, they arrive at the most plausible explanation: evaporation. In this context, abduction is a useful inferential process for exploring and understanding uncertain phenomena, favoring active and authentic learning.

In the field of computing, this form of inference has inspired the development of models that integrate pattern matching, tacit knowledge, and learning from experience to generate 'intuitive' results, i.e. plausible but not logically guaranteed (Pedwell, 2023; Ignatiev, Narodytska, & Marques-Silva, 2019). These models form the basis of a concept now known as 'artificial intuition', which is closely linked to the use of abductive inference in computational systems. Here, machines are designed to generate plausible hypotheses based on previous experience, incomplete patterns, and tacit knowledge like human intuition (Zhang et al., 2023).

Artificial intuition is the ability of a computational system to generate non-deductive, yet plausible, solutions or hypotheses based on past experiences, incomplete patterns, or implicit knowledge in ambiguous situations or when dealing with partial data. It is an emerging concept in the fields of explainable AI (XAI), machine reasoning, and cognitively inspired models. Rather than merely computing, these models enable

machines to 'reason' by approximation, much like humans when faced with an uncertain problem. Such systems can swiftly generate adequate responses in situations of incomplete information, thereby mimicking the dynamics of System 1, as described by Kahneman (2011).

Artificial intuition is particularly useful in situations where time or cognitive resources are limited, making systematic deduction impractical. In education, for instance, this manifests as systems offering immediate feedback and dynamic, personalized recommendations within adaptive environments that anticipate students' requirements and challenges. However, these systems often operate in an opaque way. The inferences they generate cannot be explained according to formal rules, and their effectiveness relies more on perceived coherence than logical transparency. This raises one of the crucial challenges of training: how can we educate people to use artificial intuition critically? How can we help students to discern between a 'plausible' answer and a 'well-founded' one, or between effective intuition and systematic error? One answer lies in promoting 'didactics of uncertainty', which values hypothetical reasoning, formulating alternatives, and reflecting on the conditions of an explanation's validity.

Recent educational literature encourages us to recognize the value of intuition, especially when combined with critical reflection and metacognitive analysis, rather than demonizing it. Studies such as that of Chen et al. (2023) demonstrate that human intuition can act as an epistemic filter when interacting with AI, helping to evaluate the relevance, accuracy, and limitations of algorithmic inferences. In school and university contexts, this highlights the urgent need to design activities that stimulate logical-formal thinking and the ability to formulate hypotheses, evaluate explanations, and consider alternatives. In light of recent research (Luckin, 2018) highlighting the role of intuition as an epistemic filter in interaction with artificial intelligence, it is important to promote a didactic approach that integrates intuitive thinking with critical reflection and metacognitive awareness.

Activity title	Grade Level	Description	Objectives
AI-supported source analysis lab	Secondary / University	Students evaluate AI responses on text or images, deciding whether they are relevant and correct.	Critical evaluation, metacognition, use of epistemic intuition
Debate on AI-generated hypotheses	Secondary / University	Starting from an educated hypothesis proposed by AI, students prepare a debate for and against.	Critical thinking, argumentative skills, epistemic reflection
Alternative scenarios imagined from AI content	Primary / Secondary	Students imagine "what if..." starting from AI explanations, building alternative hypotheses.	Creativity, hypothetical-deductive thinking, awareness of conceptual assumptions
Metacognitive diary on interaction with AI	All levels	After an activity with AI, students reflect on what convinced them, where they had doubts, and how they evaluated.	Self-regulation, metacognition, awareness of one's own inferential process

Table 5. *Educational activities for the development of inferential thinking in the presence of AI*

In summary, abduction constitutes a crucial bridge between human and artificial intelligence: it is the inferential form that most closely aligns with intuition, creativity, and the formulation of plausible hypotheses in uncertain contexts. Precisely for this reason, it demands a structured educational approach aimed at fostering awareness, sharing interpretive strategies, and critically managing inferential processes. This is particularly relevant in scenarios where AI systems engage not only in logical reasoning, but also in the emulation of affective and intuitive dimensions, as demonstrated by recent studies on generative artificial intelligence and its capacity to interpret emotional content from visual stimuli (Bilotti et al., 2023).

### **Towards integration: neuro-symbolic models and mixed inferential training**

Contemporary research in artificial intelligence is increasingly moving towards hybrid architectures, particularly so-called neuro-symbolic ones, to overcome the limitations of purely deductive models and the opacity of statistical models. These architectures combine the robustness of symbolic logic with the flexibility of sub-symbolic learning (Valiant, 2020). They integrate the generalization ability of neural networks with formal and interpretable logical structures.

This integrated approach enables knowledge to be represented hierarchically, moving from experiential patterns (intuitive, context-sensitive inferences) to more stable symbolic structures (deductive inferences), thus creating intelligent systems capable of reasoning, learning, and explaining. In the field of education, this development has significant implications, suggesting the need to transcend the traditional dichotomy between rational and intuitive thinking and promote a mixed inferential approach that fosters interaction between different modes of cognition.

An educational model inspired by these architectures should be able to:

- cultivate sensitivity to patterns and contextual significance, through exploratory activities, simulations, case studies, and epistemic games;
- gradually formalize intuition, helping students to transform spontaneous hypotheses into structured arguments and generalizable rules;
- explain the underlying rules and schemes, stimulating metacognitive awareness and the ability to analyze one's inferential processes;
- integrate explainable AI (XAI) tools into teaching, to make machine inference paths visible and compare them with human ones (Gunning et al., 2019).

This educational perspective aligns with Polanyi's (1966) definition of "integrated knowledge", which encompasses both tacit and formal dimensions. It involves a dynamic balance between intuition and explication and encompasses both "knowledge by eye" and systematized knowledge. In school and university contexts, developing this type of integrated inferential competence in students involves equipping them with the cognitive tools necessary for navigating complex, ambiguous, and AI-assisted environments, and fostering the development of technical, epistemic,

and ethical skills.

Finally, neuro-symbolic models can inspire adaptive, cognitively augmented learning environments in which AI actively collaborates with students to build inferences, explain alternatives, and refine thinking, rather than merely evaluating or suggesting content. In this sense, AI can become an ally of educational inference if education equips individuals with the ability to recognize, question, and use its logic consciously.

## **Inference, ai, and educational challenges**

Analyzing inferential models (deduction, induction, and abduction) and their applications in AI highlights a fundamental epistemic tension: the need for formal, logical, and explainable systems versus the urgency of responding to complex, uncertain, and open situations, which require flexibility, intuition, and adaptation. Current computational architectures, particularly neuro-symbolic ones, attempt to reconcile these opposing needs by integrating rigor and experiential learning, but they also present new challenges to educational thinking.

Contemporary education is therefore faced with a dual challenge: to teach students to construct robust, coherent, and reasoned inferences, and to enable them to comprehend, evaluate, and engage in dialogue with the inferences produced by intelligent systems, which are frequently neither linear nor easily explainable. This is especially urgent in light of the potential cognitive, ethical, and social risks that artificial intelligence may introduce into learning environments—risks related to bias, automation dependency, and a lack of transparency in decision-making processes (Zanetti et al., 2020).

Effective inferential education must consider the epistemic uncertainties generated by algorithms. Recent contributions by Suresh and Guttag (2024) highlight the fact that the recommendations of intelligent systems are not always transparent.

From a pedagogical perspective, this requires overcoming disciplinary fragmentation and promoting environments that stimulate critical thinking and reflection, as well as the ability to apply different inferential strategies depending on the problem at hand. As a cognitive activity with high epistemic intensity, inference should become a cross-curricular objective, guiding learning between disciplinary rigor and openness to complexity, and connecting subjects such as science, language, mathematics, philosophy, and art, where processes of meaning construction are present.

Furthermore, the increasing interaction between students and AI, whether in learning analytics, adaptive tutoring systems, or automatic correctors, makes it imperative to educate individuals in algorithmic interpretation. This will enable them to not only use but also interrogate, understand, and contextualize the inferences produced by machines. In this context, inferential education becomes an act of civic literacy that is fundamental to responsible and critical digital citizenship. From this perspective, inferential education is intertwined with the ethics of artificial intelligence, as

highlighted by Floridi (2023), which encourages us to educate citizens who can recognize the normative implications of algorithmic decisions.

Therefore, it is important to explore concrete educational strategies that promote inferential processes in the age of artificial intelligence. These strategies should combine rigor and openness, rationality and intuition, as well as autonomy and human-machine collaboration, and be implemented through curricular approaches, digital tools, and teaching practices.

## **Strategies to promote inferential processes in educational settings with ai**

### *Rethinking Training in the Age of Machine Inference*

The increasing use of artificial intelligence in education—through personalized learning platforms, automated assessment tools, adaptive environments, and virtual assistants—is profoundly transforming the ways in which knowledge is accessed, constructed, and validated. However, this transformation is far from neutral; it entails a significant reconfiguration of the inferential processes involved in learning. Students are now called not only to understand disciplinary content, but also to critically evaluate the inferences produced by AI systems, assessing the validity of explanations, the plausibility of recommendations, and the coherence of algorithmic decisions. This evaluative competence is particularly relevant in higher education, where AI-based tools are increasingly integrated into teaching processes, sometimes even replacing traditional instructional roles (Triberti et al., 2024).

In this context, education must go beyond the mere transmission of technical skills related to AI use. It should assume a broader epistemic function aimed at cultivating the ability to reason, evaluate, explain, and infer within environments that are increasingly digital, automated, and opaque. Recent studies (Chen et al., 2023; Holmes, Bialik, & Fadel, 2019) highlight the need for pedagogical interventions designed to engage both System 1 (intuitive and fast) and System 2 (reflective and analytical) thinking. This is crucial, given that many AI systems operate according to principles of 'intuitive intelligence': rapid in their responses yet often lacking transparency.

To address this challenge, it is necessary to integrate AI as an object, tool, and context of inferential education:

- As an object, AI must be studied in its inferential mechanisms, in its operating logics, in its epistemic limits;
- As a tool, AI can support superior thought processes, aid in the construction of inferences, and offer adaptive explanations;
- As a context, AI represents a cognitive ecosystem that changes ways of thinking, deciding, and learning, and therefore requires new reflective skills.

Several concrete educational strategies can effectively encourage inferential processes in interactions with AI. These strategies can be divided into three levels:

- Curriculum and educational planning;

- Teaching methodologies, strategies, and practices;
- Educational technologies and intelligent learning environments.

All of the strategies presented below aim to enhance the complementarity between deduction, intuition, and abduction, and to stimulate critical, flexible, and conscious reasoning in students in the presence of, or collaboration with, intelligent systems.

Level	Strategy	Educational objective	Operational examples
<b>1. Curriculum and design</b>	1. Include inference as a cross-cutting skill	Develop metacognitive and logical awareness	Modules on critical thinking, logic, cognitive biases, AI epistemology
<b>2. Teaching methodologies, strategies and practices</b>	2. Didactics of explanation and justification	Promote argumentative reasoning and inference control	Debate, multiple arguments, peer evaluation
	3. Simulations and open problems	Stimulate abduction, flexibility and intuition	AI decision-making scenarios, real-world case resolution
	4. Metacognitive activities on the use of AI	Develop awareness of the inferential limits of AI	Reflective journals on AI interaction, discussion of automatic results
<b>3. Digital technologies and environments</b>	5. Integration of Explainable AI (XAI) into Learning	Making the AI inferential process transparent	Process visualizations, logic maps, tools with inference tracking
	6. Co-design augmented learning environments	Promoting human-machine collaborative reasoning	Adaptive platforms with human interventions, guided personalization

Table 6. *Strategies to Promote Inference in AI-Powered Education Contexts*

#### *Include inference as a cross-cutting skill*

Incorporating inference as a cross-cutting skill calls for a curriculum redesign that goes beyond the purely logical-formal dimension to also encompass digital competencies and artificial intelligence literacy. These are fundamental in a cognitive society.

In the age of AI, one of the most urgent structural actions to promote inferential processes is the explicit integration of inference into school and university curricula as a cross-cutting cognitive skill. This means treating inference not as a mere technical or logical-mathematical skill, but as a fundamental epistemic skill that crosses all disciplines, from philosophy to science and technology to languages. Inference can provide a framework through which to understand curricular content. For instance, a scientific problem requires formal deductions and abductive hypotheses, while a literary text stimulates implicit and interpretative inferences, and a historical comparison requires causal and analogical inferences. Educating students in inference therefore means educating them in reasoning, analysis, and justifying their conclusions, regardless of the subject matter.

In the context of AI, this skill takes on added significance: students must be able to understand and evaluate the inferences produced by intelligent systems,

distinguishing between correct and convincing answers and between coherent deductions and cognitive shortcuts. Therefore, introducing curricular modules on cognitive biases, AI logic, and the epistemology of algorithmic decision-making is essential for fostering fully informed digital citizenship.

Recent studies on the development of digital skills and AI highlight the importance of strengthening AI literacy among teachers (Zhang, 2024) and the conscious use of AI tools in primary and secondary schools. In Italy, some innovative schools are already experimenting with educational paths dedicated to critical digital literacy and artificial intelligence education. For example, the "A. Einstein" High School in Rome has launched a pilot project that includes modules on cognitive biases, logic and functioning of algorithms, and ethical reflections related to AI. Students, through interdisciplinary workshops, learn to recognize the differences between "plausible" and "correct" answers produced by intelligent systems, developing metacognitive skills and critical thinking. This type of initiative, often in collaboration with universities and research centers such as CINI (National Interuniversity Consortium for Informatics), aims to train aware digital citizens, able to dialogue with emerging technologies and actively participate in the digital society without passively suffering their dynamics.

Some experiences already active at the international level (e.g., AI + Ethics in K-12 Education; UNESCO, 2019) show how it is possible to integrate this content into interdisciplinary paths, educational workshops, PCTO projects, and digital citizenship courses. The goal is not to train programmers, but inferential thinkers, capable of reasoning with, about, and against AI.

This approach aligns with the European Framework DigCompEdu (Redecker & Punie, 2017), which promotes the professional development of teachers in a digital and inferential way, highlighting the need to integrate the critical use of smart technologies into teaching. In addition, the reference to Key Competences for Lifelong Learning (Council of the EU, 2018) reinforces the centrality of critical thinking, digital competence, and learning to learn transversal dimensions to be cultivated through hybrid inferential strategies.

### *Vertical progression of inferential education*

For inferential training to be truly effective, it must be designed with vertical continuity in mind throughout the school period. From primary school onwards, pupils can be introduced to basic forms of inference through play, storytelling, and exploring the world. This stimulates them to formulate simple hypotheses, search for cause-and-effect relationships, and justify their answers. In lower secondary school, these processes can be consolidated through activities requiring generalization (induction) and comparison between alternative explanations (abduction), with the support of digital tools where possible. At secondary school and university levels, inferential meta-reasoning can be introduced explicitly: the critical analysis of premises, explanation models, and argumentative strategies, including interaction with artificial

intelligence systems. This progressive approach enables learners to develop multilevel inferential competence based on the integration of intuition, logic, and epistemic awareness, fostering autonomous, reflective, and resilient learning throughout their educational journey.

### *Didactics of explanation and justification*

A fundamental teaching strategy for developing inferential competence is to promote an approach to teaching that focuses on explanation and justification. Teaching students to 'give reasons' for their answers, explain the logical steps, and argue supportively for a position educates them in the conscious use of the inferential mind. This is particularly pertinent in the age of AI, where machine-generated outputs, while seemingly logical or persuasive, do not necessarily result from a transparent argumentative process.

In educational contexts, asking students not only what they think, but also why, encourages them to transition from passive interaction with AI, [which is] characterized by simple pattern recognition, to authentic cognitive processing. This approach can be implemented through activities such as debates, analyzing controversial cases, constructing argumentative maps, and elaborating on justified alternative responses.

Thus, the explicit justification of one's thoughts becomes an inferential exercise that forces one to identify implicit premises, recognize the validity (or fallacies) of reasoning, and compare different hypotheses. In the presence of AI, these practices enable students to interrogate automated responses, compare them with other sources, and evaluate them in [the] context of the situation. In this way, inference becomes a critical and situated activity rooted in a deep understanding of context and the purpose of cognitive action, rather than just a logical-formal process.

This form of dialogic teaching has also been successfully trialed in digital environments. Some educational AI platforms, such as intelligent tutoring systems (e.g. ASSISTments and Carnegie Learning), incorporate metacognitive questions and requests for explanations to reinforce students' reasoning. However, it is the teacher who can guide reflection on how AI arrives at certain answers and how justified they are in logical or contextual terms through didactic mediation.

In summary, promoting explanation and justification improves not only school performance, but also educates students in epistemic transparency, the critical evaluation of information, and the conscious construction of knowledge — skills that are increasingly crucial when interacting with intelligent technologies. Teaching students to explain and justify their reasoning strengthens their cognitive autonomy in the age of automation, encouraging them to question not just the 'what', but also the 'why' and the 'how'.

Against this regulatory backdrop, the new European AI Act is gaining significance, as it establishes principles of transparency, accountability, and safety in the use of artificial intelligence. It also imposes requirements for explainability and reliability that directly



impact the educational sector. In Italy, Decree No. 161 of 14 June 2022 adopted the School Plan 4.0, promoted by the Ministry of Education. This plan aims to transform classrooms into innovative, digitally advanced learning environments and opens up concrete spaces for the introduction of AI-based tools. Linking these developments to inferential practices enables the creation of an education system that can respond to pedagogical and regulatory challenges alike.

### *Simulations and open problems: training abduction and intuition*

One of the most effective teaching practices for promoting advanced inferential processes is the use of simulations, open problems, and complex scenarios that require students to do more than apply rules; they must also formulate hypotheses, evaluate alternatives, and make decisions in situations of uncertainty. These activities encourage abductive thinking, which is oriented towards constructing the best possible explanation, and stimulate rational intuition — that is, the ability to make well-founded inferences quickly. This ability also plays a decisive role in interacting with artificial intelligence systems.

In educational settings, these activities can take different forms:

- Decision-making simulations in which the student interacts with an AI system (e.g. chatbot, automatic tutor, expert system) and must decide whether to follow or correct the algorithmic proposal, motivating his or her choice;
- Open problems that do not provide a single solution, but require exploration, search for clues, subsequent inferences, and construction of plausible explanations;
- Counterfactual scenarios in which students are asked to imagine what would happen if some data changed or if the AI had access to different information.

These practices encourage nonlinear inferential thinking, highlighting the importance of abduction as a tool for understanding and adapting to real-life situations. In AI-assisted environments, these activities encourage students to question the most immediate answer, asking "why" and "how" it was produced and whether there are other possible explanations.

Rather than being an inferior form of thought, intuition represents a powerful resource, especially when accompanied by critical awareness. As Polanyi (1966) reminds us, we 'know more than we can say': valuing intuition means recognizing that many inferences are based on tacit knowledge, previous experiences, and the recognition of patterns that cannot be formally verbalized.

These activities also prepare students to face new situations and accept ambiguity, teaching them to take responsibility for decision-making in complex information environments. This is particularly relevant in a society where AI is often perceived as infallible, yet it operates based on probabilistic inferences that are sometimes biased or incomplete.

In summary, proposing simulations and open problems means more than just 'active teaching'; it involves educating students in the dynamic, flexible, and contextual use of

inference. This develops a competence that goes beyond formal correctness and is oriented towards relevance, adaptability, and responsibility.

### *Metacognitive activities on the use of AI: learning to interrogate inferences*

To encourage the conscious use of AI in educational processes, it is crucial to combine practical activities with opportunities for metacognitive reflection. The aim is not only to teach students how to use AI-based tools but also to encourage them to consider how these tools work, prompting questions such as: What data is this suggestion based on? What inferential logic underlies this answer? What do I find convincing, and what don't I find convincing?

*During a workshop, a student receives a plausible but incorrect explanation about a historical concept from an AI chatbot. He accepts it without objection. The teacher, observing the interaction, intervenes:*

*"Why do you think it's right? What evidence do you have?"*

*This simple exchange triggers a reflection: the student rereads, verifies the sources, and questions the output of the machine. It is the first step towards inferential awareness.*

Metacognitive activities aim to make visible the cognitive processes involved in interacting with AI and to develop what we could call "inferential awareness". This involves, among other aspects:

- the recognition of the premises implicit in AI suggestions;
- the ability to identify biases or fallacies in the responses generated;
- reflection on the limits of the computational context and the ethical implications of automated decisions.
- From an educational point of view, these activities can be implemented through simple but powerful tools, such as:
- reflective learning journals, in which students write down decisions made with AI and analyze the reasons;
- inference evaluation sheets, which guide students in the structured analysis of algorithmic proposals;
- meta-dialogical discussions, focused on the comparison between human and artificial reasoning;
- guided self-explanation, in which students must justify or question the solutions suggested by an intelligent assistant.

These activities address the findings of recent studies on human-machine interaction. According to Chen et al. (2023), human intuition can serve as a safeguard against blindly trusting automatic responses, provided it is honed through reflective and intentional experiences. Integrating AI into educational programs without encouraging critical thinking is akin to reinforcing automatism.

Furthermore, from an educational perspective, metacognition concerning AI encompasses not only the cognitive dimension but also the ethical and social dimensions. It involves training individuals to understand and evaluate the epistemic power of algorithms, and to question their role in knowledge, decision-making, and

responsibility.

In summary, metacognitive activities concerning the use of AI are essential for transforming the relationship with intelligent technologies into an opportunity for reflective learning and for fostering an educational culture based on analysis, doubt, and interpretative responsibility.

### *Explainable AI (XAI) integration to make inferences visible*

One of the most significant issues regarding the adoption of AI in education is the opacity of algorithmic decision-making processes. Many machine learning systems, particularly those based on deep neural networks, operate like 'black boxes': they produce correct (or seemingly correct) results, but do not explain how they arrived at them. This severely limits educational possibilities because it prevents students — and sometimes even teachers — from understanding and evaluating the underlying inferential reasoning. To address this issue, Explainable AI (XAI) has emerged as a research area in recent years. XAI aims to develop intelligent systems that can produce understandable, traceable, and interpretable inferences. As Miller and Davis (2023) also highlight, XAI plays a vital educational role by fostering algorithmic transparency and trust in the inferential process. In an educational context, integrating these tools could be a turning point, not only because it increases transparency, but also because it provides an opportunity to teach students how to read and evaluate algorithmic reasoning.

Applications of XAI in education can include:

- Visualizations of decision-making processes (e.g., decision trees, attention maps, logical chains), showing the inferential path followed by AI;
- Interactive interfaces, which allow the student to modify the input data and observe how the output changes, stimulating causal and logical reflection;
- Explanatory feedback, which accompanies the AI's responses with reasons, sources, and conditions of validity;
- Systems of "algorithmic justification", in which the system shows the symbolic rules or probabilities supporting its decision (Gunning et al., 2019).

From an educational point of view, these features make technologies more transparent and educate inference through technology. Explainable AI thus becomes an external model of reasoning that students can analyze, imitate, criticize, or improve. This enables the transition from instrumental to educational use of AI, transforming interaction with the machine into an opportunity to develop logical, deductive, and metacognitive thinking. Additionally, XAI has ethical value: in an increasingly algorithm-driven society, the ability to request explanations, comprehend the logic behind automatic decisions, and assess their consistency is integral to mature digital citizenship. For this reason, integrating XAI into teaching represents a cognitive, cultural, and democratic strategy.

Type of inference	Symbolic AI	Sub-symbolic AI	Education
Deductive	✓ explicit	× (weak)	logic, argumentation
Inductive	× (limited)	✓ learning	observation, generalization
Abductive	✓ Partial	✓ Plausibility	Hypotheses, creativity, problem solving

Table 7. *Comparison between inferential forms, AI models and educational implications*

### *Debate tool for critical questioning of Artificial Intelligence*

#### Description of an example of a tool

This tool is designed to facilitate a debate or structured questioning session on generative artificial intelligence (AI). The aim is to explore and test fundamental themes such as epistemic responsibility, biases, transparency, and the ethical implications of AI. The questions are organised into thematic sections to encourage critical thinking and learning for both students and teachers. The tool can be used in educational settings, training workshops or research activities.

Section	Question
<b>1. Understanding and self-evaluating AI</b>	1. Can you explain what you mean by "epistemic responsibility" and why it is important in the use of artificial intelligences?
	2. What are the main limitations and risks associated with the automatic generation of information?
	3. How can biases affect the answers you provide? Can you give an example?
	4. How do you rate the reliability of the information you rely on to generate responses?
	5. If you were asked to produce content on controversial topics, how do you manage neutrality and balance?
<b>2. Validation and critical selection</b>	6. How can a student verify that your answer is accurate and error-free?
	7. What tools or strategies would you recommend to a teacher to teach students to recognize distorted or false information?
	8. Can you point out some reliable sources or resources that students should use as references?
	9. How do you suggest balancing the use of AI with human critical thinking in learning?
	10. What criteria should a school adopt to integrate the use of AI in an ethical and responsible way?
<b>3. Bias, transparency and accountability</b>	11. What types of bias are most common in AI models and how do they manifest themselves in responses?
	12. How does the transparency of an AI model help build trust from users?
	13. What is the role of developers and educators in ensuring AI's epistemic responsibility?
	14. How should the functioning and limitations of an AI model be communicated to users?
	15. What ethical responsibilities derive from the widespread use of generative AI in education?
<b>4. Ethical and epistemic implications</b>	16. In your opinion, what are the main ethical implications related to the use of AI in education?
	17. How can AI help reduce disinformation and what risks does it pose?
	18. Can you comment on Floridi's (2023) reflections on epistemic responsibility?
	19. What lessons can be learned from Suresh & Gutttag's (2024) studies on biases in AI models?
	20. How would you imagine the future relationship between artificial intelligence and human knowledge, as suggested by Hassani et al. (2024)?
<b>5. Reflection and self-criticism</b>	21. What improvements do you expect in upcoming AI models to increase your epistemic responsibility?
	22. How can users help make you a more reliable and transparent tool?
	23. What are your current limitations and how can they influence a debate or learning process?
	24. How do you handle any conflicts between conflicting data or differing opinions?
	25. What would you recommend to those who use AI not to fall into an uncritical or passive use of information?

Table 8. *Structured tool to guide students and teachers in the critical analysis of generative AI systems, through five thematic sections: understanding, validation, bias, ethical implications and metacognitive reflection.*

## Epistemic responsibility in the age of artificial intelligence

Epistemic responsibility is the moral and intellectual obligation to acquire, evaluate, and disseminate knowledge accurately and critically. In an educational context that is becoming increasingly dependent on generative artificial intelligence (AI), this

responsibility is crucial to avoid the spread of incorrect, distorted, or incomplete information.

It is about empowering students in the selection and validation of information generated by AI. Students must develop epistemic skills that make them able to:

- *Critically evaluate sources*: teach how to recognize the characteristics of reliability of a source, distinguishing between verified information, opinions, and disinformation;
- *Verifying information*: encouraging the use of fact-checking techniques and comparison between different sources to validate data and claims provided by AI;
- *Reflect on the limitations of AI tools*: explain that generative models are based on historical data and may reflect biases, omissions, or errors, so their output is not an absolute truth;
- *Develop an "active responsibility" mentality*: promote awareness that every user of information must avoid the propagation of unverified content, even when generated by apparently sophisticated systems.

Useful teaching methods include practical exercises to evaluate AI-produced texts, critical class discussions, and the use of rubrics to measure the reliability of information.

Supporting students requires that teachers themselves be prepared to evaluate the bias and transparency of generative models and to:

- *Understand the technical basis of generative models*: not necessarily at an engineering level, but enough to recognize how they work and where potential biases come from;
- *Recognize implicit biases*: know that models may reflect biases present in the training data, which result in biased or discriminatory responses;
- *Assess the transparency of AI tools*: know the aspects related to "explainability", i.e. the ability of a model to justify its responses, and prefer tools with greater transparency;
- *Complementing teaching with ethical discussions*: helping students to reflect on the role of AI in society and the consequences of uncritical or irresponsible use of technologies.

Training paths may include workshops and specific online modules on AI and bias, as well as communities of practice for sharing experiences and resources.

The growing use of generative AI raises significant epistemological and ethical concerns. Floridi (2023) emphasizes that 'epistemic responsibility' concerns not only the individual but also institutions, which must guarantee access to accurate information and transparent technologies. Therefore, educating people in epistemic responsibility is a social imperative to counteract disinformation and promote informed digital citizenship.

Cheraghi et al. (2025) highlight that biases in AI models are not merely technical errors; they reflect deeper social and cultural dynamics and require a multidisciplinary approach to assess and mitigate them. Therefore, the training of teachers and students

must integrate technical knowledge, ethical reflection, and critical thinking skills.

Hassani et al. (2024) propose the concept of 'AI epistemology', recognizing artificial intelligence as an 'epistemic partner' with which to collaborate, providing critical and responsible control over knowledge generation processes maintained. This requires a rethink of traditional educational models to incorporate the ability to negotiate the meaning and validity of information in hybrid human-machine contexts.

In summary, epistemic responsibility in the digital age is not merely an individual exercise, but a collective project involving educators, students, AI developers, and policymakers. All are called upon to work together to build a reliable, transparent, and fair information ecosystem.

Among the main pedagogical and technical challenges is the integration of explainable AI (XAI): tools must be accessible and understandable for students and teachers, avoiding excessive technocentrism and overload. Metacognitive activities, if not linked to disciplinary tasks, risk becoming abstract or demotivating. It is therefore essential to foster reflection through concrete tools such as journals, checklists, and guided discussions.

The teaching of argumentation must also be adapted to students' levels and supported by examples of AI-generated responses to be critically analyzed. Moreover, the use of simulations and open-ended problems requires well-designed environments, with progression in complexity and clear evaluation criteria. To be effective, these approaches must be based on collaboration among teachers, technologists, and researchers, with particular attention to infrastructural inequalities and the centrality of educational needs. Rather than discouraging innovation, these challenges should encourage conscious experimentation, supported by ongoing training, flexible environments, and a focus on cognitive and ethical values.

## **Conclusions**

The use of explainable AI (XAI) represents a turning point in contemporary education, as it enables students to observe, understand, and question decision-making processes of machines. This transformation fosters active cognitive development, enhancing logical, deductive, and metacognitive skills. Beyond the technical aspect, XAI plays a fundamental ethical and social role: in a world increasingly governed by algorithms, the ability to ask for explanations, understand underlying logic, and critically assess automated decisions is essential for informed digital citizenship. Integrating AI explainability into education involves not only enhancing digital skills but also reinforcing democratic values and fostering a participatory culture that places humans at the core of their relationship with technology.

## References

- Ane, T. A., & Nepa, T. (2024). Emerging artificial intelligence prediction model for educational knowledge representation through learning performance. *Research on Education and Media*, 16(2). <https://doi.org/10.2478/rem-2024-0011>.
- Bilotti, U., Campitiello, L., Todino, M. D., & Sibilio, M. (2023). Emulation and understanding the emotion according to generative artificial intelligence - case study of emotional component extracted from visual artworks. *Journal of Inclusive Methodology and Technology in Learning and Teaching*, 3(4), 1-9
- Bruner, J. S. (2016). *Il processo dell'educazione: Dopo Dewey* (1960). Roma: Armando.
- Chen, J., Liao, Q. V., Vaughan, J. W., & Bansal, G. (2023). Understanding the Role of Human Intuition on Reliance in Human-AI Decision-Making with Explanations. *arXiv:2301.11692*. <https://doi.org/10.48550/arXiv.2301.07255>
- Cheraghi, R., Mahfoozi, A. M., Zolfaghari, S., Shabani, M., Ramezani, M., & Rabiee, H. R. (2025). Epistemic uncertainty aware recommendation systems via Bayesian deep ensemble learning [arXiv preprint arXiv:2504.10753]. <https://doi.org/10.48550/arXiv.2504.10753>.
- Davis, R., & Lenat, D. B. (1982). *Knowledge-Based Systems in Artificial Intelligence*. New York: McGraw-Hill.
- Di Tore, P. A. (2023). Intelligenza Artificiale e processi educativi secondo l'Intelligenza Artificiale. *QTimes – webmagazine*, 2(1), 464-481.
- Eco, U. (1979). *Lector in fabula: La cooperazione interpretativa nei testi narrativi*. Milano: Bompiani.
- Feigenbaum, E. A., & Earnest, L. D. (1977). *The art of artificial intelligence: I. Themes and case studies of knowledge engineering*. Stanford: Stanford University, Department of Computer Science, Technical Reports.
- Floridi, L. (2023). *The Ethics of Artificial Intelligence in Education*. Oxford: Oxford University Press.
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI-Explainable artificial intelligence. *Science robotics*, 4(37), eaay7120. doi:10.1126/scirobotics.aay7120. PMID: 33137719.
- Hassani, H. & Silva, E.S. (2024). Predictions from Generative Artificial Intelligence Models: Towards a New Benchmark in Forecasting Practice. *Information*, 15(6), 291.
- Hassani, H., & Silva, E. S. (2023). The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field. *Big Data and Cognitive Computing (BDCC)*, 7(2), 62, 1-16. 10.3390/bdcc7020062.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston: The Center for Curriculum Redesign.
- Ignatiev, A., Narodytska, N., & Marques-Silva, J. (2019). Abduction-Based Explanations for Machine Learning Models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1), 1511-1519.



<https://doi.org/10.1609/aaai.v33i01.33011511>.

- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Mahwah: Lawrence Erlbaum Associates.
- Kahneman, D. (2011). *Thinking, fast and slow*. London: Penguin Books.
- Luckin, R. (2018). *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. London: UCL IOE Press.
- McCarthy J., & Hayes, P. J. (1969). *Some philosophical problems from the standpoint of artificial intelligence*. Stanford: Computer Science Department Stanford, University Stanford.
- Miller, T., & Davis, G. (2023). *Explainable AI In Education: Current Trends, Challenges, And Opportunities*. SoutheastCon 2023, Orlando, FL, USA, 2023, pp. 232-239, doi: 10.1109/SoutheastCon51012.2023.10115140.
- Pedwell, C. (2023). Intuition as a “trained thing”: Sensing, thinking, and speculating in computational cultures. *Subjectivity*, 30, 348–372. <https://doi.org/10.1057/s41286-023-00170-x>.
- Peirce, C. S. (1931–1958). *Collected Papers of Charles Sanders Peirce* (Vols. 1–8), ed. C. Hartshorne, P. Weiss & A. Burks. Cambridge, MA: Harvard University Press.
- Polanyi, M. (1966). *The tacit dimension*. London: Routledge & Kegan Paul.
- Redecker, C., & Punie, Y. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Luxembourg: Publications Office of the European Union.
- Suresh, H., & Gutttag, J. (2021). A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. EAAMO 2021: Equity and Access in Algorithms, Mechanisms, and Optimization, *arXiv:1901.10002v5*. <https://doi.org/10.48550/arXiv.1901.10002>
- Triberti, S., Di Fuccio, R., Scuotto, C., Marsico, E., & Limone, P. (2024). “Better than my professor?” How to develop artificial intelligence tools for higher education. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/FRAI.2024.1329605/FULL>.
- UNESCO. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. Paris: UNESCO.
- Valiant, L. G. (2020). *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World* (2013). New York: Basic Books.
- Zanetti, M., Rendina, S., Picci, L., & Peluso Cassese, F. (2020). Potential risks of Artificial Intelligence in education. *Form@re - Open Journal Per La Formazione in Rete*, 20(1), 368–378. <https://doi.org/10.13128/form-8113>.
- Zhang, J. (2024). AI in teacher education: Unlocking new dimensions in teaching support, inclusive learning, and digital literacy. *Journal of Computer Assisted Learning*, 40(4), 1871-1885. <https://doi.org/10.1111/jcal.12988>.
- Zhang, X., Guo, Y., Stepputtis, S., Sycara, K., & Campbell, J. (2023). Explaining Agent Behavior with Large Language Models. *arXiv*, 2309.10346. <https://arxiv.org/abs/2309.10346>.