DIGITAL COGNITIVE CAPITAL AND GENERATIVE AI: RAG SOCRATE AS A METHODOLOGICAL ENVIRONMENT FOR EPISTEMIC AWARENESS

Monica Di Domenico[1]
University of Salerno
modidomenico@unisa.it

Pio Alfredo Di Tore
University of Cassino and Southern Lazio
pioalfredo.ditore@unicas.it

Double Blind Peer Review

Citation
Di Domenico, M., & Di Tore, P. A. (2026). Digital cognitive capital and generative AI: RAG Socrate as a methodological environment for epistemic awareness. *Italian Journal of Health Education, Sport and Inclusive Didactics, 10*(4).

Doi: https://doi.org/10.32043/gsd.v1i4.1848

Copyright notice:
© 2023 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
ISSN: 3103-5744

ABSTRACT
If AI is understood as an epistemic infrastructure redefining the production and validation of knowledge, and LLMs as statistical language simulators, the issue is not only what they answer, but when and why their answers are accepted as knowledge. The study draws on media as cognitive environments reorganising perception, language and epistemic authority (McLuhan, 1964), and on digital cognitive capital as critical, interpretive and metacognitive resources that orient subjects in informational environments (de Kerckhove, 2025). Generative AI is examined as a socio-technical network of subjects, models, sources and interfaces (Latour, 2005), and as a threshold between cognitive extension and delegation (Clark, 2008). The contribution presents the design and preliminary exploratory findings of a pedagogical study conducted with Socrate, an orchestrated RAG system, as a methodological environment for making observable and discussable the conditions that contribute to the authority and plausibility of AI-generated answers.

KEYWORDS
Generative artificial intelligence; digital cognitive capital; retrieval-augmented generation; epistemic awareness; cognitive delegation.

Received 2026-07-08
Accepted 2026-10-28
Published 2026-10-30

[1] Author contribution: The article is the result of the joint work of the authors, who shared the conceptual framework, structure, and methodological orientation of the study. M. Di Domenico wrote the manuscript under the scientific supervision of P. A. Di Tore.

1. Introduction

Generative Artificial Intelligence is entering educational and research contexts with increasing intensity. Its interfaces make it possible to access information, produce texts, summarise content, reformulate materials, and adapt resources to different educational needs. In this first operational dimension, generative AI appears as a tool of mediation: it can modify the language of a resource, reduce its complexity, reorganise its structure, and make it more accessible. The literature on Artificial Intelligence in Education has already highlighted these possibilities, together with the need to analyse the pedagogical, ethical, and epistemic implications connected to the use of intelligent systems in educational processes (Holmes et al., 2019; Kasneci et al., 2023; UNESCO, 2024).

The specific feature of Large Language Models lies in their ability to simulate natural language and generate fluent, coherent, and contextually plausible answers. This linguistic quality influences the way users attribute reliability and epistemic value to generated outputs. A well-structured answer, clearly formulated and organised according to a recognisable argumentative logic, may be perceived as reliable even when the user has limited access to the conditions that shaped its production. The plausibility of the answer therefore becomes a relevant pedagogical object, because it involves the relationship between language, trust, source, and validation (Bender et al., 2021; Bommasani et al., 2021; Kasneci et al., 2023).

This dynamic places the interaction with generative AI within an educational space where expectations of competence, communicative proximity, caution towards the artificial, and the possibility of automatic reliance on the system's answer are closely connected. The educational issue concerns the development

of a stance capable of critically engaging with the relationship with the artificial, avoiding both the reduction of AI to a threat and its acceptance as a self-sufficient source of knowledge (Long & Magerko, 2020; UNESCO, 2024).

This paper starts from this threshold: the moment in which a generated answer is questioned, verified, or accepted as valid. The generated answer appears to the user as a fluent and complete text, while the conditions that guide its production often remain opaque: sources, selection criteria, instructions, parameters, and configurations of the AI system. This gives rise to the pedagogical need to design educational situations in which these conditions can be recalled, observed, and discussed, so that the plausibility of the output is connected to practices of verification and validation (Bender et al., 2021; Bommasani et al., 2021; Lewis et al., 2020).

Within this perspective, the RAGTIME laboratory is positioned as a setting designed to transform the generated answer from a product to be accepted into a process to be questioned. The laboratory uses Socrate, an orchestrated RAG system developed within Tiresia, the Inter-University Team for Educational Research on Artificial Intelligence. The intentional use of Socrate makes it possible to work on a controlled documentary corpus and to make some components of the generative process observable, including retrieved sources, selected documentary portions, system instructions, generation parameters, and final output (Lewis et al., 2020).

The paper presents the design and the first exploratory evidence of a pedagogical study conducted through RAGTIME. The theoretical framework is organised around four main areas: trust in the output and linguistic simulation; digital cognitive capital and epistemic awareness; generative AI as a socio-technical network; and the boundary between cognitive extension and cognitive delegation. Starting from these references, Socrate is used within an

exploratory design to observe the practices through which participants question, interpret, verify, and validate generated answers (Clark, 2008; de Kerckhove, 2025; Latour, 2005; McLuhan, 1964).

2. Generative AI, linguistic simulation, and trust in the output

The relationship between the human and the artificial, in contemporary educational contexts, can be observed through the dialogic forms taken by generative systems. These systems produce continuous answers, modulate tone, organise explanations, and offer an experience of communicative availability that can influence learning and research practices. Linguistic simulation thus becomes a space in which expectations of competence, reliability, and proximity take shape, with relevant effects on the ways in which users attribute trust to the output of AI systems (Bender et al., 2021; Kasneci et al., 2023).

In the case of LLMs, the reference to simulation concerns the production of texts that appear to have coherence, order, and argumentative intentionality. Users may perceive the answer as the expression of competence, although it is the result of statistical procedures, technical configurations, and documentary conditions. At this point, the pedagogical issue of trust in the output emerges: a convincing answer may generate reliance before the subject has reconstructed the process that supports it (Bender et al., 2021; Bommasani et al., 2021).

Trust in the output has both a cognitive and a relational dimension. It is cognitive because it concerns the criteria through which users evaluate the completeness, coherence, relevance, and grounding of the answer. It is relational because it arises within an interaction: the system answers, reformulates, deepens, adapts the text to the question, and offers a dialogic

continuity that may be perceived as availability. Even when the focus of the study is not affectivity in a strict sense, trust in the generated answer meets a relevant threshold of the educational relationship mediated by the artificial (Kasneci et al., 2023; UNESCO, 2024).

From this perspective, epistemic awareness is a necessary component of education about AI. Educating users to work with generative systems also means educating them to recognise the conditions that make an output credible. The issue concerns the passage from receiving the answer to validating it: users must be able to distinguish between trust supported by linguistic form and trust supported by verification practices. This distinction is crucial in educational contexts, where knowledge is constructed, discussed, and argued (Long & Magerko, 2020; UNESCO, 2024).

Linguistic plausibility therefore becomes a pedagogical object because it guides user behaviour. A fluent answer may reduce the perceived need for control. An answer that makes limits, sources, and conditions explicit may instead encourage a more questioning stance. The design of AI-based learning environments must therefore consider the form of the output, the transparency of sources, the possibility of comparison, and the presence of spaces in which the answer can be critically discussed (Bender et al., 2021; Holmes et al., 2019).

This interpretation makes it possible to connect the problem of validation with the theme of digital citizenship. Interactions with generative systems involve interpretative habits, forms of trust, criteria of authority, and responsibility of judgement. The management of trust becomes an educational competence: it concerns the way in which the subject decides when to suspend acceptance,

when to consult sources, when to reformulate the question, and when to take the answer as part of a broader knowledge process (Long & Magerko, 2020; UNESCO, 2024).

In this direction, RAGTIME is positioned as a laboratory of epistemic awareness. Its aim is to create a situation in which trust in the output can be observed in practice: accepting, verifying, reformulating, comparing, consulting sources, and recognising differences between answers produced from the same question. The interest of the laboratory concerns the possibility of seeing how participants move from the answer to its evaluation, and which cognitive, critical, and metacognitive resources they activate in doing so.

3. Digital cognitive capital, AI literacy, and epistemic awareness

The passage from the answer to the process brings into play the concept of digital cognitive capital, understood as a dynamic set of critical, interpretative, and metacognitive resources that allow subjects to orient themselves in information environments (de Kerckhove, 2025). This concept does not coincide only with technical competence in the use of an interface. It concerns the way in which the subject recognises technological mediations, evaluates sources, compares answers, interprets the conditions of production of the output, and keeps their own cognitive responsibility active.

The reference to McLuhan makes it possible to read media as environments capable of reorganising perception, language, and forms of epistemic authority (McLuhan, 1964). From this perspective, generative AI can be considered as an environment that modifies the actions through which information is searched for, a question is formulated, an answer is interpreted, and trust in the

produced knowledge is constructed. Digital cognitive capital therefore indicates the ability not to stop at the surface of the answer, but to question the conditions that make it possible.

In the case of generative AI, this capital is expressed when the user understands that the answer is not a neutral or self-sufficient content, but the result of a technical, documentary, and interactional configuration. The final form of the output depends on the way the question is formulated, on the available sources, on the model used, on the instructions that guide generation, and on the conditions through which information is selected and reorganised. Knowing how to question this configuration means transforming the interaction with the system into a practice of awareness (Bommasani et al., 2021; Lewis et al., 2020).

This perspective is connected to studies on AI literacy, which underline the need to understand, critically evaluate, and use AI systems in an aware way (Long & Magerko, 2020; UNESCO, 2024). In working with generative systems, AI literacy cannot be limited to knowledge of the available functions or to the ability to produce effective prompts. It also includes the ability to recognise the mediated nature of the output, distinguish between formal coherence and epistemic reliability, evaluate the origin of information, and responsibly decide whether to accept or revise an answer.

Epistemic awareness represents the point at which digital cognitive capital and AI literacy meet. It concerns the ability to question an answer, compare different outputs, recognise the role of sources, reformulate the prompt, evaluate the limits of the system, and decide whether to accept or revise the produced content. These practices imply a stance towards generated

knowledge: the subject does not simply receive a text but submits it to a process of verification and attribution of value (Long & Magerko, 2020; UNESCO, 2024).

From this perspective, epistemic awareness also has a value for digital citizenship. Managing trust in algorithmic environments means being able to distribute authority responsibly among oneself, the system, the sources, and the interface. In educational contexts, this ability becomes particularly relevant because it concerns the formation of subjects who are able to use generative AI without fully delegating to the system the judgement on the validity of the knowledge produced (UNESCO, 2024).

4. Generative AI as a socio-technical network

Generative AI can also be observed as a socio-technical network, in which human and non-human actors take part in the construction of the answer. From this perspective, the output cannot be attributed to a single actor, because it takes shape through the relationship between the user who formulates the question, the authors of the sources, the documentary corpus, the language model, the interface, the system instructions, and the parameters that guide generation (Bommasani et al., 2021; Latour, 2005).

The socio-technical perspective makes it possible to read the authority of the output as the effect of a configuration. The answer appears as a unified text, but it is produced by a plurality of elements that contribute to its construction. Evaluating an answer therefore means reconstructing, as far as possible, the way in which these elements enter into relation: which sources are recalled, how they are retrieved, what form the answer takes, which margins of uncertainty remain visible, and which criteria the user adopts to accept or revise it (Latour, 2005; Lewis et al., 2020).

This interpretation is particularly useful in educational contexts, because it makes it possible to shift attention from AI as an autonomous object to AI as an environment of mediation. The answer generated by a language model should be understood as the outcome of a network that includes materials, instructions, parameters, design choices, and interpretative practices. Making this network visible means opening a space for discussion on the responsibility of judgement and on the conditions for validating knowledge (Latour, 2005; UNESCO, 2024).

From this perspective, the RAGTIME laboratory assumes generative AI as an environment to be observed, in which the answer becomes the point of access to a broader process. The pedagogical interest concerns the possibility of making discussable what, in ordinary interactions with general-purpose chatbots and with many available RAG systems, often remains in the background: the origin of information, the criteria for selecting documentary fragments, the effect of instructions, the weight of technical configurations, and the interpretative position of the user (Bender et al., 2021; Bommasani et al., 2021; Lewis et al., 2020).

5. Extension, cognitive delegation, and formative stance

The reference to the extended mind makes it possible to articulate the relationship between technological support and cognitive activity (Clark, 2008). AI tools can extend cognitive activity when they support processes of orientation, memory, calculation, synthesis, comparison, and reformulation. In this sense, a generative system can expand the possibility of accessing information, exploring perspectives, producing initial textual structures, and connecting contents.

Cognitive extension, however, requires an active subject, able to integrate technological support into their own thinking process and to maintain control over the criteria of validation. Cognitive delegation emerges when the subject also transfers to the system the judgement on the validity of the output. In this case, the answer is accepted because it appears coherent, fluent, or aligned with the user's expectations, while the plurality of the socio-technical network is absorbed into the surface of the generated text (Bender et al., 2021; Clark, 2008).

The passage between extension and delegation cannot be decided in abstract terms, because it becomes visible in concrete practices of interaction. It can be observed in the way the user formulates the question, consults sources, compares answers, modifies the prompt, and decides whether to accept the generated content. The pedagogical issue therefore concerns the construction of environments in which these passages can be recognised, discussed, and made the object of training (Clark, 2008; Long & Magerko, 2020).

From this point of view, RAGTIME proposes a formative stance placed between suspicion towards the artificial and automatic reliance on the answer produced by the system. The laboratory proposes using AI as an environment for observation, in which the generated answer becomes an opportunity to exercise verification, comparison, reformulation, and responsibility of judgement. This aspect appears particularly relevant in the training of teachers, researchers, and subjects engaged in advanced educational contexts (UNESCO, 2024).

6. Socrate as a methodological device

Retrieval-Augmented Generation represents a particularly significant context for observing processes of output validation, because it connects linguistic generation to a documentary corpus through information retrieval procedures. In general terms, this architecture integrates the parametric memory of the model with a non-parametric memory made of retrieved documents, making it possible to produce answers anchored to reference materials (Lewis et al., 2020).

However, the methodological relevance of RAG does not depend only on documentary anchoring, but also on the degree of observability and configurability of the environment used. In many applications, even when the use of a documentary base is declared, the user mainly accesses the final answer or partial documentary references. The conditions that guide generation often remain opaque: the way the question is interpreted, the documents retrieved, the selected fragments, the system instructions, and the parameters that influence the form of the output.

Socrate is an orchestrated RAG system developed within Tiresia, the Inter-University Team for Educational Research on Artificial Intelligence. In the present study, it is assumed as a methodological device because it allows users to observe the infrastructure that contributes to the generation of answers and to intervene, from the editor side, on some of the conditions that guide this process. The environment makes it possible to work on a controlled documentary corpus, observe the portions of text, or chunks, retrieved from the documentary base, modify the number of documents recalled by the retrieval process, intervene on system instructions, vary generation parameters, and compare the generation of different language models.

This configurability makes it possible to propose the same task again under different conditions, keeping the question stable and progressively modifying some elements of the generative process. The comparison between outputs makes visible what often remains implicit in ordinary interactions: the answer does not depend only on the prompt, but also on the way the system retrieves information, selects documentary fragments, interprets instructions, and modulates generation. In this sense, the output is treated as the situated, mediated, and partial result of a technical, documentary, and interpretative configuration, in which the person questioning the system is both editor and user.

In the design of the study, Socrate is the tool consciously adopted by the authors to make the path that connects the question to the answer analysable. Through its intentional use, the conditions of generation become an object of pedagogical observation: they can be compared, discussed, and related to the practices of trust, verification, and validation activated by the participants.

7. Methodological design of the RAGTIME laboratory

Study design

The RAGTIME laboratory was designed as an exploratory setting to observe the epistemic practices that emerge in the interaction with a generative Artificial Intelligence system and to make visible the passage from using the answer to questioning it. The design includes two successive conditions experienced by the same participants: a first phase of interaction with Socrate in an opaque condition, in which the interaction takes place through a chatbot with locked parameters, and a second phase of interaction in a configurable condition, in which the conditions of generation are made observable and modifiable. The

structure of the study integrates descriptive data and qualitative materials, according to an exploratory logic consistent with mixed methods approaches aimed at understanding complex educational phenomena (Creswell & Plano Clark, 2018).

Research questions

The laboratory design is guided by four research questions, formulated to explore the transformation of epistemic practices in the interaction with generative Artificial Intelligence systems.

RQ1 – How does interaction with generative Artificial Intelligence systems transform the epistemic function of the question in participants' practices?

RQ2 – Which epistemic criteria guide participants in accepting answers generated by Artificial Intelligence as legitimate knowledge?

RQ3 – How do forms of cognitive delegation and the exercise of cognitive capital emerge in participants' interaction with generative Artificial Intelligence systems?

RQ4 – To what extent does the visibility of the operating conditions of a generative Artificial Intelligence system influence participants' epistemic practice?

The questions do not aim to measure the effectiveness of the system in terms of performance, but to describe the practices through which a generated answer is questioned, accepted, verified, or reformulated.

Context and participants

The RAGTIME laboratory format was progressively tested between March and June 2026 in academic, seminar, and advanced training

contexts. These contexts included, among others, the AI-MCE Winter School at the University of Catania and the Seminar of the National PhD Programme in *Teaching & Learning Sciences: Inclusion, Technologies, Educational Research and Evaluation* at the University of Macerata. The documented core of the exploratory corpus includes the activities carried out with 58 doctoral students in these two contexts. Other implementations of the RAGTIME laboratory involved teachers and researchers in academic and training settings; however, they are considered here as part of the broader methodological pathway rather than as a statistically comparable sample. Participants were selected through convenience sampling, as they were attending the activities in which the laboratory was proposed. The composition of the documented group makes the observation of trust, verification, and validation practices particularly relevant, since it involves participants already familiar with practices of study, scientific argumentation, and use of sources. At the same time, this profile and the non-probabilistic selection of participants limit the generalisability of the findings, which are therefore interpreted as exploratory indications.

Task and documentary corpus

The documentary base chosen for the RAGTIME laboratory concerns DNA extraction. The corpus loaded into Socrate consisted of five documentary sections on DNA extraction, identified in the system as DNA1, DNA2, DNA3, DNA4, and DNA5. These sections were indexed in the environment and made available to the retrieval process as textual chunks. The same corpus was maintained in both phases of the laboratory. This scientific field, intentionally

distant from the pedagogical and educational domains more familiar to the participants, responds to a specific methodological need. The use of pedagogical materials could have activated previous knowledge, theoretical orientations, interpretative expectations, and dynamics of disciplinary confirmation. The choice of content external to this domain instead makes it possible to reduce the weight of thematic familiarity and to focus the observation on practices of interaction with the system: formulation of the question, which remains the same in the two phases of interaction with Socrate, attribution of trust, consultation of sources, comparison between outputs, reformulation, and validation of the answer.

The stability of the question and of the corpus makes it possible to compare the answers produced in the two phases of the laboratory. In the first phase, participants evaluate the answer generated in opaque mode; in the second, the same question is proposed again by progressively modifying some conditions of generation. The scientific content is therefore not the disciplinary focus of the research, but functions as a controlled field within which to observe the relationship between question, system, and validation.

Laboratory procedure

The first phase involves interaction with Socrate in opaque mode. Participants access the system through a public link that allows them to use the chatbot, without the possibility of viewing or modifying generation parameters. In this condition, the documentary corpus, retrieval, system instructions, parameters, and language model remain locked and are not accessible from the user side. The situation reproduces an ordinary mode of interaction with a generative system: the participant formulates the question, receives an answer, may ask for further clarification, and decides when to stop.

The second phase introduces participants to the infrastructure of the Socrate RAG system. The interaction takes place through an editor environment in which it is possible to observe and modify some conditions that guide the generation of the answer. In line with the logic of Retrieval-Augmented Generation, the laboratory makes it possible to examine the relationship between question, documentary retrieval, and linguistic generation of the output (Lewis et al., 2020). Participants use the same question as in the previous phase and progressively modify one element at a time, following a guided procedure, observing how the generated answer changes when the conditions that produce it vary.

Operationally, in the opaque phase Socrate was used with a fixed configuration on the controlled documentary corpus. Retrieval was enabled, the Top-K value was set to 10, the temperature was set to 0.25, and the maximum output length to 512 tokens. The language model used in the initial configuration was LLaMA 4 Maverick Turbo, and conversation history was enabled. In the configurable phase, the same question and the same documentary corpus were maintained, while selected conditions were progressively modified through the editor environment. The modified conditions concerned the breadth of retrieval, the retrieved chunks, the epistemic stance of the system message, the temperature, and the language model.

Table 1. Parameters and conditions modified in the configurable phase

Parameters and conditions modified in the configurable phase

Element	Brief description	Function in the laboratory
Top-K	Number of documentary fragments recalled by the retrieval system and made available to the LLM.	Observe how the breadth of documentary retrieval affects perceived completeness and exhaustiveness.
Chunk	Textual fragments retrieved from the documentary corpus.	Make visible which parts of the corpus documents enter the generative process.
System message	Instructions that guide the behaviour of the model and the stance of the answer.	Compare assertive, doubtful, or neutral answers and observe their effect on trust.
Temperature	Parameter that regulates the degree of variability of generation. Lower values tend to produce more stable answers, more closely aligned with the retrieved documentary context; higher values may support more variable formulations, less strictly bound to the fragments provided by retrieval.	Observe how the output changes when the degree of generative control varies, with attention to the relationship between adherence to the documentary corpus, linguistic fluency, and plausibility of the answer.
Language model	LLM used to generate the answer. The comparison follows a logic of progressive reduction of structural complexity, understood in relation to the architectural scale of the model and to the number of parameters that regulate its functioning.	Observe how the structure of the model affects the output when the question, corpus, and main conditions remain the same.

Data collection tools

Data collection was carried out through two Google Forms, one for each phase, structured as semi-structured instruments combining closed items, multiple-choice questions, and open-ended questions. The first form accompanied the opaque interaction and recorded the evaluation of the acceptability of the answer: satisfaction, perceived reliability, completeness, need for verification,

requested follow-up questions, reasons for stopping the interaction, consultation of the documentary base, and final evaluation of the system's reliability. The second form accompanied the configurable phase and recorded the perception of the variations produced by the modification of generative conditions: comparison between Top-K values, stances of the system message, temperature, and language models. The full questionnaires are not reproduced in the article; their main observed dimensions, sample items, and methodological functions are summarised in Table 2.

Table 2. Data collection tools and sample items

Phase and condition	**Observed dimensions**	**Sample items**	**Methodological function**
Phase 1 – Socrate in opaque mode, access through public link, free interaction with the chatbot, locked parameters.	Acceptability, satisfaction, reliability, completeness, need for verification.	"How reliable do you consider the first answer?"; "After the first answer, how much did you feel the need for further verification?"	Observe the threshold of answer acceptance in an ordinary condition of interaction.
	Requested follow-up questions, decision to stop, consultation of the documentary base.	"How many follow-up questions did you ask before stopping?"; "Why did you decide to stop?"; "Did you consult the documentary base?"	Detect spontaneous practices of control, stopping, and verification.
Phase 2 – Access to the infrastructure and modification of Top-K.	Effect of retrieval on perceived completeness and exhaustiveness.	"Compared to Top-K = 10, the answer with Top-K = 2 is…"	Observe the relationship between the breadth of documentary retrieval and the evaluation of the output.
Phase 2 – Comparison between stances of the system message.	Trust, clarity, need for verification, preference between assertive and doubtful answers.	"For studying, I prefer an answer that is: assertive / doubtful / indifferent."	Observe the effect of the epistemic stance on trust in the answer.

Phase and condition	Observed dimensions	Sample items	Methodological function
Phase 2 – Variation of temperature.	Fluency, perceived completeness, stability or variability of the answer.	"Which answer seems more complete to you?"	Observe the effect of generative variability on the perception of the output.
Phase 2 – Comparison between LLM models according to a logic of decreasing structural complexity.	Completeness, reliability, clarity, perceived differences.	"Which model seemed overall more convincing to you?"; "Did you perceive clear differences between the models?"	Observe variations in the output when the question, corpus, and main conditions remain the same.

Analysis strategy

The data collected are treated as exploratory evidence, useful for examining whether and how the laboratory can make epistemic practices observable in the interaction with generative AI. The closed answers in the Google Forms are considered in descriptive form, with attention to the distribution of evaluations concerning satisfaction, perceived reliability, completeness, need for verification, preference between configurations, and perception of differences between outputs. The open answers, the prompts, possible reformulations, and the motivations provided by the participants are instead analysed through a qualitative reading aimed at identifying thematic recurrences, in line with the approach of thematic analysis (Braun & Clarke, 2006).

The dimensions considered concern both the evaluation of the output and the way in which participants relate the answer to the conditions of its production. In particular, the analysis examines perceived clarity and completeness, the trust attributed to the answer, the consultation of sources, the perception of

variations between configurations, the reference to chunks, retrieval, and parameters, and the reformulation of the question. Taken together, these dimensions support a reading of the interaction as a process in which criteria of control, verification, and validation can gradually become explicit.

Attention is therefore directed to the criteria for accepting the answer, to signs of cognitive delegation, and to the practices through which participants compare, verify, or reformulate the generated output. At this stage, the analysis is intended to build a first interpretative mapping of the observed practices, to be further developed through a systematic coding of the materials collected, without producing generalisable results. *Methodological function of the design*

The organisation of the laboratory into an opaque phase and a configurable phase makes it possible to observe the passage from the evaluation of the answer to its interrogation. In the first phase, the output is considered as a textual product: satisfactory, reliable, complete, sufficient, or in need of verification. In the second phase, the answer is related to the conditions that guide its production, making the relationship between question, sources, retrieval, parameters, and linguistic generation more explicit.

The stability of the question and the controlled variation of the generative conditions make it possible to connect the differences between outputs to the technical, documentary, and interpretative configuration from which they take shape. RAGTIME therefore assumes a specific methodological function: creating conditions in which forms of reliance on the answer and emerging epistemic stances can be observed and discussed. .

Methodological limitations

The study has some methodological limitations. First, participants were selected within specific academic and advanced training contexts through convenience sampling; therefore, the sample cannot be considered statistically representative. Second, the laboratory was conducted with a controlled documentary corpus focused on DNA extraction, which limits the extension of the findings to other disciplinary domains. Third, the same participants experienced the opaque phase before the configurable phase; for this reason, the design does not allow the observed changes to be attributed exclusively to configurability, independently of possible familiarisation effects and progressive learning of the task. For these reasons, the findings are interpreted as exploratory indications. They are useful for discussing the methodological potential of RAGTIME and for identifying epistemic practices that may be further analysed in subsequent research, but they do not support generalisable claims about the effects of the system on participants' awareness.

8. First exploratory evidence and developments of the analysis

The materials collected in the RAGTIME laboratory are discussed as first exploratory evidence. The aim is to describe the theoretical and methodological function of the laboratory and to show how it makes some epistemic practices observable in the interaction with generative AI. A systematic analysis of prompts, reformulations, source consultations, and open answers will be developed in a later phase of the research, in line with an exploratory mixed methods design (Creswell & Plano Clark, 2018).

A first indication concerns the passage from attention to the answer to attention to the conditions of its production. In the initial phase, participants mainly evaluate visible features of the text, such as clarity, coherence,

completeness, organisation, and relevance. These criteria suggest that linguistic form may influence the perception of reliability and confirm the need to distinguish between formal coherence and epistemic reliability (Bender et al., 2021).

In the configurable phase, the collected materials suggest possible signs of a shift in attention. . The possibility of visualising chunks, modifying Top-K, comparing system messages, and varying temperature makes more evident the relationship between answer, retrieved sources, and conditions of generation. The output is therefore perceived less as a self-contained text and more as the result of a configuration (Lewis et al., 2020).

A second indication concerns trust. In the first interactions, trust seems to be supported mainly by linguistic plausibility and apparent completeness. When sources, chunks, and parameters become visible, the collected responses suggest that trust may be articulated through additional criteria: participants are invited to consider where the answer comes from, which elements of the corpus are retrieved, and why the text changes under different conditions (Long & Magerko, 2020; UNESCO, 2024).

Further indications concern verification practices and the perception of variation. The laboratory makes observable the situations in which participants accept the answer, ask for further details, consult sources, compare outputs, or reformulate the question. At the same time, the use of the same question under different conditions allows them to observe that the output is not stable in an absolute sense, but may change in length, tone, level of caution, selection of information, and argumentative structure. These elements help to make the relationship between system and answer less opaque (Lewis et al., 2020).

These initial indications points to dimensions relevant to the study of digital cognitive capital: questioning, comparison, verification, reformulation, attention to sources, and awareness of parameters. They provide an initial basis for a later systematic analysis of the transformation of questions, the role of sources, the criteria for accepting the output, and the different forms of cognitive delegation or extension observable in the interaction. The qualitative analysis of open answers may be developed through thematic coding procedures consistent with the exploratory orientation of the study (Braun & Clarke, 2006).

9. Discussion

The RAGTIME laboratory makes it possible to address the issue of generative AI from a pedagogical perspective centred on awareness and validation. The generated answer becomes an object through which practices of trust, verification, and attribution of authority can be observed. This approach makes it possible to place the use of generative AI within the broader problem of the formation of judgement in digital environments (Long & Magerko, 2020; UNESCO, 2024).

The reference to digital cognitive capital makes it possible to understand that interaction with generative AI requires resources related to the control of the knowledge process: formulating questions, reading answers critically, recognising the importance of sources, understanding the function of parameters, and distinguishing between linguistic coherence and epistemic reliability. RAGTIME seems to offer a context in which these resources can be exercised and observed, making visible actions that often remain implicit in ordinary interactions with AI (Bender et al., 2021; de Kerckhove, 2025).

The issue of simulation concerns the ability of generative systems to produce answers that appear competent, ordered, available, and reliable. Trust in the output is constructed within this experience of plausibility. Educating users to interact with AI therefore also means educating them to manage trust: to recognise when it arises from the linguistic form of the answer and when it is supported by verification practices (Bender et al., 2021; Kasneci et al., 2023).

In this sense, epistemic awareness can be read as a component of the broader issue of digital citizenship. Interaction with generative systems involves interpretative habits, expectations of availability, forms of reliance, and criteria of authority. The proposed work does not directly analyse the emotional dimension of the interaction, but it addresses a fundamental threshold: the moment in which the user decides to trust the answer. This decision has a cognitive, epistemic, and relational dimension, because it concerns the way in which the subject distributes trust among themselves, the system, the sources, and the interface (UNESCO, 2024).

The relationship between cognitive extension and cognitive delegation represents one of the main theoretical outcomes of the pathway. The use of Socrate can support forms of cognitive extension when it helps the user access sources, compare answers, understand the effect of settings, and reformulate their own question. Delegation emerges when the user accepts the answer as valid because it appears formally adequate, while the plurality of elements that determined its production remains in the background. The laboratory makes it possible to observe this threshold and to make it an object of reflection (Clark, 2008; Latour, 2005).

From this point of view, RAGTIME offers a formative stance placed between suspicion towards the artificial and automatic reliance on the answer produced by the system. The laboratory does not ask users to accept AI as an authority

or to reject it as a threat. Instead, it proposes using AI as an environment for observation, in which the system's answer becomes an opportunity to exercise verification, comparison, reformulation, and responsibility of judgement. This aspect appears particularly relevant in the training of teachers, researchers, and subjects engaged in advanced educational contexts, because it makes it possible to work on the epistemic stances that accompany the use of generative AI (UNESCO, 2024).

The choice of an observable and configurable RAG system therefore appears coherent with the pedagogical aim of the study. Socrate makes it possible to construct a situation in which the answer is related to a set of observable conditions. This passage has methodological relevance because it provides a setting for observing practices of questioning and validation. Its formative value should be considered as a hypothesis to be further explored through subsequent analyses (Lewis et al., 2020).

10. Conclusion

The contribution presented the design and the first exploratory evidence of a pedagogical study conducted through the RAGTIME laboratory, using Socrate as a methodological device to observe the conditions that regulate the production of generative output. The theoretical framework interpreted generative AI as a cognitive environment, a socio-technical network, and a threshold between cognitive extension and cognitive delegation. Starting from this framework, the laboratory was designed to make visible the practices through which participants question, interpret, verify, and validate generated answers (Clark, 2008; Latour, 2005; McLuhan, 1964).

The preliminary indications suggest that the configurability of the environment may support a shift of attention from the answer as a finished product to the

conditions that contribute to its production. The observation of chunks, retrieved sources, system instructions, and parameters makes a different relationship with the output possible. The answer is treated as an object of comparison and discussion, and its authority can be related to the process that supports it (Lewis et al., 2020).

The RAGTIME laboratory offers an initial setting for exploring possible indicators of digital cognitive capital in the interaction with generative AI. It makes it possible to document practices of trust, verification, reformulation, and validation, opening a research trajectory useful for understanding how users inhabit generative environments that are increasingly present in educational contexts (de Kerckhove, 2025; Long & Magerko, 2020; UNESCO, 2024).

The systematic analysis of the data collected represents the next development of the research. Future analyses may focus on the classification of prompts, on the transformations between the opaque phase and the configurable phase, on practices of source consultation, on the criteria for accepting or revising the output, and on signs of cognitive delegation. From this perspective, the present study contributes to defining the theoretical and methodological conditions for investigating generative AI as an environment of epistemic awareness and as a formative space for the validation of knowledge.

References

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., … Liang, P. (2021). *On the opportunities and risks of foundation models*. arXiv. https://doi.org/10.48550/arXiv.2108.07258

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195333213.001.0001

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

de Kerckhove, D. (2025). *L'uomo quantistico. Mente, società, democrazia: Dove ci porterà la prossima rivoluzione digitale*. Rai Libri.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press. https://doi.org/10.1093/oso/9780199256044.001.0001

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*, 9459–9474.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

McLuhan, M. (1964). *Understanding media: The extensions of man*. McGraw-Hill.

UNESCO. (2024). *AI competency framework for teachers*. United Nations Educational, Scientific and Cultural Organization. https://doi.org/10.54675/ZJTE2084