Learning and Analytics in Intelligent Systems 59

Dominik May
Michael E. Auer *Editors*

2025 Yearbook Emerging Technologies in Learning

Springer

Chapter 14
Socrates AI: An Ethical Model of Artificial Intelligence for Inclusive and Personalized Education

Maria Amata Garito and Marco Breda

Abstract The introduction of generative intelligence into educational contexts calls for an immediate reconsideration of the risks and ethical responsibilities associated with its use. The focus is on the impact of generative intelligence on the construction of autonomous thinking and the need to preserve inclusive conditions of access and participation. This chapter develops a systemic analysis of the vulnerabilities emerging from the use of generative models in education, articulating a taxonomy that includes ethical, psychological, cultural, social, normative and technological dimensions. From this, guidelines are proposed to steer AI integration practices consistent with the principles of accountability, safety and educational sustainability. Within this framework is *Socrates AI*, a dialogic agent designed to support critical and reflective learning through adaptive interactions oriented toward personalization, inclusiveness, recognition of individual differences and ethical management of dialogue. Socrates AI builds on its own certified knowledge base, on the maieutic approach, and on effective questioning models, within a transparent and flexible architecture. Preliminary evidence from field deployment suggests its potential to mitigate cognitive risks and foster epistemic responsibility. The chapter presents the theoretical and design foundations of the proposal, contributing to the debate on the conscious use of artificial intelligence in education and citizenship, while opening up the prospect of an experimental evaluation of the cognitive impact of AI through psychopedagogical neuroscientific tools.

M. A. Garito (✉)
Professor Emeritus of Psychotechnologies, UNINETTUNO International Telematic University, Roma, RM, Italy
e-mail: garito@uninettunouniversity.net

M. Breda
Director of the AI Laboratory, UNINETTUNO International Telematic University, Roma, RM, Italy
e-mail: marco.breda@uninettunouniversity.net

D. May and M. E. Auer (eds.), *2025 Yearbook Emerging Technologies in Learning*,
Learning and Analytics in Intelligent Systems 59,
https://doi.org/10.1007/978-3-032-09058-4_14

14.1 Introduction

Generative artificial intelligence is profoundly changing the ways in which human beings construct knowledge, make it accessible, and explore its meanings. This mutation deeply involves education and our relationship with knowledge [1]. In this framework, the design of responsible generative AI systems emerges as a major cultural challenge, with implications that precede and condition technological choices [2, 3]. In particular, integration in education requires reflection on its inclusive potential and the ability to adapt interaction to different student profiles, enhancing personalized and accessible educational pathways [4].

To this vision responds a long and coherent path of research that has taken shape, well in advance, at the International Telematic University UNINETTUNO. Today's results are the fruit of more than 30 years of research initiated by Rector Maria Amata Garito, beginning with the pioneering studies she carried out when she was professor of psychotechnology at the University of Rome "*Sapienza*" and directed the *CATTID* (Center for the Applications of Television, Information Technologies and Education) [5]. This research has been consolidated over time through studies on the evolutions of psychopedagogical models applied to distance teaching and learning processes [6, 7].

Today, with the advent of generative artificial intelligence, the International Telematic University UNINETTUNO has directed its research activities toward the creation of Socrates AI: an emblematic example of artificial intelligence designed to develop interaction models based on the Socratic method, with the aim of fostering critical thinking. It is neither a search engine nor a simple virtual assistant, but a real digital training model inspired by the socratic maieutics, based on the dialogic interaction between human mind and AI [8]. This innovative approach has made it possible to create a dynamic learning environment that encourages students to explore and compare ideas, thus developing a deeper and more critical understanding of content.

The Socrates AI project, which is available to students but is continually evolving, aims to support an ethical, self-aware, inclusive, and person-centered use of artificial intelligence in digital educational contexts. The goal is not to automate response, but to foster processes of questioning, comparing and reworking that empower students to actively participate in the construction of their own knowledge. Guided interaction takes the form of a space in which questions, hypotheses, connections between disciplinary fields and open-ended reflections emerge, far beyond the mere transmission of content. The ambition is to transform online learning into an accessible, engaging experience that can be modulated to individual characteristics, without sacrificing intellectual depth and plurality of cognitive pathways. In this sense, Socrates AI represents, in addition to a technological achievement, above all an educational vision that translates into method: an artificial intelligence model designed to train generations capable of governing technologies, not being governed by them. With this project, the development of a *new educational paradigm* based on ethical and humanistic values, capable of responding to the challenges posed by the unconscious use of AI, is realized. In this way, a coherence is affirmed between an institutional educational vision of universities and the innovative practices that guide their evolution [3].

The dialogic model at the heart of Socrates AI is unique, grounded in the latest cognitivist and connectionist approaches, and capable of stimulating the development of critical thinking through focused questioning, dialogic strategies, and guided questioning, transforming artificial intelligence into a knowledgeable educational ally. It is designed to critically interrogate knowledge. The system adapts to the specific characteristics of each student, optimizing the interaction and personalizing the learning path. The AI is capable of both proposing cues and analyzing students' responses and providing immediate feedback, helping them to continuously improve their skills and knowledge, pushing them to reason to the limits of their capabilities, and even beyond. Socrates AI thus enables students to have an engaging, interactive and personalized learning experience, and to be active participants in their own growth.

UNINETTUNO University's experience with Socrates AI demonstrates how artificial intelligence can be integrated into educational contexts in an ethical and pedagogically grounded way. The project highlights the potential of artificial intelligence in making education more equitable, accessible, personalized, inclusive, and accountable, if AI is guided by sound educational models. Socrates AI also suggests the possibility of extending similar approaches in other educational contexts, enhancing the potential of artificial intelligence as a lever to improve the quality and inclusiveness of learning at scale. In a context of profound anthropological transformation, artificial intelligence should not replace human thinking, but stimulate in students the ability to interrogate knowledge, make connections between knowledge, and maintain a critical dialogue with emerging technologies. Socrates AI fully embodies this balance, combining the wisdom of tradition with the dynamism of innovation.

After the introductory (Sect. 14.1), in which a concise *general framework* of the transformations induced by generative AI in educational contexts is outlined, the following sections develop an articulated path of analysis and proposal. In Sect. 14.2, the theme of user protection in the era of generative intelligence is addressed, with an examination of the main risks–psychological, cultural, social, regulatory and technological—and the elaboration of a model of operational *ethics* based on five axes of mitigation. Section 14.3 delves into the *cognitive risks* associated with the use of generative models, investigating the effects on autonomy of thought, decision-making processes, and learning patterns. Finally, Sect.14.4 presents *Socrates AI*, an ethical and experimental application of generative intelligence in the service of education. Analyzed here are the theoretical foundations (Sect. 14.4.1), the *architecture* of the system (Sect. 14.4.2), the integration into the *cyberspace* of education (Sect. 14.4.3), the functionality of *certified information* (Sect. 14.4.4), the *maieutic approach* to educational interaction (Sect. 14.4.5), and the emergence of *questioning models* as a new cross-curricular competence (Sect. 14.4.6).

14.2 Risks and Ethics: Protecting the User in the Age of Generative Intelligence

The ethical issue related to generative artificial intelligence is central and cannot be addressed only on an abstract or normative level: it takes the form of a concrete responsibility to protect the user. In this context, designing ethical systems means first of all recognizing potential risks and building devices capable of actively preventing them, putting the safeguarding of the person at the center. Thus, AI ethics requires design oriented toward protecting decision-making autonomy, preventing adverse effects, and ensuring transparency, traceability, and the possibility of human intervention at every stage of the process.

The risks associated with generative models are many and manifest themselves on different levels, often intertwining with each other in complex and not immediately obvious forms. A first critical area concerns the *quality and reliability of content*: the automatic generation of truthful but unfounded texts, the spread of hard-to-detect errors, and the emergence of misleading or manipulated information pose a major epistemic challenge. In these cases, the danger lies not in overt falsehood, but in the *deceptive plausibility* that makes it difficult to distinguish true from false, inducing confusion and informational disorientation. Such phenomena can be amplified by the speed with which content spreads, the lack of transparency in generative processes, and the difficulty of verifying its origin, undermining trust in digital ecosystems.

Alongside these critical issues are risks affecting the *social, cultural and value sphere*. Generative AI models, trained on large amounts of pre-existing and not always filtered data, tend to replicate and amplify systemic biases, gender stereotypes, cultural biases, and geopolitical imbalances. Minority or controversial perspectives can be silenced or reduced to the sidelines, while discriminatory or polarizing content can also be unintentionally generated, contributing to the radicalization of public discourse. Added to this is the user's vulnerability to implicit persuasive dynamics, which can influence opinions, orientations, and behavior without his or her full awareness. In this context, the apparent neutrality of AI proves illusory, and it becomes urgent to question what worldviews are implicitly conveyed through its responses.

A further set of risks is of *normative and legal nature*. The use of generative models raises complex questions about intellectual property – who owns the rights to automatically generated content? – the protection of personal data, the possibility of implicit copyright infringement, and the difficulty in determining legal liability for damage, manipulation, or misuse. The collection and use of data without consent, lack of tracking of interactions, non-compliance with sectoral regulations in regulated domains, and lack of transparency in decision-making mechanisms all contribute to a situation of widespread legal ambiguity, where the existing regulatory framework often appears inadequate to address the implications of these systems.

These are compounded by *technological and operational* risks, which manifest themselves in the form of structural vulnerabilities in the models and systems in which they are embedded. The expansion of the cyber attack surface, the malicious use of AI for purposes of disinformation, fraud, or cybercrime, the possibility of forcing ethical system constraints through *jailbreak* or *prompt injection* techniques, the unauthorized extraction of data through *model stealing*, and the inadvertent release of sensitive information represent real threats. These scenarios highlight how generative AI can help amplify existing vulnerabilities, especially in the absence of complete sharing of standards.

But the most sensitive, and at the same time most overlooked, level is the *subjective and cognitive* one: it concerns the way generative AI affects users' mental processes, conditioning thinking, learning, memory, decision-making and the construction of a sense of reality. This type of risk, less visible but extraordinarily relevant, requires specific design attention. Inadequately oriented generative systems can foster cognitive passivity, dependence on automated responses, the replacement of real social relationships with artificial simulations, and even the undue assumption of therapeutic or emotional support roles in the absence of professional context. In particularly vulnerable individuals, forms of emotional upset, desensitization, or undue influence on sensitive personal decisions may be observed, with tangible behavioral consequences.

From this perspective, AI ethics takes the form of an *practice of epistemic and formative protection*, geared not only to prevent harm, but to actively strengthen the user's capacities, promote reflection, and solicit critical awareness. The integration of generative AI in educational, work and communication environments makes it urgent to define design strategies capable of protecting cognitive autonomy. This entails the need for architectures that include critical use signals, source tracking systems, and interactive models that do not merely provide answers, but stimulate questions, contradictions, and alternative paths. Only in this way will it be possible to integrate generative AI in an ethically sustainable way, preserving the human space of decision-making, learning, and free thinking.

14.3 Cognitive Risks: the Effect of Generative AI on Human Thinking

The reflections we have made so far regarding the effect of generative AI on cognitive processes have focused on possible effects on autonomy of thought, uncertainty management, and critical judgment. But how much of this actually happens? To what extent does interaction with a generative model affect mental functioning? Some recent studies are beginning to provide documented answers.

To gain a more concrete understanding of the nature and extent of these effects, it is useful to dwell on one of the earliest and most significant pieces of experimental evidence: a study conducted at the MIT Media Lab in 2024 [9], significantly titled *Your Brain on ChatGPT*, which investigated the consequences of using generative models within a classic training task such as *writing an essay*. The investigation focused on the cognitive and neurological impact of interacting with an LLM versus other modes of support, integrating *neurophysiological methodologies*, *language analysis* and *qualitative evaluations*. This experimentation is also analyzed here in view of the research perspectives activated at the International Telematic University UNINETTUNO, which is organizing to adopt a similar approach, based on the joint use of neuroscience and psychology, with the aim of systematically investigating the impact of different artificial intelligence techniques applied to education on the human mind, in its cognitive, motivational and emotional processes.

The MIT study involved a total of 54 participants, recruited from university students, researchers and young professionals with experience in academic essay writing. The subjects were between 22 and 38 years old, with a balanced gender distribution and heterogeneous disciplinary backgrounds, including humanities, science and engineering fields. Selection was through a preliminary voluntary application phase, followed by a questionnaire to assess language skills, aptitude for using digital tools, and absence of neurological conditions that could interfere with the use of the EEG device. All participants provided informed consent and were aware that they would be monitored with brain sensing devices throughout the experiment. Prior to the start of the sessions, an introductory brief aimed at explaining the general objectives of the study was given, but without influencing spontaneous behavior during the task. The balance in group composition and the variety of profiles allowed for comparable and meaningful data, representative of different cognitive styles and levels of familiarity with AI.

The results that emerged from the *three sessions* of writing were particularly revealing. EEG analysis showed a clear differentiation in brain connectivity between the groups. The *Brain-only group* showed a more extensive and active network, with strong connections between prefrontal, occipito-parietal, and temporal areas, consistent with high cognitive involvement and autonomous handling of conceptual processing. The *Search Engine group* showed intermediate neural activity. The *LLM group*, on the other hand, manifested a weaker connectivity network, reporting reduced brain engagement in the alpha and beta bands, generally associated with semantic processing, working memory and strategic thinking.

Data collected in the *fourth session* further reinforced this reading. Participants in the *Brain-only group*, once they switched to using LLM, activated a very large neural network, comparable to that of the Search Engine group, particularly in the occipito-parietal and prefrontal nodes. In contrast, those who had used ChatGPT in the first three sessions, *LLM group* and then switched to unsupported writing showed reduced brain activation, indicating a possible form of *cognitive impairment*. These data suggest that repeated use of an LLM can lead to a progressive reduction in neural activation, with residual effects even once use of the tool has ceased. This is, at least, if it is not used judiciously.

On the *language plane*, the NLP analysis showed greater homogeneity in the essays produced by the *LLM group*, with repetitiveness in the n-grams, named entities, and topics covered. In contrast, the essays from the *Brain-only group* showed greater originality, semantic diversity and ability to articulate independently. In post-activity interviews, participants in the LLM group showed significantly less sense of "ownership" (ownership of the text) than the other groups, even struggling to recall or quote newly written content. This disconnect between production and appropriation suggests shallow and unreflective interaction with AI-generated content.

The most alarming aspect that emerged was the tendency for *unconscious incorporation* of the responses provided by the generative model. Participants in the *LLLM group* showed a lower propensity to reconsider their opinions in the presence of contradictory elements, suggesting that AI responses could function as *cognitive anchors* that could silently, but pervasively, direct the entire reflective process. This effect was all the more relevant because it was subjectively denied by the participants themselves, who declared themselves convinced that they were not influenced, despite EEG, linguistic and behavioral data showing otherwise.

In *training environments*, these dynamics pose a crucial pedagogical question. If generative models used without guidance reduce the activation of deep mental resources, while at the same time generating formally coherent but cognitively poor texts, then we risk seeing a progressive erosion of the user's autonomous initiative. This aspect, in educational contexts, is particularly relevant, since these are the places where the development of the ability to argue, manage uncertainty and construct personal judgments is an integral part of the learning process. The speed, fluidity and apparent effectiveness of AI responses risk discouraging the exercise of slowness, error and plural reflection. A system like *Socrates AI* represents a possible response to these risks, and the final sections of this paper will illustrate how mitigation strategies can be operationalized in diverse educational contexts.

In light of these findings, the ethics of artificial intelligence assume a decisive architectural function. Designing responsible systems means building cognitive environments that enhance judgment and stimulate mental activity in its most complex forms. A system such as *Socrates AI* is conceived precisely in this direction, while arising independently of the evidence that has emerged subsequently: its dialogic model does not aim to provide definitive solutions, but rather to generate questioning tensions, stimulate divergent thinking, and solicit metacognitive processes. Its conversational strategies are designed to counter, in a structural way, the anchoring effect noted in the MIT study, returning the user to the space of questioning, contradiction, and autonomous construction of meaning.

Along the same lines, the new research model set out by UNINETTUNO represents a structured evolution of these perspectives, with the aim not only of investigating the impact of artificial intelligence on mental processes, but also of redefining the methods of psychopedagogical research by adopting experimental protocols based on the direct analysis of cognitive, motivational, and emotional dynamics activated in AI-mediated learning contexts.

14.4 Socrates AI: an Ethical Application for Education

Among the many possible applications of generative models, the *Socrates AI* project stands out for its explicit orientation to training and the way it embodies, in an ongoing evolutionary roadmap, the principles of a so-called *operational ethics*.

The reflection conducted thus far shows how addressing the ethical risks of generative artificial intelligence implies, in addition to identifying critical issues, establishing operational conditions for an active, responsible and sensitive response to the contexts in which these systems are deployed. Ethics, in this perspective, does not represent an externally applied constraint a posteriori, but a dimension to be incorporated into the design logics, modes of use and forms of governance that govern their interaction with humans. To speak of *operational ethics* means, therefore, to question how to concretely intervene, at the systemic level, to orient generative AI toward modes of operation that are more transparent, more equitable, and more aware of the variety of domains in which it fits. Among these, the educational domain, where *Socrates AI* operates, represents only one of many possible fields of application, but it is emblematic in making visible the cognitive, relational and cultural implications of employing generative systems, and in highlighting the urgency of integrated strategies for risk mitigation.

As we will see in more detail in the following sections, the system integrates within itself, in a coherent way, a number of principles that guide its use from an ethical perspective. First, it is based on the use of *certified information* from selected, up-to-date and transparent sources, such as approved educational materials, academic resources and specialized databases. Added to this is support for a *maieutic approach*, which structures the dialogue with the user as a process of exploration and activation of critical thinking, drawing inspiration from the Socratic method and encouraging the emergence of questions, hypotheses and reflective paths. Also central is the promotion of correct *questioning models*, still under study, for use with the system, which encourage metacognitive and contextually appropriate practices, accompanying the student in the formulation of clear queries and critical analysis of the answers received.

In this sense, Socrates AI represents a first step toward concrete and advanced implementation on what can be seen as the five main axes on which the operational ethos moves: the adoption of an *architectural mitigation* through the separation of generative component and control modules; the promotion of a *formative mitigation* through conscious questioning training; the implementation of a *contextual mitigation* by dynamically adapting one's communicative behavior; the exercise of a *mitigation moderation* through intelligent monitoring of dialogical exchanges; and finally, the assumption of a form of *mitigation epistemic* by making explicit the criteria and sources on which it bases its responses. Through this conceptual architecture, Socrates AI is proposed as an evolutionary prototype of dialogic artificial intelligence with a formative vocation, capable of supporting learning and promoting forms of cognitive citizenship, respectful of the complexity of situations, individual differences and the need for open, reflective and participatory knowledge.

14.4.1 Theoretical Foundations

The Socrates AI project is rooted in a vision of educational technology that, far from being instrumental or ancillary, is proposed as an integral part of a profound transformation of the relationship between knowledge, subject and society. Its conceptual architecture takes shape from the theoretical elaboration of Maria Amata Garito [5], whose contributions marked a turning point in the way educational interaction is conceived in light of artificial intelligence. In continuity with her psychopedagogical model [3, 10–15], adopted by the International Telematic University UNINETTUNO, Socrates AI was born to inhabit that frontier zone where the digital meets subjectivity, and where the machine does not impose itself as a substitute for the human, but as its possible epistemic ally. These theoretical orientations find a point of synthesis in the idea of an artificial intelligence capable of supporting inclusion, responding to the variety of educational needs and facilitating deep personalization of the learning experience [16].

Underlying this is a historical reflection on transitions in the educational relationship. Garito identifies at least five pedagogical revolutions that have marked the evolution of knowledge transmission: the first with the passage *from family education to schooling*; the second with the *introduction of writing*, which transformed the communication of knowledge from oral to textual; the third with the *invention of printing* and the possibility of archiving and reproducing knowledge on a large scale; the fourth with *electronic and multimedia technologies*, which introduced interactive modes of learning; and finally the fifth, still ongoing, characterized by the adoption of *artificial intelligence* methodologies, which promises to redefine not only the means of education, but also its ends, its cognitive models, and its relational dynamics [5].

Socrates AI fits into the evolutionary path that has seen, over time, profound transformations in the ways through which people construct knowledge, thanks to the progressive adoption of the most advanced means available. From writing to printing, from electronic technologies to digital networks, each step has redefined the dynamics of learning, expanding the possibilities for access, interaction and personalization. Generative artificial intelligence today represents the latest frontier of this transformation, and Socrates AI explores its potential and defines its proper use within learning environments, where dialogue becomes a space for inquiry, reflection and the development of thought.

To gain a deeper understanding of the technological-pedagogical approach of *Socrates AI*, it is useful to recall the frame of reference from which it is inspired: that of the *Intelligent Educational System* (IES), of which it proposes a reinterpretation in light of the potential of generative artificial intelligence. Four basic components can be distinguished in this model: the *expert module* (disciplinary knowledge), the *tutorial module* (instructional strategies and diagnostic skills), the *student model* (dynamically constructed during interactions), and the *interaction module* (interface and dialogue management). Added to these is the *metacognitive dimension*, which emerges as a transversal horizon: interaction with the system aims not only at learning notions, but at developing the ability to reflect on one's learning, critically evaluate one's strategies, negotiate meanings, and transform error into a formative opportunity. In the specific case, *Socrates AI* implements these modules within a learning environment designed according to the psychopedagogical model of the International Telematic University UNINETTUNO: the expert module accesses certified content already integrated into the platform; the tutorial module relies on dialogic questioning models; the student model evolves dynamically as a function of interactions; and the interaction module is embodied in an intelligent textual interface equipped with adaptive conversational strategies.

Formative dialogue, in this perspective, is a co-construction. As Garito points out, learning is a strategic and interactive process, requiring careful design of cognitive objectives, initial diagnosis of the student's profile, selection of the most appropriate strategies, and formative assessment as continuous adjustment of the process. A distinctive feature of the Socrates AI project is precisely its ability to operate within an open epistemic logic. The system does not impose knowledge, but builds conditions for its emergence. The Socratic model, from which it takes its name, represents more than a metaphor: it is the deep structure of interaction. Through questions, raises, counterexamples, displacements and reformulations, the agent stimulates divergent, critical and self-reflective thinking in the student. The dialogue is not aimed at converging on a correct answer, but at exploring the plurality of possible interpretations, valuing uncertainty as an opportunity for intellectual growth.

The integration of AI into training systems has long since begun, but it constitutes a field of experimental research that has become new again with generative AI. Intelligent systems for training, which have become even more complex, are not just tools, but objects of inquiry themselves. The introduction of generative systems changes contexts, dynamics, roles, and makes it necessary to rethink the figure of the teacher as well. It is not a matter of replacing him or her, but of redefining his or her professionalism in light of a lifelong learning society. In a context in which content circulates in increasingly unstable and fragmented forms, training must aim to develop individuals capable of critically analyzing information, questioning its premises and reworking it independently. Socrates AI is proposed as a tool to enable this form of *agency*, preserving the centrality of the student in the construction of knowledge, but also opening spaces for a new alliance between human and artificial, teacher and agent, individual and community.

In the context of educational interaction, Socrates AI stands out for the way it makes the intelligence of mediation visible, translating it into actions that foster reflection, the opening of interpretive scenarios, and the continuous reformulation of knowledge. In dialogue, it stimulates pauses, detours, and questions that set in motion processes of rethinking. In this sense, it acts as a facilitator of learning that does not close in automatisms, but remains open to the unexpected, to error, to complexity. Interaction with Socrates AI thus becomes a permanent laboratory for the development of critical thinking, epistemic awareness and the ability to learn to learn. Within this framework, the project presents itself as a complex educational response to the challenge of the fifth pedagogical revolution: not just a tool, but an actor in the educational ecosystem, oriented to promote a vision of education as a transformation of mind, relationship, and world.

14.4.2 Architecture of Socrates AI

Socrates AI is designed as a modular generative artificial intelligence system, built to support scalable and customizable interactive learning environments. Its architecture adopts a flexible approach, designed to evolve over time and progressively accommodate new components, both in terms of function and performance, ensuring compatibility with the continuous updating of standards in the field of generative AI.

At the core of the system is a generative engine based on large language models (LLMs), integrated into a query-and-response pipeline built according to the *Retrieval-Augmented Generation* (RAG) paradigm [17]. In this architecture, the model's text outputs are guided by contextual information that is dynamically retrieved from a semantically indexed document base. The system uses a hybrid retrieval strategy, combining semantic similarity (via embeddings) [18] with traditional keyword search, in order to improve both coverage and accuracy in finding relevant sources. All information is stored in a *vector database*, optimized for searching within high-dimensional representations. A web-controlled *browsing* module is also under testing, which will enable selective access to external content, while still meeting the strict validation standards required in educational environments. The system is designed to prioritize certified internal sources, always providing clear references for the information returned, and ensuring consistency, traceability, and reliability of the generated content–especially important in learning contexts that require explicit validation. When relevant information is not available, the system refrains from generating unsupported content, thereby avoiding the phenomenon known as hallucination.

The semantic retrieval module is based on a vector representation of documents, obtained by advanced embeddings, which transform preloaded texts or from certified repositories into high-dimensional vectors. User queries are in turn converted into embeddings and compared with the semantic index hosted in the vector database to identify, with hybrid similarity criteria, the most relevant text segments. The system supports dynamic document management, with update, normalization, and quality control processes to ensure source traceability.

The interaction between the generative and retrieval modules is mediated by an orchestration logic that allows each component to be replaced, enhanced, or refined according to technological developments, while keeping the principles of transparency, modularity, and robustness unchanged. The overall architecture of Socrates AI is configured as an open and evolving platform in which each module is designed to be interoperable, upgradeable, and replaceable, with the goal of ensuring reliability, control, and responsiveness to emerging challenges in the educational use of generative systems. It is prepared to support the integration of language filters, moderation rules, and semantic evaluation modules for generated content, as well as user interaction management tools adaptable to different educational contexts. The entire system is designed to support flexible educational communication, capable of responding to different levels of competence and fostering inclusive pathways through progressive personalization of interaction.

To better understand the risks and opportunities, it is useful to clarify the operation of generative AI models a little more generally. The basis of word stream generation is statistical prediction: the system selects the most appropriate word based on linguistic correlations learned during training. Advanced architectures such as *multi-head self-attention* [19, 20] simultaneously evaluate multiple word relationships, capturing semantic nuances and syntactic dependencies with impressive power comparable to that of humans. Alongside this probabilistic mechanism, additional logics can also be introduced to improve logical consistency, argumentative soundness, or meeting criteria of completeness and accuracy. Such components, however, do not eliminate the fact that the model is devoid of understanding: it does not know why an answer is correct, if it is, nor does it possess intentionality or semantic awareness, let alone consciousness. Its effectiveness derives solely from the depth of language training, the power of mathematical correlations that the algorithms can grasp, and the richness of cooperation among algorithms with different tasks.

It is precisely this effectiveness, untethered from real understanding, that generates a *training paradox*: the systems do not understand what they produce, but they manage to do so with such consistency and variety that they are credible. This makes it difficult for the user, or student, to distinguish between genuine understanding and statistical reproduction. The illusion of completeness and reliability risks flattening curiosity, reducing the propensity for verification, and discouraging personal reworking of knowledge. Only training that makes these limitations explicit and exposes them to critical reflection can counteract the tendency toward a passive and de-empowering use of AI. Promoting a conscious and reflective training posture is necessary in order not to succumb to the seduction of ready-made knowledge, rediscovering the formative value of uncertainty, slowness, and argumentative thinking in the age of automation.

In this scenario, the adoption of structured questioning models becomes a central component. Systems such as Socrates AI do not aim to replace the educational process, but to support it through interaction logics that prompt the user to formulate better questions, reflect on their cognitive path, explore alternatives, and question what seems obvious. The technical architecture of the system is thus accompanied by a philosophy of use based on the enhancement of doubt, metacognitive activation and dialogic construction of knowledge.

14.4.3 Socrates AI in Educational Cyberspace

The integration of Socrates AI in the UNINETTUNO Teaching Cyberspace is structural and transversal: the dialogic system is accessible from multiple strategic points in the portal. Students can initiate the interaction: (a) from the *Home Page*, via a dedicated icon; (b) from the *Learning Environments* of individual teachings; (c) from the *Library* section, as a support for the consultation of materials (Fig. 14.1).

Once activated, Socrates AI dynamically calibrates the level of deepening depending on the cognitive profile and the progress of the conversation. Each response aims to clarify concepts and, at the same time, stimulate reflection and analysis, turning uncertainty into a formative opportunity. Thanks to the integration with the University's certified knowledge base, the interaction preserves scientific consistency even in personalized pathways.

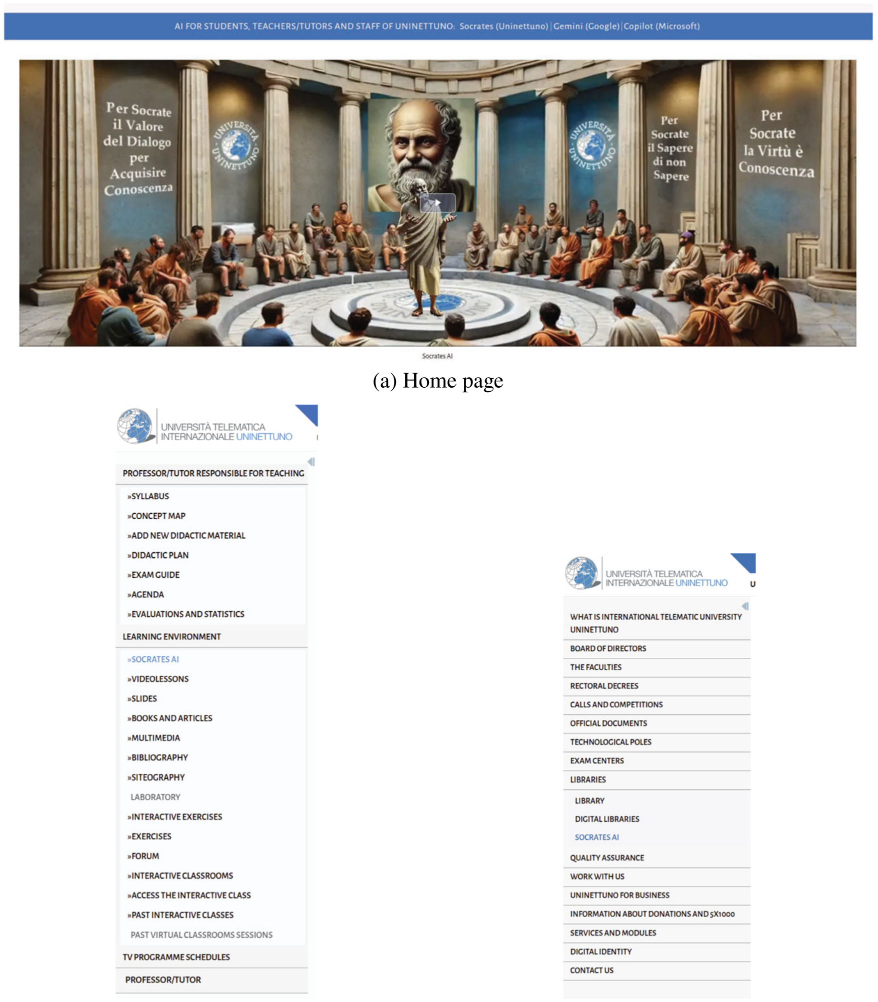
The platform thus promotes a progressive questioning model, adaptable to the disciplinary context and student characteristics. Socrates AI can assist in understanding materials, preparing for exams, navigating resources, and independently constructing study paths. In this way, it acts as an active and reflective presence in the digital learning community.

14.4.4 Certified Information

Socrates AI is designed to critically interrogate knowledge, offering answers grounded in scientifically verified content from the UNINETTUNO Teaching Cyberspace. The certified teaching materials are neatly structured and consist of more than 100,000 h of video lectures and millions and millions of pages: scientific articles, books, forums, bibliographies and reasoned sitographies, exercises and interactive classes, etc. created by professors from many prestigious Italian universities and from different countries of the world and organized according to the University's specific psychopedagogical model, inspired by cognitive and connectionist theories.

Since the human-machine interface of Socrates AI is dialogic, the interaction is developed in an adaptive way: it does not follow a rigidly predetermined course, but is built through the dialogue itself, enhancing the reliability of sources and promoting student autonomy in formulating meaningful questions. One of the main advantages of this approach lies in its ability to make the fruition of static content dynamic: although starting from a certified and pre-loaded information base, information is explored, selected and returned along personalized narrative paths, shaped by the intentions and needs expressed during the interaction. The vast amount of information available can thus be easily accessed and reorganized in real time, in a way that is optimized and consistent with the student's interests and goals, as emerged and negotiated in the dialogue, carried out according to the correct models. The engine proposes certified content, precisely tailored to the student's requests and comments, which they learn to pose in the most effective ways. This approach aims to develop a critical and reflective capacity in the use of artificial intelligence, while always maintaining consistency with the University's educational paths.

The quality and traceability of the answers, together with the absence of hallucinations, constitute a distinctive element of the system, making it a particularly reliable tool in the context of university education. This reliability, based on a generative mechanism that relies on certified and verifiable information, opens up a broader vision: the effectiveness of the model can extend beyond the single local context of use, enabling new forms of cooperation between institutional, academic and technological actors. It is in this horizon that the proposal of a *alliance for knowledge*, promoted by the International Telematic University UNINETTUNO, is placed: a


AI FOR STUDENTS, TEACHERS/TUTORS AND STAFF OF UNINETTUNO: Socrates (Uninettuno) | Gemini (Google) | Copilot (Microsoft)
Per Socrate il Valore del Dialogo per Acquisire Conoscenza
UNIVERSITÀ UNINETTUNO
Per Socrate il Sapere di non Sapere
Per Socrate la Virtù è Conoscenza
Socrates AI
UNIVERSITÀ TELEMATICA INTERNAZIONALE UNINETTUNO
PROFESSOR/TUTOR RESPONSIBLE FOR TEACHING
»SYLLABUS
»CONCEPT MAP
»ADD NEW DIDACTIC MATERIAL
»DIDACTIC PLAN
»EXAM GUIDE
»AGENDA
»EVALUATIONS AND STATISTICS
LEARNING ENVIRONMENT
»SOCRATES AI
»VIDEOLESSONS
»SLIDES
»BOOKS AND ARTICLES
»MULTIMEDIA
»BIBLIOGRAPHY
»SITEOGRAPHY
LABORATORY
»INTERACTIVE EXERCISES
»EXERCISES
»FORUM
»INTERACTIVE CLASSROOMS
»ACCESS THE INTERACTIVE CLASS
»PAST INTERACTIVE CLASSES
PAST VIRTUAL CLASSROOMS SESSIONS
TV PROGRAMME SCHEDULES
PROFESSOR/TUTOR
UNIVERSITÀ TELEMATICA INTERNAZIONALE UNINETTUNO
WHAT IS INTERNATIONAL TELEMATIC UNIVERSITY UNINETTUNO
BOARD OF DIRECTORS
THE FACULTIES
RECTORAL DECREES
CALLS AND COMPETITIONS
OFFICIAL DOCUMENTS
TECHNOLOGICAL POLES
EXAM CENTERS
LIBRARIES
LIBRARY
DIGITAL LIBRARIES
SOCRATES AI
QUALITY ASSURANCE
WORK WITH US
UNINETTUNO FOR BUSINESS
INFORMATION ABOUT DONATIONS AND 5X1000
SERVICES AND MODULES
DIGITAL IDENTITY
CONTACT US

(a) Home page

(b) Learning environment

(c) Library

Fig. 14.1 Socrates AI—platform integration

collaborative network between universities and research centers at a global level, aimed at building a common base of validated, accessible and shared scientific content, capable of supporting the training and responsible use of generative artificial intelligence [21]. In this perspective, *Socrates AI* represents a concrete first step: a dialogic assistant designed to operate on reliable academic content, and at the same time capable of embodying a collaborative model of knowledge construction, in which knowledge becomes common heritage and collective good.

14.4.5 Maieutic Approach

The Socratic method, already evoked in the previous passages, finds in this section a space for closer examination. Reviewing its original structure, as outlined in the Platonic dialogues and its evolutions, allows for a better understanding of how it informs the conceptual and operational choices underlying Socrates AI. Digital maieutics allows the system to enhance the uniqueness of each learner's journey, promoting personalized interaction that takes into account individual times, inclinations and modes of expression.

The Socratic method, as it emerges from Plato's dialogues and the Apology of Socrates, represents one of the purest forms of *training through dialogue*. It did not originate as a codified technique, but as a philosophical attitude and practice of searching for truth, based on the living interaction between master and disciple. In the ancient texts, the Socratic method is articulated through three basic moments: (a) erotesis (ἐρώτησιζ, "question"): Socrates initiates the conversation by asking simple, seemingly naive questions that set in motion the interlocutor's reasoning; (b) exetasis (ἐζέτασιζ, "examination"): starting from the answers received, Socrates critically analyzes the statements, asking for clarifications, more precise definitions, and bringing the discourse to its logical consequences; (c) elènchos (ἔλεγχ, "refutation"): through a series of questions and answers, Socrates leads the interlocutor to uncover internal contradictions in his own thinking, dismantling unfounded certainties and encouraging recognition of his own ignorance. This method does not aim to convey ready-made content, but to stimulate awareness of one's own ignorance (ἀπορία, "impasse," "disorientation") as a necessary condition for the autonomous pursuit of knowledge. In Socrates, therefore, to educate means not so much to teach as to awaken [22].

In modern interpretation, the Socratic method has been articulated in more structured stages, with the aim of making it applicable to teaching in institutional educational settings. Several authors have contributed to systematizing its basic components, recognizing its value in the development of critical thinking, epistemic awareness and personal growth [23]. From this perspective, four main moments are distinguished: (a) aporia (ἀπορία): the student is induced to become aware of his own ignorance or inconsistencies in his own thinking; (b) elènchos (ἔλεγος): one guides the critical analysis of statements, revealing contradictions and promoting the overcoming of erroneous beliefs; (c) anamnesis (ἀνάμνησις, "awakening of knowledge"): one promotes the autonomous reconstruction of knowledge, based on personal insights and reasoning; (d) positive maièutics, maieutiké (μαιευτική, "maieutic art"): one accompanies the student to the synthesis and validation of the knowledge attained, in an act of birth of truth that comes from his or her own thinking. This modern systematization does not betray the original method, but makes explicit its implicit structure, making it suitable for complex educational contexts such as contemporary ones. It emphasizes how paideia (παιδεία, "formation of the soul"), the integral formation of the human being, is not transmission of knowledge, but care of the soul and critical development of reason.

From this systematization, it is possible to describe in more detail the internal dynamics of the Socratic-Maieutic dialogue, as it can be reinterpreted in a contemporary training context or modeled in an interaction with an artificial agent. The sequence of steps should not be understood as rigidly linear, but as a recursive and adaptive progression, in which each step prepares the next and can be reactivated if understanding is not yet consolidated (Fig. 14.2).

The first stage, corresponding to *aporia*, aims to elicit in the student an authentic awareness of the limits of his or her own knowledge. Such awareness is not humiliating, but generative: it constitutes the condition of possibility for cognitive openness. It is elicited through exploratory questions (“What do you mean by X?”), terminological clarifications (“Can you explain in your own words what this concept means?”),

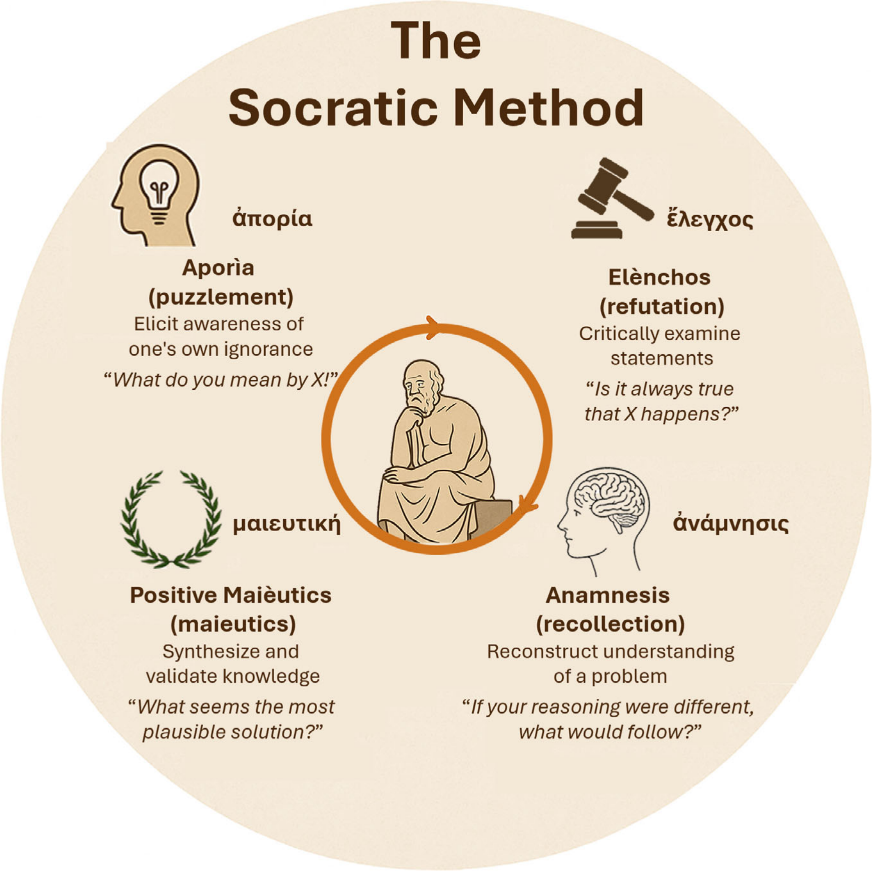


Fig. 14.2 Structured interpretation of the Socratic Method in educational dialogue. The sequence includes aporia, elènchos, anamnesis and positive maieutics, modeled as recursive and adaptive steps

or through the introduction of cases that challenge implicit definitions ("If X means Y, then how do you explain case Z?").

This is followed by the *elènchos* stage, in which thought is examined in its internal coherence. The lecturer or dialogic agent guides the student through a process of problematization: asking questions that bring out latent contradictions ("If you believe X is true, how do you reconcile that with the fact that Y exists?"), that distinguish between what appears and what is ("Is it always true that X happens this way?"), or that interrogate causal relations assumed to be obvious ("Why do you think X is the cause of Y? Are there other possible explanations?").

When the ground has been cleared of inconsistent assumptions, the *anamnesis* phase can begin: the student is stimulated to reconstruct, through guided intuitions, a new understanding of the problem. The interlocutor can propose hypothetical variants ("If your reasoning were different at this point, what would come out of it?"), suggest general reformulations ("Can we express this concept more abstractly?"), or offer heuristic metaphors ("Imagine that X is like–does it help you see the problem differently?").

We then reach the stage of *positive maièutics*, in which the emerging knowledge is synthesized and validated. Here the goal is not to provide a definitive truth, but to consolidate an understanding reached independently. The questions become conclusive and applicative: "In light of what we have discussed, what seems the most plausible solution?", "Can you give me a real-world example where this principle applies?", "Can you explain the concept as if you were to teach it to someone else?"

Through this dynamic, Socratic dialogue takes the form of a cognitive and ethical process at the same time: it forms thinking through shared speech, trains in complexity without shortcuts, and restores responsibility for one's own knowledge to the student.

In the context of Socrates AI, one of the two roles of the dialogue—that of the teacher—is played by a machine, with all the limitations that this inevitably entails. This is not simply to note a technical constraint, but a specific design choice: it is not the intention of the system to take full control of the maieutic initiative, as this would make the interaction less authentic and less formative. Instead, the emphasis is placed on the role of the student, who is empowered to exercise his or her reflective and questioning autonomy. For the maieutic method to take place effectively, it is important for the learner to develop the ability to direct the dialogue, to formulate meaningful questions, and to actively explore hypotheses and connections. It is in this perspective that the architecture of Socrates AI is designed: to support active, conscious and critical learning, without replacing the student's personal journey, but accompanying him or her in his or her development.

One of the most fascinating aspects of Socratic thought concerns his well-known distrust of writing. For Socrates, the value of speech lies not in its fixation in a text, but in the vitality of dialogue, in the possibility of being questioned, challenged, reformulated in the heat of interaction. Writing, in his view, lacks this dynamic: it is a silent word, unable to replicate, to clarify, to evolve. It freezes thought, instead of stimulating it; it conveys content, but does not activate the mind. Education, in the Socratic perspective, is not achieved through the static transmission of notions, but through a process of shared inquiry, in which questions count at least as much as answers. Yet it is precisely through writing that Socrates' thought has been transmitted to us. Plato's dialogues, even in their written form, do not merely expound doctrines, but seek to replicate the open, questioning structure of dialectical confrontation. In this apparent paradox, a new form of writing manifests itself: writing that thinks, that does not close off meaning but opens it up, that does not impose a truth but accompanies the reader on a path of discovery.

Today, generative artificial intelligence introduces a further transformation. The text, which for Socrates was inert, becomes capable of responding, interacting, proposing alternatives. But this vitality, far from being human, is produced by a computational system. The interlocutor is not endowed with consciousness or intentionality. This poses new educational challenges: critical attention to interaction, conscious planning in the use of these tools, and above all active responsibility on the part of the student must be developed. The generated text is no longer dumb, but neither is it truly dialogic, if it is not interpreted and guided within an intentional formative context.

Socrates AI is in this line of continuity and innovation, as an attempt to put dialogue back at the center of learning, in a digital environment. It does not aim to replicate traditional orality, nor does it merely simulate human confrontation: it proposes a new form of written interaction that stimulates reflection and accompanies thinking. To get even closer to the original educational experience of dialogue, the project also includes a return to voice. Voice interaction introduces a more natural and embodied dimension, in which rhythm, tone, intonation and listening contribute to the quality of the relationship. The word thus returns to being spoken, heard, experienced.

In this evolution, the ancient ideal of παιδεία, understood as the formation of the human being through logos, finds a new expression: digital writing, once silent, becomes interlocutive; artificial intelligence, if guided with awareness, can open spaces for authentic thought and contribute to a renewed practice of formative dialogue.

14.4.6 Questioning Models: A Skill Being Formalized

In the framework outlined so far, in which interaction with Socrates AI is conceived as an open and reflective formative space, the quality of the questions formulated by the student continues to be a crucial node. It is not simply a matter of obtaining answers, but of activating a dialogical process in which questioning becomes a tool for exploring, connecting and reworking knowledge.

Already recalled in the maieutic framework, the centrality of questioning finds a more operational articulation here: formulating effective questioning requires awareness, cognitive intentionality and mastery of appropriate dialogical patterns. Therefore, within the Socrates AI project, work is underway to study and formalize questioning models designed to support students in the critical exercise of their intellectual autonomy.

A questioning model should not be understood as a rigid list of formulas, but as a flexible dialogic grammar, capable of adapting to context and cognitive goals. It can include: (a) *exploratory questions*, which open up the field of inquiry; (b) *relational questions*, which connect concepts and phenomena; (c) *critical questions*, which problematize statements or assumptions; (d) *epistemic questions*, which ask for clarification of the status of the knowledge offered; (e) *applicative questions*, which test the usefulness of knowledge in concrete contexts.

In addition to the linguistic form, each question carries with it a *cognitive intentionality*: seeking to understand, testing, connecting, expanding, doubting. It is on these intentionalities that the maturation of thought is built. The reference to Socratic maieutics is decisive here: it is not about getting an answer right, but about initiating a process in which the student recognizes his or her limitations, explores the problem, and reformulates knowledge. It is a dynamic that requires patience, vigilance and assumption of responsibility.

In certain areas of knowledge, this dynamic takes on particularly complex characteristics. Knowledge is never entirely neutral or unambiguous: every interpretation implies a position, every datum is situated, every theory is historically determined. Consequently, questions must become more subtle, more self-conscious. The student does not merely seek answers, but to explore conceptual frames, to compare visions, to interrogate the relationship between knowledge and context. Questions may concern: (a) the existence of *alternative readings* of a phenomenon; (b) the *theoretical matrix implicit* in a definition; (c) how a classification reflects a certain *view of the world*; (d) the possibility of *minority or neglected perspectives*; and (e) the *normative implications* associated with a given position. To train students to formulate questions of this kind is to equip them with tools for navigating complex social systems in which answers are not already given, and truth is constructed through confrontation. From this perspective, questioning is also a political act: it decides what can be questioned, what can emerge, who has a voice.

The Socrates AI project, while not imposing a single model, aims to recognize and enhance these questioning behaviors, promoting a dialogical style that does not end with the return of information, but accompanies the student in the conscious construction of his or her own cognitive path. Formalizing such models thus means strengthening epistemic autonomy, enhancing critical thinking and contributing to the formation of a subjectivity capable of questioning the world.

14.5 Conclusions

UNINETTUNO's experience with Socrates AI shows how artificial intelligence can be integrated into educational contexts according to ethical, inclusive and pedagogically grounded criteria. The project fits within a vision in which AI is used to enrich the learning experience without reducing the centrality of reflective activity on the part of the student.

Socrates AI represents an exemplary case of the application of artificial intelligence to support cognitive autonomy and critical questioning of knowledge. The dialogic agent personalizes access to information, fosters exploration of sources, encourages the construction of individual pathways, and stimulates epistemic awareness through interaction. Students are enabled to formulate meaningful questions, compare perspectives, and interweave knowledge from different domains, thus developing a reflective and active posture in the use of generative technologies.

While this chapter does not yet include quantitative evaluation data, the system is undergoing iterative refinement, and initial informal feedback has informed the current design choices. A structured assessment phase is envisaged, aimed at exploring the cognitive and educational impact of dialogic AI through interdisciplinary methodologies. This perspective opens the way to future empirical validation, without anticipating conclusions that have not yet been systematically tested.

A possible direction for future research concerns the adaptability of the Socrates AI model to different educational levels and settings. Comparative studies across stages such as primary, secondary, and vocational education could provide valuable insights into its scalability and contextual responsiveness.

It is recommended that similar design approaches be adopted by other institutions as well, so that the use of artificial intelligence in education contributes to strengthening the equity, quality and accessibility of learning. In a scenario of accelerated transformation, the priority remains to provide students with tools capable of strengthening critical capacity and responsibility in interacting with generative systems. With this in mind, UNINETTUNO is equipping itself with theoretical and experimental tools to investigate, through integrated neuroscience and psychology approaches, the cognitive impact of different AI techniques on mental activity, in order to guide future design choices with greater awareness.

References

1. M.A. Garito, *New Models of University for the Digital Society* (International Telematic University UNINETTUNO, Roma, 2020)
2. M.A. Garito, *L'università nel XXI secolo tra tradizione e innovazione* (McGraw-Hill Education, Milano, 2015). ISBN 978-88-386-6845-6
3. M.A. Garito, The university of the 21st century, between tradition and innovation, in *Socrates Almanac – Prime Business Destination* (Europe Business Assembly, Oxford, 2015)
4. M.A. Garito, Teaching and learning on the internet: a new model of university, in *The International Telematic University Uninettuno* (Academic Star Publishing Company, 2022), pp. 167–181
5. M.A. Garito, Artificial intelligence in education: evolution of the teaching-learning relationship. Br. J. Educ. Technol. **22**(1), 41–7 (1991)
6. M.A. Garito, *La Television dans les Processus d'Enseignement Apprentissage, in Images and Scientific Education in Europe* (European Science and Technology Forum, CNRS, Paris, 1998)
7. M.A. Garito, Teaching and learning on the internet: a new model of university, in *the International Telematic University UNINETTUNO* (Academic Star Publishing Company, 2023)
8. M.A. Garito, M. Breda, Generative AI, a new model for knowledge memorization: the specific case of the international telematic university UNINETTUNO, in *2024 Yearbook of Emerging*

Technologies in Learning ed. by M.E. Auer, D. May (Cham: Springer, 2025), pp. 44–XX. https://doi.org/10.1007/978-3-031-80388-8. ISBN: 978-3-031-80387-1 (Hardcover), 978-3-031-80390-1 (Softcover), 978-3-031-80388-8 (eBook)

9. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task, 2025. *arXiv preprint* arXiv:2506.08872. Available at: https://arxiv.org/abs/2506.08872
10. M.A. Garito (a cura di), *La Multimedialità nell'Insegnamento a Distanza* (Garamond, Roma, aprile 1996)
11. M.A. Garito, *Tecnologie e Processi Cognitivi: Insegnare e Apprendere con la Multimedialità* (Franco Angeli, Milano, 1997)
12. M.A. Garito, Collaborative learning and virtual laboratories. A new way of teaching and learning on the internet, in *EDULEARN18 Proceedings* (Palma de Mallorca, Luglio 2018), pp. 3582–3587. ISBN: 978-84-09-02709-5, ISSN: 2340-1117
13. M.A. Garito, Reinventing university in the XXI century: new theories and new psycho-pedagogic models for teaching and learning in the internet, in *ICERI 2018 Proceedings*, 11th International Conference of Education, Research and Innovation (Siviglia, Nov 2018), pp. 5249–5256. ISBN: 978-84-09-05948-5, ISSN: 2340-1095
14. M.A. Garito, Reinventing university in the XXI century: the internet, the 'new buildings' of the universities and new psycho-pedagogic models. US-China Educ. Rev. B, Educ. Theory **8**6, 241–247 (2018)
15. M.A. Garito, Teaching and learning on the internet: a new model of university, in *The International Telematic University Uninettuno* (Academic Star Publishing Company, 2022), pp. 167–181
16. M.A. Garito, Universities in dialogue in a world without distance, in *Education Landscapes in the 21st Century: Cross-cultural Challenges and Multi-disciplinary Perspectives* (Cambridge Scholars Publishing, 2008), pp. 35–368
17. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Reizniece, E., Petrov, M., Gurevych, I., Riedel, S., Retrieval-augmented generation for knowledge-intensive NLP tasks. Adv. Neural Inf. Process. Syst. **33**, 9459–9474 (2020). https://papers.nips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf
18. Neelakantan, A., Wang, L., Chen, A., Zhu, J., Chowdhery, A., Mishra, V., Wang, L., Zhou, D., Kaplan, J., Text embedding models: a systematic review, 2024. arXiv preprint arXiv:2406.01607v2. Retrieved from https://arxiv.org/abs/2406.01607v2
19. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, in *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates, Inc., 2017)
20. P. Jin, B. Zhu, L. Yuan, and S. Yan, MoH: multi-head attention as mixture-of-head attention, 2024 . *arXiv preprint* arXiv:2410.11842
21. M.A. Garito, Alliances for knowledge: a strategy for building the future of university in the digital society. Adv. Soc. Sci. Res. J. **10**(7) (2023)
22. Plato, *Apology*, in *Plato: Complete Works*, ed. by J.M. Cooper, trans. by G.M.A. Grube (Indianapolis/Cambridge: Hackett Publishing Company, 1997)
23. L. Nelson, *Socratic Method and Critical Philosophy: Selected Essays* (Dover Publications, New York, 1965)