Article

Not peer-reviewed version

Dialogical Learning Support in RAG-Based E-Learning

Asya Toskova *, Kosta Georgiev, Todorka Glushkova

Posted Date: 12 February 2026

doi: 10.20944/preprints202602.1047.v1

Keywords: retrieval-augmented generation; dialogical learning; e-learning; educational chatbots; large language models; knowledge-grounded dialogue; web based learning platforms; CCA modeling

Article

Dialogical Learning Support in RAG-Based E-Learning

Asya Toskova *, Kosta Georgiev and Todorka Glushkova

Plovdiv University, Bulgaria
* Correspondence: asya_toskova@uni-plovdiv.bg

Abstract

The increasing use of large language models in e-learning environments has created new opportunities for natural language interaction between learners and intelligent systems. However, it has also raised important questions about reliability, transparency, and pedagogical appropriateness. Generative AI systems often depend on heterogeneous and insufficiently validated data sources, which can be problematic in educational contexts where accuracy and trust are essential. This paper presents a web-based learning platform built on a Retrieval-Augmented Generation (RAG) architecture, designed to support dialogical learning. Rather than treating learner questions as isolated inputs, the platform views learning as an ongoing dialogue, preserving context across interactions and grounding responses in curated and validated educational materials. The system uses a modular web-based design that clearly separates content management, retrieval and generation, and dialogue handling. This modularity enables the integration of various generative models – both open-source and commercial – and supports deployment in real e-learning environments without requiring local installation. A representative use case demonstrates how the platform can support learning at the university level. Overall, the study shows how dialogically grounded RAG-based systems can improve transparency, contextual coherence, and pedagogical value in AI-supported e-learning.

Keywords: retrieval-augmented generation; dialogical learning; e-learning; educational chatbots; large language models; knowledge-grounded dialogue; web-based learning platforms; CCA modeling

---

1. Introduction

In recent years, artificial intelligence (AI) has become an increasingly influential factor in the development of e-learning, finding wide application in intelligent tutoring systems, adaptive learning platforms, and virtual assistants. The emergence of large language models (LLMs) has further expanded the possibilities for natural language interaction between humans and systems, enabling learners to ask questions freely, receive explanations, and support their independent learning processes. In this context, dialogical learning has gained importance as a pedagogical approach in which learning unfolds through active and sustained dialogue.

Rather than relying on isolated question–answer exchanges, dialogical learning emphasizes the clarification of concepts, the formulation of follow-up questions, and the construction of meaning through interaction. Learners are viewed as active participants in the learning process rather than passive recipients of information. To support such a learning model, an AI-based system must be capable of maintaining dialogue context over time and providing explanations that are aligned with the educational content and instructional goals.

Despite this potential, the use of generative models in educational settings raises a number of challenges related to information reliability and the quality of the learning process. One of the most frequently discussed issues in LLM-based systems is the generation of so-called "hallucinations"— responses that appear plausible but are factually incorrect or insufficiently grounded. In educational environments, such behavior is particularly problematic, as it may lead to the acquisition of inaccurate knowledge and hinder the formation of correct conceptual understanding. Moreover, many existing AI-based chatbots prioritize rapid answer generation, often without supporting deeper processes of comprehension, reasoning, and incremental knowledge building that are essential for effective learning.

While hallucinations are commonly identified as a major risk to the reliability of generated responses, they can be seen as a surface manifestation of a deeper limitation related to the origin and validity of the knowledge on which these models are trained. Even in the absence of explicit hallucinations, LLM-based systems may produce responses that are conceptually inaccurate, outdated, or based on sources of uncertain credibility. Large language models are trained on vast and heterogeneous text corpora that include internet content, encyclopedic resources, forums, and other publicly available materials. This heterogeneity makes it difficult to distinguish between scientifically validated knowledge, popular interpretations, and subjective opinions. As a result, the model lacks an internal mechanism for epistemic evaluation of information and instead reproduces statistical patterns that reflect frequency and contextual co-occurrence rather than objective truth. This issue is particularly critical in educational contexts, where generated responses are often perceived as authoritative. Learners may accept the provided information without critical reflection, especially when it is presented in a confident and coherent manner, increasing the risk of forming misconceptions that are difficult to correct at later stages of learning.

One promising approach to addressing these challenges is Retrieval-Augmented Generation (RAG), which combines the generative capabilities of language models with mechanisms for retrieving information from a predefined document collection. Through this approach, system responses are grounded in specific educational materials, significantly reducing the risk of hallucinations and improving transparency. In this sense, the goal of RAG is not to guarantee absolute truth, but to establish a controlled and traceable connection between generated responses and a curated set of learning resources that have been selected and validated within the educational process.

Within the context of the proposed platform, this means that the objective is not to replace the instructor or impose an "objective truth," but to support the learning dialogue through verified sources of knowledge and contextual explanation. The system is thus positioned as a mediator in the learning process rather than as an autonomous authority. Although RAG architectures have been successfully applied in document-based retrieval and question–answering systems, their use in dialogical learning contexts remains relatively underexplored.

This paper presents a web-based platform built on a RAG architecture, designed to support dialogical learning in AI-assisted e-learning environments. The platform enables natural language dialogue between learners and the system, grounding responses in specific educational sources. Core processes for content processing, indexing, and updating are implemented as background services, while user interaction is provided through a web interface accessible via a standard browser. This approach facilitates deployment in real educational settings and allows the platform to be used without local installation or specialized client configuration.

The proposed platform extends standard RAG solutions by adapting them to an educational context in which the primary objective is not merely answer delivery, but supporting understanding and learning dialogue. The system is designed for sequential interaction, where previous questions and responses are considered at each subsequent step, and generated explanations are aligned with both the instructional content and the learning context.

This study analyzes the potential of the platform to enhance the reliability and pedagogical value of AI-based e-learning systems. In addition, a representative validation scenario is presented, and possible directions for future development are discussed, including personalization, learning analytics, and multimodal extensions.

To ensure a higher level of verification and validity of the developed processes and interactions in the preliminary development stage, we use modeling through the formal semantics of Calculus of Context-Aware Modeling (CCA).

The remainder of the paper is structured as follows: Section 2 reviews related work on RAG, document-based interaction, and educational chatbots; Section 3 presents the architecture and main components of the proposed platform; Section 4 analyzes how the system supports dialogical learning; Section 5 describes process modeling, implementation details, and a representative usage scenario; Section 6 discusses the results and outlines future research directions; and Section 7 concludes the paper.

2. Related Work

Research in artificial intelligence and e-learning over the past decade has shown increasing interest in systems that support learning through natural language interaction. These systems include a wide range of solutions, from intelligent tutoring systems and educational chatbots to recent architectures based on large language models and hybrid approaches for information retrieval and generation. This section reviews the main research directions relevant to the present study, emphasizing their capabilities and limitations in the context of dialogical learning.

2.1. Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation is an architectural approach that combines the generative capabilities of language models with mechanisms for retrieving relevant information from a predefined document collection. The main idea of RAG is to ground generated responses in specific sources, enabling greater reliability and traceability of the information used.

RAG architectures are widely used in question answering and information access systems that operate over large document collections, including technical documentation, corporate knowledge bases, and regulatory texts. The classical RAG approach introduced by Lewis et al. [1] combines the retrieval of relevant passages with generative models, grounding responses in concrete document sources (so-called document-centric RAG). Similar architectures are also used in document-based question answering systems, where dense passage retrieval and semantic search techniques are applied [2–5].

In industrial contexts, the RAG paradigm underlies many corporate assistants designed to operate on internal documents and organizational knowledge, including systems for navigating procedures, policies, and technical information (e.g., Microsoft Copilot, IBM Watsonx Assistant, Salesforce Einstein GPT). The goal of these solutions is to provide contextually relevant responses grounded in trusted internal sources, increasing their practical value in organizational environments. Although such systems demonstrate high effectiveness in information provision, they are generally not designed with a focus on pedagogical aspects of learning or dialogical interaction [6,7]

Dialogical aspects of information access have also been examined in research on conversational search and conversational question answering, where maintaining context across dialogue turns is central. Theoretical frameworks for conversational search describe retrieval as an interactive process of progressively refining the information need [8], while tasks and datasets for conversational question answering show how classical QA systems can be extended into dialogical contexts [9].

Despite their advantages, most existing RAG solutions focus on providing accurate answers to specific queries, without considering the broader context of learning or dialogical interaction. In many cases, dialogue is treated as a sequence of relatively independent requests, which limits opportunities for in-depth explanation, reflection, and gradual knowledge construction – key characteristics of effective educational processes.

2.2. Educational Chatbots and Dialogical Learning Systems

Educational chatbots and intelligent tutoring systems represent a long-standing research direction aimed at supporting learning through interactive engagement between learners and computer-based systems. Early intelligent tutoring systems (ITS) were built around predefined knowledge models, rules, and pedagogical strategies, enabling them to provide targeted feedback and adaptive instruction [10,11]. Despite their pedagogical precision, these systems require substantial manual effort to model educational content and are limited in flexibility and scalability.

With advances in natural language processing technologies and the emergence of dialog interfaces, there has been growing interest in educational chatbots that offer more natural language interaction and lower barriers to accessing learning content. Analyses of existing solutions show that such chatbots are mainly used to support self-directed learning, provide administrative assistance, and answer frequently asked question [12]. However, in many cases, they function as general-purpose conversational agents without clear alignment to specific learning objectives or pedagogical models.

One of the main challenges in using generative chatbots in education is the lack of control over the structure of the learning dialogue. Although these systems can generate persuasive and linguistically correct responses, they rarely consider learners' prior knowledge, difficulty levels, or the need for gradual conceptual development. This often results in dialogue that remains superficial and fails to serve as an effective instructional tool.

Research on dialogical learning emphasizes the importance of sustained and goal-oriented dialogue for achieving deeper understanding. Analyses of real learning interactions show that effective learning is supported by clarification questions, argumentation, and reflection that develop within the dialogue [13]. In this context, dialogue is viewed not as a series of independent questions and answers, but as a process of joint knowledge construction.

Recent research on conversational question answering demonstrates how classical question–answering tasks can be extended into dialogical contexts that maintain interaction history and contextual continuity across dialogue turns [9]. Despite this progress, the integration of dialogically oriented pedagogical design with architectures that provide a controlled and verifiable source base remains limited.

This highlights the need for educational platforms that combine dialogical interaction with mechanisms for grounding responses in concrete and validated educational documents. In this sense, RAG-based architectures provide a promising foundation for building systems that go beyond the functionality of traditional educational chatbots and support deeper and more meaningful learning.

The present study is situated at the intersection of these research directions by proposing a web-based RAG platform designed to support dialogical learning. Rather than treating dialogue as a sequence of independent queries, the platform conceptualizes it as a sequential learning process in which context, source grounding, and pedagogical objectives play a central role.

3. Proposed RAG-Based Learning Platform

In response to the limitations of existing RAG solutions and educational chatbots, this section presents a web-based platform designed to support dialogical learning through controlled interaction based on specific sources. The platform combines a Retrieval-Augmented Generation architecture with a dialogically oriented design, treating the learning process as a sequential interaction rather than an isolated series of questions and answers.

3.1. Overall Architecture and Design Principles

The proposed platform is a web-based system with a modular architecture that allows flexible extension and adaptation to various educational contexts. The main design principle is a clear separation between processes related to educational content management and those responsible for dialogical interaction with learners. This separation enables the platform to be used for both self-directed learning and as a complement to formal e-learning courses.

The system architecture is conceptually divided into three main layers (Figure 1):

- knowledge management layer
- response management layer
- dialogical interaction layer

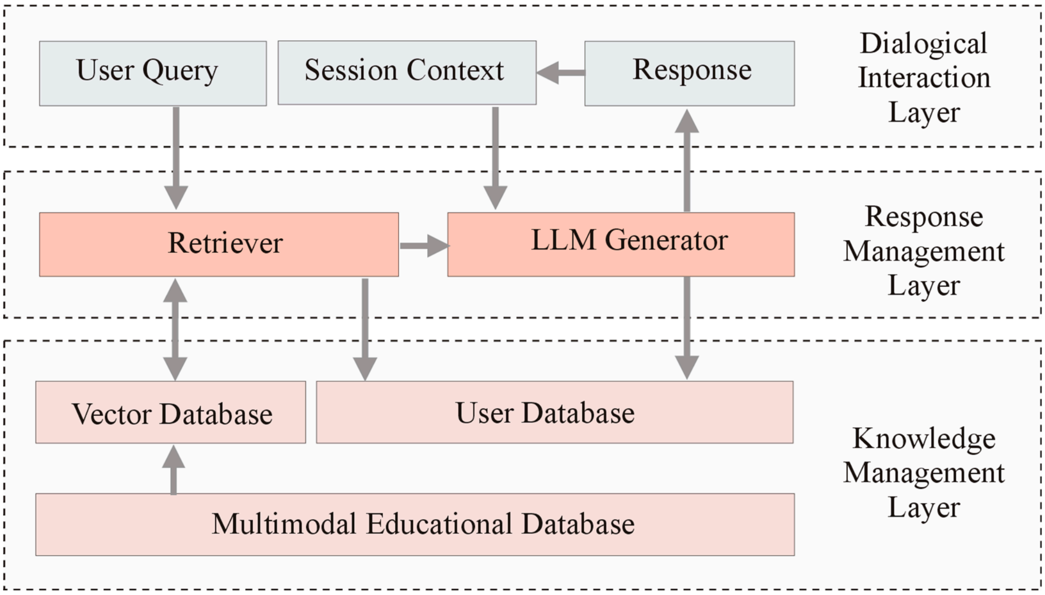

Figure 1. Architecture of the proposed RAG-based e-learning platform for dialogical learning support.

The knowledge management layer handles the storage, processing, and indexing of educational materials. Learning content can be organized into thematic collections corresponding to specific disciplines or courses. Before becoming available for dialogical interaction, the content undergoes preprocessing, including text segmentation and the creation of vector representations, enabling efficient semantic retrieval.

The response management layer implements the core RAG functionality of the platform. For each user query, the system retrieves relevant passages from the educational knowledge base and uses them as context for response generation. Thus, responses are based not only on the language model's abstract knowledge but also on concrete educational materials, increasing the reliability and pedagogical value of the information provided.

The dialogical interaction layer maintains context across dialogue sessions and implements the web interface through which learners interact with the system. Dialogue is modeled as a sequential process in which previous questions and answers are considered when interpreting each new query. This enables the system to provide clarifying, elaborative, or explanatory responses aligned with the progression of the learning dialogue.

The platform supports multiple user profiles through authentication and session management mechanisms. Each learner has an individual profile that stores interaction history and learning progress data. This architectural foundation allows clear separation of context between users and creates conditions for personalized learning support.

The response generation component allows selection of a language model from a predefined list. This approach provides flexibility in using different generative models, including both open-access and commercial solutions. As a result, the platform is not tied to a specific model and can be adapted to the requirements of a given educational environment, available resources, or institutional policies.

The web-based approach facilitates deployment of the platform in real educational settings and enables its use without local installation. In addition, it creates opportunities for integration with

 doi:10.20944/preprints202602.1047.v1

other e-learning systems and for future extensions related to personalization, learning progress tracking, and educational content recommendation.

3.2. Knowledge Base Construction and Content Management

A central component of the proposed platform is the knowledge base, which forms the foundation for dialogical interaction and response generation. Unlike general-purpose language models, which derive their knowledge from large-scale, heterogeneous data collections, the knowledge base in the proposed platform is built from clearly defined and deliberately selected educational materials. This enables for better control over content and alignment with specific educational objectives and contexts.

The learning content is organized into thematic collections that may correspond to individual disciplines, courses, or modules. Each document included in the knowledge base undergoes preprocessing, which involves segmenting the text into logically coherent units. This step is essential for effective retrieval of relevant information and for maintaining semantic coherence during response generation.

After segmentation, textual fragments are transformed into vector representations that enable semantic search and retrieval of content conceptually related to the learner's query. Thus, the platform does not rely solely on keyword matching but also considers the context and meaning of learners' questions. The retrieved fragments serve as context for subsequent response generation and provide a clear link between the dialog outcome and specific educational sources.

In addition to textual documents, the knowledge base can be expanded in a multimodal way with figures, diagrams, and other visual learning resources, which play an important role in understanding complex concepts. These resources can be linked to specific topics or textual fragments and used as complementary context within dialogical interactions. A multimodal knowledge base requires models and frameworks capable of performing semantic search and alignment across different modalities within a shared representational space [14]. Contemporary models such as CLIP [15] enable joint encoding of text and images, supporting effective image–text retrieval and semantic indexing of visual educational materials. Vision–language models such as BLIP [16,17] further extend this functionality through automatic generation and interpretation of image descriptions, which is particularly useful when working with educational figures and illustrations. More general multimodal language models, such as Flamingo [18] and Kosmos-1 [19], demonstrate capabilities for reasoning over combined textual and visual inputs, supporting contextual explanation and example-based learning. In industrial and research contexts, frameworks such as NVIDIA NeMo [20,21] provide infrastructure for building scalable multimodal pipelines that combine text, speech, and image processing, making them suitable for integration into RAG-based educational platforms.

Content management is designed as a separate functional component, allowing flexible updating and expansion of the knowledge base. Instructors or administrators can add new learning materials, update existing documents, or organize content into different thematic structures. When content is modified, the corresponding parts of the knowledge base can be reprocessed and reindexed without disrupting the overall operation of the system. Modules for automatic generation of tests and assessment tasks can also be added to support diagnostic evaluation of learners' levels.

This approach enables the platform to be adapted to diverse educational scenarios, ranging from university courses to specialized training and self-directed learning. Furthermore, the clear separation between the language model and the educational content creates conditions for greater transparency and pedagogical control, as system responses are grounded in concrete, preselected sources rather than in an ill-defined volume of external data.

3.3. Retrieval and Answer Generation Workflow

The workflow for retrieval and answer generation is designed to ensure a clear connection between the learner's query, the educational source, and the generated response. The primary objective is to guarantee that the dialogue is grounded in relevant and contextually appropriate educational content, rather than in abstract or uncontrolled knowledge of the language model.

When a learner submits a question, the system first interprets it within the context of the ongoing dialogue, considering previous questions and answers as well as the current topic. This prevents treating each query as an isolated input and creates the conditions for a coherent and continuous learning conversation.

Next, the system performs semantic retrieval of relevant fragments from the knowledge base. Instead of relying on exact keyword matching, retrieval is based on the semantic similarity between the query and the educational content. The selected fragments provide the most appropriate context for the current question and serve as the basis for answer generation.

After retrieval, the relevant passages are provided to the generative language model as contextual input for forming the response. In this way, the model draws on concrete educational sources from the knowledge base, enabling responses that are more accurate, better aligned with the learning materials, and more suitable for educational purposes.

Answer generation is performed by a selected generative model, with the choice of model configurable depending on the specific usage scenario. Responses are formulated in an explanatory style consistent with the dialogical nature of the interaction. When necessary, they can be further clarified or expanded through subsequent dialogue turns.

The entire workflow – from question interpretation to answer generation – is designed to be transparent and traceable. This allows both learners and instructors to clearly understand the sources on which a given response is based and supports the development of trust in the system as a learning tool.

3.4. Dialog Interaction and Context Handling

Dialogue management is a key component of the proposed platform and plays a central role in its application within an educational context. Dialogical context is maintained over time through the storage and analysis of interaction history. Previous questions, responses, and referenced educational sources are considered when interpreting each new query, enabling the system to recognize references, clarifications, and continuations of previously discussed topics.

This allows learners to ask questions naturally without needing to restate the full context each time. The platform is designed to encourage clarification and deeper understanding. When a query is ambiguous or too general, the system may propose clarifying questions or guide the learner toward more specific aspects of the topic. This approach supports active learner participation and aligns with the principles of dialogical learning, in which knowledge is constructed through interaction.

System responses are adapted not only to the content of the knowledge base but also to the progression of the dialogue. This enables explanations that build upon previous responses, avoid unnecessary repetition, and adjust to the required level of detail at any given moment. As a result, the dialogue acquires a structure similar to pedagogical interaction between an instructor and a learner.

Maintaining dialogical context also creates opportunities for subsequent analysis of the learning process. Interaction history can be used to identify difficulties, frequently asked questions, or topics that require further explanation. This opens possibilities for future extensions of the platform with functionalities for personalized learning support.

By integrating mechanisms for dialogue and context management, the proposed platform goes beyond the functionality of standard RAG systems and serves as a tool for learning support. In this way, the technological architecture is directly aligned with the pedagogical objectives of dialogical learning.

3.5. Regulatory and Ethical Constraints in Educational Use of Language Models

The use of language models and intelligent tutoring systems in education within the European Union is primarily regulated by the Artificial Intelligence Act (AI Act), adopted in 2024 and gradually implemented from 2025 to 2026 [22]. The regulation introduces a risk-based approach by classifying AI systems according to their potential impact on the fundamental rights of citizens.

High-Risk Applications of Language Models in Education

According to the AI Act, AI systems used in education are classified as high-risk when they:

- perform automated assessment or certification of learners;
- influence access to education or qualifications;
- profile learners to predict behavior.

In this context, language models are not permitted to make autonomous decisions that have academic or legal consequences for learners without mandatory human oversight. Automated assessment and decision-making in education require a human-in-the-loop and additional measures to ensure transparency and accountability.

Permissible Applications of Language Models in Education

The European Commission permits the use of language models in education when their role is supportive rather than decisive. Permissible scenarios include cases in which language models:

- support learning through explanations, examples, and guidance;
- participate in dialogical learning and reflective interactions;
- are used for self-study and practice.

Under these conditions, such systems are considered low-risk or limited-risk, provided that interaction with AI is clearly disclosed and transparency regarding the data and sources used is ensured.

European regulations explicitly prohibit inferring sensitive characteristics, such as cognitive abilities or academic potential, without informed consent and human oversight [23–25]. Systems must clearly inform users that they are interacting with AI and must avoid creating the impression that the system provides authoritative or final decisions. In this sense, a RAG-based platform for dialogical learning that uses language models to support learning complies with current European regulatory and ethical requirements.

4. Dialogical Learning Support

The primary objective of the proposed platform is to support the learning process through dialogue. This process is seen as sequential and active, with learners constructing understanding by asking questions, clarifying concepts, and gradually building their knowledge. The architectural decisions described in the previous section are specifically designed to support this type of interaction.

The platform supports dialogical learning by maintaining context over time and tailoring responses to the progression of the learning dialogue. Building on previously discussed topics fosters deeper understanding of the material and prevents fragmented information acquisition.

A key aspect of dialogical support is the ability to provide clarification and guidance. When a learner poses an overly general or unclear question, the platform may respond with an explanation or suggest a more specific focus. This approach resembles a pedagogical strategy in which an instructor directs attention to essential elements of the content rather than providing a direct and final answer. In this way, learners are encouraged to participate actively in the dialogue, formulate more precise questions, or independently arrive at the answer.

System responses are formulated to support understanding rather than simply provide factual information. This includes using an explanatory style, linking new concepts to previously introduced

doi:10.20944/preprints202602.1047.v1

ones, and avoiding unnecessarily complex terminology when it is not needed. Within the dialogue, the system can provide different levels of detail depending on the conversation's development and the learner's needs.

An important element of dialogical learning is the connection between generated responses and the educational sources used. Learners can trace the relationship between their questions and the corresponding learning materials, enhancing the transparency of the learning process and trust in the system.

Through these mechanisms, the platform creates conditions for meaningful and coherent learning. In this context, it serves as a tool for supporting learning dialogue that complements the role of the instructor and traditional e-learning resources.

5. Modeling and Implementation of Learning Scenario

Educational platforms are context-aware in various aspects, meaning they dynamically consider changes in the educational environment and use this information to adapt and provide appropriate learning resources and services to the user [26].

5.1. CCA Modeling of Learning Scenario

When developing a real educational platform that integrates heterogeneous, dynamically changing modules, the preliminary modeling process is particularly important. The formalization of the context-aware system can be represented and modeled using the formal mathematical notation CCA (Calculus of Context-aware Ambients), which fully realizes the process of the π-calculus [27]. A CCA ambient is an identity through which a specific object or component is associated. Each ambient has a name, boundaries, and can contain other ambients within itself. Ambients can change their location and communicate with other ambients. There are three possible relationships between two ambients: parent, child, and sibling. Each ambient communicates with other ambients. The message exchange process is carried out using handshaking. In CCA notation, when two ambients exchange a message, three notations are used: "::" for interaction between sibling ambients; "↑" and "↓" for child-parent and parent-child communication; "<>" indicates sending a message, and "()" indicates receiving a message. Processes P are a basic syntactic category in CCA modeling. The process of each ambient defines its behavior and communication with other ambients in the system. Process 0 does nothing and terminates immediately. Process P|Q indicates that process P is running in parallel with process Q [28].

The main ambients used in the CCA model of the learning scenario are presented in the following Table 1.

Table 1. CCA Ambients.

Notation	Description
PA	Personal Assistant of Student
RT	Retriever Component
SC	Session Context Module
LLMG	LLM Generator
KM	Knowledge Manager
VDB	Vector Database, child of KM ambient
UDB	User Database, child of KM ambient
MEDB	Multimodal Educational DB, child of KM ambient

The PA, RT, SC, LLMG, and KM ambients are siblings in the ambient hierarchy, and VDB, UDB, and MEDB are children of the KM ambient. This basic learning scenario can be represented by the following CCA model as processes of the participating ambients:

$$P_{PA} \equiv \begin{pmatrix} RT ::< PAi, UserQuery > .0| \\ SC ::< PAi, UserQuery > .0| \\ LLMG :: (Response).0 \end{pmatrix}, \quad (1)$$

$$P_{RT} \equiv \begin{pmatrix} PA :: (PAi, UserQuery).KM ::< PAi, UserQuery > .0| \\ KM :: (PAi, LearningKnowlege). \\ LLMG ::< PAi, UserQuery, LearningKnowlege > .0 \end{pmatrix}, \quad (2)$$

$$P_{LLMG} \equiv \begin{pmatrix} RT :: (PAi, UserQuery, LearningKnowlege). \\ SC :: (PAi, LearningContext). \\ PA ::< Response >.RT ::< PAi, Response > .0 \end{pmatrix}, \quad (3)$$

$$P_{SC} \equiv \begin{pmatrix} PA :: (PAi, UserQuery). \\ LLMG ::< PAi, LearningContext > .0 \end{pmatrix}, \quad (4)$$

$$P_{KM} \equiv \begin{pmatrix} RT :: (PAi, UserQuery). \\ VDB \downarrow< PAi, UserQuery >. \\ UDB \downarrow< PAi, UserQuery > .0| \\ VDB \downarrow (PAi, LearningKnowlege). \\ RT ::< PAi, LearningKnowlege > .0| \\ LMMG :: (PAi, Response).UDB \downarrow< PAi, Response > .0 \end{pmatrix}, \quad (5)$$

$$P_{MEDB} \equiv (VDB ::< EducationalResources > .0|), \quad (6)$$

$$P_{VDB} \equiv \begin{pmatrix} KM \uparrow (PAi, UserQuery). \\ MEDB :: (EducationalResourses). \\ KM \uparrow< PAi, LearningKnowlege > .0 \end{pmatrix}, \quad (7)$$

$$P_{UDB} \equiv \begin{pmatrix} KM \uparrow (PAi, UserQuery).0| \\ KM \uparrow (PAi, Response).0 \end{pmatrix}. \quad (8)$$

The ccaPL programming language is a machine-readable version of CCA syntax. Let h be a conversion function from CCA to ccaPL; that is, for a process P in CCA, h(P) is the corresponding ccaPL notation. The ccaPL interpreter is developed as a Java application. Therefore, a simulator can be created to verify the scenario described above, which can be used for preliminary testing and verification of the designed processes. To assist specialists in the preliminary modeling of processes and scenarios in various context-dependent systems, we have created a visual CCA editor through which the designed services can be modeled, analyzed, and optimized before real development begins [29].

5.2. Implementation of Learning Scenario

The proposed platform is designed for practical application and validation in a real e-learning environment, without requiring specific client-side configuration or local installation. Its web-based approach enables learners and instructors to interact with the system through a standard web browser, facilitating deployment in university courses and educational platforms.

From an implementation standpoint, the platform's main functional components are organized as separate services. The processes for processing and indexing educational content operate independently of the user interface and can be activated when learning materials are added or updated. The dialog interface provides access to the RAG functionality and maintains the history of interactions with learners.

Within the platform, the response generation component allows selection of a generative language model from a predefined list. This design enables the system to operate with different models, including both open-access and commercial solutions, depending on the specific usage scenario, available resources, or institutional policies. As a result, the platform is not limited to a single model and can be adapted to diverse requirements related to quality, cost, or data governance.

5.3. Example Scenario

To demonstrate the practical application of the proposed platform, a prototype was developed based on the RAG-based design principles described in previous sections. The prototype allows learners to upload their own educational materials for a specific university-level course and interact with an AI assistant strictly within the scope of the uploaded content.

Educational materials, such as lecture presentations and PDF documents, are linked to a dedicated course profile that stores contextual information, including the course name, the AI assistant's role, and additional configuration parameters (Figure 2). These profile-level settings provide contextual grounding for the interaction and ensure that generated responses remain aligned with the specific learning scenario.

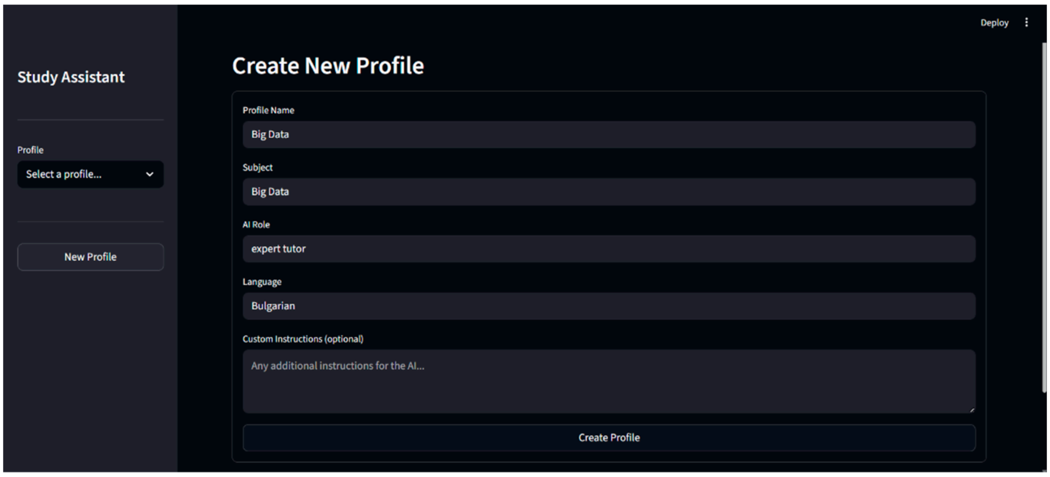

Figure 2. Creation of a course-specific profile used to define contextual parameters.

After materials are uploaded, they are converted to text and processed for semantic indexing (Figure 3). The content is segmented into overlapping text fragments and transformed into vector representations using a sentence embedding model. The resulting vectors are stored in a vector index to enable efficient semantic retrieval.

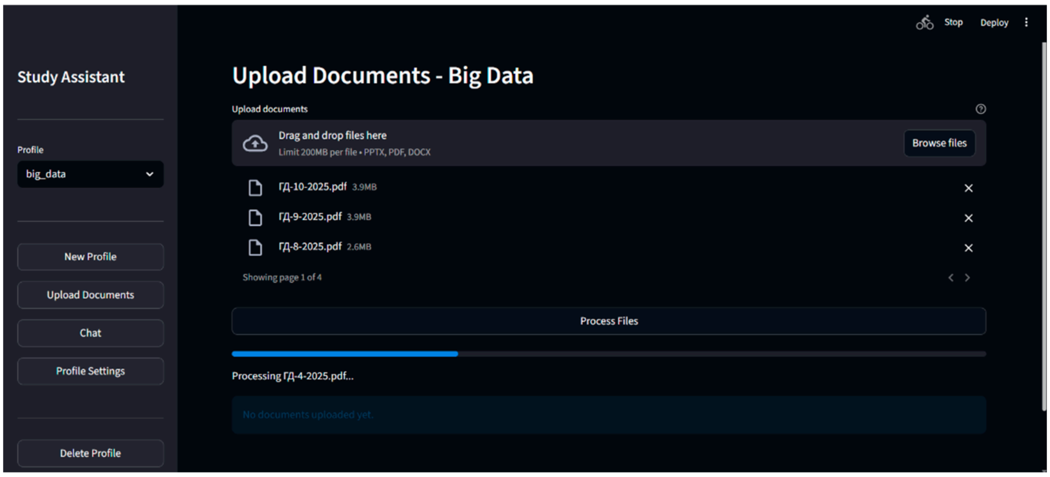

Figure 3. Upload and preprocessing of course materials.

During interaction, learner queries are processed using a hybrid retrieval strategy that combines semantic similarity search with keyword-based retrieval. The retrieved content fragments, along with

 doi:10.20944/preprints202602.1047.v1

the user query and relevant contextual information, are provided as input to a large language model, which generates responses grounded exclusively in the uploaded educational materials (Figure 4). This ensures that the interaction remains limited to validated course content and supports consistent, context-aware dialogue.

In the illustrated example, the dialogical interaction is conducted in Bulgarian, as the responses are generated exclusively from the uploaded learning materials. However, the platform itself is language-independent and can operate in different interaction languages depending on the selected model and configuration.

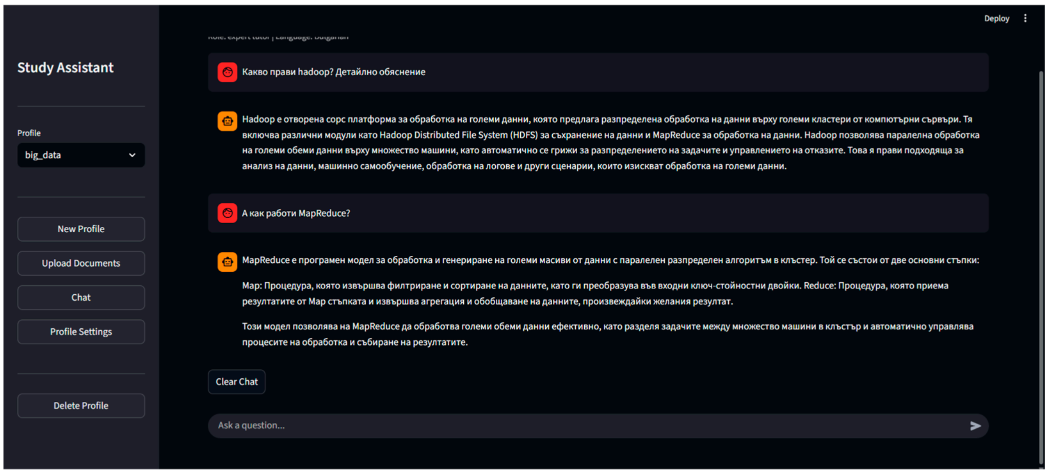

Figure 4. Example of contextual dialogical interaction between a learner and the AI assistant.

The prototype is implemented entirely in Python and is accessible through a web-based user interface built with the Streamlit framework, allowing access via a standard browser without local installation. Vector indexing and semantic retrieval are implemented using FAISS, while textual embeddings are generated with the all-mpnet-base-v2 model from the Sentence Transformers library. The hybrid retrieval process combines semantic search with the BM25 algorithm. Response generation is performed through the OpenAI API using the GPT-3.5-turbo model.

This prototype demonstrates the feasibility of the proposed architecture and validates that a RAG-based approach can be effectively applied to support dialogical learning in real educational scenarios.

The provided implementation details are intended to support reproducibility of the proposed approach in comparable educational settings.

5.4. Practical Considerations

The proposed approach enables the platform to function both as a standalone tool for learning support and as a complement to existing e-learning systems. The web interface supports various usage scenarios, from individual self-directed learning to support for courses delivered in face-to-face or distance learning formats.

This implementation shows how the architectural and dialogical principles described in the previous sections can be applied in practice to support real learning processes, without undermining the pedagogical role of the instructor or the structure of the course.

6. Discussion and Future Directions

This paper presents a web-based platform built on a Retrieval-Augmented Generation (RAG) architecture, with the primary objective of supporting dialogical learning in an e-learning environment. Unlike traditional RAG solutions and general-purpose educational chatbots, the proposed system is specifically designed to focus on sequential learning dialogue, interaction context,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.1047.v1

and grounding in specific educational sources. In this approach, technological solutions are subordinated to pedagogical goals rather than the reverse.

The design of the learning scenario was additionally supported by formal modeling using the Calculus of Context-aware Ambients (CCA), which provided a structured view of component interactions and informed the system design, without constraining the pedagogical flexibility of the platform.

A key conclusion of this study is that the use of RAG architectures in education should not be seen as a guarantee of the "truthfulness" of generated content. Even when the source base is limited, the quality and validity of knowledge depend on the selection and currency of the educational materials used. Thus, the platform's role is not to replace the instructor or established educational practices, but to provide transparent and controlled mediation between learners and learning content. The platform is designed in accordance with current European regulatory and ethical guidelines, positioning language models as tools to support dialogical learning.

A major advantage of the proposed platform is its modular and flexible architecture, which allows for the selection of a generative model based on the specific usage scenario. This enables adaptation to different institutional policies, data protection requirements, and available resources, as well as future integration of new models without substantial architectural changes. Such an approach is especially important in educational contexts, where sustainability and long-term applicability of systems are critical.

The knowledge base is designed to extend beyond textual documents to include figures, diagrams, and other educational resources associated with specific topics and concepts. A multimodal RAG approach, in which visual resources are indexed and retrieved alongside textual content, enables richer explanations of learning materials, particularly in disciplines with a strong visual component.

Despite the platform's potential, its limitations must be acknowledged. The effectiveness of dialogical learning depends not only on architectural decisions, but also on the quality of educational content, the formulation of questions, and the active participation of learners.

Several directions for future development are possible. One planned functionality is the ability for instructors to generate tests and assessment tasks in multiple variants based on the educational content included in the platform. Such tests may be used both for knowledge checking and as a source of information about learners' current levels. The accumulated results from dialogical interactions and level diagnostics could serve as a basis for personalization, adjustment of explanation granularity, and recommendation of appropriate learning content. Additionally, analysis of dialogical interactions may help instructors identify difficult topics and frequently encountered learning challenges.

7. Conclusions

In conclusion, the proposed platform demonstrates how RAG architectures can be adapted and used as a foundation for dialogically oriented educational systems. By combining controlled educational sources, maintenance of dialogical context, and a web-based design, the platform offers a balanced approach between the technological capabilities of contemporary language models and the pedagogical requirements of effective learning.

There are different directions for the future development of the system. We believe that expanding into a cyber-physical and social (CPSS) educational platform, in which services and interactions take into account dynamic changes in the physical, virtual, and social worlds, can provide a more natural and effective environment for learning and self-training [30]. One direction for development is the use of such a personalized cyber-physical RAG platform in the education of students with special educational needs (especially students from the autistic sector), as well as in interest clubs, where interaction with the surrounding physical world and dialogic learning can significantly increase activity, motivation, and effectiveness of training in a multidisciplinary aspect.

Author Contributions: Conceptualization, A.T.; methodology, A.T.; software, K.G.; validation, K.G.; formal analysis, T.G.; investigation, A.T.; writing—original draft preparation, A.T., K.G., T.G.; writing—review and editing, A.T., T.G; funding acquisition, T.G.

Funding: This study is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Re-public of Bulgaria, project № BG-RRP-2.004-0001-C01 and supported by the project FP25-FMI-010 "Innovative interdisciplinary research in informatics, mathematics and educational pedagogy" at the Plovdiv University "Paisii Hilendarski".

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. K. H. Goyal, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel and D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks" Computer Science, Computation and Language, 2020, https://doi.org/10.48550/arXiv.2005.11401.
2. D. Chen, A. Fisch, J. Weston and A. Bordes, "Reading Wikipedia to Answer Open-Domain Questions" Computer Science, Computation and Language, 2017, https://doi.org/10.48550/arXiv.1704.00051.
3. V. Karpukhin, B. Oğuz, S. Min, P. Lewis, L. Wu, S. Edunov, D. Chen and W. Yih, "Dense Passage Retrieval for Open-Domain Question Answering" Computer Science, Computation and Language, 2020, https://doi.org/10.48550/arXiv.2004.04906.
4. K. Guu, K. Lee, Z. Tung, P. Pasupat and M.-W. Chang, "REALM: Retrieval-Augmented Language Model Pre-Training" Computer Science, Computation and Language, 2020, https://doi.org/10.48550/arXiv.2002.08909.
5. G. Izacard and E. Grave, "Leveraging Passage Retrieval with Generative Models" Computer Science, Computation and Language, 2021, https://doi.org/10.48550/arXiv.2007.01282.
6. E. Dinan, S. Roller, K. Shuster, A. Fan, M. Auli and W. J., "Wizard of Wikipedia: Knowledge-Powered Conversational Agents" Computer Science, Computation and Language, 2019, https://doi.org/10.48550/arXiv.1811.01241.
7. A. Radford, J. Wu, R. Child, D. Luan, D. Amodei and I. Sutskever, "Language Models are Unsupervised Multitask Learners" OpenAI Technical Report, https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_-learners.pdf, 2019.
8. F. Radlinski and N. Craswell, "A Theoretical Framework for Conversational Search" Proceedings CHIIR '17A, pp. 117-126, https://doi/10.1145/3020165.3020183, 2017.
9. S. Reddy, D. Chen and C. Manning, "CoQA: A Conversational Question Answering Challenge" Computer Science, Computation and Language, p. https://doi.org/10.48550/arXiv.1808.07042, 2019.
10. K. VanLehn, "The Behavior of Tutoring Systems" International Journal of Artificial Intelligence in Education, 16(3), 2006, 227-265, https://doi/10.5555/1435351.1435353.
11. B. P. Woolf, Building Intelligent Interactive Tutors: Student-centered Strategies for Revolutionizing E-learning, Morgan Kaufmann, ISBN 0080920047, 9780080920047, 2010.
12. R. Winkler and M. Söllner, "Unleashing the Potential of Chatbots in Education: A State-of-the-Art Analysis" Academy of Management Annual Meeting Proceedings (1), 2018, https://doi.org/10.5465/AMBPP.2018.15903abstract.
13. A. Graesser, N. Person and J. Magliano, "Collaborative Dialogue Patterns in Naturalistic One-to-One Tutoring" Applied Cognitive Psychology, 1995, https://doi.org/10.1002/acp.2350090604.
14. R. E. Mayer, „Multimedia Learning (3rd ed.)" Cambridge University Press., ISBN 9781107187504, https://doi.org/10.1017/9781316941355 , 2020.
15. A. Radford, J. Kim, C. Hallacy, R. A. G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger and I. Sutskever, "Learning Transferable Visual Models From Natural Language Supervision (CLIP)" Computer Science: Computer Vision and Pattern Recognition, https://doi.org/10.48550/arXiv.2103.00020, 2021.

16. J. Li, D. Li and C. H. S. Xiong, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation." Computer Science: Computer Vision and Pattern Recognition, https://doi.org/10.48550/arXiv.2201.12086, 2022.
17. J. Li, D. Li, S. Savarese and S. Hoi, "BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models" Computer Science: Computer Vision and Pattern Recognition, https://doi.org/10.48550/arXiv.2301.12597, 2023.
18. J. Alayrac, J. Donahue, P. Luc, A. Miech, I. Barr, Y. Hasson and K. Simonyan, "Flamingo: A Visual Language Model for Few-Shot Learning" Computer Science: Computer Vision and Pattern Recognition, p. 54, https://doi.org/10.48550/arXiv.2204.14198, 2022.
19. S. Huang, Y. Zhang, Y. Wang, L. Zhao, Z. Li, Y. Chen and J. Gao, "Language Is Not All You Need: Aligning Perception with Language Models" Computer Science: Computation and Language, https://doi.org/10.48550/arXiv.2302.14045, 2023.
20. O. Kuchaiev, B. Ginsburg, I. Gitman, V. Lavrukhin, C. Case, P. Micikevicius and Y. Zhang, "NeMo: A Toolkit for Building AI Applications Using Neural Modules" Computer Science: Machine Learning, p. 6, https://doi.org/10.48550/arXiv.1909.09577, 2019.
21. NVIDIA Corporation, "NeMo Framework for Multimodal Conversational AI" Technical Report, pp. https://docs.nvidia.com/nemo-framework/user-guide/latest/overview.html, 2023.
22. European Union, "Artificial Intelligence Act" https://artificialintelligenceact.eu/, 2024.
23. European Commission, "Ethical Guidelines on the Use of Artificial Intelligence and Data in Teaching and Learning" Publications Office of the European Union,. https://education.ec.europa.eu/news/ethical-guidelines-on-the-use-of-artificial-intelligence-and-data-in-teaching-and-learning-for-educators, 2022.
24. European Commission, "Proposal for a Regulation laying down harmonised rules on artificial intelligence" https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence, 2021.
25. European Commission, "Ethics Guidelines for Trustworthy AI" https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, 2019.
26. A. Dey, "Understanding and Using Context" Personal and Ubiquitous Computing Journal, vol. 5, pp. 4-7, https://doi.org/10.1007/s007790170019, 2001.
27. J. Riely and M. Hennessy, "A typed language for distributed mobile processes" in: Proc. 25th Annual ACM Symp. on Principles of Programming Languages, pp. 378–390, https://doi.org/10.1145/268946.268978, 1998.
28. F. Siewe, H. Zedan and A. Cau, "The calculus of context-aware ambients" Journal of Computer and System Sciences, vol. 77, no. 4, pp. 597-620, https://doi.org/10.1016/j.jcss.2010.02.003, 2011.
29. T. Glushkova, K. Rusev and D. Stoyanov, "Multi-Context Modeling of Processes and Services in Cyber-Physical Educational Space" In 2023 International Conference on Big Data, Knowledge and Control Systems Engineering (BdKCSE), pp. 1-8, doi: 10.1109/BdKCSE59280.2023.10339749, 2023.
30. S. Stoyanov, T. Glushkova, A. Stoyanova-Doycheva, J. Todorov, A. Toskova (2020). A Generic Architecture for Cyber-Physical-Social Space Applications. In: Jardim-Goncalves, R., Sgurev, V., Jotsov, V., Kacprzyk, J. (eds) Intelligent Systems: Theory, Research and Innovation in Applications. Studies in Computational Intelligence, vol 864. Springer, Cham. https://doi.org/10.1007/978-3-030-38704-4_14