IJIRIS: International Journal of Innovative Research in Information Security
Volume 12, Issue 08, November 2026
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2026.v1208.01

COGNILEARN: A Multimodal Adaptive AI Learning Assistant for Personalized and Interactive Education

Dr Manjushree C V
Department of Information Science and Engineering
Vemana Institute of Technology
Bengaluru, India
manjushree.vit@gmail.com

Nachiketh K
Department of Information Science and Engineering
Vemana Institute of Technology
Bengaluru, India
nachikethkrishna.e2004@gmail.com

Prajwal S
Department of Information Science and Engineering
Vemana Institute of Technology
Bengaluru, India
prajwal0210@gmail.com

Ranjith Reddy A
Department of Information Science and Engineering
Vemana Institute of Technology
Bengaluru, India
ranjithreddy0845@gmail.com

Sagar M D Nadig
Department of Information Science and Engineering
Vemana Institute of Technology
Bengaluru, India
sagarmdnadig@gmail.com

***Abstract*—Traditional digital learning platforms often provide largely uniform content and limited interaction, making it difficult to address differences in learner pace, prior knowledge, strengths, and weaknesses. This paper presents COGNILEARN, a proposed multimodal adaptive AI learning assistant intended to provide personalized and interactive educational support. The platform is designed to accept text-, image-, audio-, and video-based learning inputs and combine Natural Language Processing (NLP), machine learning, computer vision, speech processing, and generative AI. Planned capabilities include AI-assisted doubt clarification and concept explanation, adaptive recommendations, quiz and assessment support, progress tracking, and learning analytics. The proposed architecture comprises a learner-facing interface, application and API services, educational-data storage, modality-specific processing, an adaptive learning engine, and an analytics dashboard. The supplied Phase-1 report describes the design and expected functionality but does not provide a reproducible benchmark dataset, measured performance metrics, or a completed experimental evaluation. Accordingly, this paper presents the architecture and an evaluation plan without attributing unverified accuracy, latency, or learning-gain figures to the system.**

***Index Terms*—Multimodal learning, adaptive learning, intelligent tutoring, generative AI, learning analytics, NLP, personalized education.**

I. Introduction

A. Background and Motivation

Online education and digital learning platforms make educational material widely accessible, but many systems still depend on fixed learning sequences and limited forms of interaction. Students differ in their preferred learning formats, understanding of a topic, and pace of progress. A platform that considers only one type of input or a small number of assessment scores may not provide enough context to personalize support.

COGNILEARN is described in the supplied project report as a multimodal adaptive AI learning assistant. It is designed to bring multiple learning-content formats and AI-assisted tutoring together in one platform. The proposed system aims to let learners ask questions, receive explanations, practise through quizzes, access recommendations, and monitor their progress.

B. Problem Statement

Conventional learning platforms may not adequately adapt learning content and support to individual student needs. They can provide limited support for mixed media, immediate doubt clarification, identifying weak areas, and recommending the next learning activity based on learner progress. COGNILEARN addresses this design problem by proposing a unified interface for multimodal educational interaction, AI tutoring, adaptive recommendations, and learning analytics.

C. Objectives

The aim is to design a personalized learning platform that supports text-, image-, audio-, and video-based learning and provides AI-driven assistance, adaptive recommendations, and progress monitoring. The project objectives are to:

1) Design a multimodal learning interface for educational material and learner questions.
2) Provide AI-assisted doubt clarification and concept explanations.
3) Use learner interactions and assessment history to guide personalized recommendations.
4) Track quiz performance and learning progress through analytics and visual dashboards.
5) Provide a modular architecture that can accommodate additional processing and deployment services.

D. Scope and Contributions

The supplied report describes the intended system and its components, rather than a verified production deployment. This paper therefore covers the proposed architecture, workflow, major modules, and a practical evaluation plan. Features described as planned or optional should be confirmed against

IJIRIS: International Journal of Innovative Research in Information Security
Volume 12, Issue 08, November 2026
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2026.v1208.01

the implementation before being represented as completed capabilities.

The defensible project-level contributions are the integration of multimodal educational interaction, a design for AI-assisted tutoring, a proposed use of learner performance and interaction history for recommendations, learning analytics within the study workflow, and a modular architecture whose components can be tested and extended. These are design contributions, not claims of research priority or experimentally established superiority.

II. Related Work and Research Gap

The supplied project report surveys multimodal learning analytics, adaptive educational assistants, transformer-based multimodal learning, and engagement estimation. These research themes motivate COGNILEARN. The discussion summarizes them as represented in the report; bibliographic details should be verified against the original publications before submission.

A. Multimodal Learning Analytics

Multimodal learning analytics considers several types of learner-related data rather than relying only on clicks, logins, or quiz scores. Potential modalities discussed in the report include text, audio, video, facial expressions, gestures, and interaction logs. This motivates processing educational content and learner interactions through modality-appropriate components [1], [2].

B. Adaptive and Personalized Learning

Fixed learning systems may not meet every learner's needs. An adaptive assistant can use available performance and interaction history to identify topics that need reinforcement and recommend relevant study material or practice activities. COGNILEARN proposes an adaptive learning engine and recommendation panel for this purpose [3].

C. Transformer Models and AI Tutoring

The report discusses transformer-based approaches as a foundation for processing language and relating information across modalities. It also describes large language models as potential sources of natural-language explanations, question answering, and generated learning content. These capabilities are intended to support COGNILEARN's interactive tutor. The exact model, prompts, safeguards, and integration must be documented from the implementation [4].

D. Engagement and Learning Analytics

The supplied report surveys research themes in estimating engagement and monitoring learning progress. Proposed analytics include performance tracking, engagement analysis, behavioural analytics, progress visualization, and report generation. These functions should be treated as planned unless the implementation and collected data demonstrate that they are operational. Observable measures, such as quiz scores and time spent, should be distinguished from inferred states, such as attention or engagement [5], [6].

E. Research Gap

The project report motivates an integrated platform that brings multimodal content, AI tutoring, personalization, recommendations, and progress tracking together. The proposed contribution is this integrated design. A claim that COGNILEARN is the first system of its kind, or that it outperforms prior work, cannot be established from the Phase-1 report alone and would require a verified literature review and comparative evaluation.

III. System Architecture

A. Proposed Multimodal Workflow

Figure 1 summarizes the proposed learning workflow. Learners interact through a web interface and provide questions, learning materials, or assessment responses. The application validates inputs and routes them to the relevant processing component. An AI tutor can generate explanations using the configured model and available educational context. Learner interactions and assessment history may then inform recommendations and progress summaries. This is a conceptual architecture synthesized from the project report, not a claim that each component is implemented.

B. Shared Learner State and Traceability

The proposed learner record may contain profile information, questions and interactions, completed resources, assessment attempts and scores, recommendations, and progress summaries. A traceable system should connect each recommendation to the observable information that informed it and make clear which measures are recorded data and which are inferred. The report does not define a formal learner-state representation or a data schema.

C. Implementation Modules

The report identifies the following proposed modules:

- **Front end:** login and registration, student dashboard, AI chat interface, quiz and assessment pages, recommendations, navigation, and notifications.
- **Back end and APIs:** request handling, authentication, sessions, data processing, and server-side application logic.
- **Data layer:** user profiles, learning history, quiz records, recommendations, and interaction logs.
- **AI and modality processing:** NLP, machine learning, language models, computer vision, speech processing, and video processing where implemented.
- **Adaptive learning and analytics:** weak-area identification, study recommendations, performance tracking, progress visualization, and report generation.

D. Technology Requirements

The report lists HTML, CSS, JavaScript, React.js, Tailwind CSS, and Chart.js as possible front-end technologies; Python with Flask, Django, or FastAPI and optionally Node.js for back-end services; and MySQL, MongoDB, or PostgreSQL

IJIRIS: International Journal of Innovative Research in Information Security
Volume 12, Issue 08, November 2026
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2026.v1208.01

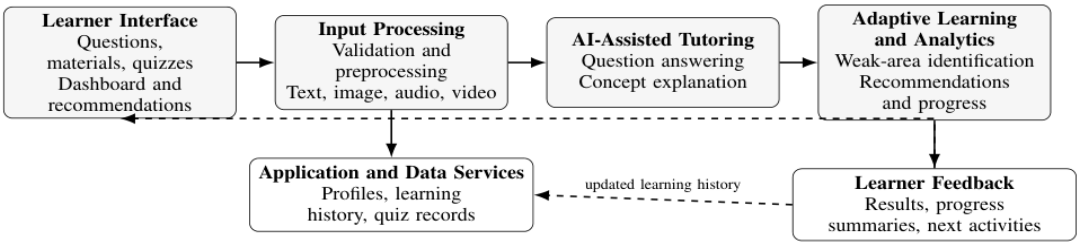


Fig. 1. Conceptual COGNILEARN workflow based on the supplied Phase-1 report. Modality paths, persistence, and feedback depend on the final implementation.

for data storage. OpenCV, Whisper AI, and OCR are mentioned as possible processing tools. AWS, Firebase, Google Cloud, Docker, and GitHub appear as possible deployment or supporting services. These are alternatives in the report, not a confirmed technology stack. The final paper should identify only technologies present in the project source code or deployment configuration.

IV. Methodology

A. Input Acquisition and Preprocessing

The system is intended to accept learner interactions and educational materials in the formats supported by the implementation. Inputs should be validated for type, size, encoding, and format. Text may be normalized; images may require resizing or OCR; audio may require transcription; and video may require frame extraction or audio transcription. The actual preprocessing steps must be documented separately for each implemented modality.

B. AI-Assisted Tutoring

A learner's question is routed to the configured AI tutor, which should produce an explanation relevant to the question and available educational context. Reliability can be improved by constraining responses to provided or approved learning materials where possible, handling model failures, and making uncertainty clear. The supplied report does not specify a particular foundation model, retrieval pipeline, or measured hallucination rate.

C. Learner Representation

A learner profile may use observable indicators such as topic-wise assessment scores, attempt counts, completed resources, and recent activity. These signals can support a simple, explainable estimate of which topics need additional practice. The project report does not define a formal learner-state vector or mathematical update rule, so no specific learned model is asserted here.

D. Recommendation Logic

A basic recommendation workflow could map a low score or repeated incorrect answers in a topic to introductory resources and practice questions for that topic. A topic with consistently strong performance may lead to more advanced material. Recommendations should be recorded with their reason, and learners should be able to revisit or dismiss them. This is a suggested operationalization of the report's recommendation objective, not a verified description of existing code.

E. Progress Feedback Loop

After a learner uses a resource or completes a quiz, the application can update the available learning history and refresh progress summaries. If implemented, the recommendation engine can use the updated history to select subsequent activities. The feedback loop should be tested for correct persistence, consistent scoring, and appropriate recommendations.

V. Evaluation Plan and Expected Outcomes

A. Evaluation Scope

The supplied Phase-1 report describes expected system behaviour but does not include test logs, benchmark data, participant studies, or numerical measurements. This section therefore defines evaluation measures and expected functionality rather than reporting completed experiments. A successful upload by itself should not be counted as evidence that a modality has been understood correctly.

B. Metrics

For a quiz with C correct answers out of N scored questions, the percentage score can be defined as

$$S_{\text{quiz}} = \frac{C}{N} \times 100, \qquad N > 0. \tag{1}$$

The aggregation method, treatment of unattempted questions, and minimum number of attempts should be defined in the implementation.

Recommendation quality can be evaluated using binary relevance labels. Precision at k is

$$\text{Precision@}k = \frac{\text{relevant items among the top } k}{k}. \tag{2}$$

This metric requires a labelled evaluation set. Response latency can be measured as elapsed time from request submission to display of the answer:

$$T_{\text{response}} = t_{\text{completion}} - t_{\text{submission}}. \tag{3}$$

Median and 95th-percentile latency should be reported over documented requests, separating model response time from

IJIRIS: International Journal of Innovative Research in Information Security
Volume 12, Issue 08, November 2026
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2026.v1208.01

upload and preprocessing time where possible. For each modality, functionality coverage can be measured as

$$\text{PassRate}(\%) = \frac{\text{passed test cases}}{\text{executed test cases}} \times 100. \quad (4)$$

To assess learning outcomes, a controlled or carefully designed pre-test/post-test study should report participant count, question design, duration, baseline, outcome measure, and uncertainty. A change in quiz score alone does not establish causation without an appropriate study design.

C. Expected Functional Outcomes

The intended outcomes are a unified interface for supported educational input types; AI-assisted explanations and question answering; a mechanism to use available assessment and interaction history for recommendations; a dashboard for recorded quiz results and progress; and a modular design that can be extended with additional processors and deployment services. Each capability must be checked against the implementation and tested before it is described as operational.

D. Evaluation Table

Table I lists the principal measurements required to evaluate the proposed system. The current evidence column reflects the supplied report, which provides no measured results.

E. Interpretation

The principal value of the proposed system is its integrated approach: multiple content modalities, AI tutoring, adaptive recommendations, and analytics are intended to operate within one learning environment. Integration alone does not establish that every modality works reliably or that learning outcomes improve. Implementation verification and user-centred evaluation are necessary before quantitative performance claims can be made.

VI. Limitations, Ethics, and Reproducibility

A. Implementation and Evaluation Limitations

The report lists several alternative frameworks, databases, and cloud platforms, so the final technology stack must be confirmed from source files. Numerical performance and learning improvement are not demonstrated. The report does not provide implementation details, test cases, or user-study results sufficient for reproducing an evaluation. Accordingly, this paper describes a proposed system and evaluation methodology rather than verified end-to-end capabilities.

B. AI Reliability and Learner Safety

Generated explanations can be incorrect or incomplete. Response-quality checks, clear uncertainty handling, and human review may be needed for high-stakes educational content. Recommendations should be explainable and should not prevent learners from accessing other suitable material. The system should distinguish observed measurements from inferred learner states and avoid treating engagement estimates as definitive indicators of attention.

C. Privacy, Fairness, and Accessibility

Learner histories and uploaded media may contain personal information and should be collected and retained only as necessary. A deployed system should document consent, authentication, authorization, secrets management, transport security, access control, retention, backup, and deletion practices. Recommendations should be checked for unequal performance across learner groups and compatibility with assistive technologies. Facial, behavioural, or physiological signals should not be treated as definitive measures of attention without validated evidence and informed consent.

D. Reproducibility Requirements

An implementation-based evaluation should include source code and dependency files; a list separating implemented features from planned features; the exact model/API, database, back-end framework, and deployment platform; test inputs and raw results; trial counts and environment details; and ethics, consent, privacy, and user-study details where applicable. Benchmark reports should document the evaluation set, task definitions, scoring procedures, and limitations.

VII. Conclusion and Future Enhancements

A. Conclusion

COGNILEARN is a proposed multimodal adaptive AI learning assistant intended to make digital learning more interactive and learner-centric. The supplied Phase-1 report describes an architecture that brings together multiple educational input formats, AI-assisted tutoring, adaptive recommendations, and learning analytics. Its intended workflow allows learners to ask questions, access explanations, practise through assessments, and review progress through a dashboard. The design provides a basis for integrating educational interaction and personalization in one learning environment.

The available report does not establish quantitative performance or measured learning gains. Future work should verify implemented features, document the exact technology stack, and evaluate answer quality, modality coverage, latency, recommendation relevance, privacy, and educational outcomes using reproducible experiments.

B. Future Enhancements

Future work should implement and document end-to-end pipelines for each supported modality; add retrieval grounded in approved course materials, source references, and response-quality checks for tutoring; define an explainable learner model and compare recommendation strategies against a baseline; and conduct controlled evaluations of usability, answer quality, modality coverage, and learning outcomes. Privacy and security should be strengthened through clear consent, role-based access, encryption, retention limits, and audit logging. Deployment work should include monitoring, error handling, backup, and documented configuration management.

IJIRIS: International Journal of Innovative Research in Information Security
Volume 12, Issue 08, November 2026
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2026.v1208.01

TABLE I
PROPOSED COGNILEARN EVALUATION PLAN

Evaluation item	**Measurement**	**Current evidence**	**Required next step**
AI answer quality	Expert rubric or correctness rate	Not reported	Prepare a labelled question set and scoring rubric
Response latency	Median and 95th percentile	Not reported	Collect timestamped test runs
Text processing	Pass rate on representative cases	Not reported	Run documented text tests
Image processing	Pass rate or OCR quality	Not reported	Test supported image formats and content
Audio processing	Transcription quality or pass rate	Not reported	Test representative audio and noise conditions
Video processing	Pass rate on defined tasks	Not reported	Define the task and evaluate sample videos
Recommendation quality	Precision@k or expert rating	Not reported	Label recommendations against learner needs
Learning outcomes	Pre/post assessment change	Not reported	Conduct a documented learner study

REFERENCES

[1] García et al., "Research Trends in Multimodal Learning Analytics: A Systematic Mapping Study," in *Proc. IEEE CHILECON*, 2023. Bibliographic details transcribed from the supplied report; verify before submission.
[2] L. Chen and P. Singh, "Framework for Designing Multimodal Learning Analytics Systems," in *Proc. IEEE ICALT*, 2024. Bibliographic details transcribed from the supplied report; verify before submission.
[3] H. Zhang et al., "AI-Enabled Intelligent Assistant for Personalized and Adaptive Learning," *IEEE Transactions on Learning Technologies*, 2023. Bibliographic details transcribed from the supplied report; verify before submission.
[4] Y. Li and T. Brown, "Multimodal Learning with Transformers: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. Bibliographic details transcribed from the supplied report; verify before submission.
[5] K. Ahmed et al., "Detecting Drowsy Learners Using Multimodal Learning Analytics," *IEEE Access*, 2021. Bibliographic details transcribed from the supplied report; verify before submission.
[6] S. N. Karimah and S. Hasegawa, "Automatic Engagement Estimation in Smart Education/Learning Settings," *Smart Learning Environments*, 2022. Bibliographic details transcribed from the supplied report; verify before submission.