August 2025

NeuroQuest: A Multi-Agent AI Framework for Adaptive Learning Through Intelligent Knowledge Creation

Ravi Varma Kumar Bevara
University of North Texas, ravivarmakumarbevara@my.unt.edu

Mr. Nishith Reddy Mannuru
University of North Texas, nishithreddymannuru@my.unt.edu

Thuan L. Nguyen
University of North Texas, thuan.nguyen@unt.edu

NeuroQuest: A Multi-Agent AI Framework for Adaptive Learning Through Intelligent Knowledge Creation

Completed Research Full Paper

Ravi Varma Kumar Bevara
University of North Texas
ravivarmakumarbevara@my.unt.edu

Nishith Reddy Mannuru
University of North Texas
nishithreddymannuru@my.unt.edu

Thuan Luong Nguyen
University of North Texas
Thuan.Nguyen@unt.edu

Abstract

The increasing integration of artificial intelligence in education has led to the development of adaptive learning systems capable of enhancing knowledge delivery and comprehension. This paper proposes a multi-agent AI framework, NeuroQuest, advancing the domain by synthesizing real-time academic content into structured, multimodal formats, including PDFs, presentations, and podcasts. Unlike conventional AI-based tools primarily focusing on information retrieval, NeuroQuest employs a collaborative agentic approach using OpenAI's Swarm framework to search, refine, and generate highly structured educational outputs. A user evaluation study comprising students and professors demonstrated high satisfaction levels in content clarity, readability, and multimodal accessibility. While users highlighted the platform's strengths in organizing and synthesizing knowledge effectively, response speed emerged as a trade-off requiring further optimization. These findings establish NeuroQuest as a promising AI-driven academic tool that bridges the gap between information retrieval and structured knowledge synthesis.

Keywords

Agentic AI, AI Agents, Multi-Agent AI System, Adaptive Learning, Multimodal Learning, Educational AI, Knowledge Synthesis, AI-Driven Content Generation.

Introduction

Integrating artificial intelligence (AI) into education fuels adaptive learning systems designed to personalize instruction based on learner needs, enhancing engagement and comprehension, which can enhance learning outcomes (Zawacki-Richter et al., 2019). However, many current systems face significant limitations: they often rely on predefined content, offer restricted personalization, and struggle to adapt to new information or diverse learning styles.

Another critical limitation is the predominant focus on text-based delivery, failing to leverage multimodal formats like structured summaries, visual presentations, or spoken explanations, which are crucial for varied learning preferences (Li et al., 2024). While generative AI promises personalized, multimodal learning, most tools still yield static, text-heavy outputs, limiting pedagogical flexibility.

To address these shortcomings, this paper introduces NeuroQuest, a multi-agent AI framework designed to support adaptive learning through intelligent knowledge synthesis. Unlike systems that focus solely on retrieval or personalization, NeuroQuest leverages a task-specialized, agentic architecture that performs real-time academic search, validating, multimodal formatting, and transforming unstructured data into study-ready content. The framework orchestrates agents responsible for retrieval, synthesis, and publishing. The framework also enables agents to generate organized PDFs, slide decks, and podcast transcripts aligned with diverse user needs.

NeuroQuest's innovation is its orchestration layer that integrates existing tools like GPT-4 and DuckDuckGo into an intelligent workflow optimized for education. By tackling static adaptivity, limited modality, and the synthesis deficit, NeuroQuest provides a scalable solution for structured, AI-powered learning. Key contributions include the NeuroQuest system itself, its demonstration of modular agent collaboration, and a comparative study validating its superiority in generating structured, academically focused content.

Related Work

Artificial intelligence (AI) integration into adaptive learning systems (ALS) enhances individualized education with meta-analyses confirming medium-to-large cognitive improvements compared to static approaches (Wang et al., 2024). These systems effectively optimize content distribution and evaluation using advanced algorithms. However, significant limitations persist. Many platforms excel at retrieving pre-structured content but lack robust capabilities for real-time knowledge synthesis from diverse sources, hindering comprehensive understanding (Gligorea et al., 2023). Furthermore, despite the benefits of multimodal delivery (presentations, visuals) for varied learning preferences, most systems remain heavily reliant on text-based formats (Kabudi et al., 2021).

Addressing these shortcomings, multi-agent systems (MAS) offer a compelling architectural solution for e-learning. MAS utilize autonomous agents for specialized tasks like learner profiling and content sequencing, enhancing system responsiveness and adaptability (Wang et al., 2025). Notably, some MAS models support real-time synthesis, improving content coherence (Almohammadi et al., 2017), while AI techniques advance learning style detection for personalization (Ezzaim et al., 2024). Building on these foundations while tackling persistent issues like reactivity and bias (Xie et al., 2019), NeuroQuest is proposed. It employs collaborative MAS architecture for real-time synthesis of current academic content from trusted sources. NeuroQuest distinctively features agentic modularity and generates structured, adaptable multimodal outputs, aiming to bridge the synthesis and modality gaps in existing ALS and provide scalable, pedagogically aligned learning tools that move beyond simple retrieval and static presentation.

Design of Multi AI-Agent System (MAS) Architecture

NeuroQuest is an advanced multi-agent AI system designed to automatically retrieve, synthesize, and generate structured educational content for students and educators. It operates within OpenAI's Swarm framework, which enables distributed agents to collaborate efficiently, allowing specialized agents to work in parallel while maintaining a seamless workflow. Swarm, developed as an agentic infrastructure for large-scale automation, allows independent AI agents to communicate, coordinate, and refine outputs dynamically (OpenAI, n.d.). By leveraging this framework, the system ensures efficient task execution, modularity, and adaptability to diverse academic needs (See Figure 1).

The high-level system architecture of this framework is shown in Figure 1, outlining the interaction between key system components. The user, either a student or a professor, initiates a query through the Orchestration Agent, which serves as the system's central coordinator. This agent determines the appropriate processing pipeline based on the nature of the request and routes it to the multi-agent channel, where three specialized agents—the Web Search Agent, Researcher Agent, and Publishing Agent—work collaboratively to retrieve, analyze, and structure information. The final output is then made accessible through Streamlit UI, providing users with structured content in text documents, presentations, or audio-based materials. At the core of the system is the multi-agent processing channel, where the Web Search Agent is responsible for retrieving relevant information from external sources. The search results are then forwarded to the Researcher Agent, which functions as the intelligence layer of the system. The processed content is then passed to the Publishing Agent, which formats and prepares the final output. This agent ensures that the content is well-structured, clear, and aligned with the requested format, whether it be a text summary, a PowerPoint presentation, or a podcast transcript.

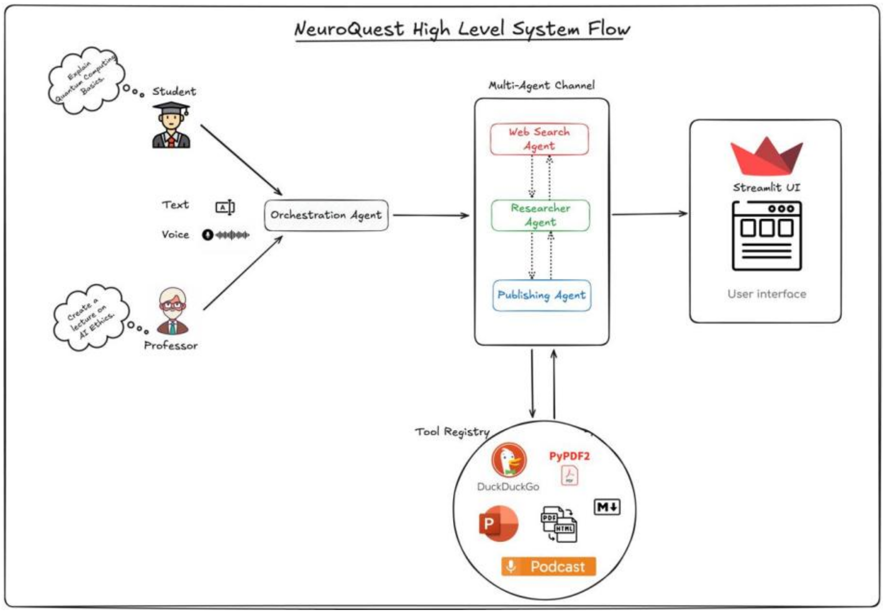

Figure 1: High-Level System Architecture of NeuroQuest

The platform's modularity allows it to cater to multiple user groups, particularly students and professors, who interact with the system differently. A student seeking concise study materials on quantum computing would receive a structured summary optimized for learning, while a professor preparing a lecture on AI ethics would be provided with a well-organized presentation alongside a textual outline. This adaptability enables the system dynamically to adjust content based on user intent, making it a versatile tool for academic study and instructional preparation. This agentic framework, built upon Swarm's autonomous agentic collaboration, ensures knowledge retrieval and processing are efficient and scalable. The system maintains high responsiveness by distributing tasks across specialized agents, while producing structured, high-quality educational content.

System Implementation and Workflow

NeuroQuest uses a modular multi-agent architecture where each agent performs a specific role—retrieving, synthesizing, and formatting content—in a dynamic, coordinated workflow. Rather than creating novel foundational models, the system integrates proven tools like GPT-4 and DuckDuckGo within a scalable orchestration layer. This design enables real-time academic content generation that is structured and adaptable to diverse educational needs. The framework's strength lies in how agent collaboration and task delegation are optimized for structured learning support rather than altering the underlying models' behavior(See Figure 2).

This architectural framework ensures systematic knowledge acquisition from external sources, followed by rigorous verification, organization, and transformation into diverse output formats, making the system particularly advantageous for educational applications across various academic levels. As illustrated in (Figure 2), these agents engage in complex interactions with external resources, including search APIs, advanced language models, and supplementary engineering components, forming an integrated ecosystem for educational content generation.

The workflow initiates with the Web Search Agent's comprehensive analysis of external knowledge sources. This agent connects with the DuckDuckGo API to retrieve extensive search results encompassing scholarly articles, academic papers, and relevant educational materials. However, these initial results typically present in an unstructured format with redundant entries, necessitating sophisticated processing by the

Researcher Agent. The Researcher Agent implements a series of fine-tuned procedures and algorithms to eliminate duplicate entries and ensure content quality through a prompt-driven validation pipeline. Specifically, the Researcher Agent uses a language model instructed to: (1) remove redundant content and duplicate URLs, (2) prioritize recent and credible sources, (3) verify consistency across multiple search results, and (4) extract key facts, statistics, and quotes with proper attribution. It also merges related themes and flags contradictory information during synthesis. By organizing content into logical sequences and preserving key contextual relationships, the agent enhances both the coherence and academic integrity of the structured output. These rules are encoded in the agent's instruction set and executed dynamically by the LLM-based model, enabling scalable and context-aware filtering without hard-coded heuristics.

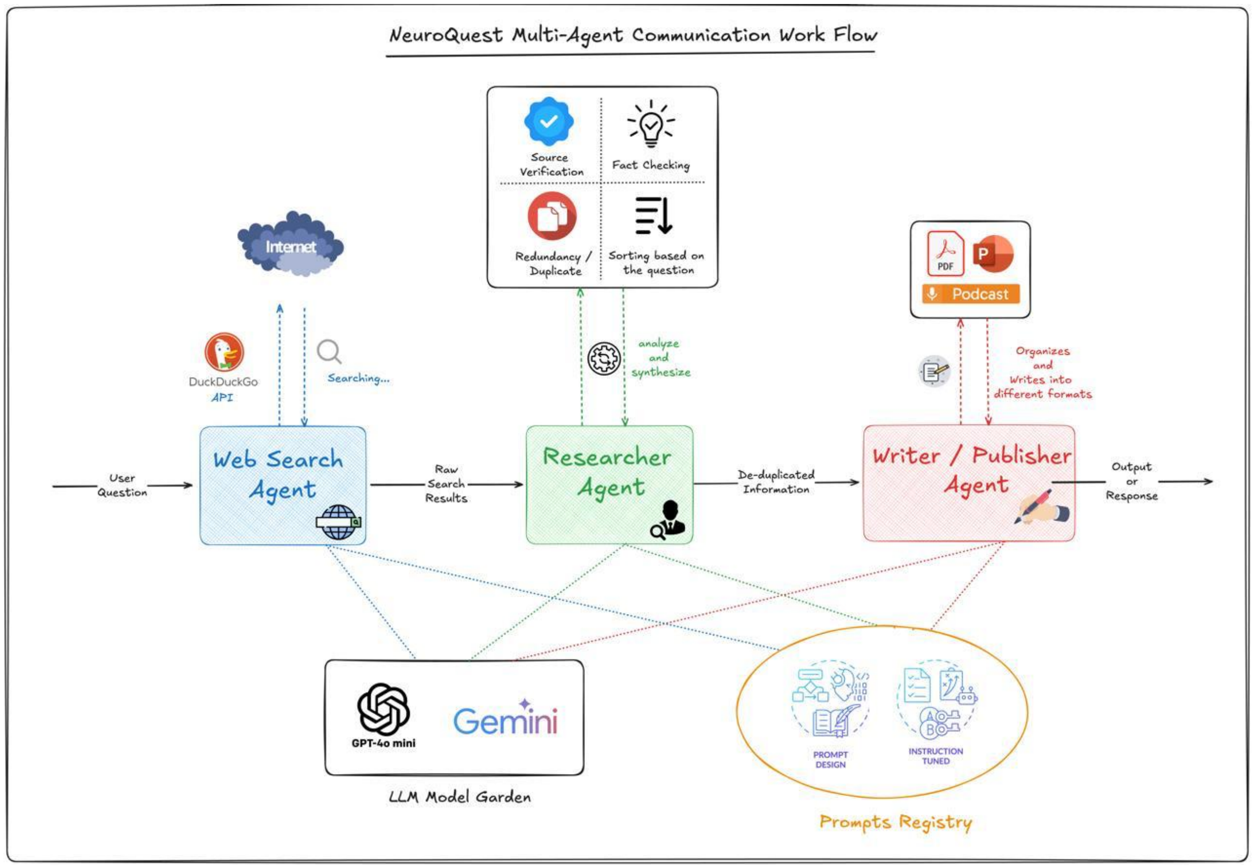

Figure 2: NeuroQuest's multi-agent workflow for content retrieval and structuring.

Following the content refinement phase, the Writer/Publisher Agent organizes the validated information into structured, pedagogically appropriate outputs. The formatting procedure demonstrates significant flexibility, being specifically tailored to accommodate various user-requested delivery formats. In the case of text documents, the content undergoes a systematic conversion to Markdown format before being processed through WeasyPrint technology to generate professional-grade PDF files, ensuring consistent formatting and structural integrity throughout the document. For lecture presentations, the system leverages advanced GPT-4 technology to create thoroughly segmented slides with precisely aligned visual displays, enhancing the educational value of the content.

As demonstrated in Figure 3, the content processing pipeline incorporates specialized handlers designed to optimize different output formats through distinct processing pathways. The PDF Handler processes structured text received from the Writer Agent through a sophisticated pipeline, utilizing HTML formatting conventions and WeasyPrint technology for final document production. Additionally, PyPDF2 technology facilitates advanced content extraction capabilities for supplementary processing requirements. The PPT Handler employs specialized algorithms to convert organized content into logical data structures before formatting it into coherent slide presentations, ensuring optimal information distribution and pedagogical flow across slides. This architectural approach provides substantial advantages through efficient work distribution, with each agent assigned with specific responsibilities within the system through precise task delineation. The clear separation of concerns enables remarkable system scalability across multiple academic domains while facilitating the seamless integration of additional tools and processing components in subsequent system iterations. The methodology guarantees consistent automated content generation through systematic information retrieval, verification, and structuring processes while maintaining rigorous contextual processing integrity throughout the pipeline.

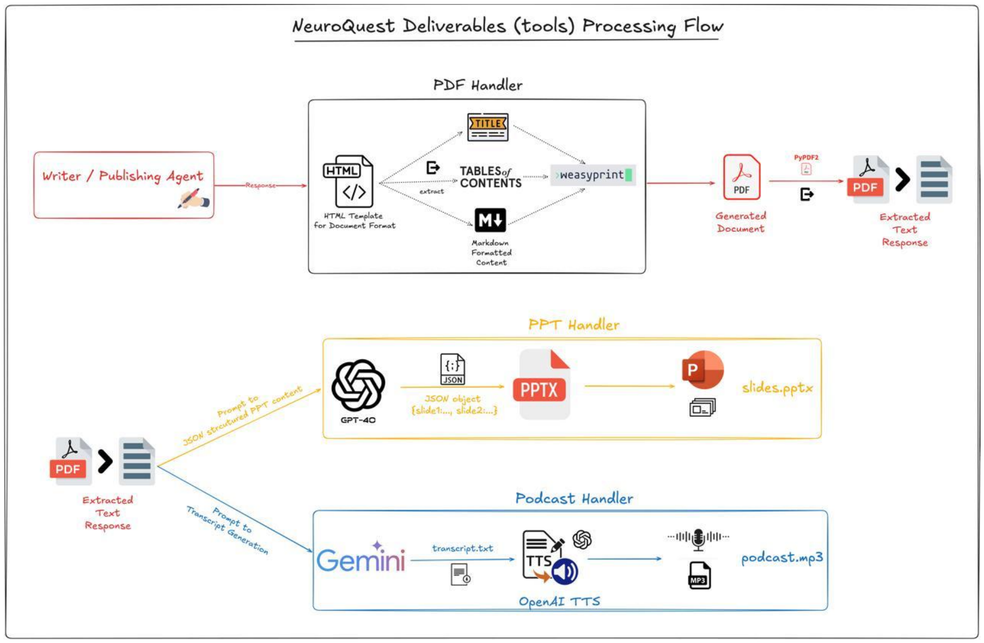

Figure 3: Output processing pipeline for PDFs, slides, and podcasts.

The multi-agent architecture's modular design allows for continuous system enhancement and optimization, with each agent capable of independent upgrades and improvements without disrupting the overall system functionality. This flexibility ensures the system can adapt to evolving educational requirements and technological advancements while maintaining its core functionality. Furthermore, the system's robust verification processes and structured output generation capabilities make it particularly valuable for educational institutions seeking to automate and standardize their content creation processes while maintaining high academic standards. The implementation of this methodology demonstrates significant potential for transforming educational content generation across various academic contexts. By combining automated content retrieval with sophisticated processing and formatting capabilities, the system addresses the growing need for efficient, scalable educational content development solutions while maintaining the necessary academic rigor and pedagogical effectiveness required in modern educational environments.

Results and Discussion

The evaluation of NeuroQuest was assessed across different performance metrics, including content quality, readability, efficiency, coherence, multimodal adaptation, speed, and user satisfaction. The evaluation involved 17 participants from diverse subject groups, including students, professors, and others, who were selected voluntarily. Data was collected using a Qualtrics-based survey that included five Likert-scale questions comparing NeuroQuest and ChatGPT across accuracy, logical organization, readability, usefulness, and structural clarity. Two open-ended questions were also included to collect qualitative insights: “What was the most useful aspect of NeuroQuest for you?” and “What would you improve in NeuroQuest?” This evaluation focused on descriptive statistics and user perceptions without applying inferential statistical tests. In addition to the comparison questions, participants responded to detailed Likert-scale items assessing NeuroQuest across seven core metrics. Content Quality measures perceptions of factual accuracy, relevance to the query, and the system's ability to filter redundant or conflicting information. Readability evaluated ease of comprehension and the clarity of explanations for complex topics.

Efficiency focused on perceived time savings, future utility, and practical academic support. Coherence assessed the logical structure, topic flow, and preservation of context. Multimodal adaptability captured the perceived usefulness of PDF, slide, and podcast formats. Speed was evaluated through user perceptions of generation time. Satisfaction reflected overall impressions and the likelihood of recommending NeuroQuest. These targeted items allowed for nuanced insights into the system's effectiveness across academic contexts. Participants, including students and professors, contributed feedback and qualitative opinions on the system's effectiveness.

The findings demonstrate that NeuroQuest provides organized, research-ready content across multiple formats, achieving very high user satisfaction, but there are some concerns about response speed. The structured output of NeuroQuest ensures clarity and logical organization, making it highly effective for complex topics such as quantum computing. As shown in Figure 4, the system-generated content maintains a well-defined structure, offering a summary, key definitions, and fundamental principles in a format suitable for both academic and professional use. This structured approach enhances readability and usability, distinguishing NeuroQuest from conventional AI-generated responses (See Figure 4).

Figure 4: Example of NeuroQuest-generated content on quantum computing.

	Mean Score	Standard Deviation
Content Quality	4.53	0.52
Readability	4.73	0.46
Efficiency	4.40	0.63
Coherence	4.80	0.56
Multi-Format	4.73	0.46
Speed	3.40	1.24
Satisfaction	4.87	0.35

Table 1. NeuroQuest Evaluation Summary

Table 1 and Figure 5 show that users provided positive scores to NeuroQuest in essential categories, with the content's quality averaging 4.53, readability 4.73, and coherence 4.80. The findings affirm that users perceive NeuroQuest's knowledge as accurate and well-organized, hence providing clarity in complicated topics.

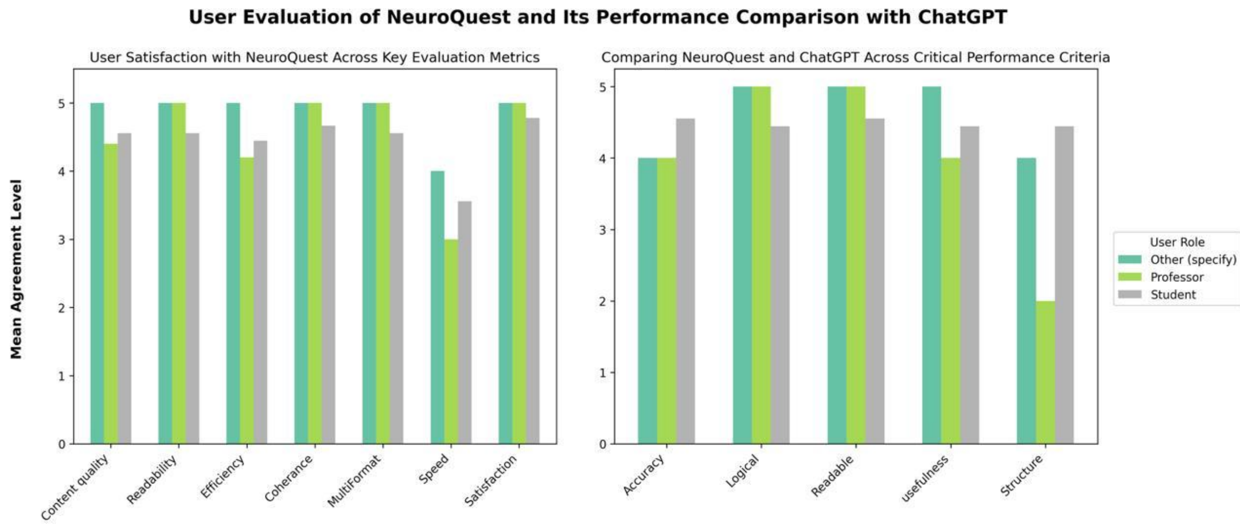

Figure 5: Mean Agreement Levels of User Evaluations of NeuroQuest and Comparison with ChatGPT.

Furthermore, the capability for multimodal adaptability, enabling content conversion into PDFs, presentations, and podcasts, received an overall positive review, achieving a score of 4.73. The highest-rated variable was overall satisfaction (4.87), signifying that users predominantly perceived NeuroQuest as valuable for academic and research projects. However, response speed was considered as a limitation, attaining a reduced score of 3.40, reflecting of the tradeoff between content structure and prompt response. The reduced speed score can be attributed to the system's layered content generation pipeline, which involves retrieving information from the web, synthesizing and verifying content, and formatting it into PDFs, slides, or podcasts.

While critical to delivering high-quality, structured, and multimodal outputs, these processes introduce latency. The tradeoff between structured richness and responsiveness is a natural limitation in systems focused on academic rigor. Future work will explore architectural and engineering optimizations to improve runtime efficiency without compromising output quality. Additionally, while the participant pool included both students and professors, the study did not segment feedback by user type or domain expertise. A breakdown of responses across academic roles would provide deeper insights into the system's utility and is planned for future evaluations. A comparative assessment between NeuroQuest and ChatGPT reveals notable advantages in organized content generation. Table 2 and Figure 5 demonstrate the performance comparison, indicating that NeuroQuest achieved competitive scores in accuracy (4.33) and logical coherence (4.67), whereas structural clarity received a lower rating (3.60).

	Mean Score	Standard Deviation
Accuracy	4.33	0.62
Logical	4.67	0.62
Readable	4.73	0.59
Usefulness	4.33	0.62
Structure	2.60	1.30

Table 2. Comparison of NeuroQuest vs ChatGPT

These descriptive statistics indicate that although ChatGPT may be regarded as more adaptable in producing conversational replies, NeuroQuest's advantage resides in its ability to enhance and organize content into more structured formats. NeuroQuest's organized layout guarantees users access to thoroughly validated, logically ordered information, reducing redundancy and inconsistency sometimes found in ChatGPT's dynamic, real-time replies. Although ChatGPT is more adept at quick, engaging conversation, NeuroQuest specializes in producing polished, research-ready results.

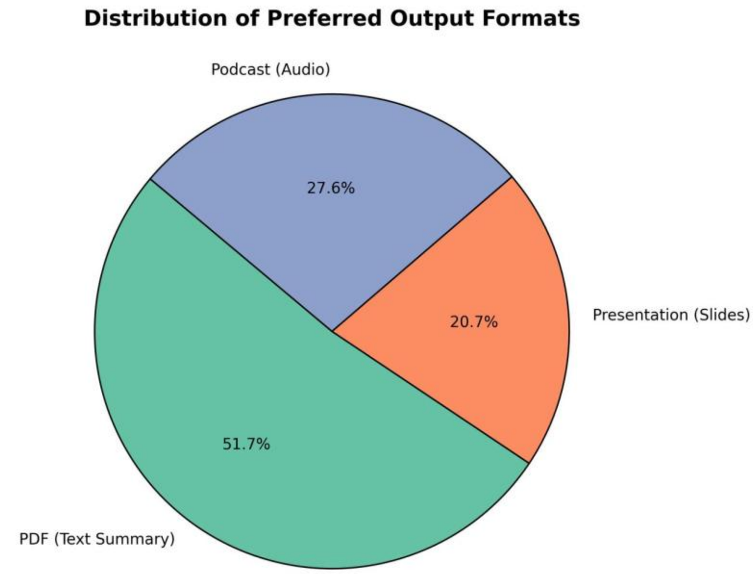

Figure 6: User Choice Distribution of the Output Format(s).

The system's multimodal functionality is an additional distinctive trait. Figure 6 depicts the distribution of preferred content forms, with 51.7% of users favoring PDFs, so highlighting the significance of text-based structured summaries in academic and research contexts. Podcasts comprised 27.6% of user choices, underscoring the importance of audio content for accessibility and passive learning. Presentations were chosen by merely 20.7% of users, indicating that although they fulfill a significant function, they may want additional refinement to improve participation. Future enhancements may include interactive or dynamically created presentations to augment their functionality.

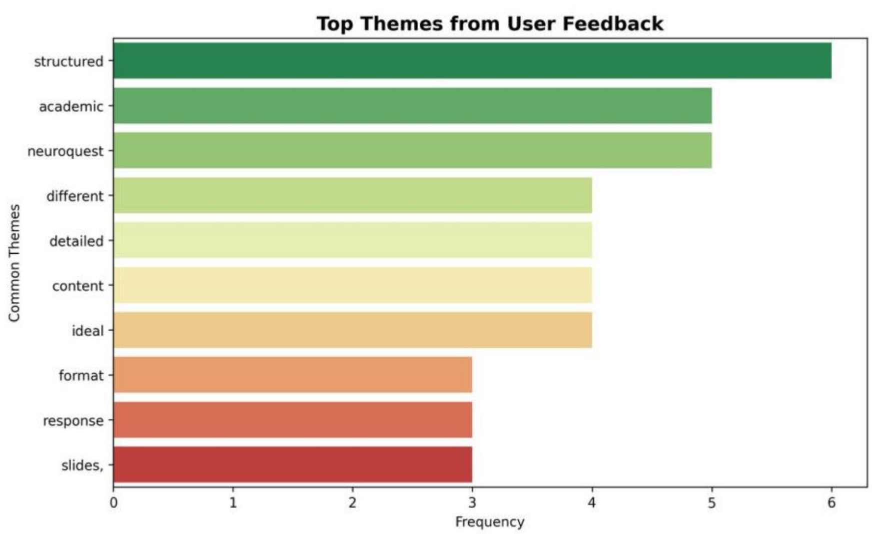

Figure 7: Dominant Themes from Open Ended Responses of the Users Survey on NeuroQuest Evaluation.

Beyond structured rating scales, qualitative insights from users provide additional depth to the evaluation. Figure 7 captures dominant themes from open-ended responses, where frequently mentioned terms such as "structured," "academic," "content," and "format" affirm NeuroQuest's effectiveness in delivering well-organized information. These themes suggest that users value the tool primarily for its ability to synthesize complex information into clear, structured outputs. However, Figure 8 reveal a consistent concern regarding processing speed. Words like "response," "faster," "enhance," and "time" frequently appeared in user feedback, reinforcing that while NeuroQuest outperforms traditional AI assistants in content organization, users expect faster output generation.

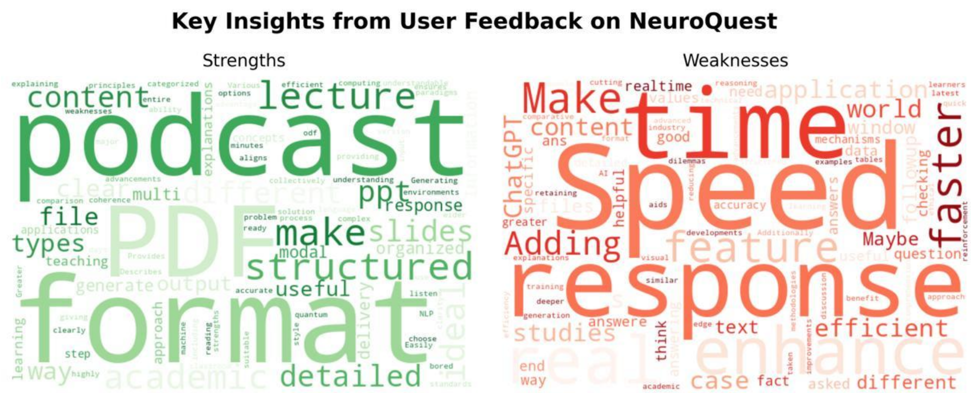

Figure 8: Word Cloud Analysis demonstrating Strengths and Weakness of NeuroQuest.

This tradeoff is essential in prospective advancements as feasible alternatives like asynchronous mode content generation or sequential result representations may enhance real-time responsiveness without affecting quality. The findings indicate that NeuroQuest provides a superior alternative to conventional AI-driven academic tools, especially for users seeking organized, verified, and multimodal outputs. Although its response speed requires improvement, its capacity to refine, validate, and present information in numerous formats stands out. The evident user preference for organized information demonstrated through quantitative evaluations and thematic analysis, underscores the necessity for AI-driven academic tools emphasizing accuracy, logical coherence, and accessibility. Continuing initiatives to enhance the throughput while safeguarding the integrity of structured information generation will further solidify NeuroQuest's prominence as a top AI-driven collaborator.

Conclusion

This study presents NeuroQuest as a structured, multimodal, and intelligent learning system that redefines how academic knowledge is retrieved, synthesized, and delivered. Leveraging a modular multi-agent architecture built on OpenAI's Swarm framework, NeuroQuest coordinates agents for retrieval, refinement, and output generation, producing study-ready content in formats such as PDFs, slides, and podcasts. Students and professors reported high satisfaction with the system, particularly in content clarity, coherence, and multimodal accessibility. A key limitation identified through user feedback was response speed, attributed to the system's real-time search and formatting processes. While this tradeoff affects responsiveness, it enhances the final output's depth, accuracy, and structure. Future work will focus on improving system efficiency and exploring retrieval-augmented generation (RAG) and visual content integration. Although the system utilizes existing models like GPT-4 and DuckDuckGo, its innovation lies in orchestrating these tools through intelligent, role-specific agents optimized for educational delivery, distinguishing it from more generic AI workflows.

To improve robustness, future development will address edge cases such as vague queries or sparse search results. These challenges can impact content quality, especially in slides and podcasts. Planned enhancements include automated query refinement and fallback strategies to maintain output reliability in less optimal scenarios. Additionally, a formal comparative analysis with other adaptive learning and multi-agent frameworks is planned to benchmark NeuroQuest's performance, scalability, and educational impact more rigorously. While this study evaluated user perceptions, future iterations will incorporate pre- and post-assessments to measure direct learning gains. Ethical considerations remain central. NeuroQuest does not collect or store user data, ensuring privacy-preserving usage. Users are transparently informed that output is AI-generated. Although explicit bias mitigation mechanisms are not yet implemented, future work will explore filtering techniques and fairness strategies to support equitable learning experiences. NeuroQuest offers a scalable path forward for intelligent educational systems grounded in structured knowledge delivery and user-centered design by addressing these current limitations and evolving toward more adaptive, context-aware content generation.

References

Almohammadi, K., Hagras, H., Alghazzawi, D., & Aldabbagh, G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. Journal of Artificial Intelligence and Soft Computing Research, 7(1), 47-64.

Ezzaim, A., Dahbi, A., Aqqal, A., & Haidine, A. (2024). AI-based learning style detection in adaptive learning systems: a systematic literature review. Journal of Computers in Education, 1-39.

Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: a literature review. Education Sciences, 13(12), 1216.

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. Computers and Education: Artificial Intelligence, 2, 100017.

Li, H., Xu, T., Zhang, C., Chen, E., Liang, J., Fan, X., ... & Wen, Q. (2024). Bringing generative AI to adaptive learning in education. arXiv preprint arXiv:2402.14601.

OpenAI. (n.d.). Swarm: Educational framework exploring ergonomic, lightweight multi-agent orchestration [Python software]. GitHub. Retrieved February 20, 2025, from https://github.com/openai/swarm

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. Journal of educational psychology, 106(2), 331.

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2024). AI-driven adaptive learning for sustainable educational transformation. Sustainable Development.

Wang, X., Huang, R. T., Sommer, M., Pei, B., Shidfar, P., Rehman, M. S., ... & Martin, F. (2024). The Efficacy of Artificial Intelligence-Enabled Adaptive Learning Systems From 2010 to 2022 on Learner Outcomes: A Meta-Analysis. Journal of Educational Computing Research, 07356331241240459.

Wang, T., Zhan, Y., Lian, J., Hu, Z., Yuan, N. J., Zhang, Q., ... & Xiong, H. (2025). LLM-powered Multi-agent Framework for Goal-oriented Learning in Intelligent Tutoring System. arXiv preprint arXiv:2501.15749.

Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. Computers & Education, 140, 103599.

Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. British Journal of Educational Technology, 50(6), 3119-3137.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 1-27.