Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

LLM-Augmented Academic Analytics: A Retrieval-Augmented Generation Framework for Real-Time Student Performance Intelligence in Higher Education

Vamshi Vanguri
PhD in Information Technology University of the Cumberlands, Williamsburg, Kentucky, USA
Vvanguri59826@ucumberlands.edu

ARTICLE INFO

Received: 18 Dec 2024
Revised: 10 Feb 2025
Accepted: 28 Feb 2025

ABSTRACT

Higher education institutions face a persistent challenge in translating large volumes of student performance data into timely, actionable intelligence for faculty and academic advisors. Conventional learning analytics dashboards rely on pre-aggregated historical datasets and static visualizations that are incapable of addressing context-specific, real-time queries. This paper proposes the Retrieval-Augmented Academic Analytics Framework (RAAF) — a novel architecture that integrates Retrieval-Augmented Generation (RAG) with existing institutional BI infrastructure to enable natural language querying of live student performance data. Drawing on foundational work in enterprise RAG-BI integration (Pathoori, 2025), this study extends the paradigm to the specific constraints and governance requirements of higher education environments, including FERPA compliance, multi-source academic record systems, and the unique challenge of hallucination risk in high-stakes advising contexts. The RAAF framework is validated through a conceptual case study at a private liberal arts university, demonstrating a projected 64% reduction in time-to-insight for academic advisors and a structured approach to LLM-generated output governance. Findings suggest that RAG-based academic analytics represents a meaningful evolution beyond static dashboards and carries significant implications for early intervention systems, retention analytics, and institutional decision support.

Keywords: Learning Analytics, Retrieval-Augmented Generation, Large Language Models, Higher Education BI, Student Performance Intelligence, Academic Dashboard, LLM Governance, FERPA Compliance, Conversational Analytics, Institutional Data Systems

1. INTRODUCTION

The volume of data generated within higher education institutions has grown substantially over the past decade. Student information systems, learning management platforms, financial aid databases, course registration systems, and advising records collectively produce datasets of considerable scale and complexity. Yet the translation of this data into timely academic intelligence — intelligence that reaches faculty, advisors, and administrators in the moment it is needed — remains an unsolved institutional problem. Most institutions rely on periodic reports, static dashboards, and scheduled data extracts that introduce latency between data generation and institutional response.

Learning analytics as a discipline has attempted to address this gap since its emergence in the early 2010s. Siemens (2013) defined learning analytics as the measurement, collection, analysis, and reporting of data about learners and their contexts for the purpose of understanding and optimizing learning and the environments in which it occurs. Systems such as Purdue University's Course Signals (Arnold & Pistilli, 2012) demonstrated early proof that predictive indicators drawn from course management data could support real-time intervention workflows. However, these systems remained constrained by the same fundamental limitation as conventional BI platforms: they surfaced pre-

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

defined views of pre-aggregated data, requiring analysts to anticipate questions in advance and encode them into static visual structures.

The emergence of large language models (LLMs) and the Retrieval-Augmented Generation (RAG) architecture offers a structurally different approach. Rather than querying pre-built views, a RAG-enabled analytics system allows users to pose natural language questions against live data, with the LLM handling retrieval, synthesis, and response generation in real time. Pathoori (2025) introduced the RAG-BI framework in the enterprise context, demonstrating that retrieval-augmented dashboards could enable context-aware analytics through LLM integration, fundamentally shifting the query model from passive visualization to active, conversational intelligence. That work established both the architectural blueprint and the governance considerations for deploying RAG within enterprise data environments.

The application of RAG-based analytics to higher education, however, introduces domain-specific constraints that the enterprise BI context does not fully address. Academic data is governed by the Family Educational Rights and Privacy Act (FERPA), which imposes strict access controls on student records and prohibits disclosure without consent. Advising conversations informed by AI-generated output carry consequences — incorrect or hallucinated information about degree requirements, course prerequisites, or financial aid eligibility can materially affect student outcomes. The multi-tenant nature of institutional data systems, where student records span registrar, financial, academic, and co-curricular platforms with minimal integration, adds further architectural complexity.

This paper introduces the Retrieval-Augmented Academic Analytics Framework (RAAF) — a domain-specific adaptation of the RAG-BI paradigm for higher education institutional analytics. The framework addresses the retrieval architecture, LLM integration layer, governance controls, and compliance requirements specific to academic environments. A conceptual case study demonstrates the framework's application at a private liberal arts institution and evaluates its projected impact on advising efficiency, early alert capabilities, and data governance outcomes.

The remainder of this paper is organized as follows: Section 2 reviews relevant literature across learning analytics, enterprise BI evolution, and RAG architectures. Section 3 presents the RAAF framework design. Section 4 describes the conceptual case study methodology. Section 5 reports results and discusses findings. Section 6 concludes with implications for research and practice.

2. LITERATURE REVIEW

2.1 Learning Analytics and Its Limitations

Learning analytics emerged as a formal discipline following the First International Conference on Learning Analytics and Knowledge in 2011 and has since produced a substantial body of research addressing predictive modeling, engagement measurement, and at-risk student identification (Siemens, 2013; Romero & Ventura, 2010). Early implementations demonstrated measurable improvements in retention when analytical signals were surfaced to advisors in a timely manner. Arnold and Pistilli (2012) reported that Purdue's Course Signals system, which used predictive modeling to assign traffic-light risk indicators to students, contributed to improved retention rates among participating students.

However, subsequent research identified consistent limitations in deployed learning analytics systems. Brown (2011) noted that the third wave of educational technology would require analytics platforms that moved beyond reporting toward genuine decision support. The structural constraint has been the architecture of the underlying platforms: conventional learning analytics dashboards, like their enterprise BI counterparts, are built around pre-defined questions. Analysts must anticipate which data views faculty and advisors will need, encode those views into visualizations, and accept that novel queries will fall outside the system's scope until a new dashboard is built and deployed.

This limitation is not incidental but architectural. Static dashboards are optimized for distribution — they deliver consistent, governed views of data to many users simultaneously. They are not designed for the exploratory, context-sensitive queries that characterize actual advising conversations, where an advisor might need to know, in the

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

moment, whether a specific student's combination of incomplete coursework, financial aid standing, and declared major creates a particular graduation risk profile.

2.2 Enterprise BI and the Emergence of Conversational Analytics

Enterprise BI has followed a parallel evolution. Chaudhuri et al. (2011) described BI technology as encompassing data warehousing, reporting, and analytical processing, noting that the field had matured significantly but remained centered on structured, pre-defined analytical workflows. The introduction of self-service BI tools reduced the technical barrier to dashboard creation but did not resolve the fundamental constraint: queries still had to be encoded as visual structures rather than expressed as natural language.

The integration of natural language processing into BI interfaces has been explored incrementally, but prior approaches — based on keyword extraction and predefined query templates — lacked the semantic flexibility to handle complex, contextual questions. The deployment of transformer-based LLMs changed this calculus significantly. Zhao et al. (2023) surveyed the landscape of large language models and documented their capability for complex reasoning, multi-step inference, and domain-adapted response generation, establishing the theoretical foundation for LLM integration within analytical systems.

Pathoori (2025) synthesized these developments into a coherent enterprise architecture, introducing the RAG-BI framework as a mechanism for enabling context-aware analytics through LLM integration with BI platforms. The framework establishes a retrieval layer that queries live enterprise data in response to natural language inputs, passes the retrieved context to an LLM for synthesis, and returns a governed, auditable response. The work demonstrated that this architecture could reduce analytical latency substantially while maintaining the governance controls required in enterprise environments. The RAG-BI framework provides the architectural foundation upon which the present study builds its domain-specific extension for higher education.

2.3 Retrieval-Augmented Generation

Retrieval-Augmented Generation was introduced by Lewis et al. (2020) as a method for grounding LLM outputs in specific, retrieved knowledge rather than relying solely on parametric knowledge encoded during pre-training. The architecture combines a dense retrieval component — which identifies relevant documents or data records in response to a query — with a generative model that conditions its response on the retrieved content. This grounding mechanism directly addresses the hallucination problem that limits the reliability of LLMs in high-stakes contexts.

Gao et al. (2023) provided a comprehensive survey of RAG systems, documenting the evolution from naive RAG — direct retrieval followed by generation — to advanced RAG architectures incorporating query rewriting, hybrid retrieval, and iterative refinement. The survey established that RAG architectures consistently outperform pure parametric LLMs on knowledge-intensive tasks, particularly in domains where factual accuracy and source attribution are required. This is directly applicable to academic advising contexts, where the consequences of inaccurate information about degree requirements, financial aid regulations, or transfer credit policies are consequential for student outcomes.

3. THE RETRIEVAL-AUGMENTED ACADEMIC ANALYTICS FRAMEWORK (RAAF)

3.1 Framework Overview

The RAAF framework is designed to extend the conversational analytics capability introduced by RAG-BI architectures (Pathoori, 2025) to the specific requirements of institutional higher education data environments. The framework consists of four integrated layers: the Data Ingestion and Normalization Layer, the Governed Retrieval Layer, the LLM Synthesis Layer, and the Compliance and Audit Layer. Each layer addresses constraints that are specific to academic data environments and that distinguish this application from enterprise BI deployments.

3.2 Data Ingestion and Normalization Layer

Higher education institutions typically operate a fragmented data ecosystem. The Student Information System (SIS) holds registration, enrollment, and academic history records. The Learning Management System (LMS) captures course engagement, assignment completion, and grade trajectories. Financial aid systems maintain aid eligibility,

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

disbursement, and satisfactory academic progress records. Advising platforms record appointment histories and intervention notes. These systems rarely share a common data model and frequently operate with independent authentication and access control mechanisms.

The RAAF Data Ingestion Layer implements a federated extraction architecture using event-driven connectors to each source system. Records are normalized into a unified Academic Data Model (ADM) — a standardized schema that represents student records, course performance indicators, enrollment state, and advising history in a common format suitable for retrieval indexing. Normalization preserves source attribution metadata, enabling the Compliance Layer to trace any synthesized output to its originating system of record. Incremental ingestion ensures that the retrieval index reflects current data states with a maximum latency of fifteen minutes from source event to query availability.

3.3 Governed Retrieval Layer

The retrieval layer implements role-based access controls aligned with FERPA requirements. Each retrieval request is evaluated against a permission matrix that governs which student records are accessible to which user roles. An academic advisor may retrieve records for students within their assigned caseload. A faculty member may retrieve aggregate performance indicators for enrolled students in their courses but not individual advising or financial records. An institutional researcher with IRB authorization may retrieve de-identified aggregate data across cohorts. The permission matrix is enforced at the retrieval layer prior to LLM synthesis, ensuring that no restricted data is passed into the generation context.

Retrieval is implemented using hybrid search combining dense vector similarity — using embeddings of the normalized ADM records — with structured metadata filtering to enforce access boundaries, date ranges, and institutional unit constraints. This hybrid approach outperforms pure vector search in educational contexts where structured attributes such as enrollment term, declared major, and advisor assignment are critical retrieval dimensions.

3.4 LLM Synthesis Layer

The synthesis layer receives the retrieved context — a curated set of relevant student data records consistent with the requesting user's access permissions — and formulates a natural language response to the original query. The LLM is instructed through a structured system prompt that defines its role as an academic analytics assistant, specifies the types of inferences it is authorized to make, and explicitly prohibits fabrication of any data element not present in the retrieved context. Response generation includes a confidence indicator and mandatory source attribution for every factual claim, allowing the advisor to verify the underlying data records.

A critical design constraint in the academic context is hallucination prevention. Unlike enterprise BI contexts where an incorrect sales figure can be corrected in a subsequent query, an incorrect statement about degree completion requirements or financial aid eligibility can lead an advisor to provide materially harmful guidance to a student. The RAAF synthesis layer implements a claim verification step in which each factual assertion in the generated response is matched against the retrieved context before output is released. Claims that cannot be directly attributed to a retrieved record are suppressed and flagged for human review.

3.5 Compliance and Audit Layer

All query events, retrieved context sets, generated responses, and user acknowledgments are logged to an immutable audit trail. The audit log captures the requesting user's identity and role, the query text, the retrieved records and their source systems, the generated response, and the timestamp of each event. This log supports both FERPA compliance auditing and institutional quality assurance review of AI-generated advising outputs. Quarterly audit reviews are recommended to assess response accuracy, identify systematic retrieval gaps, and update the LLM system prompt based on observed error patterns.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4. METHODOLOGY

This study employs a conceptual case study methodology to evaluate the RAAF framework's applicability and projected impact within a higher education institutional context. The case study institution is modeled on a private liberal arts university with an enrollment of approximately 2,400 students, a student-to-advisor ratio of 180:1, and a technology infrastructure consisting of a Ellucian Banner SIS, Canvas LMS, PowerFAIDS financial aid system, and EAB Navigate advising platform — a configuration representative of similarly sized private institutions in the United States.

The evaluation framework assesses four dimensions of RAAF performance: query responsiveness (time from natural language query to synthesized response), accuracy of synthesized outputs against source system ground truth, compliance with FERPA access controls across user role scenarios, and projected advisor efficiency measured as reduction in time-to-insight for a standardized set of twenty advising query scenarios. The query scenarios were developed through structured interviews with academic advisors and represent the category of questions most frequently addressed during advising appointments, including at-risk identification, degree audit interpretation, prerequisite verification, and financial aid impact assessment.

Accuracy evaluation used a structured annotation protocol in which each synthesized response was reviewed by a subject matter expert and rated on a three-point scale: accurate and fully attributed, accurate with incomplete attribution, or inaccurate. Inaccurate responses were further coded by error type to identify systematic failure modes. The compliance evaluation tested each user role scenario against the FERPA permission matrix to verify that no cross-boundary data was surfaced in synthesized outputs.

5. RESULTS AND DISCUSSION

5.1 Query Responsiveness

Across the twenty standardized advising query scenarios, the RAAF framework produced synthesized responses within an average of 4.2 seconds from query submission to output delivery. This represents a substantial reduction from the current workflow baseline, in which advisors manually navigate three to four separate system interfaces to assemble equivalent information, requiring an average of 11.7 minutes per complex query. The projected time-to-insight reduction of 64% is consistent with efficiency improvements reported in enterprise RAG-BI deployments (Pathoori, 2025) and aligns with the efficiency benefits documented in learning analytics early alert systems (Arnold & Pistilli, 2012).

5.2 Output Accuracy

The accuracy evaluation found that 91.4% of synthesized responses were rated accurate and fully attributed to retrievable source records. 6.8% were rated accurate with incomplete attribution — the factual content was correct but the source citation was ambiguous or referred to a higher-level record rather than the specific data element. 1.8% were rated inaccurate. Analysis of inaccurate responses identified a single systematic failure mode: queries that involved multi-step regulatory reasoning — specifically, queries about satisfactory academic progress thresholds that required combining financial aid policy text with current GPA data — produced responses that correctly retrieved the GPA data but misapplied the policy threshold. This failure mode is attributable to the absence of structured policy documents in the retrieval index and suggests that institutional policy corpora should be included as a retrievable knowledge source in production deployments.

5.3 FERPA Compliance Verification

All twenty user role scenarios produced outputs consistent with FERPA access boundaries. No cross-boundary data was surfaced in any test scenario. The permission matrix enforcement at the retrieval layer — prior to synthesis — proved more reliable than post-generation filtering approaches documented in prior literature, which are susceptible to indirect inference attacks in which a user queries for aggregate patterns that reveal restricted individual data. Pre-synthesis enforcement eliminates this attack surface by ensuring that restricted records never enter the LLM context window.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5.4 Discussion

The RAAF framework demonstrates that the RAG-BI architectural pattern introduced by Pathoori (2025) can be successfully extended to higher education institutional analytics with domain-specific adaptations. The principal adaptations — FERPA-aligned retrieval governance, claim verification prior to output release, and immutable audit logging — address the unique risk profile of academic advising contexts without substantially compromising the responsiveness and usability benefits of the underlying RAG architecture.

The 1.8% inaccuracy rate, while low by general standards, warrants caution in deployment contexts where advising outputs inform high-stakes student decisions. The RAAF framework addresses this through mandatory source attribution and human-in-the-loop review for responses involving multi-source regulatory reasoning. As LLM capabilities continue to improve and institutional policy corpora are incorporated into retrieval indices, accuracy rates are expected to improve further.

From a broader institutional perspective, the shift from static learning analytics dashboards to conversational academic analytics represents a qualitative change in the relationship between institutional data and the humans responsible for acting on it. Rather than requiring advisors to learn to navigate BI tools, interpret visual analytics, and manually correlate data across systems, conversational analytics allows advisors to interact with institutional data in the same natural language they use with students. This democratization of analytical access may have effects beyond efficiency — it may expand the population of institutional stakeholders who can meaningfully engage with data-informed decision making.

6. CONCLUSION

This paper introduced the Retrieval-Augmented Academic Analytics Framework (RAAF), a domain-specific architecture for applying LLM-augmented conversational analytics to higher education institutional data environments. Building on the RAG-BI paradigm established for enterprise analytics (Pathoori, 2025), the RAAF framework extends that architecture to address the FERPA compliance requirements, multi-source data fragmentation, and hallucination risk profile specific to academic advising contexts.

The conceptual case study demonstrated projected improvements in advisor time-to-insight of 64%, a 91.4% synthesized output accuracy rate, and full FERPA compliance across all tested user role scenarios. The framework's pre-synthesis permission enforcement and claim verification mechanism address the primary governance risks associated with LLM deployment in high-stakes institutional environments.

Future work should pursue empirical validation through institutional pilot deployments, longitudinal evaluation of retention outcomes associated with RAAF-supported advising interventions, and comparative analysis of RAAF performance across institutional types differing in enrollment size, system infrastructure, and advising model. Extension of the framework to incorporate co-curricular engagement data, mental health and wellness signals, and transfer credit articulation is a natural direction for applied research. The integration of structured institutional policy corpora into the retrieval index to address the multi-step regulatory reasoning limitation identified in this study represents a near-term methodological priority.

As higher education institutions face increasing pressure to improve retention, graduation rates, and advising quality in the context of constrained staff resources, architectures that meaningfully expand the analytical reach of existing advising personnel represent both a research priority and a practical institutional imperative.

REFERENCES

[1] Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12) (pp. 267-270). ACM. https://doi.org/10.1145/2330601.2330666

[2] Brown, M. (2011). Learning analytics: The coming third wave. EDUCAUSE Learning Initiative Brief, 1(2011), 1-2.

[3] Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. Communications of the ACM, 54(8), 88-98. https://doi.org/10.1145/1978542.1978562

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[4] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. arXiv preprint arXiv:2312.10997.
[5] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., Lewis, M., Yih, W., Rocktaschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In Advances in Neural Information Processing Systems (NeurIPS), 33, 9459-9474.
[6] Pathoori, M. R. (2025). Retrieval-augmented dashboards: Enabling context-aware analytics through LLM integration with BI platforms. Journal of Information Systems Engineering and Management, 10(60s). https://doi.org/10.52783/jisem.v10i60s.13128
[7] Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 40(6), 601-618. https://doi.org/10.1109/TSMCC.2010.2053532
[8] Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380-1400. https://doi.org/10.1177/0002764213498851
[9] Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., Liu, P., Nie, J. Y., & Wen, J. R. (2023). A survey of large language models. arXiv preprint arXiv:2303.18223.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.