Integrating AI and Learning Analytics for Data-Driven Pedagogical Decisions and Personalized Interventions in Education

Ramteja Sajja[1] · Yusuf Sermet[1] · David Cwiertny[2,3,4,5] · Ibrahim Demir[1,6]

Abstract

This research explores the conceptualization, development, and deployment of an innovative learning analytics tool, leveraging OpenAI's GPT-4 model to quantify student engagement, map learning progression, and evaluate diverse instructional strategies within an educational context. By analyzing critical data points such as students' stress levels, curiosity, confusion, agitation, topic preferences, and study methods, the tool provides a comprehensive view of the learning environment. It also employs Bloom's taxonomy to assess cognitive development based on student inquiries. In addition to technical evaluation through synthetic data, feedback from a survey of teaching faculty at the University of Iowa was collected to gauge perceived benefits and challenges. Faculty recognized the tool's potential to enhance instructional decision-making through real-time insights but expressed concerns about data security and the accuracy of AI-generated insights. The study outlines the design, implementation, and evaluation of the tool, highlighting its contributions to educational outcomes, practical integration within learning management systems, and future refinements needed to address privacy and accuracy concerns. This research underscores AI's role in shaping personalized, data-driven education.

1 Introduction

In the rapidly transforming realm of education, the confluence of learning analytics (LA) and artificial intelligence (AI) signifies a critical paradigm shift. The transformative capabilities of these technologies and their potential to revolutionize the educational landscape have spurred significant research (Ouhaichi et al., 2023). Interdisciplinary insights from learning sciences provide an invaluable foundation to probe and comprehend human learning processes, and the amalgamation of this knowledge with AI technologies holds

Extended author information available on the last page of the article

the potential to enhance pedagogical methodologies (Luckin & Cukurova, 2019). Recent advancements in AI, particularly the development of large language models (LLM) like GPT-4, have showcased promising applications in various educational contexts (Pursnani et al., 2023). These models have demonstrated improved performance in standard examinations, highlighting the importance of noninvasive prompt modifications and revealing remarkable progress in their mathematical capabilities. Such advancements pave the way for AI's potential to address complex challenges in engineering, health (Sermet & Demir, 2018, 2021), and education.

AI's evolving role in education is further exemplified by the development of tailored AI-enabled educational assistants that integrate with Learning Management Systems (LMS) to provide personalized learning experiences, particularly in fields requiring complex data interactions (Sajja et al., 2024). These innovations underscore the growing impact of AI in enhancing student engagement and comprehension in specialized domains such as environmental sciences. However, despite these advancements, a significant gap persists in the integration of AI with pedagogical practices. Many AI developers exhibit limited familiarity with learning sciences research, resulting in tools that are technically sophisticated but pedagogically misaligned (Luckin & Cukurova, 2019). This disconnect highlights the imperative need to foster collaborative relationships among AI developers, educators, and researchers to ensure that technological innovations are grounded in educational theory and practice.

Current LA tools often rely on quantitative metrics such as attendance, grades, and completion rates, which, while useful, fail to capture the nuanced, multidimensional nature of student learning. For instance, these tools frequently overlook qualitative aspects such as emotional engagement, contextual understanding, and real-time feedback, which are critical for fostering meaningful learning experiences (Smith et al., 2022). Moreover, many existing AI-driven educational tools prioritize algorithmic efficiency over pedagogical relevance. While AI models can predict student performance with high accuracy, they often lack mechanisms to provide timely, constructive feedback to students, limiting their utility in fostering self-regulated learning (Ouyang et al., 2023). This gap between AI capabilities and pedagogical needs is further exacerbated by the limited collaboration between AI developers and educators, resulting in tools that are technically advanced but educationally misaligned.

In real-world scenarios, these limitations manifest in several ways. For example, in online learning environments, instructors often struggle to identify disengaged students until it is too late, as traditional LA tools fail to capture subtle indicators of disengagement, such as changes in participation patterns or affective states. Similarly, in specialized fields like environmental sciences, where complex data interactions are common, existing tools lack the sophistication to provide personalized, context-aware feedback, hindering students' ability to grasp intricate concepts (Sajja et al., 2024). These challenges underscore the need for LA tools that not only leverage AI's predictive capabilities but also integrate insights from learning sciences to provide holistic, real-time support for both students and educators.

With the steady integration of AI within the education sector (Sermet & Demir, 2020), it is crucial to comprehend its impact and application. Alam (2021) offers a particularly insightful exploration of AI's evolution from computer technology to intelligent online education and embedded computer systems. AI not only enhances the efficiency and quality of instruction but also tailors individual learning pathways by adapting curriculum to meet each student's unique needs. In this context, Ouyang et al. (2023) provides a noteworthy

examination of AI performance prediction models within education. The study identifies a gap where existing AI models tend to prioritize algorithm development and optimization, often overlooking the need for timely, continuous feedback to students. The collective insights from these studies underscore the motivation behind our research to harness the potential of AI within LA to optimize educational outcomes while addressing the gaps in current tools and practices.

Our study aims to achieve several primary objectives. Firstly, we aim to develop an AI-empowered LA tool designed to effectively measure and analyze student engagement. This technology's goal is to monitor and quantify various aspects of student involvement, enabling the delivery of pedagogical interventions based on real-time, actionable insights. Secondly, we seek to construct a comprehensive AI-based system to monitor students' learning progression. This objective relies on harnessing AI's predictive modeling capabilities to map individual learning pathways, providing educators with a nuanced understanding of each student's educational journey. Lastly, we aim to establish a mechanism to assess students' affective states. By leveraging AI capabilities, we aim to gain insights into the emotional context of learning, recognizing its significant role in shaping cognitive processes and educational outcomes.

Our research boundaries are explicitly outlined, with a central focus on the integration of AI into LA for improved educational outcomes. In addressing the challenges of educational opportunities, our tool is designed with a focus on inclusivity, ensuring that insights derived from AI analytics are accessible and beneficial to all student populations. The specific elements under consideration include the measurement of student engagement, tracking of learning progression, and assessment of affective states. Evaluated metrics comprise, but are not limited to, participation rates, progression speed, depth of understanding, emotional response, and learning adaptation.

The structure of this paper is as follows: Sect. 2 delves into the existing literature, emphasizing gaps in the current knowledge. Section 3 articulates the methodology, detailing the design choices and implementation of the LA tool. Section 4 presents our findings and interprets their significance. Finally, Sect. 5 discusses the strengths and limitations of our study, potential future directions, and summarizes our contributions to the field.

2 Related Work

LA, an emerging interdisciplinary field, stands at the crossroads of machine learning, data science, education, and cognitive psychology. Initiated with the intent to leverage the vast data produced through technological integration in education, LA has grown into a community with substantial impacts on research, practice, policy, and decision-making (Baker & Inventado, 2014; Gašević et al., 2017). This composite field draws on theoretical underpinnings, design considerations, and data science methodologies to offer innovative perspectives for enhancing learning (Reimann, 2016; Siemens, 2013). Despite being in its growth phase, LA has demonstrated potential for transforming teaching practices, informing learning research, and reshaping the educational landscape (Clow, 2013; Dawson et al., 2019).

The expansion of LA has been fueled by an explosion of accessible learner data and concurrent advancements in data management methodologies. This data availability enables a more nuanced comprehension of student behavior and performance, aiding educators in effi-

cient resource allocation and improving teaching outcomes (Clow, 2013). LA encompasses a wide range of data types, each requiring unique analytical techniques and methodologies. Applications range from eye-tracking and automated online dialog analysis to ecosystem surveys and log data analysis at individual and collaborative levels, extending even to visual LA applied to IoT data (Nistor & Hernández-García, 2018).

Recent advancements in deep learning have significantly contributed to knowledge tracing, a critical component of LA. Deep Knowledge Tracing (DKT) uses data to forecast student performance, facilitating early identification of potential learning hurdles and enabling precise intervention strategies (Casalino et al., 2021). The field employs diverse methods, including visual data analysis, social network analysis, and educational data mining, to support targeted course offerings, curriculum development, personalized learning experiences, and instructor performance improvement. However, challenges persist in data tracking, evaluation, and analysis, alongside ethical and privacy concerns. Moreover, there is a pressing need for stronger integration with learning sciences to optimize learning environments (Avella et al., 2016).

Recent studies have made significant strides in understanding and measuring student engagement, a central focus of LA research. Bowden et al. (2021) propose a holistic framework for analyzing student engagement, examining involvement and expectations as antecedents to engagement and assessing its impact on student and institutional outcomes. While their approach provides a comprehensive theoretical foundation, it lacks practical tools for real-time engagement measurement and intervention. Similarly, Goldberg et al. (2021) employs machine learning to assess visible engagement during classroom instruction, validating a manual rating system using gaze, head pose, and facial expressions as engagement indicators. Although their work demonstrates the potential of machine vision-based approaches, it is limited to physical classroom settings and does not address the complexities of online or hybrid learning environments.

In the realm of e-learning, Bhardwaj et al. (2021) utilize deep learning algorithms to track students' real-time emotional responses. By integrating facial landmark detection, emotional recognition, and survey results, they compute a Mean Engagement Score (MES) to enhance digital learning methodologies. While their approach offers valuable insights into affective states, it relies heavily on visual data, which may not be feasible or ethical in all educational contexts. Furthermore, Nkomo et al. (2021) provides a systematic review of the impact of digital technologies on teaching and learning practices over the past decade, highlighting notable gaps in understanding and measuring student engagement with digital technologies. Their findings underscore the need for more adaptable and scalable LA tools that can address diverse learning environments and student populations.

Despite these advancements, several gaps remain in the LA landscape. Current methods often lack integration with pedagogical theory and teaching practice, making it challenging for educators to apply insights effectively. Additionally, many LA tools prioritize quantitative metrics over qualitative and affective dimensions, which are crucial for understanding student engagement. Scalability and generalizability across different educational settings also pose significant challenges, as do ethical and privacy concerns related to data collection and analysis.

Our study contributes to the field of LA by developing an AI-enhanced tool that integrates pedagogical theory, leverages multimodal data sources, and provides real-time, actionable insights for educators. This approach addresses the limitations of existing tools while pri-

oritizing inclusivity, scalability, and ethical considerations. By leveraging advanced AI techniques such as natural language processing (NLP) and affective computing, our tool can analyze textual and behavioral data to provide a more holistic understanding of student engagement and affective states, addressing the limitations of visual-data-centric approaches. Furthermore, our tool is grounded in learning sciences research, ensuring that its design and functionality align with established pedagogical principles. This integration enables educators to apply insights from the tool in ways that are both theoretically sound and effective.

Building on the holistic framework proposed by Bowden et al. (2021), we aim to operationalize their theoretical insights into a practical tool capable of real-time engagement measurement and intervention. Unlike Goldberg et al. (2021), our approach is designed to be adaptable to both physical and online learning environments, leveraging multimodal data sources beyond visual cues to ensure inclusivity and scalability. Drawing on the affective computing techniques employed by Bhardwaj et al. (2021), we incorporate emotional recognition into our tool but extend its application to include textual and behavioral data, addressing ethical concerns and broadening its utility. Furthermore, our tool is designed to provide actionable insights to educators, bridging the gap between LA research and teaching practice. By incorporating advanced AI and machine learning techniques, we aim to analyze and respond to student data in real-time, offering immediate, personalized feedback and support for students while providing educators with nuanced insights into learning progression and affective states, which can address the limitations of existing LA tools.

3 Methodology

This research centers around the development of a novel LA tool, conceptualized as an extension to existing educational platforms such as educational intelligent assistants. The tool's primary purpose is to enhance the capacity of these platforms to collect, process, and analyze a wealth of data, extending beyond the information traditionally gathered by LMS. VirtualTA, an AI-augmented intelligent educational assistant, plays a pivotal role in this study. VirtualTA is specifically designed to address logistical questions about courses (Sajja et al., 2023a) and provide responses to course-content related inquiries. Additionally, it can generate flashcards and quizzes (Sajja et al., 2023b).

While the development and application of the tool proposed in this study is contextualized within the VirtualTA environment, it is designed with flexibility and adaptability in mind, making it compatible with a multitude of educational chatbots, and smart assistants such as Instant Expert (Sermet & Demir, 2019) or LMS. The overarching aim of the tool is to offer educators a comprehensive understanding of students' performance, engagement, and learning patterns. This, in turn, enables them to make data-driven decisions to enrich the overall learning experience. The research proceeds systematically, encompassing stages of data collection, processing, analysis, and finally, tool development and deployment.

3.1 Theoretical Framework: Design Science Approach

Design Science Research (DSR) is a well-established methodology in information systems and educational technology that focuses on the development and evaluation of artifacts

designed to address practical problems (Hevner et al., 2004; Peffers et al., 2007). This study adopts the DSR approach to guide the development of an AI-enhanced Learning Analytics Tool, ensuring that the research maintains both theoretical rigor and practical applicability. By following the structured DSR process, the study aims to create a system that is adaptable to student needs, responsive to emotional states, and mindful of ethical and privacy considerations while providing personalized and actionable insights for educators and learners.

The research begins with problem identification and motivation, recognizing the increasing need for AI-driven learning analytics that deliver real-time, adaptive feedback to support both students and instructors. Traditional learning analytics systems often fail to account for emotional and behavioral learning patterns, limiting their ability to provide meaningful, personalized interventions. The objective of this study is to design and develop a tool that integrates cognitive, affective, and behavioral learning indicators, ensuring a comprehensive and context-sensitive approach to learning analytics.

The design and development phase of the research focuses on constructing the system architecture to process learning data using machine learning techniques. The tool incorporates real-time analytics, predictive modeling, and visualization dashboards to enhance the educational experience. Special attention is given to ethical considerations, ensuring privacy, security, and transparency in data handling. The architecture is developed with flexibility in mind, allowing continuous refinement based on user feedback.

The demonstration phase, currently in progress, involves the testing of the prototype with instructors who interact with the system and provide qualitative feedback on its usability and effectiveness. Through hands-on engagement, educators assess the system's user interface, predictive accuracy, and responsiveness, allowing for initial insights into the tool's practical impact on teaching and learning. These interactions will inform subsequent refinements to enhance the tool's usability and effectiveness before broader deployment.

The evaluation phase is an ongoing and future-oriented component of the research, designed to assess both technical performance and user-based outcomes. The technical evaluation will focus on key performance indicators such as system efficiency, predictive accuracy, and response time. Simultaneously, user feedback from instructors and students will be gathered to evaluate the practical effectiveness and usability of the system. This iterative evaluation process will ensure that the tool continues to evolve based on real-world needs, refining its predictive capabilities and user experience over time.

Finally, the findings from this research will be communicated to the academic and practitioner community through publications, workshops, and educational conferences. By sharing insights from both the development and evaluation phases, this study aims to contribute to the broader advancement of AI-driven learning analytics and inform best practices for adaptive educational technologies. The integration of the Design Science framework ensures that this study not only contributes theoretically to the field of AI-enhanced education but also delivers a practical solution with tangible benefits for educators and learners.

3.2 Data Collection and Pre-processing

The effectiveness of the AI-enhanced learning analytics tool depends on a structured approach to data collection and pre-processing, ensuring accurate and meaningful insights. The system captures student-chatbot interactions, emotional states, academic topics, and engagement metrics to track learning behaviors and identify areas requiring support. Both

quantitative data (e.g., engagement metrics, quiz scores) and qualitative data (e.g., sentiment analysis, open-ended feedback) are integrated for a holistic understanding of student learning. To ensure data reliability, pre-processing involves cleaning, structuring, and analyzing interactions using NLP and cognitive assessment models. This step refines raw data into actionable insights, enabling educators to monitor student progress and personalize interventions effectively.

3.2.1 Data Collection

At its core, the system records raw dialogues between students and the chatbot, creating a detailed log of queries and responses. This log serves as the foundation for further analysis, enabling the tool to track the evolution of student inquiries over time and identify recurring themes or challenges. Emotional analysis is performed to infer affective states such as stress, curiosity, confusion, and engagement. By analyzing the tone, language, and context of student messages, the system provides valuable insights into the emotional context of learning, allowing educators to identify moments when students may need additional support or encouragement.

Another critical component of the data collection process is topic mapping. The tool identifies and categorizes the topics of student questions and messages, creating a detailed map of academic interests and areas of difficulty. This feature helps educators understand which subjects or concepts students find most challenging or engaging, enabling them to tailor their teaching strategies accordingly. For example, if a significant number of students are asking questions about a particular topic, it may indicate that the topic requires further clarification or additional resources.

Engagement metrics are also a key focus of the data collection process. The tool tracks interactions with features such as flashcards and quizzes, capturing data on time spent on questions, quiz duration, and outcomes. These metrics provide insights into how students engage with different learning materials and activities, highlighting patterns in study habits and preferences. For instance, a student who spends a significant amount of time on a particular quiz question may be struggling with the underlying concept, while a student who completes quizzes quickly may benefit from more challenging material.

To ensure a comprehensive view of student behavior, the tool integrates both quantitative and qualitative data. Quantitative data includes metadata such as assignment scores, interaction frequency, and time spent on learning materials, while qualitative data encompasses instructor-conducted surveys and open-ended feedback. This dual approach allows for a nuanced understanding of student engagement and performance, combining the objectivity of numerical data with the depth of subjective insights.

3.2.2 Data Pre-processing and Analysis

Raw data, while abundant and rich in potential insights, requires rigorous pre-processing to ensure its accuracy, reliability, and usability. The pre-processing pipeline begins with data cleaning, where irrelevant, duplicated, or incorrect data entries are identified and removed. For example, interactions such as "user clicked on a button" are filtered out, as they provide no meaningful insights into student behavior or learning progression. Similarly, incomplete

quiz attempts, where a student starts but does not finish, are excluded to avoid inaccurate time measurements and misleading results.

Once the data is cleaned, it is structured into a standardized format to ensure compatibility with downstream analysis tools. The tool uses JSON (JavaScript Object Notation) for this purpose, a lightweight and flexible data-interchange format that is widely supported across various platforms and programming languages. JSON's hierarchical structure allows for efficient organization of complex data, such as nested student interactions or multi-layered emotional analysis results.

The next phase of pre-processing involves emotional and cognitive analysis, where advanced NLP techniques are employed to infer affective states from student communications. By analyzing the tone, language, and context of student messages, the tool can identify emotional states such as stress, curiosity, confusion, or engagement. For instance, a student's message expressing frustration or uncertainty might be flagged as indicative of stress or confusion, prompting the system to recommend targeted interventions.

In addition to emotional analysis, the tool assesses cognitive development by analyzing student questions using Bloom's Taxonomy, a widely recognized framework for categorizing educational goals. Questions are classified according to their cognitive complexity, ranging from lower-order thinking skills (e.g., recalling facts) to higher-order thinking skills (e.g., analyzing or evaluating concepts). This analysis helps educators understand the depth of understanding and identify areas where students may need additional support or more challenging material.

The pre-processed data is then updated in real-time on the instructor's dashboard, providing educators with actionable insights into student engagement and learning progression. This step ensures that the insights presented are derived from a clean, consistent data set, free from the noise and inconsistencies that often plague raw data.

3.3 Ethical Considerations

Ethical considerations, particularly regarding student data privacy, are central to the design and implementation of the LA tool. The tool adheres to the principles of "privacy by design," ensuring that student data is handled responsibly and transparently. This approach prioritizes the protection of student confidentiality while fostering a secure environment for engagement and learning.

To achieve this, the tool implements strict data anonymization protocols. Raw text data, such as student messages and queries, is discarded immediately after analysis, ensuring that only computed metrics including emotional state indicators (e.g., stress, curiosity), engagement scores, and learning progression data are retained. Additionally, student data is not associated with personally identifiable information (PII). Instead, a unique, system-generated identifier is assigned to each user, allowing for longitudinal tracking while preserving anonymity.

The tool also employs temporal data collection mechanisms, ensuring that information is recorded only during active interactions. For instance, data is captured when a student engages with the chatbot or completes a quiz, but no passive tracking occurs beyond these sessions. This approach minimizes unnecessary data retention, mitigating potential risks associated with unauthorized access or data misuse while maintaining analytical utility.

Access to stored data is strictly restricted. The database is secured with encryption protocols and can only be accessed by authorized personnel with role-based permissions. These permissions ensure that only designated researchers or system administrators can retrieve data necessary for analysis, and no individual user's data can be accessed without appropriate authorization. Additionally, access logs are maintained to monitor any interactions with the database, providing an audit trail for accountability.

Beyond technical safeguards, the system is designed to encourage trust and openness in student interactions. By eliminating the storage of raw conversational data and implementing identity-abstraction measures, students are more likely to engage without concerns about privacy intrusion. This is particularly important in educational settings, where students may hesitate to seek academic support if they perceive a risk of their interactions being permanently recorded or analyzed beyond their intended use.

These ethical measures are not supplementary considerations but rather foundational elements of the tool's architecture. By embedding privacy protections at every stage of development, the system upholds ethical data practices while ensuring that learning analytics can be leveraged to enhance educational outcomes in a secure and responsible manner.

3.4 Metrics and Indicators

The tool employs a suite of metrics and indicators to provide a holistic view of student interactions and experiences. These metrics are categorized into three key areas:

- Engagement Metrics: Measures such as participation rates, interaction frequency, and time spent on tasks provide insights into student interest, commitment, and active learning behaviors.
- Affective States: Emotional analysis techniques infer states such as stress, curiosity, and confusion, offering a window into the emotional context of learning and identifying potential areas where students may require additional support.
- Learning Progression: Metrics such as quiz performance, topic mastery, and cognitive complexity based on Bloom's taxonomy are tracked over time, providing a measure of academic improvement and knowledge depth.

A detailed breakdown of these key metrics and their specific indicators is provided in Table 1, which outlines the methods used to assess engagement, affective states, and learning progression. While these metrics provide valuable insights, they are not diagnostic tools and should be interpreted in context, considering qualitative factors and additional student inputs. To ensure clarity and transparency, the tool includes the following disclaimer where applicable: *"These metrics are estimates and should be used as reference data only, not as definitive measures of student performance or engagement."*

3.5 Development and System Architecture of the Learning Analytics Tool

The LA tool is designed to enhance student learning experiences by leveraging AI-driven analytics, sentiment analysis, and cognitive progression tracking. Built on OpenAI's GPT-4 (OpenAI, 2023), the tool seamlessly integrates with Virtual Teaching Assistants (VTAs) and LMS to provide real-time, data-driven insights into student engagement, emotional

states, and cognitive progression. Unlike conventional LMS platforms that primarily collect behavioral data, the LA tool offers a more comprehensive analysis of student interactions by incorporating sentiment assessment, topic identification, and learning progression tracking, allowing for a deeper understanding of learning behaviors and challenges.

The system architecture of the LA tool, as illustrated in Fig. 1, follows a structured workflow to ensure seamless integration with LMS environments, virtual assistants, AI-based analytics, and instructor dashboards. The process begins with the LMS Environment, where the tool connects to various learning platforms such as Schoology, Blackboard, Moodle, Google Classroom, and Canvas. This enables the retrieval of course resources and interaction logs, forming the foundation for deeper learning analytics.

Next, the Virtual Assistant acts as the primary interface through which students engage with the system. The tool records students' questions, tracks the time taken per question and quiz duration, and monitors their preferred study methods, such as flashcards, quizzes, and brief summaries. This data provides an initial layer of engagement analysis, offering insights into how students interact with educational content.

The collected raw data is then sent to the Framework Foundation, where it undergoes data pre-processing, analysis, storage, and security validation. During pre-processing, raw interaction logs are cleaned and structured for analysis, ensuring that irrelevant or incorrect entries are removed. The cleaned data is then analyzed to identify patterns in student queries, engagement behaviors, and emotional states. Strict anonymization protocols are implemented to protect student identity, and only necessary information is retained for further processing.

A key component of the system is GPT-4 Enhanced Analytics, where AI models process refined data to generate real-time insights. One of its primary functions is sentiment analysis, which assesses students' emotional states, including stress, confusion, curiosity, and agitation. By analyzing language patterns and interaction behaviors, the system can provide contextual emotional insights, helping educators identify moments when students may be struggling or disengaged. Additionally, the tool conducts user analytics to evaluate overall learning behaviors and engagement levels. By tracking interaction frequency, preferred study methods, and participation trends, the system identifies patterns in student engagement, enabling instructors to tailor support strategies accordingly.
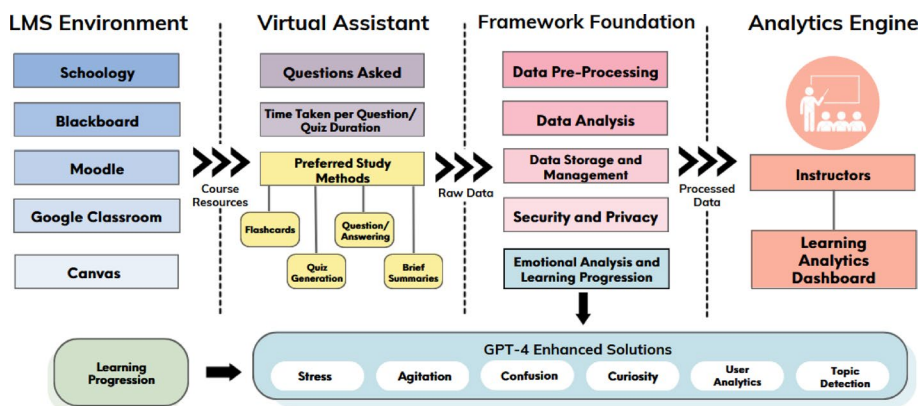

Fig. 1 System architecture diagram and components

Once the data processing is complete, insights are delivered to the Analytics Engine, where they are made available to instructors through a dedicated learning analytics dashboard. This dashboard allows educators to monitor student engagement levels in real-time, providing immediate visibility into participation trends and learning behaviors. Additionally, instructors can identify students who may be struggling, based on cognitive and emotional indicators such as repeated confusion signals or lack of progression in learning complexity. By leveraging these AI-driven insights, educators can refine instructional strategies, adjusting course materials, assessments, and support mechanisms to align with student needs.

This end-to-end workflow ensures that data from LMS platforms is captured, analyzed, and transformed into actionable insights. By maintaining a privacy-conscious and ethically responsible AI framework, the system allows educators to make informed pedagogical decisions while respecting student confidentiality. The integration of AI-enhanced solutions not only enables early intervention for struggling students but also fosters a more personalized, adaptive, and responsive learning experience.

3.6 Implementation and Deployment

The LA tool is designed for seamless integration with various educational platforms, including VTAs and LMS. The primary focus during development and deployment was to transform student interactions into meaningful data points, allowing for real-time analysis of engagement, affective states, and learning progression. The system ensures compatibility with multiple educational chatbots and LMS platforms, making it a scalable and adaptable solution for diverse learning environments.

3.6.1 Natural Language Inference

A key component of the LA tool is Natural Language Inference, which enables the system to analyze student queries for emotional and cognitive insights. The tool utilizes OpenAI's GPT-4, a cutting-edge language model, to assess stress levels, agitation, curiosity, and confusion in student interactions. This ability to infer subtle emotional cues enhances the tool's effectiveness in identifying learning difficulties and engagement levels.

The choice of GPT-4 over other models was driven by several critical factors. First, GPT-4 offers superior contextual understanding, outperforming models like BERT and T5 in recognizing nuanced emotional cues and broader contextual intricacies. This makes it ideal for analyzing student interactions. Second, GPT-4's few-shot learning capabilities allow it to quickly adapt to new educational contexts with minimal training, reducing development time and computational overhead. Finally, custom prompt engineering enhances GPT-4's accuracy in detecting emotional states and cognitive complexity, ensuring reliable and meaningful insights. Alternative models, such as BERT and T5, were considered but rejected due to their limited contextual depth and higher dependency on fine-tuning.

To derive these essential emotional and cognitive metrics, the LA tool employs few-shot learning in combination with prompt engineering, configured with a temperature setting of 1.0. Few-shot learning enables the model to generalize across diverse contexts with minimal labeled examples, making it highly adaptable without requiring extensive dataset-specific retraining. Prompt engineering, on the other hand, refines the model's interactions, ensuring that responses align precisely with the educational domain's unique requirements. The tem-

perature setting of 1.0 plays a crucial role in this process, balancing response diversity and contextual relevance. A higher temperature generates more varied and creative responses, whereas a lower temperature leads to more deterministic outputs. By setting the temperature to 1.0, the model explores a wide range of possible responses while maintaining logical coherence, allowing it to better capture subtle emotional nuances in student queries.

Additionally, the tool maintains a structured conversation history to ensure continuity in student interactions. Given GPT-4's token limitation (maximum of 8,192 tokens, approximately 6,000 words), the tool employs a first-in-first-out (FIFO) strategy, ensuring that the most recent interactions remain prioritized. This method enables longitudinal tracking of student learning progression, allowing the system to assess knowledge retention and cognitive development based on Bloom's Taxonomy.

3.6.2 Software Development and Architecture

The LA dashboard has been designed using vanilla JavaScript, ensuring a user interface that is both straightforward and efficient for instructors. To aid in visualizing student data patterns and offering a more interactive experience, the tool integrates Chart.js, a library known for its capability to produce dynamic, real-time charts. This allows instructors to easily discern trends and insights, facilitating data-driven decision-making.

The technical foundation of the system relies on the synergistic combination of Express.js and Node.js in the backend. This duo not only streamlines backend operations but also optimizes performance, ensuring that the system remains responsive even with high data loads. One of the essential tasks of the backend infrastructure is to handle and process data from the educational intelligent assistant, which in this context is provided by VirtualTA, as well as to interface directly with OpenAI's GPT API. Communication with these educational assistants is facilitated through APIs of the LA tool, ensuring a seamless transfer of data. All the information sourced from these chatbot interactions is structured in the JSON format, chosen for its inherent flexibility and ease of integration. After thorough processing and analysis, the data is relayed to the dashboard, ensuring that instructors have timely and easy access to the insights they need.

3.7 Case Study Design

The AI-driven learning analytics tool was evaluated through a survey distributed to the teaching faculty at the University of Iowa. The primary aim was to gather feedback on the tool's potential effectiveness in enhancing educational outcomes and its compatibility with existing LMS such as Canvas. This evaluation focused solely on faculty perceptions of the tool's potential to improve student engagement and provide valuable real-time insights for instructional decision-making. It is important to note that the faculty did not interact directly with the tool, and the feedback gathered was based on their understanding of its described features. Feedback was collected through an anonymous online survey provided to the participating faculty members. The survey was designed to capture insights across several key areas:

- Importance of Classroom Insights: Faculty were asked to select all relevant insights they considered important for understanding student behavior and performance in their

classrooms.

- Student Privacy and Data Security: Participants shared their concerns regarding student data privacy and security when using new AI-driven tools.
- Accuracy of Tool-Provided Insights: Faculty were asked to rate the importance of receiving accurate insights from new tools, ensuring reliability in their teaching practices.
- Primary Teaching Area and Role: Instructors indicated their primary teaching area and their role within the institution, to better contextualize their feedback.
- Additional Comments and Suggestions: An open-ended question invited faculty to provide any additional comments or suggestions regarding tools designed to improve student learning and engagement.

4 Results and Discussion

This section presents the findings from our evaluation of the AI-driven learning analytics tool, focusing on its capabilities and the insights gathered from both synthetic test data and survey responses from the teaching faculty at the University of Iowa. The results are structured into two parts: the first evaluates the tool's performance using synthetic data, demonstrating its potential functionalities, while the second presents feedback from a survey of faculty, offering their perspectives on the tool's practical applications and perceived value in educational settings.

We aim to assess how effectively the tool could track student engagement, monitor emotional states, and provide actionable insights for educators. This analysis explores the technical performance of the tool based on synthetic data, alongside faculty perceptions of its potential usefulness and ease of integration into existing pedagogical practices.

4.1 Synthetic Data Analysis

The LA tool employs various metrics to analyze students' learning experiences, including emotional states, topic preferences, learning progression, and study methods. By integrating these diverse factors, it offers a holistic perspective on student engagement and cognitive development, surpassing traditional assessment methods. To evaluate the tool's capabilities while ensuring privacy, synthetic test data was generated using ChatGPT, simulating real-world learning behaviors without using actual student data.

The synthetic data was designed to mirror authentic student interactions with VirtualTA across different learning scenarios. Inquiry-based learning was a key focus, with simulated students posing questions ranging from basic factual inquiries, such as "What is Newton's third law?" to complex problem-solving discussions like "How does Newton's third law apply in a zero-gravity environment?" This allowed the tool to be assessed on its ability to classify queries according to Bloom's Taxonomy and detect shifts in cognitive complexity.

To evaluate sentiment analysis, student responses were constructed to represent emotional states such as stress, curiosity, confusion, and agitation. These emotions were embedded in academic contexts, such as a stressed student expressing frustration with physics equations or a curious student questioning the gravitational pull on ocean tides. The tool was tested on its ability to accurately classify these emotional states and provide insights into psychological engagement and learning challenges.

Study habits and quiz interactions were also simulated, with students engaging in activities like flashcard creation, quizzes, and content summarization. For instance, students requested summaries of historical events or definitions for scientific terms, allowing the tool to recognize study preferences and adapt to different learning styles. By analyzing these interactions, the system demonstrated its ability to tailor feedback and learning recommendations effectively.

4.1.1 Stress, Curiosity, Confusion, and Agitation Measurement

The LA tool employs GPT-4 for sentiment analysis to evaluate students' emotions during their interactions with VirtualTA, focusing on stress, curiosity, confusion, and agitation. These emotions are scored on a scale of 1 to 10 to reflect intensity. For example, difficulty-related queries may indicate high stress, while exploration questions suggest curiosity. This real-time emotional analysis helps track students' emotional journeys, highlighting trends and spikes. By capturing these emotional shifts, educators can provide tailored responses, enhancing support and improving the learning environment.

The dashboard shows both class-wide (Fig. 2a) and individual (Fig. 2b) stress levels over time, correlating them with key academic events such as quizzes and assignments. These visualizations help educators understand how emotional states align with the academic schedule, enabling timely interventions to alleviate stress and support student well-being.

4.1.2 Topic Analysis and Student Questioning Patterns

The LA tool uses GPT-4 for topic analysis and tracking questioning patterns in student interactions with VirtualTA. It identifies and records the main topics of student queries, mapping out areas of interest or difficulty over time. This enables the identification of patterns, like frequently discussed subjects or challenging topics. Alongside, the tool analyzes questioning patterns, including question types, sentiments, and sequencing. This reveals a student's learning journey, such as progressing from basic to complex questions within a topic, offering insights into their understanding. These analyses allow educators to tailor their teaching and resources more personally, and the system to adapt its responses to each.
student's learning path.

Figure 3 presents an organized tabular representation of unique topics that students in the class have inquired about up until the point of data collection. For the sake of clarity and efficient use of educators' time, we've ensured that duplicate entries are carefully
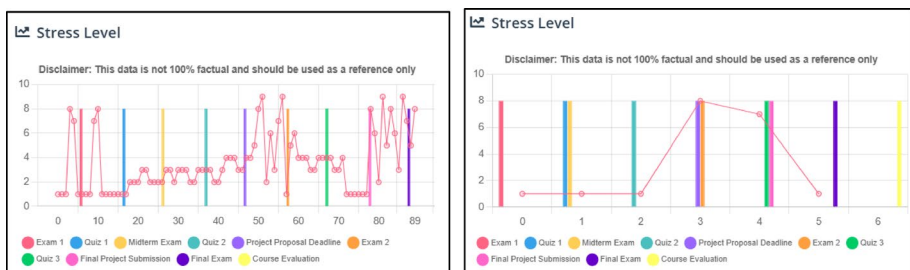

Fig. 2 Screenshots from the instructor's dashboard **a** stress level graph for overall students, **b** stress level graph for an individual student

Fig. 3 Topics asked table for overall students in the class
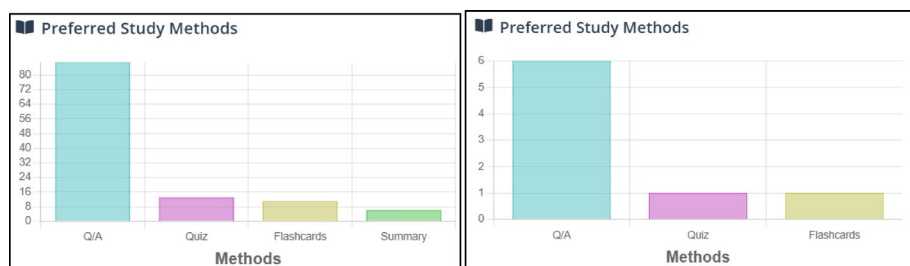

Fig. 4 Screenshots from the instructor's dashboard **a** preferred study methods for overall students in the class, **b** preferred study methods for a specific student in the class

removed from the table. Each row corresponds to a unique topic, providing educators with a straightforward view of the breadth of subject's students have engaged with throughout the course. This streamlined, tabulated format provides a quick and comprehensive view of the range of topics discussed, aiding them in identifying key areas of student interest or potential difficulties.

4.1.3 Study Methods and Preferred Learning Approaches

The LA tool tracks student interactions with VirtualTA, recording activities like asking questions, creating flashcards, taking quizzes, or requesting summaries. This data helps identify individual and class-wide study preferences, such as a preference for active recall through flashcards or learning by testing via quizzes. These insights allow the system to tailor responses and resources to align with each student's habits, enabling a more personalized learning experience. Educators can use this data to adjust their teaching strategies, promoting more effective and diverse study methods.

The graphs show class-wide (Fig. 4a) and individual (Fig. 4b) study methods, highlighting the most-used interaction modes questions, quizzes, summaries, and flashcards. This data helps educators understand student preferences, enabling them to tailor instruction and enhance student engagement.

4.1.4 Tracking and Analyzing Student Interactions

The LA tool tracks student quiz interactions via VirtualTA, recording quiz topics, completion times, time spent on each question, and quiz results. This data helps analyze students'

interests, time management, comprehension speed, and areas of difficulty. The tool allows for personalized learning by adjusting quizzes to match individual strengths and improvement areas. Quiz results also provide insights into student understanding and topic mastery.

The bar graph (Fig. 5a) shows time spent on each quiz question, highlighting question complexity and potential challenges. Figure 5b shows quiz outcomes, distinguishing correct and incorrect responses. These insights help educators identify areas needing more focus, supporting targeted teaching strategies and student improvement.

4.1.5 Learning Progression and Bloom's Taxonomy

The LA tool also incorporates a novel method of tracking students' learning progression by leveraging the few-shot learning capabilities of GPT-4. Each interaction a student has with the VirtualTA, including all messages and questions asked during a session, is analyzed through the lens of Bloom's Taxonomy (Bloom & Krathwohl, 2020), a widely accepted framework that categorizes learning objectives into different levels of complexity and specificity.

By using the GPT-4 model, we extract cognitive level indicators from student queries, allowing us to track changes in the depth and complexity of their understanding over time. For instance, questions indicative of recall or comprehension suggest an early stage of learning, while questions demonstrating application, analysis, synthesis, or evaluation suggest more advanced understanding. This progression mapping is based on digital planning verbs derived from bloom's taxonomy (Churches, 2010; Nikolić & Dabić, 2016). The digital planning verbs used for the learning progression are listed in Table 2. By integrating these verbs within the GPT-4 analytical framework, we can offer nuanced insights into the cognitive journey of the students. By observing these trends, educators can tailor their teaching strategies to better meet the changing needs of their students, offering a more personalized and effective learning experience.
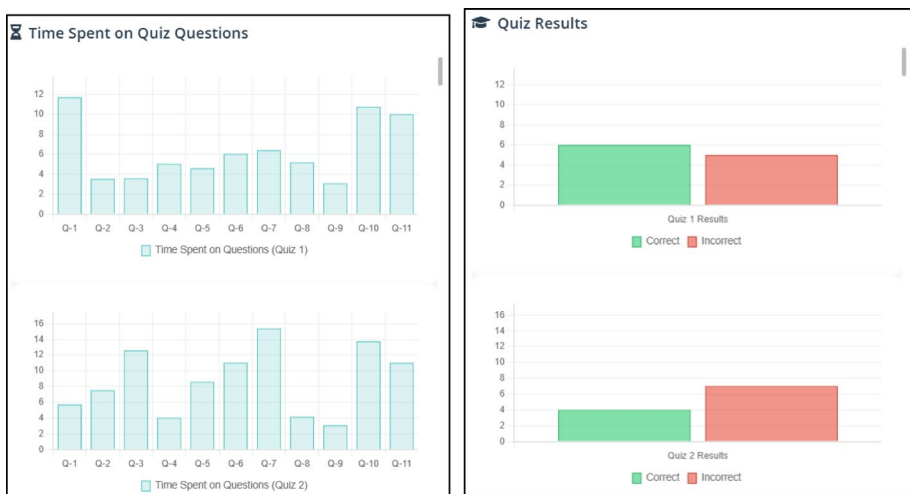
Fig. 5 Screenshots from the instructor's dashboard **a** time spent on quiz questions and **b** quiz results of the respective quiz

Figure 6 presents a detailed evolution on the progression and transformation of students' inquiries. Such a comprehensive understanding is invaluable, as it can facilitate the detection of learning trends, highlight topics that might necessitate additional clarification or emphasis, and consequently, guide the development of more effective, tailored teaching strategies. A disclaimer is included with the table, visible only when the table is hovered over, ensuring the layout remains clean and focused.

4.1.6 Interactive Holistic Report

As we transition into discussing the visualization of analytics data, it's worth noting the essential role that clear, impactful data visualization plays in the realm of LA. While collecting and analyzing comprehensive data is the foundation of any analytics tool (Sit et al., 2021; Xu et al., 2019), the true utility of such tools lies in their ability to translate this complex data into a form that is both digestible and actionable visual representations (Alabbad et al., 2022). With educators as the primary end-users of these insights, it becomes paramount to present the data in a manner that readily aids in decision-making, enhances understanding, and improves the learning experience for students. Presented in Fig. 7 is the comprehensive instructor's dashboard, designed to provide a broad or focused view of student LA. Instructors can conveniently navigate through this interface, opting to delve into the learning patterns of a specific student or obtain an overarching understanding of the class's performance. This flexibility equips educators with a valuable tool to guide instructional decision-making based on individual or group LA.

4.2 Case Study Results: Faculty Feedback on the LA Tool

A case study was conducted with the teaching faculty at the University of Iowa to gather feedback on the real-world applicability of the LA tool. An anonymous survey was administered to evaluate the tool's effectiveness, ease of integration, and concerns around privacy and data security. Since the survey questions were optional, there was a slight discrepancy in the number of responses for each question, with 1–2 responses varying.

As shown in Fig. 8, respondents came from a diverse range of disciplines, with a significant proportion (38%) representing STEM fields such as Math, Science, and Engineering. This group's familiarity with technology influenced their feedback, emphasizing accuracy and reliability. Other respondents came from Professional Studies (14%), Humanities (11%), Social Sciences (8%), and Arts (2%), reflecting a range of pedagogical approaches. Notably, 27% of the respondents selected "Other," representing specialized fields like Nursing, Health Sciences, Medicine, Pharmacy, and Education, indicating that faculty in healthcare-

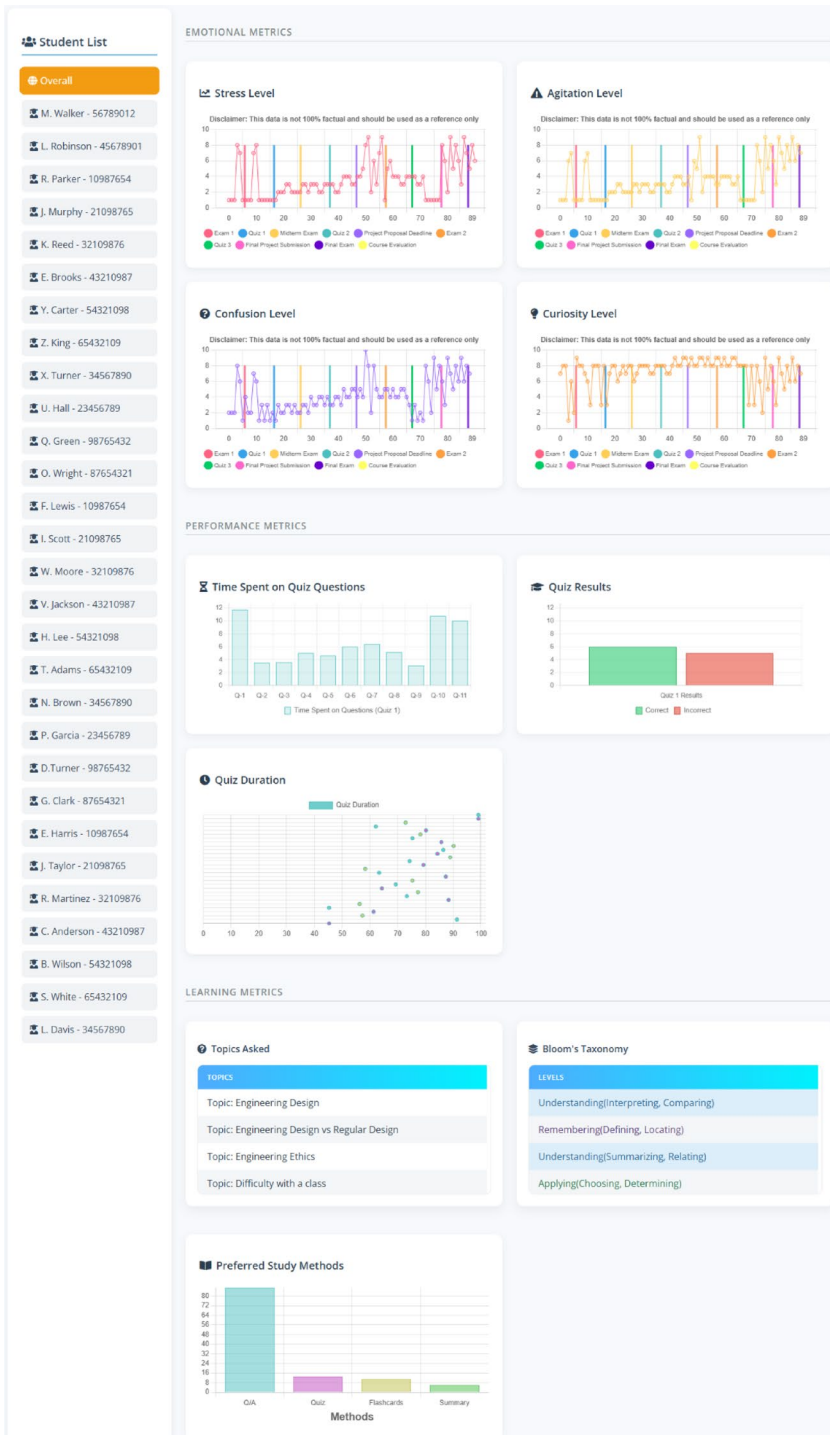**Fig. 6** Student query progression using Bloom's taxonomy

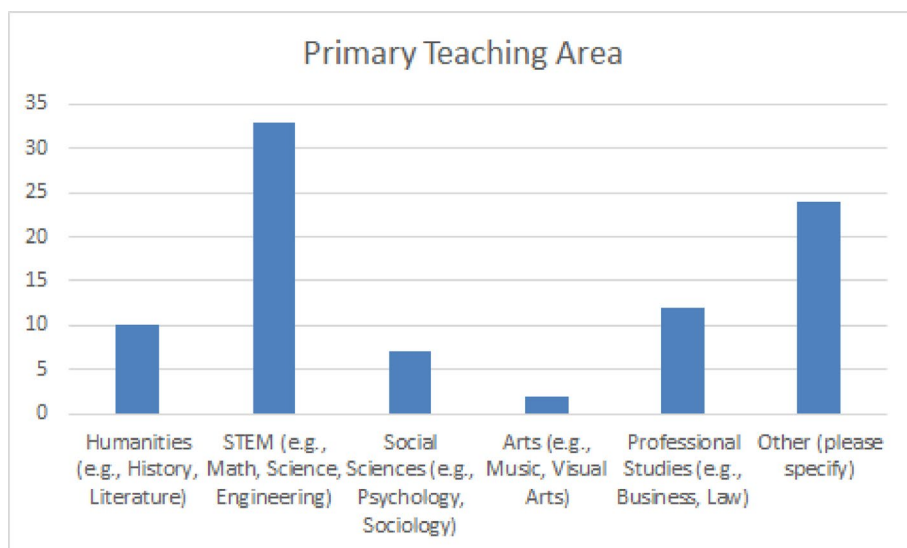Fig. 7 Instructor view of learning analytics dashboard—proof of concept with synthetic data

Fig. 8 Primary teaching area of the faculty (n = 88)

related disciplines may have different expectations for AI tools due to the sensitive nature of their subject matter.

As shown in Fig. 9, regarding roles, Assistant Professors (30%), Professors (22%), and Associate Professors (22%) made up most respondents, bringing seasoned perspectives based on long-term teaching experience. Graduate Assistants (7%) and Lecturers (3%) also contributed feedback, highlighting needs around ease of use and integration, while the 15% of the respondents selecting "Other" included adjunct instructors and postdoctoral researchers, adding further diversity to the perspectives on AI use in education.

4.2.1 Importance of Classroom Insights

The data in Fig. 10 highlights the key insights faculty find important for understanding student engagement and learning in the classroom. The highest priority for faculty was identifying frequently asked questions (80%) and tracking the progression of student learning over time (75%). These insights suggest that instructors value understanding both common areas of confusion and student learning trajectories, enabling them to tailor their teaching more effectively. Additionally, identifying common misconceptions (73%) was also a significant priority, further emphasizing the need for tools that help instructors address knowledge gaps.

Understanding students' preferred study methods (43%) also plays a role, though it is not as highly prioritized as other metrics. This suggests that while understanding how students approach their studies is important, it is secondary to more direct measures of learning progression and common misunderstandings. Less emphasis was placed on identifying student engagement with online tools (34%) and tracking timelines such as milestones (42%), indicating that while useful, these insights are not viewed as essential compared to other metrics.
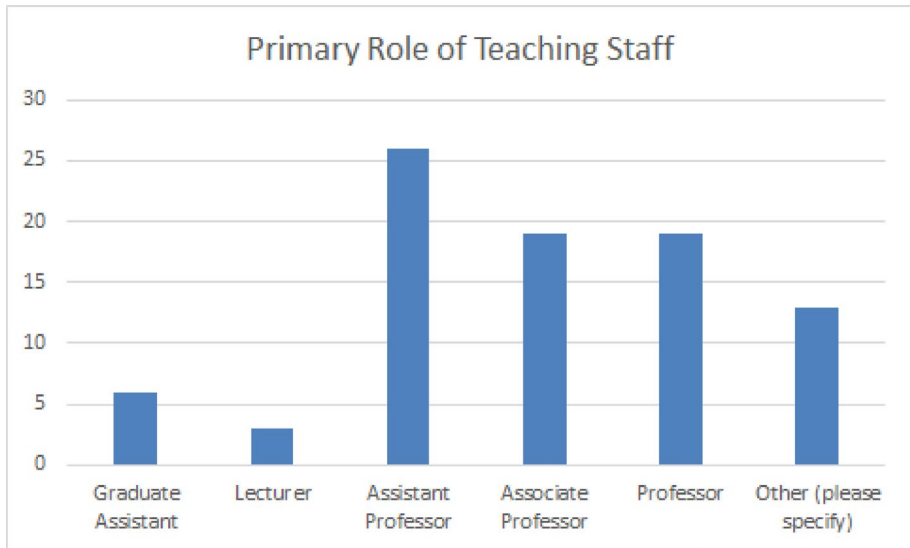
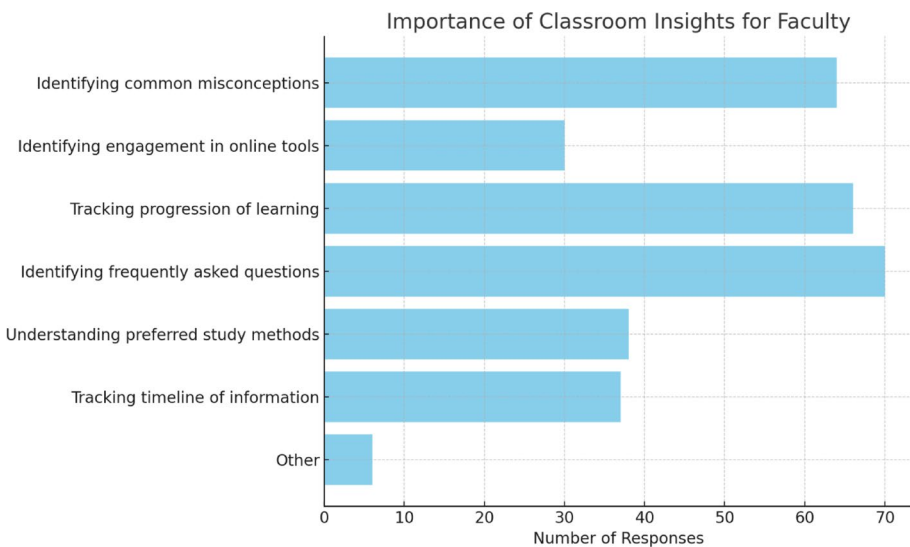**Fig. 9** Primary role of the teaching staff (n=86)

Fig. 10 Importance of classroom insights (n=88)

To further analyze faculty responses, we examined differences across disciplines. STEM faculty showed a stronger preference for identifying misconceptions (81%) and tracking student progression (79%), due to the structured, hierarchical nature of STEM learning. In contrast, the Humanities and Social Sciences faculty placed slightly greater emphasis on preferred study methods (49%), possibly reflecting a more diverse range of learning strategies within these fields. Assistant Professors and Graduate Assistants, who often juggle

multiple responsibilities, focused on ease of use and time-saving features, highlighting the need for tools that integrate seamlessly into their workflows.

Overall, the data shows that faculty prioritize insights that provide actionable information on student learning and areas of difficulty, which can directly inform their teaching practices.

4.2.2 Faculty Concerns: Accuracy of Insights vs. Privacy

Figure 11 presents faculty concerns regarding two critical factors influencing AI adoption: student privacy and data security, and the accuracy of AI-generated insights. Both concerns were rated highly, with 51% of faculty expressing serious concerns about privacy and 59% indicating strong concerns about the accuracy of AI-driven insights.

Concerning privacy stems primarily from uncertainties about how student data is stored, processed, and protected. This was particularly pronounced among faculty from healthcare-related disciplines, where ethical considerations regarding data confidentiality are stringent. Addressing these concerns requires institutions to implement transparent privacy policies, robust encryption methods, and clear guidelines on data handling.

Similarly, concerns about accuracy highlight skepticism regarding AI's reliability in assessing student learning. Many faculty members hesitated to rely on AI-generated insights without validation mechanisms, emphasizing the need for human oversight in interpreting data. Notably, the STEM faculty displayed higher levels of concern regarding accuracy (64%), due to the precise nature of their subjects, while Social Sciences and Humanities faculty showed marginally lower concern (53%), reflecting a greater tolerance for interpretive variability.

These concerns underscore the necessity for continued refinement of AI tools, ensuring that insights are both accurate and ethically handled. Institutional support, including faculty
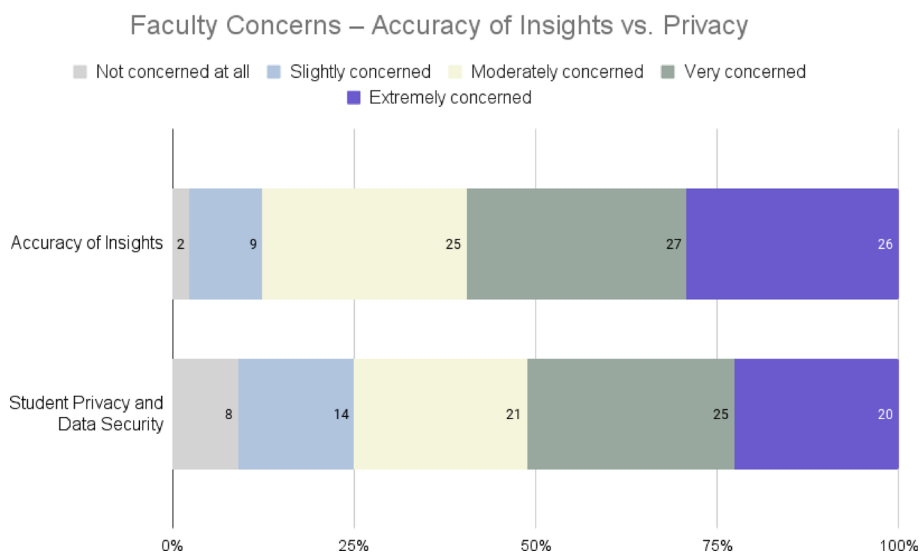

Fig. 11 Faculty concerns—accuracy of insights vs. privacy (n=89 & n=88)

training and transparent communication on AI's limitations, will be critical for fostering trust in these systems.

4.2.3 Additional Feedback and Suggestions

The open-ended feedback provided by participants revealed a cautious yet optimistic stance toward the integration of AI in education. Many participants saw potential benefits, particularly in saving time for students and aiding in deeper study processes. For example, AI could help students approach problems more thoughtfully, much like guidance from teaching assistants during office hours, rather than simply providing answers. This perspective highlights a desire to use AI in a way that enhances learning without replacing critical thinking.

However, concerns about the potential negative impacts of AI on student learning also surfaced. Some participants worry that over-reliance on AI could stifle critical thinking, emphasizing the need for students to engage in reasoning and problem-solving independently. Additionally, the fear that AI could be a barrier to developing skills in identifying misinformation was raised, suggesting a need for AI tools to focus on fostering students' ability to discern and reason with reliable information.

Data privacy and security were recurring concerns. Several participants questioned how AI handles sensitive student data, drawing attention to the necessity of transparent privacy policies and independent evaluations of AI tools. Some have referenced recent high-profile privacy violations in technology sectors, underscoring the importance of ethical data management in educational tools.

The comments also indicate practical concerns that could affect the tool's adoption. Many educators feel pressed for time and would appreciate easier ways to incorporate AI into their teaching without needing to completely overhaul their courses. Additionally, some faculty are waiting for AI technology to mature, reflecting skepticism toward current systems like ChatGPT due to reliability issues.

Participants also offered suggestions for improvement. Many recommend providing clear documentation and training on how the tool works, its limitations, and best practices for use. Ethical considerations were a recurring theme, with faculty emphasizing the importance of transparent privacy policies and independent evaluations of the tool's impact on student learning. These insights highlight the need for a user-centered approach to AI tool development, ensuring that the tools meet the practical and ethical needs of educators and students alike.

4.3 Discussions

By comparing the results from both synthetic data and faculty feedback, several key insights emerge. The synthetic data highlights the tool's technical capabilities, while faculty feedback emphasizes real-world applicability and concerns, particularly around data security and the accuracy of insights. One of the key features of the tool is its ability to measure both active and passive participation, providing nuanced insights into students' emotional states and learning progression. These capabilities enhance educators' ability to make informed, timely interventions, improving student engagement and learning outcomes. The integration of AI (GPT-4) further enriches this approach by allowing real-time sentiment analysis and

topic identification, extending the tool's adaptability to various educational contexts and LMS platforms.

Both the synthetic data and faculty feedback underscored the value of real-time emotional analysis in improving student engagement. Faculty recognized the potential of such insights to tailor interventions and enhance classroom dynamics. However, while the tool's emotional analysis capabilities are promising, they are inherently subjective and susceptible to biases present in AI-driven sentiment detection. Emotional expression varies by cultural and individual factors, which GPT-4 may not fully account for, leading to misinterpretations. Faculty feedback also highlighted that these features may not fully capture the complexities of human emotions and external factors affecting engagement. This suggests the need for further refinement to ensure the tool delivers more accurate and comprehensive emotional insights. Future iterations should consider incorporating human validation of sentiment classifications to enhance reliability. Additionally, research should explore whether the model's sentiment detection aligns with traditional psychological assessments of student emotions.

The faculty expressed strong concerns around the accuracy of AI-generated insights and the security of student data. These concerns align with the synthetic data results, which demonstrated the tool's ability to track learning patterns and engagement but also revealed limitations in consistency and reliability. Faculty feedback emphasized that accuracy and data security are critical factors for real-world applicability, particularly in fields like healthcare and STEM, where precision is essential. The real-world concerns highlight the need for more robust privacy measures, consistent accuracy, and transparency in how data is handled. Compliance with regulations such as FERPA and GDPR will be essential in gaining faculty trust. Additionally, independent evaluations and audits of the tool's security and accuracy could strengthen its credibility.

A major limitation of the tool is its reliance solely on Bloom's Taxonomy for classifying student inquiries. Bloom's Taxonomy was originally developed as a framework for instructional design rather than for automated text analysis. The study does not rigorously assess whether AI-generated classifications align with human evaluations of cognitive complexity. While the tool categorizes student inquiries based on Bloom's digital planning verbs, it is not validated by human coders, meaning there is no manual verification to ensure AI-generated classifications align with human evaluations of cognitive complexity. Future studies should incorporate human-coded validation to assess the reliability of AI-driven cognitive classification. Furthermore, the use of Bloom's Taxonomy alone may not fully capture the complexities of student learning. Alternative models, such as Self-Regulated Learning (SRL) models, Activity Theory, or Cognitive Load Theory, could provide a more nuanced understanding of student engagement and progression. Expanding the tool to integrate these frameworks would offer a more comprehensive perspective on learning behaviors beyond hierarchical cognitive categories.

Additionally, the tool does not currently distinguish between lower-order and higher-order cognitive processes beyond keyword mapping. While Bloom's Taxonomy provides a structured framework, AI's ability to infer deep cognitive engagement from textual data remains an open question. Without manual validation, there is a risk that the classification of student inquiries may misrepresent actual cognitive complexity. Future research should explore whether AI-driven categorization aligns with traditional human assessment of cognitive skills and how discrepancies can be mitigated.

The tool's seamless integration with existing LMS platforms like Canvas was well received, but faculty emphasized the need for more intuitive data visualization to ensure ease of use. This feedback points to the necessity of refining the interface and reporting capabilities to better match the practical demands of educators, especially those who may be less familiar with data-driven tools. Additionally, the success of the tool depends on the digital literacy of both educators and students, underscoring the importance of comprehensive training to ensure effective use and interpretation of the generated data.

It is important to note that the dataset may be biased toward a more positive outlook on such technologies due to the overrepresentation of STEM faculty in the survey responses. STEM disciplines often have a higher familiarity with technological innovations, which may result in greater comfort and optimism about AI tools compared to fields like the arts, which were less represented in the sample. This skew in representation could mean that the perspectives of faculty in disciplines with less reliance on data-driven technologies, such as the humanities and arts, are underrepresented, potentially affecting the generalizability of the feedback. Future studies should aim for a more balanced disciplinary distribution to capture a wider range of perspectives.

A key limitation of this study is the absence of classroom experimentation. While the faculty survey provided valuable feedback on the potential applications of the LA tool, further evaluation is necessary in real classroom settings to fully understand its impact on student learning and engagement. Classroom experiments would offer critical insights into how effectively the tool integrates with teaching workflows, its practical benefits for students, and any unforeseen challenges that may arise during active use. Conducting controlled trials comparing classes that use the LA tool with those that do not provide empirical evidence on its pedagogical impact. This would substantiate claims regarding its ability to enhance learning outcomes through actionable insights.

Future research should address these limitations by incorporating human validation of AI classifications, testing alternative educational frameworks beyond Bloom's Taxonomy, and conducting real-world trials to assess the tool's practical effectiveness. By refining these aspects, the LA tool can become a more reliable, ethical, and pedagogically sound instrument for enhancing student learning.

5 Conclusion and Future Directions

This research developed and deployed a novel LA tool, offering significant advancements in educational technology. The tool integrates with VirtualTA, capturing and analyzing student interaction data to provide deeper insights into engagement, performance, and learning patterns. Unlike traditional learning management systems, this tool enhances pedagogical decision-making by incorporating AI-driven sentiment analysis, topic tracking, and cognitive classification based on Bloom's Taxonomy. A key innovation is its ability to measure both active and passive participation, offering a holistic assessment of student engagement. Additionally, it automates the mapping of student inquiries onto cognitive levels, enabling structured tracking of learning progression. These contributions demonstrate how AI-powered tools can extend beyond standard metrics to deliver meaningful, actionable insights.

This study builds on prior research in learning analytics and AI-driven educational tools by addressing existing gaps. While AI has been used to predict student performance

and engagement, its integration with pedagogical frameworks has been limited. This tool bridges that gap by aligning AI-driven insights with established educational theories. Unlike prior LA tools that focus on grades and attendance, this tool employs real-time sentiment analysis and cognitive classification to provide a nuanced understanding of student learning. By integrating these techniques, this research contributes to AI-enhanced education by demonstrating how data-driven insights can translate into actionable pedagogical interventions.

The tool is designed to provide real-time, personalized feedback and support for students while equipping educators with insights into learning progression and affective states. By extending traditional analytics with AI-driven sentiment detection and cognitive classification, it enables early identification of struggling students, optimization of instructional methods, and the creation of inclusive learning environments. However, its effectiveness remains to be validated in real-world classroom settings. The lack of empirical classroom-based experimentation limits the assessment of its impact on teaching workflows. Additionally, the exclusive reliance on Bloom's Taxonomy presents a constraint, as it does not incorporate alternative models of learning and engagement. Future research should explore frameworks such as SRL models, Activity Theory, or Cognitive Load Theory to provide a more comprehensive perspective on student behaviors and engagement. Further refinement of emotional analysis and sentiment detection is necessary. While the tool offers valuable insights into student emotions, AI's limitations in interpreting nuanced expressions remain a challenge. Future iterations should integrate multimodal analysis using voice tone, facial recognition, or physiological indicators to enhance emotional assessment. Cross-validation with human-labeled sentiment datasets will also help reduce biases in AI-driven detection, ensuring reliability and pedagogical relevance.

To ensure practical integration into educational settings, the tool must be embedded seamlessly within existing LMS platforms and instructional workflows. Faculty feedback emphasized the importance of ease of use, minimal disruption to teaching routines, and compatibility with familiar systems such as Canvas or Moodle. The tool's API-based architecture supports such integration by enabling real-time data exchange between LMS environments and the analytics dashboard. For classroom use, the tool could be positioned as an optional analytics companion for instructors, providing immediate visibility into student emotional states and cognitive progression without requiring significant course redesign. To support adoption, institutions should develop onboarding programs, including training modules that explain the interpretation of analytics outputs, the boundaries of AI-driven insights, and ethical data usage. Additionally, the tool's real-time alerts and engagement summaries could be integrated into early-warning systems, helping academic advisors and teaching assistants proactively support students at risk of disengagement. On a broader institutional level, aggregated insights from course-level analytics could inform curriculum development, equity interventions, and program assessment. Ensuring that the dashboard offers role-based access and customizable views will be essential to accommodate the needs of faculty, instructional designers, advisors, and administrators alike. By aligning with existing pedagogical processes and institutional priorities, the tool can serve as a scalable and sustainable component of data-informed education.

Classroom-based experimentation is crucial for evaluating the tool's effectiveness in improving engagement and learning outcomes. A mixed-methods approach combining quantitative performance metrics with qualitative feedback from educators and students will provide insights into how well the tool integrates into teaching workflows. These experi-

ments should also determine whether real-time analytics lead to meaningful pedagogical interventions and improved student learning trajectories. Comparative studies with existing LA tools will validate its effectiveness relative to other approaches.

Given concerns about the reliance on Bloom's Taxonomy, alternative frameworks should be explored. Implementing SRL models or Cognitive Load Theory could provide a richer understanding of how students engage with learning material beyond hierarchical cognitive classifications. The tool currently lacks the ability to distinguish between lower-order and higher-order cognitive processes beyond keyword mapping, raising concerns about classification accuracy. Future research should examine whether AI-driven categorization aligns with human assessments and address discrepancies through comparative studies.

Privacy and data security remain critical concerns. Faculty feedback highlighted the need for stronger anonymization techniques, local data processing options, and increased transparency in data usage policies. Compliance with FERPA, GDPR, and other regulations will be essential for fostering trust and ensuring ethical AI deployment. Future research should explore privacy-preserving AI approaches such as federated learning to enable data analysis while minimizing privacy risks.

User interface and data visualization enhancements are also necessary. Educators emphasized the need for intuitive dashboards with real-time alerts and adaptive visualization techniques to simplify data interpretation. Providing customizable reporting features will improve usability across diverse teaching environments, ensuring that insights are easily actionable.

Expanding the tool's applications to special education remains an important future direction. Investigating personalized learning pathways, differentiated assessment strategies, and assistive AI-driven interventions will broaden its applicability. Research has shown that adaptive learning models enhance outcomes for special education students, and incorporating similar mechanisms into this tool could extend its utility to diverse educational contexts. Emerging technologies such as Virtual Reality (VR) and Augmented Reality (AR) could further enrich learning analytics. These technologies could enable immersive assessments, offering deeper insights into student engagement and cognitive load. Future iterations may integrate VR-based interaction tracking to analyze student participation in virtual environments.

In summary, this research establishes a foundational framework for AI-enhanced learning analytics while highlighting the need for continued refinement and validation. Addressing identified limitations and expanding capabilities will enable the tool to evolve into a more robust and ethical solution for personalized education. By integrating insights from prior research and aligning with pedagogical frameworks, this study contributes to bridging the gap between AI development and educational practice. Interdisciplinary collaboration among AI researchers, educators, and policymakers remains crucial to ensuring AI-driven educational tools are effective and pedagogically sound.

Appendix

See Tables 1 and 2.

Table 1 Overview of key metrics and indicators

Metric	Description
Engagement	Engagement metrics encompass both active and passive aspects. Active engagement may be measured by the number of questions posed to the chatbot, the frequency of flashcard or quiz functionalities usage, and the total time spent interacting with learning materials. Passive engagement can be gauged by login frequency, session duration, and metrics as simple as page views (Yang, 2021)
Stress	Stress is inferred from certain proxy indicators, including the frequency of negative expressions and specific keywords hinting at factors such as increased workload, assignment anxiety, or concerns about exam performance in chatbot interactions (Thukral et al., 2020). It's not a direct measure but can offer valuable insights
Confusion	Confusion can be identified by the frequency of repeated questions asked to the chatbot or the number of questions on a specific topic (Saqr et al., 2022), suggesting a difficulty in understanding that subject
Curiosity	Curiosity could be measured by the variety of topics that a student inquiry about or the number of exploration questions (IBM, 2021) they ask that extend beyond the prescribed syllabus
Learning progression	Utilizing Bloom's taxonomy (Bloom & Krathwohl, 2020), we assess the depth and sophistication of student queries, allowing us to discern shifts from foundational knowledge inquiries to analytical or evaluative questions
Agitation	Identifying agitation involves tracking certain behaviors that could indicate heightened emotional states. Indicators might include the frequency of negative expressions and rapidly switching topics (IBM, 2021). These are not direct measures of agitation but can serve as proxies

Table 2 List of digital planning verbs

Bloom's categories	Verbs
Remembering	Copying, defining, finding, locating, quoting, listening, googling, repeating, retrieving, outlining, highlighting, memorizing, networking, searching, identifying, selecting, tabulating, duplicating, matching, bookmarking, bullet-pointing
Understanding	Annotating, tweeting, associating, tagging, summarizing, relating, categorizing, paraphrasing, predicting, comparing, contrasting, commenting, journaling, interpreting, grouping, inferring, estimating, extending, gathering, exemplifying, expressing
Applying	Acting out, articulate, reenact, loading, choosing, determining, displaying, judging, executing, examining, implementing, sketching, experimenting, hacking, interviewing, painting, preparing, playing, integrating, presenting, charting
Analyzing	Calculating, categorizing, breaking down, correlating, deconstructing, linking, mashing, mind-mapping, organizing, appraising, advertising, dividing, deducting, distinguishing, illustrating, questioning, structuring, integrating, attributing, estimating, explaining
Evaluating	Arguing, validating, testing, scoring, assessing, criticizing, commenting, debating, defending, detecting, experimenting, grading, hypothesizing, measuring, moderating, probing, predicting, posting, rating, reflecting, reviewing, editorializing
Creating	Blogging, building, animating, adapting, collaborating, composing, directing, devising, podcasting, wiki building, writing, filming, programming, simulating, role playing, solving, mixing, facilitating, managing, negotiating, leading

Declarations

Conflict of interest The authors declare that they have no conflicting or competing interests.

References

Alabbad, Y., Yildirim, E., & Demir, I. (2022). Flood mitigation data analytics and decision support framework: Iowa Middle Cedar watershed case study. *Science of the Total Environment, 814*, 152768.

Alam, A. (2021). Should robots aareplace teachers? Mobilisation of AI and learning analytics in education. In *2021 International Conference on Advances in Computing, Communication, and Control (ICAC3)* (pp. 1–12). IEEE.

Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning, 20*(2), 13–29.

Baker, R. S. & Inventado, P. S. (2014). Educational Data Mining and Learning Analytics. In A. J. Larusson & B. White (Eds.), Learning Analytics: From research to practice (pp. 61–75). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-3305-7_4

Bhardwaj, P., Gupta, P. K., Panwar, H., Siddiqui, M. K., Morales-Menendez, R., & Bhaik, A. (2021). Application of deep learning on student engagement in e-learning environments. *Computers & Electrical Engineering, 93*, 107277.

Bloom, B. S., & Krathwohl, D. R. (2020). *Taxonomy of educational objectives: The classification of educational goals. Book 1, Cognitive Domain*. Longman.

Bowden, J. L. H., Tickle, L., & Naumann, K. (2021). The four pillars of tertiary student engagement and success: A holistic measurement approach. *Studies in Higher Education, 46*(6), 1207–1224.

Casalino, G., Grilli, L., Limone, P., Santoro, D., & Schicchi, D. (2021). Deep learning for knowledge tracing in learning analytics: an overview. *TeleXbe*.

Churches, A. (2010). Bloom's digital taxonomy.

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education, 18*(6), 683–695. https://doi.org/10.1080/13562517.2013.827653

Dabić, T. (2016). Bloom's taxonomy revisited in the context of online tools. In *Sinteza 2016-International Scientific Conference on ICT and E-Business Related Research* (pp. 315–320). Singidunum University.

Dawson, S., Joksimović, S., Poquet, O., & Siemens, G. (2019). Increasing the impact of learning analytics. In Proceedings of the Ninth International Conference on Learning Analytics & Knowledge (LAK'19). New York, NY, USA: ACM. https://doi.org/10.1145/3303772.3303784

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learning: Research and Practice, 3*(1), 63–78. https://doi.org/10.1080/23735082.2017.1286142

Goldberg, P., Sümer, Ö., Stürmer, K., Wagner, W., Göllner, R., Gerjets, P., Kasneci, E., & Trautwein, U. (2021). Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction. *Educational Psychology Review, 33*, 27–49.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*, 75–105.

IBM. (2021). Tone analysis. IBM Cloud Documentation. https://cloud.ibm.com/docs/natural-language-understanding?topic=natural-language-understanding-tone_analytics.

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology, 50*(6), 2824–2838.

Nistor, N., & Hernández-García, Á. (2018). What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior, 89*, 335–338.

Nkomo, L. M., Daniel, B. K., & Butson, R. J. (2021). Synthesis of student engagement with digital technologies: A systematic review of literature. *International Journal of Educational Technology in Higher Education, 18*, 1–26.

OpenAI. (2023). GPT-4 Technical Report. *arXiv E-Prints*, https://doi.org/10.48550/arXiv.2303.08774

Ouhaichi, H., Spikol, D., & Vogel, B. (2023). Research trends in multimodal learning analytics: A systematic mapping study. *Computers and Education: Artificial Intelligence*. https://doi.org/10.1016/j.caeai.2023.100136

Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education, 20*(1), 1–23.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77.

Pursnani, V., Sermet, Y., & Demir, I. (2023). Performance of ChatGPT on the US Fundamentals of Engineering Exam: Comprehensive Assessment of Proficiency and Potential Implications for Professional Environmental Engineering Practice. *arXiv preprint* arXiv:2304.12198.

Reimann, P. (2016). Connecting learning analytics with learning research: The role of design-based research. *Learning: Research and Practice, 2*(2), 130–142. https://doi.org/10.1080/23735082.2016.1210198

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2023b). Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education. *arXiv preprint* arXiv:2309.10892.

Sajja, R., Sermet, Y., Cwiertny, D., & Demir, I. (2023b). Platform-independent and curriculum- oriented intelligent assistant for higher education. *International Journal of Educational Technology in Higher Education, 20*, 42. https://doi.org/10.1186/s41239-023-00412-7

Sajja, R., Sermet, Y., & Demir, I. (2024). End-to-end deployment of the educational AI hub for personalized learning and engagement: A case study on environmental science education.

Saqr, M., Jovanovic, J., Viberg, O., & Gašević, D. (2022). Is there order in the mess? A single paper meta-analysis approach to identification of predictors of success in learning analytics. *Studies in Higher Education, 47*(12), 2370–2391.

Sermet, Y., & Demir, I. (2018). An intelligent system on knowledge generation and communication about flooding. *Environmental Modelling & Software, 108*, 51–60.

Sermet, Y., & Demir, I. (2019). A Generalized Web Component for Domain-Independent Smart Assistants. *arXiv preprint* arXiv:1909.02507.

Sermet, Y., & Demir, I. (2020). Virtual and augmented reality applications for environmental science education and training. In *New Perspectives on Virtual and Augmented Reality: Finding New Ways to Teach in a Transformed Learning Environment*, (pp. 261–275).

Sermet, Y., & Demir, I. (2021). A semantic web framework for automated smart assistants: A case study for public health. *Big Data and Cognitive Computing, 5*(4), 57.

Siemens, G. (2013). Learning analytics: The emergence of discipline. *American Behavioral Scientist, 57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Sit, M., Langel, R. J., Thompson, D., Cwiertny, D. M., & Demir, I. (2021). Web-based data analytics framework for well forecasting and groundwater quality. *Science of the Total Environment, 761*, 144121.

Smith, P., Powers, E., Anderson, H., Dellinger, J. T., & Siemens, G. (2022). Broadening the NLP Agenda in Learning Analytics for the Age of AI. *Companion Proceedings of the 12th*, 182.

Thukral, S., Sangwan, S., Chatterjee, A., & Dey, L. (2020). Identifying pandemic-related stress factors from social-media posts--effects on students and young-adults. *arXiv preprint* arXiv:2012.00333.

Xu, H., Demir, I., Koylu, C., & Muste, M. (2019). A web-based geovisual analytics platform for identifying potential contributors to culvert sedimentation. *Science of the Total Environment, 692*, 806–817.

Yang, Y. (2021). Does greater engagement in online general education courses lead to better academic performance? Evidence from Chinese university students. *Open Journal of Social Sciences, 9*(6), 298–315.

Authors and Affiliations

Ramteja Sajja[1] · **Yusuf Sermet[1]** · **David Cwiertny[2,3,4,5]** · **Ibrahim Demir[1,6]**

✉ Ramteja Sajja
   rsajja@tulane.edu

   Yusuf Sermet
   ysermet@tulane.edu

David Cwiertny
david-cwiertny@uiowa.edu

Ibrahim Demir
idemir@tulane.edu

1   ByWater Institute, Tulane University, New Orleans, LA 70118, USA

2   Department of Civil and Environmental Engineering, University of Iowa, Iowa City, USA

3   IIHR Hydroscience and Engineering, University of Iowa, 300 S. Riverside Dr, Iowa City, IA 52246, USA

4   Department of Chemistry, University of Iowa, Iowa City, USA

5   Center for Health Effects of Environmental Contamination, University of Iowa, Iowa City, USA

6   River-Coastal Science and Engineering, Tulane University, New Orleans, LA 70118, USA