EAB
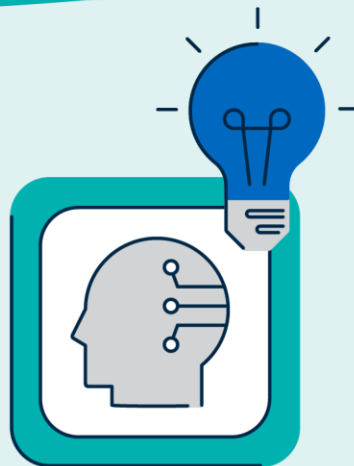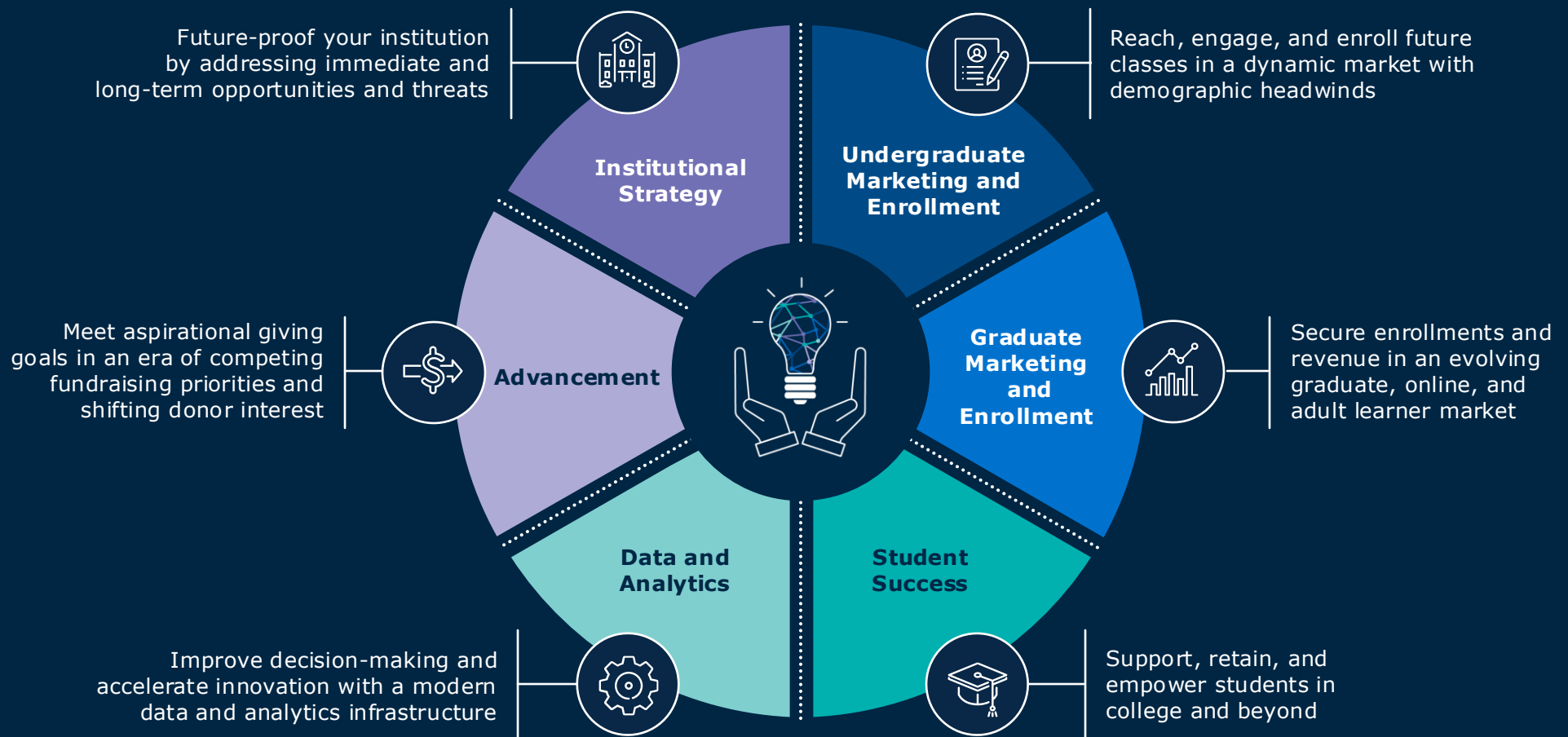
Compendium of AI Applications in Higher Education

45+ Case Studies of AI Applications for College and University Campuses

Insight-powered Solutions for Your Top Priorities and Toughest Challenges

Future-proof your institution by addressing immediate and long-term opportunities and threats

Institutional Strategy

Reach, engage, and enroll future classes in a dynamic market with demographic headwinds

Undergraduate Marketing and Enrollment

Meet aspirational giving goals in an era of competing fundraising priorities and shifting donor interest

Advancement

Graduate Marketing and Enrollment

Secure enrollments and revenue in an evolving graduate, online, and adult learner market

Data and Analytics

Student Success

Improve decision-making and accelerate innovation with a modern data and analytics infrastructure

Support, retain, and empower students in college and beyond

We partner with 2,800+ institutions to accelerate progress, deliver results, and enable lasting change.

95%+ of our partners return to us year after year because of results we achieve, together.

Project Contributors

Project Director

Abhilash Panthagani, MPP

Contributing Consultant

Benjamin Zevallos
Callie Gompf-Phillips
Antoinette Waller

Executive Director

Ann Forman Lippens
Afia Tasneem

Table of Contents

How to Use This Resource

As AI's emergence has redefined what is possible in higher education, institutions worldwide are exploring how they can deploy generative AI applications on their campus today.

This compendium serves as a **reference book of generative AI applications** designed to help higher education leaders begin experimenting with and make informed decisions about AI investments. By showcasing a broad spectrum of case studies, from low-cost and easily deployable solutions to those requiring substantial investment, institutions can make informed decisions to pursue AI applications that suit their unique contexts.

EAB sourced case studies through literature review, industry conferences, and research interviews with higher education leaders (e.g., Chief Business Officers, Chief Information Officers, Provosts). This compendium will solely focus on *generative AI* applications.

Compendium Organization

The compendium is structured into three primary sections, categorized by the type of AI application profiled and the depth of each case study:

Section	AI Applications Profiled	Quantity	Location
A Quick Start Guide for Free and Low-Cost Tools	How institutions are using free or low-cost AI tools or to drive team and department-level gains	11	p. 6
Case Studies of AI Applications	AI applications across the spectrum that institutions have either developed internally or licensed from a vendor partner	38	p. 9
In-Depth Exploration of Select AI Applications	Detailed examinations of how institutions developed their own AI applications, covering all phases from development to deployment	7	p. 38

EAB

A Quick Start Guide for Free and Low-Cost Tools

eab.com

A Quick Start Guide for Free and Low-Cost Tools

This section includes a variety of creative ways higher education leaders are using **free or low-cost AI tools** (e.g., ChatGPT, DALL-E) to drive team or department-level productivity gains. These use cases not only drive productivity gains with low to almost nonexistent implementation costs (i.e., by leveraging tools that higher education already has at its disposal), but they also demonstrate how leaders can promote and encourage AI experimentation on campus.

1

Outline and draft strategic plan components

The Executive Vice Chancellor & University Provost at the **City University of New York System Office** uses ChatGPT in her strategic planning process to help craft a compelling narrative and develop clear, consistent themes.

2

Draft policies and policy manuals

The Director of International Services at the **University of Idaho** uses Google Gemini to create initial drafts of policies and to refine their language and tone in final reviews. For example, her team employed Gemini to draft and review an internal policy manual that was approximately 100 pages long, complete with images, diagrams, and data. The manual had been delayed for years due to limited staff capacity.

3

Produce and iterate on product designs

Babson College's Chief Information Officer and faculty worked together to use DALL-E to develop new product designs for students and to continually test ideas throughout the design process.

4

Design art installations

The Chief Information Officer at the **University of Oregon** hosted an AI art competition where IT staff were encouraged to create personalized office decorations using DALL-E 2. The winners' AI-generated art was then showcased throughout the information services building.

The IT department at the **University of Nevada, Las Vegas** used DALL-E 2 to design a winning float for the homecoming parade.

5

Generate project visuals for facilities planning

Texas State University's Chief Business Officer (CBO) leveraged DALL-E to create visual mock-ups of donor names on buildings. The CBO successfully created ten mock-ups based on the needs of the Advancement Office, which were then presented to the university's president along with one from their architect. The president ranked all of the CBO's renderings higher than the architect's.

6

Produce personalized advancement outreach at scale

The Annual Giving Team at the **University of South Carolina** used ChatGPT to generate customized messages for its high-volume annual Giving Day campaign based on past donor behavior and written in the voice and tone of the giving team. They also composed 392 appeals written in the voice of Deans from over 20 departments. The team provided ChatGPT examples of previous Dean's appeals, fund descriptions, and donor history. As a result, they achieved the highest number of Giving Day donors (including 2K first-time donors) in years and saved over two weeks of staff time.

A Quick Start Guide for Free and Low-Cost Tools (cont.)

7

Copyedit and fine-tune enrollment outreach

The admissions team at **Rosemont College** uses ChatGPT-4 to copyedit and fine-tune similar outreach communications. For example, they use ChatGPT-4 to not only catch grammatical errors, but to ensure a uniform "voice" and consistent formatting in outreach. Doing so helps minimize back and forth exchanges with the marketing department. The Director of Admissions also uses ChatGPT-4 to gather insights from across the internet on the most effective times and strategies to communicate with prospective applicants.

8

Create scripts and compositions for media content

The creative marketing team at **Southern New Hampshire University** used ChatGPT-3.5 to write the script and develop the shot selection for a 30-second commercial advertising the institution. The Chief Marketing Officer noted that ChatGPT considerably reduced the time and money spent in production. The script only needed minor modifications and eight out of ten of the shots were in line with what her team would have proposed.

9

Edit marketing images

The Director of Enrollment Marketing and Communications at **Emory University** regularly uses Photoshop's AI-powered Generative Fill to edit marketing collateral. For example, she was able to swiftly remove distracting caution tape from a recruiting photo to use for wide-scale marketing.

10

Craft speeches

The president of **Wells College** used ChatGPT to write parts of his commencement address.

South Dakota State University's marketing coordinator regularly uses ChatGPT to refine and shorten his speeches.

11

Produce social media content

The Associate Director of Student Life and Leadership at the **University of Alaska Anchorage** promotes student events by creating social media content with ChatGPT-4. In particular, he uses ChatGPT-4 to refine the tone and voice of posts, ensuring the output resonates with the intended audience.

The Marketing and Social Media Coordinator for **South Dakota State University** uses generative AI tools to tackle the "copy shop" challenge in higher education marketing. For example, he employs ChatGPT as an editorial assistant for idea generation and to overcome writer's block while creating social media captions.

Source: Gardner, Lee, "Your College's New Marketing Campaign, With a Boost From AI", *The Chronicle of Higher Education*, January 22, 2024; LeBlanc, Paul, "We're Asking the Wrong Questions About AI", *Inside Higher Ed*, March 12, 2023; Moody, Josh, "The ChatGPT Commencement Address", *Inside Higher Ed*, June 29, 2023; Schmitz, Matthew, "AI at work: How some UAA staff are using AI", *The Northern Light*, October 31, 2023; Swaak, Taylor; "Admissions Offices Need More Students and Less 'Drudgery.' Is AI the Answer", *The Chronicle of Higher Education*, November 1, 2023; EAB interviews and analysis.

EAB

Case Studies of AI Applications

SECTION

2

- Teaching and Learning
- Student Success
- General Campus Support
- Multi-Assistant Platform
- Information Technology
- Human Resources

- Procurement
- Admissions
- Enrollment
- Advancement
- Research

Navigating Case Studies of AI Applications

This section includes a **range of case studies** demonstrating how institutions are implementing AI applications across different domains on campus. It features two types of applications:

1. **AI applications developed internally** and/or with vendor partners. These applications may be driven by a campus unit (e.g., IT) or individual institutional stakeholders (e.g., faculty researchers).

2. **Commercially available vendor offerings** licensed by an institution.

Case Study Organization

The case studies are organized by Functional Area (e.g., Admissions, Student Success, Enrollment). For applications that concern multiple higher education functions, the application is categorized by the functional area it impacts the most. For example, a tool that can generate job descriptions would be tagged to HR, even if another unit like IT developed and piloted the tool.

AI applications with multiple examples will be grouped under a shared category to illustrate how institutions are approaching building the same use case . Examples under the same AI Application umbrella will be organized by implementation effort and resource intensity (least to most).

Each case study is also evaluated on four variables:

- **Impact**: How significant are the measurable benefits (e.g., productivity, learning outcomes) of deploying this application on campus?

- **Implementation Effort**: How much institutional effort is required to operationalize the application, from development to deployment?

- **Resource Cost**: What is the cost (e.g., staffing, financial investment) of implementing the application?

- **Status**: Is the application readily available on campus? Is it still in development or being trialed?

The following page provides a more detailed breakdown of variable ratings.

Variable Ratings for Case Studies of AI Applications

*Variable ratings are approximated by EAB experts and based on information available at the time of the case study's publishing. For instance, while the resource costs for developing an AI application similar to one featured in the compendium may decrease over time, the compendium's rating reflects the resource costs incurred to implement the application at that specific point in time.

Impact:

Low	Produces moderate gains that marginally change existing operations (e.g., productivity boost from select faculty automating a component of their math problem grading process).
Medium	Generates sizeable gains that impacts roles and responsibilities on campus (e.g., streamlines a task that department staff used to spend weeks per year on).
High	Fundamentally transforms existing systems (e.g., business workflows, student experience).

Implementation Effort:

Low	Requires standard technical staff expertise and minimal organizational support; can be deployed in a short timeframe (i.e., under a month).
Medium	Requires some sustained technical staff commitment (e.g., a couple of IT staff working part-time) and participation from other organizational units (e.g., reviewing and providing feedback) to operationalize.
High	Necessitates a team of dedicated technical staff and support from other campus units (e.g., vetting and socializing applications); tends to require significant time (>one month) to implement on campus.

Resource Cost:

$	$0 to 2K
$$	$2K to 10K
$$$	$10K to 50K
$$$$	50K and greater

Status:

Development	The application has established viability as a proof-of-concept but has yet to be trialed with a substantial audience.
Trial	The institution and/or vendor is running a limited trial of the application.
Deployed	The application is officially deployed to its target audience.
Discontinued	The application is no longer being developed or deployed.

Teaching and Learning

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Deliver real-time assignment feedback	Provides students instant, personalized feedback, increasing faculty capacity for more valuable coaching and interventions.	**Imperial College London** (ICL) developed Lambda Feedback to provide real-time, personalized feedback on math problems. Lambda provides low-level guidance to help students overcome obstacles without giving away answers. The platform offers feedback the moment problems are submitted, helping students learn at their own pace. Lambda Feedback is a web-based platform that employs parameterized questions to ensure each student receives unique problems, preventing students from sharing answers. Feedback provided for both formative and summative assessments is manually checked. Any changes that need to be made (e.g., if an error is spotted in a question response) are applied algorithmically to the whole cohort. Students can ask questions about the feedback they receive during tutorials throughout term, or in a dedicated tutorial feedback session in January.	Trial (as of June 2024)	ICL faculty and students developed the Lambda Feedback platform in-house and first deployed it in the 2021-22 academic year. In the 2023-24 academic year, Lambda Feedback was deployed in 25+ modules across three departments. ICL is now piloting Lambda Feedback across eight departments and 1000+ students. For the 2024-25 academic year, ICL is planning to expand the platform across the institution. ICL plans to hire eight undergrad students to help adapt content for the platform as well as to help with development and data analytics.	Low	Med.	$$
Deliver real-time assignment feedback	Reduces grading time and ensures consistent feedback, while requesting human intervention when necessary.	PhD students at the **University of Birmingham** created Graide, an AI platform capable of assessing short text, essays, and even math. Graide learns an assessor's grading style to evaluate work and final answers. It also accepts written and digital submissions. Graide claims to reduce grading times by 87% and that students receive 7x feedback. When Graide is unsure of its evaluation, it proactively requests human intervention.	Deployed (as of June 2024)	Oxbridge, a blended learning platform, conducted an independent testing trial of Graide, analyzing 235 essays from English literature classes and comparing their evaluations with the Graide-scored essays. Graide demonstrated over 80% accuracy with the Oxbridge scores. The University of Kent, Aston University, and the University of Sussex all deploy Graide on campus.	Med.	Med.	$$

Source: Haynes, Peter, "How can we utilize AI in higher education?," *Higher Education Policy Institute*, October 11, 2023; University of Birmingham, "New AI platform to mark essays in all subjects – including English," *University of Birmingham*, November 1, 2023; EAB interviews and analysis.

1) Implementation Effort.

eab.com

Teaching and Learning (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Deliver real-time assignment feedback	Streamlines faculty grading and feedback for essays, while providing real-time writing support and evaluation for students.	KEATH.ai, developed by the **University of Surrey** researchers, is a personal AI Assessment Assistant that is designed to grade essays and provide detailed, real-time feedback across many fields and languages. KEATH.ai can accurately grade essays ranging from 1,000 to 15,000 words, claiming a baseline accuracy of 85%. With ongoing training and optimization, accuracy is expected to reach up to 95%. It also offers one-click generation of student performance reports.	Deployed (as of June 2024)	Academics at Surrey and other partners in the education sector tested KEATH.ai on assignments they had previously graded. The team at KEATH.ai plans to work with more faculty for ongoing trials. The team is also working with Surrey's Digital Learning Team for ongoing training and calibration.	Med.	Med.	$$
Score free-response questions for large-scale assessments	Enables instructors to employ more impactful testing mechanisms like free-response questions without creating additional grading burdens.	The **Texas Education Agency** (TEA) adopted an automated scoring system utilizing natural language processing technology similar to ChatGPT-4, for the State of Texas Assessment of Academic Readiness (STAAR). This AI scoring system reviews and scores constructed free-response questions in key subject areas.	Deployed (as of June 2024)	TEA refined their AI scoring system using 3,000 student responses that underwent two rounds of human scoring. Human scorers continue to play a role in ensuring quality control, reviewing scores flagged for low confidence. After initial AI grading, 25% of test responses are sent back to human graders for review. This is part of an effort to modernize the STAAR test and expedite the scoring process, potentially saving $15-20 million annually. The TEA plans to hire fewer than 2,000 graders in 2024, compared to 6,000 hired in 2023.	Med.	Med.	$$

Source: University of Surrey, "AI-driven academic marking tool unveiled – offering potential to enhance student learning outcomes and free up academic time for teaching and research," *University of Surrey*, November 30, 2023 ; Windes, Isaac, "Computers will score most STAAR essays this year, raising concerns at local school districts," *San Antonio Report*, March 22, 2024; Peters, Keaton, "Texas will use computers to grade written answers on this year's STAAR tests," *The Texas Tribune*, April 9, 2024; EAB interviews and analysis.

1) Implementation Effort.

eab.com

Teaching and Learning (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Supply 24/7 personalized real-time tutoring	Provides students' one-on-one academic support at times where previously not possible, enabling instructors to spend more class time on student relationships, hands-on practice, and learning.	The Stephen M. Ross School of Business at the **University of Michigan** launched a virtual Teaching Assistant (TA) pilot program in 2025 to facilitate self-paced learning. Instructors can customize the virtual TA to their specific courses and teaching style, so that it can explain complex course concepts to students and work through difficult problems without directly giving students answers. The virtual TA is powered by Google's Gemini models.	Deployed (as of May 2025)	U-M (along with several other institutions) has piloted the virtual TAs in various business courses (e.g., financial technology, operations strategy, and operations analytics and statistics). Early pilot results show increased student engagement and deeper instructor understanding of student learning patterns. The virtual TA delivers real-time data on student usage and frequently asked questions, enabling instructors to simplify and improve teaching materials. Leaders from Ross plan to incorporate insights from the pilot program into a study co-led with Northwestern University's Kellogg School of Management, involving 9,000 students, 72 courses, and 26 schools.	Med.	Med.	$$$
Supply 24/7 personalized real-time tutoring	Offers course-specific learning support, allowing faculty to devote more class time to student relationships and functions tools are incapable of (e.g., if a student becomes so lost they don't know what question to ask).	Students in **Harvard University** and **Yale University's** CS50 Introduction to Computing and Programming course now have access to a personal, around-the-clock AI teaching assistant dubbed the "CS50 duck". The CS50 duck can work students through assignments they might be stuck with, help them write code more efficiently, and even explain computer science concepts. The CS50 duck also incorporates gamification elements. For example, every time a student asks a question, they lose a heart; if they run out of hearts, the student must wait until they regenerate to ask any more questions.	Deployed (as of June 2024)	Developed by a Harvard computer science professor, the CS50 duck employs Azure and OpenAI tools coupled with the universities' vector databases that house information like lectures. Students can access the CS50 duck through a ChatGPT-like interface on the web, or though the course's Q&A platform. The CS50 duck, inspired by the practice of "rubber duck debugging," is designed to aid students interactively without delivering full solutions, encouraging independent problem-solving. This stands in contrast to tools like ChatGPT which may provide over-assistance.	Med.	Low	$

1) Implementation Effort.

eab.com

Teaching and Learning (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Supply 24/7 personalized real-time tutoring	Provides students' one-on-one academic support at times where previously not possible, enabling instructors to spend more class time on student relationships, hands-on practice, and learning.	Khanmigo is a conversational AI tutor and teaching assistant developed by Khan Academy. It supports real-time, one-on-one tutoring personalized to students' proficiency and learning style across all subjects. For example, Khanmigo can teach concepts, help code applications, provide live writing assistance, serve as a debate partner, or even let students have a conversation with a historical figure. It also supports speech-to-text interaction. Khanmigo can also assist teachers with creating lesson plans, student progress summaries, learning objectives, rubrics, and exit tickets. Teachers can even analyze class performance trends. The **First Avenue School** deployed a pilot of Khanmigo for students in grades five through eight.	Deployed (as of June 2024)	Khanmigo operates on ChatGPT-4 and has been trained on Khan Academy's collection of educational resources, including lesson plans, essays, and sample problems (e.g., Khan Academy's complete library of mathematics problems). The Khanmigo development team spent thousands of hours fine-tuning and prompt-engineering Khanmigo to increase accuracy, narrow its focus, and operate better in a learning environment. The team not only applied stronger moderation filters but trained Khanmigo to avoid giving outright answers to students when they are struggling and to always keep an encouraging tone. They also practiced "red teaming" to try and deliberately find potential vulnerabilities in how the tool responds to certain questions. The development team behind Khanmigo consists of experts in English Language Arts (ELA) and writing instruction. As of April 2024, Khanmigo has been embraced by over 53 school districts, reaching approximately 65,000 students, with plans to expand to 500,000 to one million users by fall 2024. The program covers all subjects. School districts can purchase Khanmigo for $35 per student annually (reduced from $60 in November 2023). Teachers can access Khanmigo for $9 monthly or $99 annually.	Med.	Low	$$$$

1) Implementation Effort.

eab.com

Teaching and Learning (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Provide tailored instructional feedback to educators	Supports consistency and accuracy in supervisor feedback and evaluation processes for students.	Sibme's AI Insights analyzes uploaded lesson recordings (video or audio), instructional materials, and student work to develop personalized, evidence-driven feedback on educators' teaching practice, including timestamped video observations. Coaching feedback analyzes and emphasizes key instructional metrics such as Teacher Talk Time, Student Talk Time, and Question Distribution. Leaders of **Martin Luther College's** graduate and continuing education teacher training programs implemented Sibme's AI Insights to provide teachers with instructional feedback (e.g., where lesson plans miss key objectives, where students were not engaged). Sibme's AI insights and summaries reduce instructional coach time spent on reviewing teacher instructional videos so they can spend more time on higher-level coaching and mentorship.	Deployed (as of Sept. 2025)	Sibme is an instructional coaching software company founded by a former Texas teacher in 2013. They provide technological solutions for coaching, monitoring, and assessing the teaching efficacy of K-12 educators, which higher education leaders can use to enrich their teacher training programs. Developers at Sibme partnered with engineers at AWS and integrating AWS AI technologies (specifically, Amazon Rekognition and Amazon Comprehend) to begin offering AI-powered features in 2024. Sibme's data strategy and policies align with FERPA, COPPA, and GDPR privacy regulations.	Med.	Low	$$
Provide instructors with course material design support	Saves instructor time by streamlining course and instructional materials development directly within the learning management system (e.g., course structure, test questions, assessment tasks).	The **University of Leeds** was one of the first institutions to enable the AI Design Assistant feature in the LMS Blackboard Learn Ultra when it first became available in September 2023. Based on existing content in the course, the AI Design Assistant provides instructors with AI-generated recommendations for adding new modules, generating test questions, developing authentic assessment tasks (based on Bloom's taxonomy), creating grading rubrics, and adding images to course content when prompted. For most content types, users can input topic guidance, desired difficulty level, and complexity, etc. In the first five months of its release, 95% of instructors responding to in-app feedback surveys indicated that the AI Design Assistant was saving them time, and 92% would recommend it to their peers.	Deployed (as of Sept. 2025)	Blackboard Learn Ultra's AI Design Assistant is powered by the Azure OpenAI Service. Anthology (Blackboard Learn's parent company) and Microsoft's product development staff convened 19 clients from their AI Advisory Council, including University of Leeds, to develop and test the functionality of the AI Design Assistant before its release in 2023.	Med.	Low	$

1) Implementation Effort.

eab.com

Student Success

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Provide a 24/7 assistant for all student support-related inquiries	Supplies around-the-clock answers to common queries, improving student support response time and quality of service.	In an attempt to address declines in student mental health and connectedness during the pandemic, the **University of Nevada, Las Vegas** (UNLV) released the Digital President Whitfield to answer thousands of questions across over 500 topics including academic programs, health and wellness, and mentorships. The Digital President lets students ask questions to a virtual avatar of UNLV's president through a chat interface. The chatbot is enhanced with ChatGPT-3.5 to enable more natural conversations.	Discont. (as of June 2025)	UNLV first developed a virtual avatar of their President during the pandemic. In 2023, UNLV gained early access to an enterprise version of ChatGPT-3.5 and integrated it with their Digital President with the help of AI Foundation. The CIO emphasized how integral ChatGPT-3.5 was for improving the efficacy and conversational fluidity of the Digital President. Since they were early movers, licensing for GPT-3.5 cost around $50K. While the technical development and ongoing maintenance of the Digital President are outsourced to the AI Foundation, UNLV's internal IT team was responsible for collecting training materials as well as sourcing feedback on where to implement guardrails from diverse groups on campus, such as the Student Recreation Center, Health Center, and Enrollment Management.	Low	Med.	$$$$
Provide a 24/7 assistant for all student support-related inquiries	Fields and resolves student support inquiries, reducing the demand on human resources and allowing staff to focus on more complex or sensitive student problems.	The **University of Galway** created Cara, an AI-powered student engagement platform that provides 24/7 support to students by answering common questions on topics from bike parking and classroom locations, to mental health resources. Cara leverages information already gathered by student services to free up human support staff capacity. As of October 2023, Cara has answered more than 30,000 queries, resolving 91% of questions without having to route to a human at Galway.	Deployed (as of June 2024)	The head of student engagement projects at the University of Galway partnered with the AI-startup Galvia to develop Cara. She and her team trained Cara on information gathered by student support services. In the future, Galway envisions using analytics to identify engagement trends and roadblocks in the student experience (e.g., if a student is struggling with fees or a particular module).	High	Med.	$$$$

1) Implementation Effort.

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Provide a 24/7 assistant for all student support-related inquiries	Supplies around-the-clock answers to common queries, improving student support response time and quality of service.	In an attempt to address declines in student mental health and connectedness during the pandemic, the **University of Nevada, Las Vegas** (UNLV) released the Digital President Whitfield to answer thousands of questions across over 500 topics including academic programs, health and wellness, and mentorships. The Digital President lets students ask questions to a virtual avatar of UNLV's president through a chat interface. The chatbot is enhanced with ChatGPT-3.5 to enable more natural conversations.	Discont. (as of June 2025)	UNLV first developed a virtual avatar of their President during the pandemic. In 2023, UNLV gained early access to an enterprise version of ChatGPT-3.5 and integrated it with their Digital President with the help of AI Foundation. The CIO emphasized how integral ChatGPT-3.5 was for improving the efficacy and conversational fluidity of the Digital President. Since they were early movers, licensing for GPT-3.5 cost around $50K. While the technical development and ongoing maintenance of the Digital President are outsourced to the AI Foundation, UNLV's internal IT team was responsible for collecting training materials as well as sourcing feedback on where to implement guardrails from diverse groups on campus, such as the Student Recreation Center, Health Center, and Enrollment Management.	Low	Med.	$$$$
Provide a 24/7 assistant for all student support-related inquiries	Fields and resolves student support inquiries, reducing the demand on human resources and allowing staff to focus on more complex or sensitive student problems.	The **University of Galway** created Cara, an AI-powered student engagement platform that provides 24/7 support to students by answering common questions on topics from bike parking and classroom locations, to mental health resources. Cara leverages information already gathered by student services to free up human support staff capacity. As of October 2023, Cara has answered more than 30,000 queries, resolving 91% of questions without having to route to a human at Galway.	Deployed (as of June 2024)	The head of student engagement projects at the University of Galway partnered with the AI-startup Galvia to develop Cara. She and her team trained Cara on information gathered by student support services. In the future, Galway envisions using analytics to identify engagement trends and roadblocks in the student experience (e.g., if a student is struggling with fees or a particular module).	High	Med.	$$$$

1) Implementation Effort.

eab.com

Student Success (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Offer a one-stop assistant for students' everyday tasks	Delivers personalized academic and nonacademic support to students and parents, improving engagement and educational outcomes.	**Los Angeles Unified School District (LAUSD)**, in partnership with AllHere, launched "Ed," an individualized AI personal assistant for students' everyday tasks. Ed offers timely support and guidance (e.g., upcoming test reminders, bus schedules), academic progress monitoring (e.g., grade results, attendance, graduate readiness), and personalized learning activities (e.g., individualized learning plans, book recommendations). As students progress, Ed's voice and responses will evolve with the student, and he will take on more of a teacher or mentor role.	Discont. (as of Sept. 2024)	LAUSD entered into a five-year, $6 million contract with AllHere to develop Ed. Ed is connected to multiple student information sources, from state test scores to attendance and disciplinary records. Rigorous safety protocols are also in place, including content filters and age restrictions, to ensure that interactions are appropriate and constructive. Ed boasts a 93% accuracy rate and only responds to inquiries within the district university of information. In the future, local, state and federal agencies will also have a role in monitoring the AI tool's growth. Ed is available in 100 languages and currently serves approximately 55,000 of the districts' most vulnerable students. Soon, all 540,000 students in the district and their parents or guardians will have access to Ed.	High	High	$$$$

1) Implementation Effort.

Source: Gifford, Aaron, "LAUSD Rolls Out Proprietary AI Assistant for Students, Parents," *Government Technology*, March 21, 2024; EAB interviews and analysis.

eab.com

General Campus Support

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Create a general campus information aide for staff, faculty, and student support	Provides a safe alternative to public tools and helps campus users access institutional information and resolve their first-line inquiries.	**The University of Tennessee Knoxville** (UTK) deployed UT Verse, a chat-based AI platform that can answer questions on general institutional information as well as university business and research for internal campus audiences. It is designed as a campus alternative to ChatGPT and configured to prioritize UT-related queries (related to UT-specific research or data). For example, it may add emojis or text that celebrates the Vols, UT, and Tennessee. UT Verse is built on ChatGPT-4 and is now available to everyone in the UT system. It offers both text-to-speech and speech-to-text capability. UTK recommends that users avoid sharing PII[2], PHI[3], or FERPA[4] data.	Deployed (as of June 2024)	An internal UTK Office of Innovative Technology (OIT) "tiger team" consisting of five system specialists (IT help desk leader, office of communications leadership, vice provost, and vice chancellor of research) collaborated with Microsoft to deploy UT Verse in just three weeks. OIT initially funded $60K for the project, with promises of additional support from the Provost and Research if usage increases. UTK used prompt-tuning to make the model more UTK-specific, and OIT developers customized the user interface with UT imagery and branding, using code sourced from GitHub. UT Verse also safeguards sensitive data within the UT ecosystem. The model is trained on a GPT-4 platform for Microsoft Azure OpenAI. UTK is working on adding the functionality to allow users to upload their documents to UT Verse.	Low	Med.	$$$$

1) Implementation Effort.
2) Personally Identifiable Information.
3) Protected Health Information.
4) Family Educational Rights and Privacy Act.

General Campus Support

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Create a general campus information aide for staff, faculty, and student support	Provides around-the-clock answers to queries about campus or specific courses in a streamlined mobile app interface.	In 2025, the **University of Michigan** (U-M) launched Go Blue, a mobile AI companion that specializes in U-M specific information and provides access to course-specific AI tutors built with U-M Maizey (see below). Through Go Blue, users can access a wide array of information regarding campus dining, student organizations, shuttle busses, student support resources, sports, campus news, directories information, campus events, and course materials (where applicable). GoBlue includes location services capabilities, allowing users to share their location and ask about campus directions, bus schedules, and weather. It also includes voice-based input and is available for both iOS and Android users.	Deployed (as of Sept. 2025)	Go Blue is powered by U-M's customized OpenAI's GPT-4, MiMaizey. Information and Technology Services staff maintain a Go Blue App Beta Testing Team to obtain feedback on new features and updated information before general releases.	Med.	High	$$$$
Build a platform that enables users to build personalized AI agents	Enables any campus user to create a personalized instance of a generative AI model.	The **University of Michigan** (U-M) created U-M Maizey, a no-code platform that allows campus users to build their own customized chatbots by connecting their own datasets and personal accounts (e.g., Google, Canvas). For example, U-M faculty have built classroom AI tutors using U-M Maizey, which now integrates with Canvas. Users can customize the tool's persona and define how it should respond to queries. For example, they can adjust the temperature of models (i.e., how creative or conservative responses should be). Users can add a maximum of 50 data sources.	Deployed (as of Sept. 2025)	U-M Maizey is available to campus users across all three University of Michigan campuses and Michigan Medicine. The tool is trained on specific data provided by the user, such as Google Docs, PDFs, and other unstructured, natural-language text. Once the chatbot is ready, a unique URL is generated. This URL can be shared with others in the selected MCommunity Group, allowing them to access and interact with the chatbot. The tool is approved for moderate sensitive data (i.e., information covered by FERPA). U-M Maizey was offered free of charge through September 30, 2023. Now, departments, units, or projects pay for Maizey usage through UM's departmental charge system. UM is working on developing personal Maizey bots for students. UM's AI toolsets were used by an average of 15,000 users a day as of February 2024.	Med.	High	$$$$

1) Implementation Effort.

General Campus Support (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Build a platform that extends capabilities of generative AI models to users	Enables any campus user to experiment with and customize generative AI models, reducing inequitable access to "premium" model features and ensuring that security, privacy, and accessibility aligns with institutional standards.	PhoenixAI is **The University of Chicago's (UChicago)** custom generative AI chat service offered to faculty, staff, and students. PhoenixAI provides all UChicago community members with free access to OpenAI's GPT-4.1 and o3 LLMs via a user interface that, unlike most commercial offerings, meets A11y standards and, especially, the University of Chicago Digital Accessibility Standards. By hosting PhoenixAI in the University's protected Microsoft Azure environment, user data is more protected from third-party commercial use than if they had purchased enterprise licenses from OpenAI. Users can decide which model (GPT-4.1 or o3) to use at any point via drop-down menu, employ internet search to inform model output, analyze data, write software code, customize the chat tool to their needs, and more. Additionally, within PhoenixAI, users can create, share, edit, and delete custom "AI Assistants," which use Retrieval Augmented Generation (RAG) with user-provided files to enhance model responses. Users commonly use this capability to create course-specific tutors or chatbots for use within a research group.	Deployed (as of Sept. 2025)	UChicago staff partnered with Royal Cyber to develop and launch Phoenix AI in September 2024. They built PhoenixAI in the Microsoft Azure environment and the first release used OpenAI's GPT-4o model. By working with a vendor versus exclusively using in-house staff to develop the tool, UChicago was able to launch the first version of PhoenixAI within six weeks of signing their contract. A cross-functional group of staff across the institution work on the regular update releases to fix bugs and test new functionality and/or features in collaboration with Royal Cyber. The UChicago team includes project managers, business analysts, designers, software developers, accessibility experts, and information security staff. As of August 2025, PhoenixAI allows users to toggle between GPT-4.1 and GPT-o3. UChicago pays for access to OpenAI's models via a token-based pricing scheme. To cap their expenses, they provide students, faculty, and staff with model-dependent query quotas. Users are notified when they reach their daily limit: • *GPT-o3 model*: 50 queries per day limit for faculty and 10 queries per day for all other users. • *GPT-4.1 model*: 50 prompts per day for all users. Staff at UChicago are working to update PhoenixAI to include a new security role, Personal Health Information (PHI) Restricted, which would enable university researchers to share sensitive data with PhoenixAI while complying with HIPAA data privacy standards.	Med.	Med.	$$

1) Implementation Effort.

eab.com

General Campus Support (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Create a sandbox to familiarize campus with AI tools	Gives campus users the opportunity to experiment with AI tools in a controlled environment.	**Harvard University** launched the AI Sandbox pilot program to provide a unified interface for campus members to experiment with various large language models (LLMs) platforms while ensuring data privacy and security. Platforms offered include Azure OpenAI GPT-3.5, GPT-3.5 16k, GPT-4, and GPT-4 32k; Anthropic Claude 2 and Instant; and Google PaLM 2 Bison.	Trial (as of June 2024)	Over 40 IT professionals collaborated with academic divisions to create the AI Sandbox. The pilot program included 50 faculty participants, who have incorporated these tools into their courses in a range of ways. For example, one faculty used AI tools to summarize student responses and learning based on in-class assessments while another created a virtual environment to stage a "whodunnit" murder mystery in France. Access to the AI Sandbox is currently coordinated by the schools.	Med.	Med.	$$

1) Implementation Effort.

eab.com

Multi-Assistant Platform

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Create institution-specific AI assistants for campus constituents	Streamlines access to institutional information and boosts productivity for general and institution-specific tasks.	The **University of California, San Diego** (UCSD) built TritonGPT, a suite of AI assistants, to answer university-specific questions and help campus users with general tasks like summarizing documents and creating tailored content to university-specific tasks like navigating UCSD policies and processes. Some specific AI assistants included under the TritonGPT umbrella include: • UCSD Assistant: answers queries related to policy, process, and help documentation. • Job Description Helper: generates job descriptions for hiring managers, based on over 1,300 career tracks job standard templates. • Fund Manager Coach (coming soon): elucidates UCSD's financial policies and procedures for those managing grants and other department finances.	Deployed (as of Sept. 2025)	TritonGPT is powered by Meta's Llama 3 open-source language model. IT indexed over 23 data sources to train TritonGPT, from the academic personnel website, course catalogs, to UCSD's strategic plan. TritonGPT is hosted on local infrastructure at the San Diego Supercomputer Center. USCD brought TritonGPT from idea to proof-of-concept in less than two months. It then tested TritonGPT in a two-month early user program with over 400 health sciences and campus employees to source their feedback to improve accuracy. UCSD has now expand TritonGPT all 37,000 campus employees. Development was spearheaded by the Workplace Technology Services team and the Academic Technology Services Team. This includes a student staff programmer, a UX designer, and a senior AI architect. Read a more in-depth case study in Section 3.	Med.	High	$$$$

1) Implementation Effort.

eab.com

Multi-Assistant Platform (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Create institution-specific AI assistants for campus constituents	Provides secure access to genAI tool capabilities while supporting institution-specific use cases.	**The University of California, Irvine** (UCI) launched ZotGPT Chat in January 2024, a customized generative AI platform. It is built in Azure AI and provides secure (i.e., private) access to GPT-4 Turbo to campus at no charge. ZotGPT can help users brainstorm ideas, create images, and formulate plans. It also supports voice chat, chat history, and file uploads.	Deployed (as of June 2024)	ZotGPT Chat was developed by the Office of Information Technology AI Workgroup at UCI. ZotGPT Chat offers a mobile-friendly experience, voice chat, and protects the confidentiality of personal and institutional data up to the P3 level. The P3 level is defined as institutional information whose unauthorized disclosure or modification could result in small to moderate fines, penalties, or civil actions (e.g., Student records, UCI personnel records, IT security information)[2]. The source code is available to all UC campuses to jumpstart similar efforts. UCI plans to continue adding functionality to ZotGPT. They expect to launch ZotDesk, an IT help desk agent in Fall 2024. By the end of the year, they will also launch ScholarConnect to make faculty and their research more accessible by pulling and organizing information found in faculty CVs (e.g., honors, publications). Soon, campus members will be able to create customizable chatbots using departmental data or websites. The ZotGPT team is offering AI training courses through the online learning platform Udemy to increase adoption and showcase popular use cases. As of March 2024, ZotGPT Chat had 1,796 unique users and was trending towards 200 daily users.	Med.	High	$$$$

---

1) Implementation Effort.
2) UCI's data classification matrix.

Source: Gallegos, Emma, "UC Irvine and UC San Diego launch customized AI tools," *EdSource*, March 22, 2024; UC IT Blog, "UC Irvine Introduces Innovative AI Solution: ZotGPT Chat," *UC IT Blog*, March 28, 2024; Paykamian, Brandon, "UC Irvine Develops Custom GenAI Platform 'ZotGPT Chat'," *Government Technology*, March 21, 2024; EAB interviews and analysis.

Multi-Assistant Platform (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Create institution-specific AI assistants for campus constituents	Provides community members with purpose-built AI agents optimized for university tasks in a protected environment	PatriotAI is **George Mason University's** (GMU) enterprise AI platform that grants students, faculty, and staff access to six institution-specific AI agents in a secure environment: • *PatriotChat*: serves as a conversational assistant for general George Mason information. • *DocuMate*: summarizes, derives key takeaways from, and compares scholarly documents. • *PatriotPal*: accurately answers student questions about university-specific administrative procedures, academic policies, and services. • *NourishNet*: connects students experiencing food insecurity to food access programs, offering a stigma-free experience. • *CourseMate*: helps students understand course content, analyze reading materials, and prepare for exams. • *SyllaBright*: supports course design, helping faculty create and refine course material (e.g., syllabi, planning assessments)	Deployed (as of August 2025)	PatriotAI is an instance of nebulaONE, Cloudforce's multi-modal Generative AI platform built on Azure. It is a product of GMU's partnership with Microsoft Education and Cloudforce. GMU describes this as only the "first lineup" of AI tools available on the PatriotAI platform. Community members and leaders on GMU's AI Visioning Taskforce gave significant input into the development of the first six agents currently featured on PatriotAI. Agents within PatriotAI use Brave Search API for general internet and targeted site search queries. Currently, GMU students, faculty, and staff can sign in with their university Microsoft account to access the beta version of the platform. Leaders plan to debut the platform at scale in Fall 2025.	High	Med.	$$

1) Implementation Effort.

eab.com

Multi-Assistant Platform (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Build a platform that enables users to build, experiment, and use AI-powered solutions	Encourages both novice and experienced users to engage with AI tools and build AI-powered products.	**Arizona State University**'s (ASU) AI Acceleration team produced the model- and cloud- independent CreateAI Platform to give the ASU community the opportunity to engage with and build AI tools in a secure environment. Not only can community members develop and deploy custom AI models and AI-powered products in the platform, but they can also approach the AI Acceleration team with products they have created to potentially scale further. The CreateAI Platform currently offers 21 large language models (LLMs). It also includes multiple vector databases and supports document upload extraction and ingestion.	Deployed (as of Sept. 2025)	The AI Acceleration team launched the CreateAI Platform in October 2023 with six LLMs; as of April 2024, the platform is integrated with 21 large language models. CreateAI is FERPA[2]-approved environment due to living within ASU's garden wall. The AI Acceleration team is comprised of representatives from Program and Design, Data Science, AI Development, and Data Architecture and Enrichment groups. Products that have already been built using the CreateAI platform include a 24/7 syllabus Q&A bot, tutor bot, and course builder assistant. ASU hosted the Betaland Community Showcase so that community members could share their proof-of-concept and beta products.	High	High	$$$$

1) Implementation Effort.
2) The Family Educational Rights and Privacy Act.

Source: Arizona State University, "CreateAI Platform," *Arizona State University*, 2024; EAB interviews and analysis.

Information Technology

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Triage IT help desk tickets	Speeds up response time and eliminates time spent on manual effort needed to sort IT help desk tickets.	The **University of South Florida's** (USF) AI-Powered IT Service Desk classifies incoming service tickets, automatically sorting requests based on the following categories: Status, Service Team, Issue Type, and Priority. Previously, student workers manually sorted 100K tickets annually for USF's IT Service Desk. Automating the triaging process is expected to save USF about $90,000 per year in student labor costs, and overall net savings of $68,200 per year.	Deployed (as of June 2024)	The Service Desk is built on ChatGPT-3.5 Turbo. It employs the OpenAI Embeddings API[2] to help understand and encode service team description data and ticket content and the Completion API to process and generate responses. Advanced Prompt Engineering techniques, such as "Chain of Thought," enable the AI to process complex inquiries effectively. The Service Desk also incorporates a suite of services in Azure, including a WebServices VM and a custom-developed MS Semantic Kernel Framework, which supports its integration with existing IT workflows. USF estimates that the annual cost of hosting the Service Desk will be about $22K per year. Read a more in-depth case study in Section 3.	High	Med.	$$$
Resolve IT help desk inquiries	Acts as a virtual IT help desk, providing students IT assistance and answers to allow IT teams to focus on more complex issues.	**Ithaca College's** Analytics and Special IT Projects team developed the IT and Analytics AI Agent to field and resolve campus IT help desk inquiries (e.g., I can't get into my portal to view my grades). If the agent cannot find the correct response, it will automatically generate a service request. Even if it does offer a solution, it will still ask if the user wants it to submit a ticket.	Discont. (as of February 2025)	Ithaca's CIO, in partnership with the Analytics and Special IT Projects team, built the AI Agent using OpenAI's API platform. They primarily employed RAG[3] to enhance the AI agent, using Ithaca's TeamDynamix Knowledge Base as a reference material for IT solutions. On the front end, Ithaca uses Slate to provide contextual data about students and securely deploy web portals. By building the portals on Slate, they were able to tie their AI agent to their single sign-on (SSO) authentication system. Ithaca pays $1K a month for OpenAI's API access through Microsoft Azure. They aimed to launch the IT and Analytics AI Agent in mid-2024, but due to data concerns, halted production. Read a more in-depth case study in Section 3.	Med.	High	$$$

1) Implementation Effort.
2) Application programming interface.
3) Retrieval-Augmented Generation.

Source: Educause, "Revolutionizing the IT Service Desk: GPT-Powered Ticket Classification at USF," *Educause*, October 11, 2023; EAB interviews and analysis.

Information Technology (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Monitor email security threats	Bolsters protection against email security threats and reduces need for manual email review.	**California Polytechnic State University**'s (Cal Poly) Security Operations Center (SOC) implemented SlashNext's AI-driven email security service, to review emails for potential security threats like phishing. The tool claims a 99.9% detection rate and a one-in-one-million false-positive rate. As a result, Cal Poly was able to drastically reduce time spent reviewing emails. In the first 24 hours after deployment, the team saw an 80% reduction in reported malicious emails, freeing up significant staff time for projects like the management of IPs[2] or identification of high-risk AWS[3] activity. Previously, 75% of the student SOC[4] team's time was spent conducting reviews.	Deployed (as of June 2024)	After a successful 14-day trial, SlashNext was deployed to monitor 6,500 faculty and staff inboxes, analyzing over a million emails in the first week. It detected 434 zero-hour link attacks (previously unidentified threats with no known malware signature) and 271 business email compromise (BEC) emails, significantly reducing the burden on security operations center (SOC) staff. Over a six-month period, SlashNext detected 7,984 BEC and social engineering threats, targeting 2,747 users. The average cost of a BEC incident in 2023 was $137,000.	High	Med.	$$$

1) Implementation Effort.
2) Internet Providers.
3) Amazon Web Services.
4) Security Operations Center.

Source: Educause, "AI-Powered Email Monitoring Enables SOC Students to Improve Campus Security," *Educause*, May 3, 2024; EAB interviews and analysis.

Human Resources

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Support staff with customized payroll assistance	Simplifies access to an employee's personal payroll information.	**EY**, a multinational professional services firm, developed the EY Intelligent Payroll Chatbot as part of their next-generation payroll platform in partnership with Microsoft. The chatbot will respond to payroll inquiries for both internal EY employees and their clients. For example, the chatbot will provide tax documents, pay slips, and insights to 298,000 employees across 159 countries in 49 languages. It is also engineered to dissect and interpret complex payroll data against a backdrop of global regulatory frameworks.	Trial (as of June 2024)	EY piloted the chatbot within various client environments, including airlines and tech companies. It produced a 93% correct first-time answer ratio and is expected to enhance both employee satisfaction and first-contact-resolution KPIs by greater than 50%. Currently, the payroll chatbot answers more than 500 employee questions a day. The chatbot is built on ChatGPT within Azure OpenAI Service. It parses details from employee pay slips, tax regulations, and employer policies, to deliver precise answers to payroll queries. It forms an integral component of EY's Employee Experience Mobile and Web Applications. EY teams trained the chatbot by uploading data from a range of sources into the bot and asking its payroll consultants in various countries to share questions employees had recently asked. The chatbot is anticipated to achieve a more than 80% query resolution rate, aiming to halve the current costs associated with addressing payroll inquiries. The chatbot operates within EY's controlled data environments to ensure privacy and data integrity.	High	High	$$$$

1) Implementation Effort.

Source: Bach, Deborah, "'The last frontier of disruption': With its new AI chatbot, EY teams seek to take the pain out of payroll questions," *Microsoft*, June 5, 2023; EAB interviews and analysis.

eab.com

Procurement

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Guide drafting of procurement statements of work (SOWs)	Enables procurement specialists to draft SOWs quickly and consistently, addressing common pain points like lack of commodity knowledge and unclear milestones.	Procurement specialists at **California Polytechnic State University, San Luis Obispo (Cal Poly)** use ScopeBuilder: an AI-assisted tool that streamlines the development of statements of work (SOW) for contracts, bids, and RFPs. ScopeBuilder guides university procurement specialists through the SOW drafting process step-by-step (e.g., suggesting milestones, drafting work scopes in real-time) based on user input, established best practices, and relevant past contracts. ScopeBuilder ensures compliance with institutional standards, offers support for buyers with limited familiarity in a commodity area, and incorporates templates and preferred language from exemplary historical documents.	Deployed (as of Sept. 2025)	Cal Poly's Chief Procurement Officer led the development of ScopeBuilder in partnership with a team of students and Amazon Web Services (AWS) partners from the university's Digital Transformation Hub (DxHub). The DxHub is an AWS-powered innovation hub that provides students with experience solving real-world problems in collaboration with AWS. ScopeBuilder is primarily powered by Anthropic's Claude Sonnet 3.5 model via Amazon Bedrock. All functionality and services exist within the AWS ecosystem to ensure confidentiality of university procurement processes. Student developers built the front-end using React (hosted on Amazon S3 cloud object storage). ScopeBuilder also uses Amazon CloudFront for content delivery, AWS Certificate Manager for encryption, and AWS Lambda for computing functions.	Med.	Med.	$$

Source: EAB interviews and analysis.

eab.com

Admissions

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Review admission application essays	Accelerates application review and affords admissions staff extra time to focus on more substantive portions of the application review process like interviewing.	Researchers at the **University of Colorado Boulder** and the **University of Pennsylvania** developed an AI tool to analyze college application essays and identify applicant characteristics like leadership and perseverance.	Develop. (as of June 2024)	Researchers first collected more than 300,000 anonymous, 150-word essays submitted to colleges that focused on extracurricular activities and work experiences. A group of admissions officers then read and scored those essays based on seven characteristics. The researchers then trained the AI system to evaluate those characteristics in essays using the criteria set by the admissions officers. The AI system was then able to identify whether students across backgrounds demonstrated those qualities (e.g., teamwork, intrinsic motivation) in their essays.	Low	Med.	$$
Automate transcript review	Reduces manual effort and allows staff to reallocate their time towards activities like student engagement and personalized counseling.	**Maryville University** used Sia to automate transcript review, streamlining their admission process and decreasing the potential for human error. The use of Sia's tool has afforded 14 admissions counselors 50% more time to focus on building relationships with students through online and social media campaigns.	Deployed (as of June 2024)	Sia, is a proprietary Unified AI Engine, designed and developed by OneOrigin Inc. Sia connects to an institution's CRM[1] or SIS[2] and converts PDF transcripts into machine-readable formats.	Med.	Low	$$$

1) Implementation Effort.

eab.com

Enrollment

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Answer enrollment-related questions for students and families	Streamlines access to all enrollment-related information for both prospective and current students, reducing roadblocks in the student experience (e.g., confusion around financial aid) that could lead to summer melt and attrition.	**Ithaca College's** Analytics and Special IT Projects team developed Ithaca Insights, a Prospective Student AI Chatbot that provides prospective students tailored information about Ithaca College, courses, and campus (e.g., "What classes would I be taking my first year as a Physical Therapy major?"). The chatbot, which employs OpenAI's API[2] platform, can answer inquiries in the following areas: major information and requirements, financial aid, academic advising, housing, general wellbeing, campus information, and town information.	Discont. (as of February 2025).	Ithaca's CIO, in partnership with the Analytics and Special IT Projects team, built the AI Agent using OpenAI's API platform. They primarily employed RAG[3] to enhance the AI agent, specifically using information from their student data lake house and Ithaca College web pages. On the front-end, Ithaca uses Slate's CRM[4] to provide contextual data about students and securely deploy web portals. By building the portals on Slate, they were able to tie their AI agent to their single-sign on (SSO) authentication system. If approved, they are targeting a late 2024 launch. Ithaca pays $1K a month for OpenAI's API access through Microsoft Azure. They aimed to launch the agent in mid-2024, but due to data concerns, halted production. Read a more in-depth case study in Section 3.	Med.	High	$$$
Answer enrollment-related questions for students and families	Improves accessibility of enrollment-related information for current students and student success staff providing support to students.	**New York University's** (NYU) IT team is building an Enrollment Management Student Success Chatbot is designed to assist student success staff in addressing inquiries from current students about various aspects of the NYU student experience, such as academic policies and administrative processes. The agent will streamline the resolution of common student queries regarding registration, enrollment status, and academic policies (e.g., "What are the steps for students to get their GI bill benefits certified?", "What steps are required to return from a Leave of Absence?").	Develop. (as of June 2024)	NYU IT is spearheading the tool's development along with a handful of student workers. The project team used RAG to enhance OpenAI's ChatGPT-3.5 API on a 600-page NYU bulletin built by enrollment management, which contains detailed information about all 26 schools. The team also partnered with student success support staff to source 45 frequently asked questions they could begin training and optimizing the chatbot for. The chatbot will be integrated directly into NYU's student success website, making it easily accessible for students. Read a more in-depth case study in Section 3.	Med.	Med.	$$

1) Implementation Effort.
2) Application programming interface.
3) Retrieval-Augmented Generation.
4) Customer resource management.

Source: EAB interviews and analysis.

Enrollment (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Generate tailored degree recommend-ations	Helps students navigate over 900 degree offerings and find best-fit academic programming.	**Arizona State University's** (ASU) AI Degree Recommendation Engine helps prospective and current students navigate over 900 degrees offered by the institution. The engine recommends students personalized degree options based on interests students express in a multiple-choice quiz. After students complete the quiz indicating their interests and preferences, the engine uses semantic matching capabilities and conducts a similarity search to suggest potential degree options. Then the engine generates personalized degree explanations that describe why the degree would be a good match for the student.	Trial (as of June 2024)	The development of the proof of concept took six weeks. ASU's development team first gathered and cleaned information from the university's degree pages, which include general degree descriptions written by the content team. They then used this information with a RAG[2] framework to enhance the AI engine's degree recommendations. The degree engine then underwent user testing with a group of students and internal stakeholders, including support services like career services. Read a more in-depth case study in Section 3.	Low	Med.	$$
Field and resolve student registration questions	Increases speed of service while enabling staff to focus on more complex registration issues that require human intervention.	**Ohio State University** (OSU) developed a proof-of-concept chatbot trained on registration data sources to serve as a 24/7 source of information for students with registration-related queries.	Develop. (as of June 2024)	OSU's IT team developed their AI bot using foundational models accessed through Amazon Web Services (e.g., Anthropic, Meta). They trained the AI system with data extracted from the official OSU website and the institution's course catalog, primarily captured in PDF format. The team incorporated 400 distinct user intents and over 1,000 queries of commonly asked registration questions. It took five software developers eight weeks of full-time development work to produce the proof-of-concept tool. The team spent approximately $5000 on compute and is still evaluating on which AI model(s) to launch their registration bot. They aim to further develop the bot to be a one-stop-shop that can answer any institution-related inquiry. The OSU IT team is currently submitting a proposal to develop the AI bot to satisfy this vision.	Med.	Med.	$$

1) Implementation Effort
2) Retrieval augmented generation.

Source: Faller, Mary Beth, "ASU+GSV Summit tackles big questions about AI, technology, education," *Arizona State University News*, April 18, 2024; EAB interviews and analysis.

Enrollment (cont.)

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Generate personalized student recruitment campaigns	Personalizes student recruitment and messaging at scale, boosting event attendance and prospective student engagement.	**Trinity Valley Community College** (TVCC) used Element451's Bolt AI Copilot to create an automate personalized drip campaigns for prospective students. The AI-generated campaign outlined communication content and automated workflows, including step-by-step rounds of communication, based on the college's goals and student data. TVCC then set the target audience and triggered the workflow, allowing the AI Copilot to handle personalized messaging at scale. As a result, TVCC reached 15,000 students via automated personalized workflows, and saw a 40% increase in messages sent during the first week of the chatbot. Ultimately, TVCC experienced an 18% increase in graduation rate over a year and a 940% increase in event attendance.	Deployed (as of June 2024)	Bolt Copilot is powered by OpenAI's GPT-4. To create the personalized campaigns, TVCC primarily used the Campaign Copilot capabilities of the Bolt AI Copilot. In particular, they used behavioral triggers, website interactions, and demographic data to deliver targeted messages to students throughout their recruitment journey.	Med.	Med.	$$$

1) Implementation Effort.

Source: Carballo, Sirley, "Webinar Recap: How Trinity Valley is Redefining Student Success," *Element451*, 2024; EAB interviews and analysis.

eab.com

Advancement

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Solicit small gifts autonomously	Expands donor reach by using AI agents for autonomous frontline fundraising	**Texas State University** (TXST) enlisted Givzey's Version2.ai to deploy autonomous fundraising agents called Virtual Engagement Officers (VEOs). Without any staff intervention, VEOs qualify donors, build relationships, and solicit and close gifts. Within just one week, TXST's VEO, Emma, reached 925 donors, had more than five touchpoints with 70% of the assigned portfolio, and raised $4,000 from 78 gifts. Overall, Emma independently completes the work of over 20 gift officers by sending personalized messages to 2,000 donors in its portfolio, analyzing donor behavior for tailored engagement, and responding to donor inquiries in real-time.	Deployed (as of September 2025)	Version2.ai enabled TXST's advancement team to launch VEOs in just three guided weeks with a total time investment of around eight to ten hours. Staff onboard their autonomous fundraiser with institutional information and the advancement team's processes and goals.	High	Low	$$$

1) Implementation Effort.

Source: EAB interviews and analysis.

eab.com

Research

AI Application	Rationale	Capsule Description	Status	Implementation Detail	Impact	IE[1]	Resource Cost
Assist researchers in navigating grant funding regulations	Simplifies compliance with funding regulations for researchers and provides real-time guidance on allowable expenses, ensuring proper conduct throughout the research process.	**New York University** (NYU) is building a Research Financial Support Allowable Costs Chatbot that will answer questions faculty researchers, research department employees, central office administrators, or grant financial analysts may have about grant funding regulations for allowable costs. It is designed to simplify and ensure compliance with funding regulations. The chatbot will provide real-time guidance on allowable expenses and help researchers justify their expenditure decisions during audits and reviews.	Develop. (as of June 2024)	NYU IT is leading development of the chatbot on OpenAI's ChatGPT-3.5 platform. The team enhanced the chatbot using RAG[2] with six key documents based on uniform conduct guidelines from the federal government. From there, they distributed a survey to a group of Research Financial Support staff to source and test questions they would want the chatbot to answer. Success in this phase will lead to an extended trial involving more Research Financial Support staff and departmental staff within Schools, preparing the tool for a wider institutional rollout. Read a more in-depth case study in Section 3 (p. 48).	Med.	Med.	$$
Build AI tool environment secure for healthcare practice and research	Fosters efficiencies and innovation in health care practice and research while ensuring security and privacy compliance (i.e., HIPAA[3])	The **University of Pennsylvania Medical School** launched the PennAI Chat, a virtual AI assistant built on GPT-3.5 Turbo that is cleared and designed for sensitive data (e.g., HIPAA). The associated PennAI Cloud Service offers users access to open source and commercial large language models (LLMs) as well datasets from clinical and operational systems to enable research experimentation. PennAI Chat and Cloud Service are hosted within their HIPAA-covered Azure tenant and network security, meaning all data stays between Penn's firewalls.	Develop. (as of June 2024)	A broad, interdisciplinary team from Penn Medicine, Microsoft, and Databricks developed PennAI Chat and Cloud Service. There was also a large list of pilot users (11 users) that spanned disciplines from Radiology to Clinical Informatics & Health. Penn Medical School accesses OpenAI's GPT-3.5 Turbo through a partnership with Microsoft, and they plan to upgrade PennAI Chat to GPT-4.	Med.	High	$$

1) Implementation Effort.
2) Retrieval-Augmented Generation.
3) Health Insurance Portability and Accountability Act.

In-Depth Exploration of Select AI Applications

3

- Ithaca College
- University of South Florida
- Arizona State University
- University of California San Diego
- New York University
- York University
- University of Birmingham

eab.com

Navigating In-Depth Explorations of AI Applications

This section provides in-depth profiles of higher education institutions that are **developing their own AI applications**. These case studies will document the institution's journey, from technical development to deployment.

These profiles provide insight into many of the questions institutions that are considering pursuing AI applications themselves are facing:

- *How does this AI application serve the needs of our institution? What is the cost-benefit of pursuing this application?*

- *How do we develop the application from a technical standpoint to ensure accuracy and minimize risks?*

- *How do we go from proof-of-concept to production? Who do we need to involve?*

- *What are the development and hosting costs? How can we finance the project?*

Section Three Overview

Institution		AI Applications Profiled	Location
	Ithaca College	• Ithaca Insights: Prospective Student AI Chatbot • IT and Analytics AI Agent	p. 40
	University of South Florida	AI-Powered IT Service Desk	p. 44
	Arizona State University	AI Degree Recommendation Engine	p. 46
	University of California San Diego	TritonGPT	p. 48
	New York University	• Enrollment Management Student Success Chatbot • Research Finance Support Allowable Costs Chatbot	p. 52
	York University	AURA	p. 56
	University of Birmingham	StudyStash	p. 61

Ithaca College

Ithaca Builds Prospective Student AI Chatbot and IT and Analytics AI Agent

Type of Institution: Private Liberal Arts College

Location: Ithaca, New York

Enrollment: 4,619 Undergraduate, 435 Postgraduate

Operating Expenditures (In Thousands): USD 189,408

Application Overview

In 2024, Ithaca College's Analytics & Special IT Projects team developed two distinct AI-powered tools:
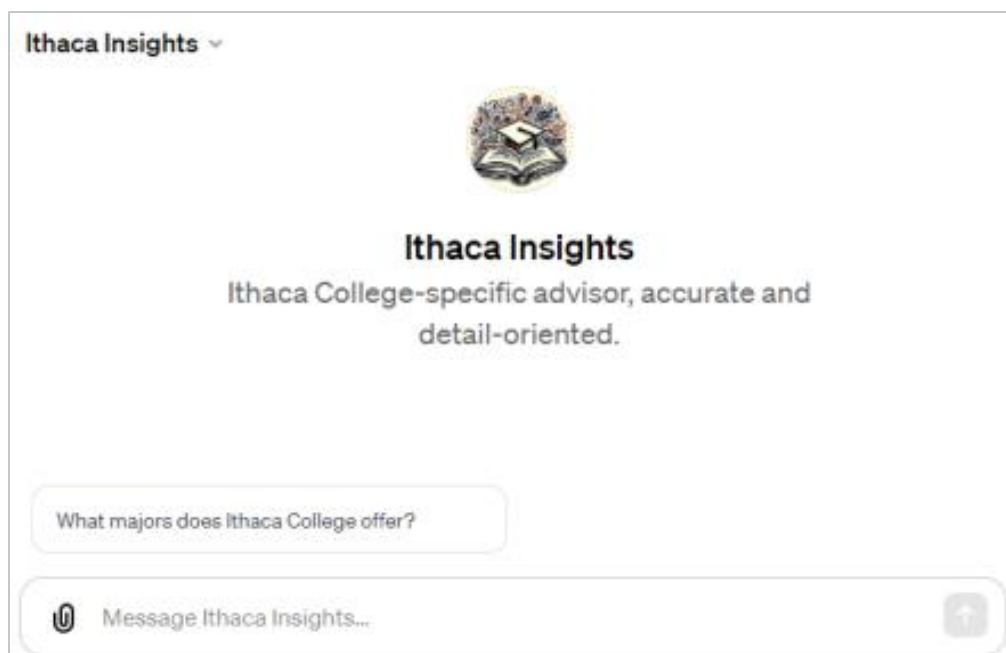- **Ithaca Insights**: Prospective Student AI Chatbot
- **IT and Analytics AI Agent**

Application Functionality

Ithaca Insights is an AI chatbot that provides prospective students tailored information about Ithaca College, courses, and campus (e.g., major information and requirements, academic advising, financial aid, housing, general wellbeing, campus information, town information).

Sample Questions:

- *"What are the requirements for a computer science major?"*
- *"Do I have to get a meal plan if I prefer to make my own meals?"*
- *"What are the deadlines for applying for financial aid?"*
- *"How do I select a roommate?"*
- *"I am nervous that I won't be successful, how will IC help me?"*
- *"As a music major, am I able to take biology as a minor?"*

Ithaca Insights ⌄

Ithaca Insights

Ithaca College-specific advisor, accurate and detail-oriented.

What majors does Ithaca College offer?

Message Ithaca Insights…

A Proof-of-Concept

Ithaca

"What are the requirements for a computer science major?"

"What is the do in Ith"

"Do I have to get a meal plan if I prefer to make my own meals?"

"How

"I am nervous that I won't be successful, how will IC help me?"

Response:

It's completely normal to feel nervous about starting college and wondering about your success. At Ithaca College, there are several resources and programs designed to support you academically and personally, ensuring you have the best chance to succeed.

1. **Center for Student Success**: This is your go-to resource for any academic challenges or questions you might have. The Center offers academic advising, tutoring services, and workshops on study skills and time management.

2. **Academic Programs**: Ithaca College offers a wide range of programs and courses that are designed to be engaging and supportive. The faculty are committed to helping students succeed and are available for one-on-one guidance.

3. **Peer Support**: The college has a strong sense of community. You can find peer support through various student organizations and mentorship programs.

4. **Wellness Resources**: The Hammond Health Center on campus provides health services, and there are also counseling and psychological services available to support your mental health and wellbeing.

5. **Center for Career Exploration and Development**: This center can help you with career guidance, internships, and job search strategies, aligning your academic pursuits with your career goals.

6. **Financial Support**: If you have concerns about financial aid, you can visit the Financial Aid office or check the FAFSA website for the latest information on financial aid options.

Remember, it's important to reach out and utilize these resources. Everyone's college journey is unique, and Ithaca College is dedicated to supporting you every step of the way. If you have any specific concerns or needs, feel free to ask, and I can provide more tailored information.

Note: Tools in our

ITHACA COLLEGE

The Ithaca team is also exploring expanding the prospective student chatbot to offer tailored career guidance advice for prospective and current Ithaca students.

A Proof-of-Concept

Career Insights: Exploring possible career paths based upon courses taken, academic performance and other factors

Response:

Given the student's focus on Physics, evident from courses such as Thermal Physics, Quantum Mechanics, and Advanced Physics, he could be interested in careers related to research, teaching, or practical applications in Physics. This might include roles as a physicist in a government lab, a university professor, a high school teacher, or a role in a private sector firm that applies physics principles such as an engineering company or technology firm. His experience with Advanced Physics Research suggests an interest and ability in conducting academic research, which could lead to a career in academia or a government research lab.

Bearing in mind his other courses such as Environmental Chemistry and Proposal and Grant Writing, he may also be considering careers at the intersection of environmental science and physics, perhaps working in policy, environmental consulting or conservation organizations. In these roles, he would leverage his physics knowledge to understand environmental issues and his grant writing skills to secure funding for projects or applications.

Additionally, the student's course in Proposal and Grant Writing could lead him to consider roles in technical writing or communications within scientific or technological corporations or organizations. It's also worth noting that he has taken several PALS courses such as Tennis, Badminton, and Springboard Diving, which could be purely for recreational purposes, however, if he has a strong passion for sports, he may also consider a career in sports science or coaching.

Remember, these are all possibilities based on observation of his course selections. The student's unique interests and experiences would be the most influential factors in his career choices.

Note: Tools in our AI Sandbox are proof-of-concepts intended to explore possible use cases and capabilities. They may or may not result in an actual deployed service.

ITHACA COLLEGE

Source: Ithaca College, Ithaca, NY; EAB interviews and analysis.

Ithaca College (cont.)

The **IT and Analytics AI agent** can independently field and respond to questions that campus members pose to the IT help desk. The AI agent is focused on improving IT services by providing real-time solutions to common IT issues faced by the campus community. If the agent cannot find the pertinent information to an inquiry, it will automatically generate a service request. Even if it does offer a solution, it will still ask if the user wants it to submit a ticket if they find the response unsatisfactory.

Information Technology and Analytics AI Agent

You: I get requests in DUO but they don't seem to work.

Agent: I'm sorry to hear that you're having trouble with Duo. This issue can sometimes occur if the Duo mobile app is not synchronized correctly with your phone. Here are steps you can take to possibly resolve the issue:

1. Open the Duo Mobile app on your phone.
2. In the list of accounts, find and tap on the Ithaca College entry.
3. Tap on the settings gear, usually located in the top right corner of the app.
4. Tap on "Account Settings".
5. Scroll down and find the "Time correction for codes" section and tap on it.
6. Tap on "Sync now".

If this doesn't help, or if you're experiencing a different kind of issue with Duo, I can open a ticket for you with the IT Service Desk. They will be able to more thoroughly investigate and resolve the issue. Would you like me to open a ticket for you?

Information Technology and Analytics AI Agent

You: I get a number to enter, and when I put it in it looks like it was accepted but I still can't get into Homer Connect to view my grades.

Agent: I understand how frustrating it can be when you're unable to access important platforms like Homer Connect. I have gone ahead and opened a high-priority ticket for you with the IT Service Desk. They will be able to further assist you with this issue.

In the meantime, please double-check that you're entering the correct Duo passcode and not the Duo Push notification code that you use to verify you're attempting to log in.

You might also want to try signing out of all accounts and signing back in, or trying a different browser or device.

For any immediate assistance, don't hesitate to contact the IT Service Desk directly at 607-274-1000. They are available to help.

Project Status

Ithaca created proof-of-concepts for both AI applications but did not bring either to production, citing that the data was not yet organized enough to confidently scale either application. However, using the lessons learned from the experience, Ithaca later launched Nebula, an AI-powered mental health application for health and wellness counselors, in 2025.

Key Participants

Executive Sponsor

• Dave Weil, CIO, proposed the applications to the Board and stewarded their development.

Analytics & Special IT Projects Team

• Five members from this team (e.g., data warehouse engineers, integration developers) led by Rob Snyder, Director, spearheaded the technical development of both applications. This team reports to the Deputy CIO and VP of Applications and Infrastructure.

Coordinator for AI Initiatives & Senior IT Project Manager (new, temporary position)

• This non-technical role helps shepherd the tools from proof-of-concept to production by collaborating with different units, such as career services, to gather input (e.g., where to implement guardrails) and secure final approval. The coordinator reports directly to the Deputy CIO and VP of Applications Infrastructure.

Development and Implementation Details

Phase 1: Setting Up the Foundation

The CIO attributed Ithaca's ability to rapidly spin up both AI applications to the **robust data management systems** they have developed over the last few years. He attributed their success to three components:

• Centralizing their data in a data lakehouse for analytics, operational reporting, and data-informed decision making streamlined their ability to rapidly access data in a consistent format for AI application purposes.
• Investing in data infrastructure (e.g., SnapLogic) enabled the project team to integrate systems and move data around more freely.
• Strengthening data governance structures instilled an institutional culture of sharing–and not hoarding–data and metadata.

Source: Ithaca College, Ithaca, NY; EAB interviews and analysis.

Ithaca College (cont.)

Development and Implementation Details (cont.)

Phase 2: Developing Applications

In December 2023, Ithaca's Analytics & Special IT Projects team began exploring AI-powered application development options. Both of Ithaca's AI applications, which use RAG[1], are built on the same foundation. The primary difference is the data sources that the applications pull from.

- On the back end, the project team built both applications on OpenAI's API[2] for ChatGPT-3.5. They also integrated Tavily, a search engine for large language models (LLM) that can be integrated with any LLM to enable it to deliver real time information from the web.

- The project team enhanced the LLM using RAG. For example, to access student data, Ithaca live accesses its data lakehouse (which houses a variety of data sources, including the student information system (SIS), degree audit system, and campus engagement system). This allows the AI tools to pull from a student's schedule and club membership if a student asks about their availability. Ithaca also uses Neo4j as a vector store for managing static content like course and program information that is published yearly (e.g., for the IT and Analytics AI agent, Neo4J would include documentation on IT services and troubleshooting).

- On the front-end, Ithaca uses Slate for securely deploying web portals and front-end interactions, providing contextual data about students. By building the portals on Slate, they were able to tie their application to their single sign-on (SSO) authentication system.

- The project team tied together the backend API, the front-end portal, and the vector store with in-house code developed primarily in Python.

Phase 3: Future Roadmap

- **IT and Analytics AI Agent**: The team is targeting a June 2024 deployment to move the IT and Analytics AI agent into production.

- **Prospective Student AI Chatbot**: The CIO, project team, and coordinator for AI initiatives are currently working with campus leaders like career services to help implement guardrails and add more functionality to the chatbot. They are discussing potentially expanding the prospective student AI chatbot to respond to more topics regarding student life, success, retention, and engagement. If approved, they are targeting a late 2024 launch.

- For both applications, the team is still evaluating AI models to use on the back end; they are considering using Claude.

Sources of Data

- **Student Data Lakehouse**: Student information system (SIS), degree audit system, and campus engagement system.
- **Neo4j**: Knowledge-based articles and course catalog data.
- **TeamDynamix Knowledge Base**: IT help desk reference material.
- **Ithaca College web pages**: Public/quasi-public data available on the Ithaca College web pages.

Project Cost and Financing

To build and run these applications, Ithaca pays 1K a month for OpenAI's API access through Microsoft Azure. IT has earmarked 2% (~200K) of its budget next year sustaining these (and more) IT pilots. While the CIO is not concerned about ongoing development costs, he is wary of the cost of token usage once users can submit queries. Ithaca's CIO is considering three strategies to mitigate costs of applications.

1. **Optimizing Query Processing with Multi-Model Approach**: Utilizing a cost-effective model like ChatGPT-3.5 to simplify and condense multiple queries into one, before passing them to more expensive, high-fidelity AI models like PaLM 2 can significantly reduce token usage per user-submitted query.

2. **Caching Similar Responses**: Using a response caching mechanism to store answers to frequently asked questions can reduce the need for the AI model to repeatedly generate responses to similar inquiries. However, this comes at the expense of response agility.

3. **Implementing Query Limits**: Implementing a cap on the number of repetitive inquiries a user can submit can help prevent escalating token costs. This safeguard would activate when repetitive inquiries reach a predetermined limit, prompting the system to offer a direct support ticket option instead.

Ithaca is most likely to employ the first and potentially third mitigation strategies above. Before seriously considering the second strategy, the team would want to find a way to ensure outputs are still personalized after the caching process.

1) Retrieval-Augmented Generation.
2) Application programming interface.

University of South Florida

The University of South Florida Builds AI-Powered IT Service Help Desk

Type of Institution: Public Research University

Location: Tampa, Florida

Enrollment: 36,452 Undergraduate, 9,242 Postgraduate

Operating Expenditures (In Thousands): USD 1,562,591

Application Overview

In 2024, the University of South Florida (USF) developed an **AI-Powered IT Service Desk** to automate the triaging of service tickets to service teams. Previously, student workers manually sorted 100K tickets annually for USF's IT Service Desk.

Application Functionality

USF's AI-Powered IT Service Desk classifies incoming service tickets, automatically sorting requests based on the following categories:

- Status: whether the request has been processed/completed
- Service Team: the relevant IT domain the request is related to
- Issue Type: the type of request (e.g., incident)
- Priority: low, medium, high

The Service Desk also provides a detailed explanation of how it classified the service ticket in each of the categories above, which service teams can explore. The system uses Azure OpenAI's API[1] for ChatGPT-3.5 Turbo.

In Action …

1) Application programming interface.

University of South Florida (cont.)

Key Participants

- Jason Hair, Senior Director, USF IT Research Computing
- Drew Oliver, Assistant Director, USF IT Research Computing

Development and Implementation Details

Service team description data is first chunked to optimize it for AI model processing (built on ChatGPT-3.5 Turbo API[1]). The OpenAI Embeddings API then helps to understand and encode service team descriptions and ticket content into vector representations.

A vector database serves as a semantic memory that stores data about service teams and their past ticket resolutions, enabling the AI model to make informed routing decisions by matching new tickets with historical data. The model can then assess each ticket and assign it to the relevant service team according to ITIL[2] standards for priority and issue type through semantic functions and similarity searches. The model uses the Completion API to process the prompt (i.e., the task description, service team information, example outputs, and ticket information) and generate a response. Advanced Prompt Engineering techniques, such as "Chain of Thought," helps the AI model process complex inquiries effectively.

Based on the AI model's output, ticket information is updated in Jira, creating a feedback loop that continually enhances its performance over time. This continuous learning mechanism is vital for improving the model's accuracy and efficiency. The Service Desk also incorporates a suite of services in Azure, including a WebServices VM and a custom-developed MS Semantic Kernel Framework, which supports the AI model's integration with existing IT workflows.

System Architecture
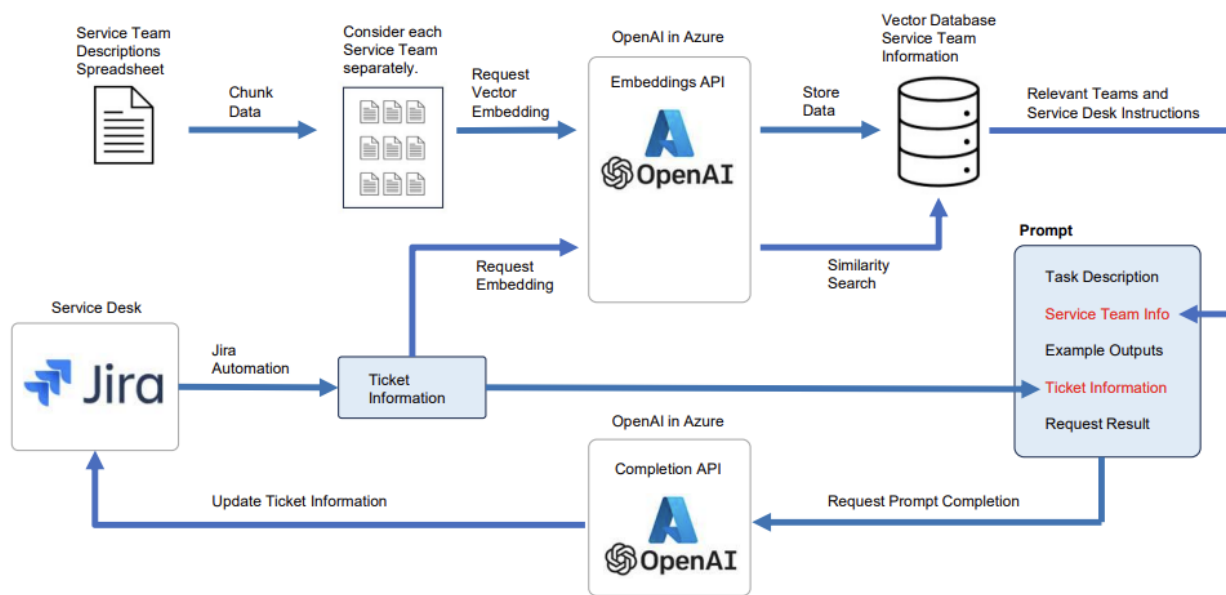
Project Cost and Financing

USF IT estimates the annual cost of hosting the AI-Powered IT Service Desk will be $21,800 per year. See the full cost breakdown below:

- $20K: Maintenance of Prompt.
- $1,300: All related costs to using Azure OpenAI.
  - Licensing API access for models GPT-3.5-Turbo-4k and the embedding model text-embedding-ada-002 that is designed to optimize the number of tokens used.
  - Token usage for 100,000 tickets (based on an average 2,500 tokens per ticket).
  - Three times the semantic functions are factored into the cost estimation.
- $500: WebServices VM.

Estimated Impact

USF estimates that the AI system will reduce the overall operating cost of the service desk from $90,000 to $21,800 annually.

1) Application programming interface.
2) Information Technology Infrastructure Library.

Arizona State University

Arizona State University Builds AI Degree Recommendation Engine

Type of Institution: Public Research University

Location: Tempe, Arizona (with additional campuses across Arizona)

Enrollment: 112,177 Undergraduate, 30,459 Postgraduate

Operating Expenditures (In Thousands): USD 3,120,385

Application Overview

In 2024, **Arizona State University (ASU)** designed the AI degree recommendation engine to help students navigate its over 900 degree offerings. The engine recommends students personalized degree options based on interests students express in a short, multiple-choice quiz.

Previously, students had to manually sift through the Undergraduate and Graduate Majors and Degrees page to find an academic path suited for them.

Application Functionality

The AI degree recommendation engine features three main components:

1. **Intake quiz:** Students take a multiple-choice quiz that asks about their academic and personal needs, interests, and preferences. This creates a a holistic profile replete with information that is not always considered when typically recommending degree options.

   Intake quiz questions:

   - Where are you in your educational journey?
   - What type of degree are you seeking?
   - How would you like to pursue your undergraduate degree?
   - What are you interested in?
   - What about "insert interest area" are you interested in?
   - What are some of your hobbies?
   - How do you want to work in the future?

2. **Degree pairing**: The recommendation engine uses semantic matching capabilities and conducts a similarity search based on user inputs to suggest three suitable degree options.
3. **Degree explanation**: The engine then retrieves degree information from public degree pages to generate a personalized explanation for each suggested degree option, detailing why it is a good match for the student based on their quiz responses.
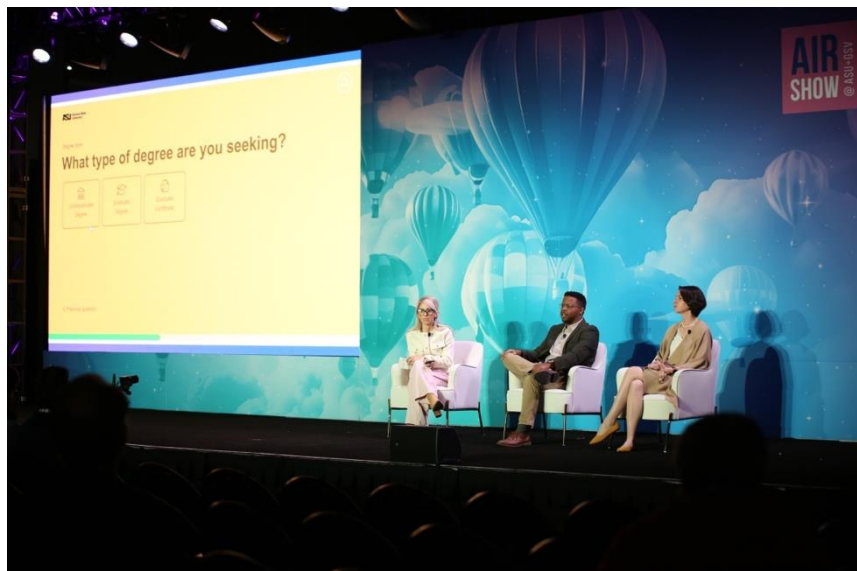
Arizona State University (cont.)

Development and Implementation Details

- The development of the proof-of-concept AI degree recommendation engine took a total of six weeks.

- Two software developers from ASU's development team created the back-end for the degree recommendation engine within a couple of weeks.

  – They first gathered and cleaned information from the university's degree pages, which include general degree descriptions written by the content team. They then used this information with a RAG[1] framework to enhance the AI engine's degree recommendations.

  – The development team chose ChatGPT-3.5 as the foundation for the engine, due to lower token costs compared to GPT-4. The team also employed multi-shot prompting to further increase accuracy and emphasize a certain tone of voice.

- Developers designed the intake quiz so that inputs submitted by students are transferred to a single HTML[2] PDF[3] and sent to ChatGPT-3.5 on the back end. The development team opted for this model keep token costs under control.

  – The team chose a multiple-choice quiz to minimize the risk open-ended text inputs could pose. Having only predetermined inputs reduces the risk of the degree recommendation engine responding to something inappropriately.

- A team of user experience researchers built the front-end user interface (e.g., animations, interface) in a month.

- The degree engine underwent user testing with a group of students and internal stakeholders, including support services like career services.

---

eab.com

University of California, San Diego (UC San Diego)

UC San Diego Develops TritonGPT, a Suite of AI Assistants

Type of Institution: Public Research University

Location: La Jolla, San Diego, California

Enrollment: 33,096 Undergraduate, 8,910 Postgraduate

Operating Expenditures (In Thousands): USD 7,800,000

UC San Diego

Application Overview

UC San Diego developed TritonGPT, a suite of AI assistants, to provide general assistance for tasks like summarizing documents and specialized assistance for university-specific questions and tasks such as navigating UC San Diego policies and processes.

Application Functionality

*The following description of TritonGPT is found on UC San Diego's TritonGPT: AI powered support web page.

TritonGPT, a suite of AI Assistants, is adept at handling university-specific questions, creating tailored content, summarizing documents, and more, thanks to its training on extensive public-facing university information.

TritonGPT operates from the secure confines of the San Diego Supercomputer Center, built with Meta Llama 3—a dialogue-optimized, open-source large language model. This secure, on-premises hosting ensures UC San Diego's complete control over data usage and sharing.

What Can You Do With TritonGPT?

- **Ask UC San Diego Related Questions**: Pose questions like "What is the policy on employee travel reimbursement?" or "What are some good restaurants on campus?" TritonGPT will provide detailed and relevant information.
- **Content Generation**: Need help with content creation? Try commands like "Generate an outline for a presentation slide deck based on <insert topic>" or "Produce an email to thank my employees for <insert what you are thankful for>."
- **Document Summarization**: Copy and paste documents or articles related to UC San Diego, then ask TritonGPT to summarize the content. It's a time-saving feature for extracting key information.
- **Content Editing**: Utilize TritonGPT for editing and refining content related to UC San Diego. It's a valuable tool for polishing emails, reports, or any written material.
- **Seek feedback and suggestions**: TritonGPT can provide feedback and suggestions to help you improve your work processes and procedures. You can ask questions like "What are some ways I can improve my communication skills in the workplace?" or "Are there any suggestions for streamlining our team's workflow?"
- **Ask for Recommendations**: Seeking recommendations for UC San Diego events, study spots, or local hangouts? TritonGPT has you covered.

TritonGPT Task-Specific Assistants: TritonGPT consists of a suite of AI Assistants, and we are actively working on expanding this lineup.

- **UC San Diego Assistant:** UC San Diego related policy, process, and help documentation is spread out over a vast array of websites. The UC San Diego Assistant brings it all together by answering your questions directly. It is also great for incorporating UC San Diego context in generating new content and brainstorming ideas. Always reference the sources cited when relying upon its answers.
- **Job Description Helper:** TritonGPT includes a Job Description Helper that will streamline the job description creation process for hiring managers. Leveraging over 1,300 career tracks job standard templates, it uses a predefined flow that engages hiring managers in a dialogue, capturing the specific requirements of the job. The AI then crafts language that not only complies with established job card standards but also accurately reflects the unique characteristics of the position. This feature reduces the time and effort involved in drafting job descriptions, ensuring they are both precise and tailored to the individual needs of the role.

Source: Attridge, Margaret, "Campus-GPT: How 2 University of California Campuses Are Designing Their Own Specialized AI Tools," *Best Colleges*, April 24, 2024; Davidson, Nikki, "Meet TritonGPT: AI That Loves Tedious University Tasks", *Government Technology*, May 24, 2024; Prior, Shannon, "TritonGPT, shaping the future at UC San Diego through student innovation", *UC IT Blog*, May 30, 2024; UC San Diego, "TritonGPT: AI powered support", *UC San Diego*, May 24, 2024; EAB interviews and analysis.

University of California, San Diego (cont.)

Application Functionality (cont.)

The following description of TritonGPT is found on UC San Diego's TritonGPT: AI powered support web page.

- **General AI Assistant:** This tool expands beyond UC San Diego's scope, accommodating larger information exchanges. It interacts with a Large Language Model for tasks like document summarization, idea generation, and creating various content such as emails and reports.
- **Fund Manager Coach:** Recognizing the crucial role of Fund Managers in overseeing grants and managing departmental finances, this assistant will enhance understanding of UC San Diego's financial policies and procedures. Fund Manager Coach is trained on documentation for developing research proposal budgets, advising faculty on contract and grant guidelines, reviewing and approving financial transactions, managing payroll, and ensuring that applicable guidelines are being followed during the course of contract and grant spending.

TritonGPT's Top Tasks: How UC San Diego Staff Use AI

Explore the most common ways UC San Diego staff are leveraging TritonGPT to streamline tasks and enhance their work, May 2024

Task	Percentage
Summarizing or rewriting text	17%
Brainstorming	14%
Other	11%
Queries about policies/procedures	10%
Job description writing	10%
Career development questions	9%
Queries about events and depts.	8%
Queries about financial reporting/procedures	7%
IT questions	6%
Queries about benefits, comp., and hiring	3%
Queries about investment and retirement	3%
Programming and code production	2%

Source: Attridge, Margaret, "Campus-GPT: How 2 University of California Campuses Are Designing Their Own Specialized AI Tools," *Best Colleges*, April 24, 2024; Davidson, Nikki, "Meet TritonGPT: AI That Loves Tedious University Tasks", *Government Technology*, May 24, 2024; Prior, Shannon, "TritonGPT, shaping the future at UC San Diego through student innovation", *UC IT Blog*, May 30, 2024; UC San Diego, "TritonGPT: AI powered support", *UC San Diego*, May 24, 2024; EAB interviews and analysis.

Retrieval-Augmented Generation.

University of California, San Diego (cont.)

Executive Sponsors

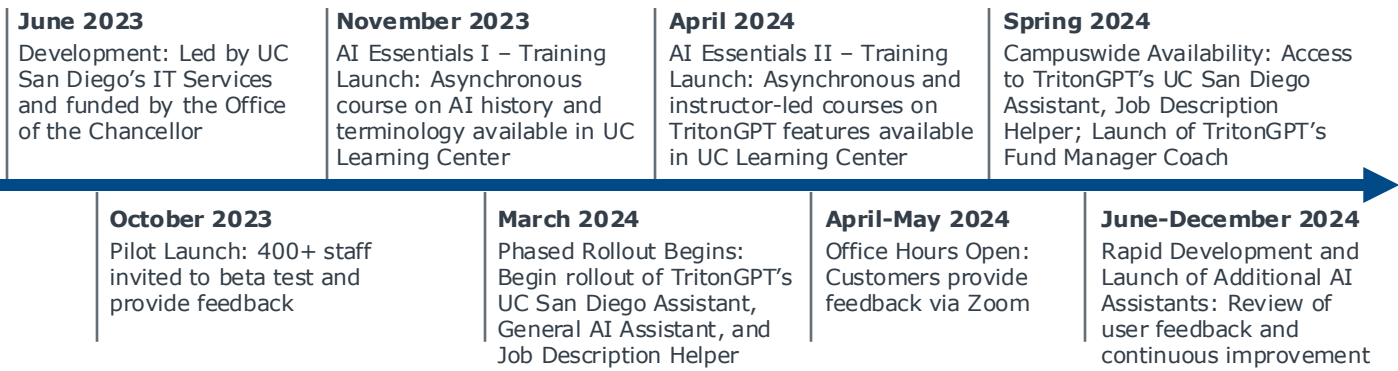- Pradeep Khosla, Chancellor
- Vince Kellen, Chief Information Officer

Key Advocates:

- Brett Pollak, Senior Director, Workplace Technology Services
- Antonio Nava, Senior Program Manager, Office of Operational Strategic Initiatives

Application Development:

- Adam Tilghman, Principal Architect, Academic Technology Services
- Jeremy Wiles, User Experience Architect
- Jack Brzezinski, Senior AI Architect
- Matthew Holland, ITS Student Engineer

Project Timeline

June 2023
Development: Led by UC San Diego's IT Services and funded by the Office of the Chancellor

October 2023
Pilot Launch: 400+ staff invited to beta test and provide feedback

November 2023
AI Essentials I – Training Launch: Asynchronous course on AI history and terminology available in UC Learning Center

March 2024
Phased Rollout Begins: Begin rollout of TritonGPT's UC San Diego Assistant, General AI Assistant, and Job Description Helper

April 2024
AI Essentials II – Training Launch: Asynchronous and instructor-led courses on TritonGPT features available in UC Learning Center

April-May 2024
Office Hours Open: Customers provide feedback via Zoom

Spring 2024
Campuswide Availability: Access to TritonGPT's UC San Diego Assistant, Job Description Helper; Launch of TritonGPT's Fund Manager Coach

June-December 2024
Rapid Development and Launch of Additional AI Assistants: Review of user feedback and continuous improvement

Development and Implementation Details

Application Development

- UC San Diego's IT Services led development of TritonGPT, from idea to proof-of-concept, in less than two months. IT Services allocated four FTEs for development, including a senior AI architect, user experience architect, and four student IT interns serving as float engineers. The initiative was funded by the Office of the Chancellor.

    – UC San Diego had a head start developing TritonGPT because they already had a relationship with the San Diego Supercomputer Center (where TritonGPT is now hosted locally) and were experimenting with AI platforms and even running clusters for students. As such, they ran early TritonGPT pilots on existing clusters, before purchasing more hardware.

- UC San Diego's Lean Six Sigma program informed the administrative use cases IT Services first developed in the form of the task-specific AI assistants (e.g., UC San Diego Assistant, Fund Manager Coach). The Lean Six Sigma governance group fielded requests from the community and selected the best use cases to develop. They are creating a similar process for academic requests.

- TritonGPT is powered by Meta's Llama 3 open-source large language model (LLM) and leverages the Danswer framework, an open-source project founded by two UC San Diego alumni, which lets organizations connect their tools and documents to LLMs. It is also integrated with the institution's single sign-on (SSO). TritonGPT initially used Llama 2 but upgraded to Llama 3 in April 2024.

    – UC San Diego's IT Services chose an open-source model to keep costs affordable. The CIO mentioned that this structure costs roughly one-tenth of the stated prices for ChatGPT Enterprise.

- IT indexed UC San Diego's key websites to enhance TritonGPT using RAG[1], enabling it to provide information and support related to the university's policies, procedures, and campus life with accuracy and efficiency. All the website content is ingested nightly, so TritonGPT has access to the most updated information. To do so, UC San Diego's IT services uses the Vespa AI vector database.

- UC San Diego is now tracking TritonGPT usage and applying reinforcement learning by allowing users to rate TritonGPT's answers using the thumbs-up or thumbs-down graphics that appear alongside its answers. This helps improve the tool for users asking similar questions in the future.

Source: Attridge, Margaret, "Campus-GPT: How 2 University of California Campuses Are Designing Their Own Specialized AI Tools," *Best Colleges*, April 24, 2024; Davidson, Nikki, "Meet TritonGPT: AI That Loves Tedious University Tasks", *Government Technology*, May 24, 2024; Prior, Shannon, "TritonGPT, shaping the future at UC San Diego through student innovation", *UC IT Blog*, May 30, 2024; UC San Diego, "TritonGPT: AI powered support", *UC San Diego*, May 24, 2024; EAB interviews and analysis.

Retrieval-Augmented Generation.

Development and Implementation Details (cont.)

TritonGPT Rollout

- In October 2023, UC San Diego launches a two-month TritonGPT early user program with over 400 health sciences and campus staff pilot users to test and source feedback on TritonGPT. These users were selected by nomination.
  - Office of Operational Strategic Initiatives hosts initial training sessions for early users. UC San Diego now offers online training modules that teach campus users how to use TritonGPT as well as core AI concepts.
- In March 2024, TritonGPT's "second wave" phased rollout begins in the Vice Chancellor-Chief Financial Officer area. During this phase, access is progressively granted to larger cohorts of campus employees. UC San Diego also begins rollout of TritonGPT's task-specific assistants like the Job Description Helper and UC San Diego Assistant.
- In May 2024, TritonGPT is rolled out to all 37,000 campus employees.
  - TritonGPT usage on campus is currently following a roughly 80/20 percent Pareto distribution, where 20 percent of campus users are responsible for 80 percent of TritonGPT queries.
- In the future, UC San Diego plans to roll out TritonGPT to the general student population.

Future Roadmap

UC San Diego IT Services plans to add the following functionality to TritonGPT:

- Enabling users to create their own GPTs with prompt engineering
- Uploading files and documents to TritonGPT's context window
- Building knowledge graph capabilities
- Supporting text to SQL so that TritonGPT can directly read out of activity hubs like grant financial activity

Sources of Data

- Academic Personnel website
- Admissions website
- Blink
- Business Analytics Hub
- Calendar of Events
- Career Center

- Chancellor website
- The Commons
- Course Catalog
- Educational Technology
- Foundation
- Housing and Dining

- Policies (UC San Diego and UCOP)
- ServiceNow Knowledge Base content (public facing)
- Strategic Plan
- Student Financial Solutions
- Transportation

- TritonLink
- UC Path website
- UC San Diego Brand
- University Centers
- University Communications
- UC San Diego Today

Source: Attridge, Margaret, "Campus-GPT: How 2 University of California Campuses Are Designing Their Own Specialized AI Tools," *Best Colleges*, April 24, 2024; Davidson, Nikki, "Meet TritonGPT: AI That Loves Tedious University Tasks", *Government Technology*, May 24, 2024; Prior, Shannon, "TritonGPT, shaping the future at UC San Diego through student innovation", *UC IT Blog*, May 30, 2024; UC San Diego, "TritonGPT: AI powered support", *UC San Diego*, May 24, 2024; EAB interviews and analysis.

New York University (NYU)

NYU Builds AI Chatbots for Students and Faculty Researchers

Type of Institution: Large Private Research University

Location: New York, Abu Dhabi, and Shanghai

Enrollment: 61,950 Total; 29,760 Undergraduates, 27,575 Postgraduates

Operating Expenditures (In Thousands) : USD 3,400,000 (FY 2020)
Washington Square campus

Application Overview

NYU IT, along with various Administrative units, is currently developing a suite of AI applications. Their goal is to provide an enterprise platform to enable administrative, instruction, and research communities within NYU to adopt generative AI (GenAI). They plan to release the following two domain-specific applications in Fall 2024:
- Enrollment Management **Student Success Chatbot**
- Research Finance Support **Allowable Costs Chatbot**

In Fall 2023, NYU IT developed the NYU RIT projects portal, which is being used as the Enterprise-Level GenAI Adoption Platform. This platform allows users to test their approved GenAI-related project ideas using text-based LLMs (large language models). These LLMs are sourced through cloud service providers under contract with NYU. This platform enables NYU to track metrics like usage and cost for the GenAI projects.

Application Functionality

The **Student Success Chatbot** is designed to assist student success staff in addressing inquiries from current students about various aspects of the NYU student experience, such as academic policies and administrative processes. The current objective of the chatbot is to help staff quickly and efficiently answering student questions. However, after thorough testing, the goal is to integrate the chatbot directly into NYU's website making it easily available for NYU student to interact with it directly.

Sample Questions:

- *"What are the steps for students to get their GI bill benefits certified?"*
- *"Where can NYU undergraduate students find and make an appointment with my academic advisor?"*
- *"What academic support resources are available at NYU?"*
- *"What steps are required to return from a Leave of Absence?"*
- *"Who can I speak with regarding my Financial Aid award letter?"*

The **Allowable Costs Chatbot** will answer questions faculty researchers, research department employees, central office administrators, or grant financial analysts may have about grant funding regulations for allowable costs. It is designed to simplify and ensure compliance with funding regulations.

Currently, a very small team of Research Finance Support (RFS) staff are responsible for answering questions related to research grants and allowable expenses during business hours for over 2,000 researchers at NYU. The chatbot will be integrated directly into the NYU RFS website, making it easily accessible to NYU faculty researchers.

Sample Questions:
- *"For research projects, what is the allowable maximum business meal cost per person?"*
- *"Can I use these funds to hire a federal work study student through the undergraduate research assistant program out of the Provost's Office?"*
- *"Can the grant cover travel expenses and/or conference registration for doctoral students or affiliate researchers?"*
- *"Are vacation payout and one-time bonus personnel costs allowable on sponsored research projects?"*
- *"Is Capital Equipment - 62300 - allowable on this project without prior approval? If approval is necessary, what are the next steps?"*
- *"Can I use Airbnb instead of a hotel?"*

New York University (cont.)

Executive Sponsor

Don Welch, CIO, NYU IT

Project Governance

Beth Davidovich, AVP, NYU IT, Institutional Solutions Group

Testing Team Leads

- John Burdick, Student Success, Enrollment Management
- Meidy Candia, Research Finance Support

Testing Team

- Trevor Ban, and Torrey Crosby: Student Success, Enrollment Management
- Josel Stoute, and Emma Xinyuan Zhang: Allowable Costs, Research Finance Support

GenAI Adoption Program Team

- Meenakshi Baker, Executive Director, Strategic Initiatives and Planning (NYU GenAI Adoption Program Lead)
- Sergey Samsonau, Artificial Intelligence Technical Lead (NYU GenAI Adoption Technical Lead)
- Utku Ege Tuluk, and Adrian Hodge: Shanghai Libraries (GenAI Chatbot Application Development Team)
- Andi Shehu, and Hazem Lashen: NYU IT Research Technology Services, (Enterprise GenAI Adoption Tool Support and Development)
- Stratos Efstathiadis, Carl Evans, Minh Lai, Taidi Chen, and Sophia Eller: NYU IT Research Technology Services (Enterprise GenAI Adoption Tool and Application Infrastructure)

Project Status

Development of the two domain-specific AI applications is ongoing. The project team is currently in Phase 2 outlined below, testing the LLMs (*last updated July 2024*).

Development and Implementation Details

NYU's project team is following the same core process outlined below to develop both domain-specific applications. They will continue to follow this general outline as they potentially develop more chatbot applications powered by LLMs in the future.

Relatedly, the project team is also identifying how to enhance the Enterprise-Level GenAI Adoption Platform with new features through discussions with the Provost Office and is creating a plan for development, testing, and implementation. Currently, the tool supports Chat, Batch, RAG[1], specific profile, and download of chat features and functions.

Phase 1: Proposal Intake and Selection (~1 month)

Since last year, business units across campus have approached the GenAI Adoption Program team with AI application ideas to develop. The adoption team evaluates and selects proposals to pursue. Once a proposal is selected, the Program Lead identifies the application development lead and business team members and secures approval for the application's development from the university's security, privacy, and enterprise architecture teams. The Technical Lead provides technical guidance, code review, and approval for publishing the application on the Posit Connect platform. The application development and LLM testing using the Enterprise GenAI Adoption Tool occurs simultaneously.

Phase 2: Application Development, Testing, and Survey Administration (~4 months)

The project team begins a cycle of iteratively developing the application and testing the LLM they are building the application on. They created the GaRAGe framework (Generalized RAG[1] Framework) for GenAI applications that guides the development of both the student success and allowable costs chatbots.

GaRAGe Framework:

1. Able to give accurate and consistent answers to prompts that have been designed from a baseline set of tested and documented questions
2. A wrapper provider for prompt engineering and system settings for custom applications
3. Developed for and hosted in Posit Connect
4. Consisting of applications that are tested with 20-50 users and optimized according to user feedback
5. Able to support multiple user roles with access to the prompts, documents

Both of NYU's AI applications use the same LLM. They use ChatGPT-3.5 because it is cost-effective and allows units to spend more on driving adoption, but the team is considering using Meta's open-source Llama 3 in the future. Both applications are also being developed using Python Streamlit and will be available through Posit Connect, a hosting platform for R and Python applications. The primary difference in the development of the AI applications lies in the data sources used to train the LLM, which utilizes RAG.

1) Retrieval-Augmented Generation.

New York University (cont.)

- **Student Success Chatbot**: Enrollment Management provided a 600-page NYU bulletin that contains detailed information relevant to all 26 schools within the university. This unified resource effectively addresses a broad spectrum of student needs. The project team chose not to have the LLM scrape public websites because they found that outdated website information led the LLM to generate inaccurate responses. On the other hand, the NYU bulletin is updated consistently every year.

- **Allowable Costs Chatbot:** The project team is training the LLM on a set of documents defining uniform conduct guidelines for research grants. This also includes a several hundred pages long sponsored project administration (SPA) handbook.

The testing team administers a survey to a select group of their clients, asking them the top five questions they would want the chatbot application to answer accurately. The testing team takes these questions and creates variations on which to test the LLM. The testers test the LLM independently and rate its responses on consistency, accuracy, and confidence and review the scores with the survey takers. The testing team works with the Gen AI Adoption Program leads to determine the parameters and configurations that should be implemented for the LLM in the application. Testing is eventually expanded to a larger number of users in phases who can test the LLM by simply applying a thumbs-up or thumbs-down feedback metric to its responses through the chatbot application. The project team will use this information to further augment the LLM by updating content or changing the parameters.

- **Student Success Chatbot:** The testing team (three Enrollment Management staff) reached out to nine student success support staff members with a survey that asked them to list the top five questions they commonly received from students. The support team shared 45 questions and subsequently recorded how well the LLM responded to them below. Contingent on the success of this testing stage, the testing team will expand testing to more support staff within Enrollment Management.

- **Allowable Costs Chatbot:** The testing team (three Research Finance Support staff) administered a survey to 13 Research Finance Support staff, asking them the top five questions they would want the chatbot application to answer accurately about post-grant allowable expenses. Contingent on the success of this testing stage, the test team will expand testing to more Research Finance Support staff and departmental staff within Schools.

Phase 3: Final Adjustments and Launch Preparation (~2 months)

The final stage before deployment lasts between 45 to 60 days, focusing on making final adjustments based on comprehensive testing outcomes. Business units must agree with terms of use outlined by the project team that dictate how they can use the application and their obligations to administering. A RACI matrix is created for distribution of administration and technical tasks.

Project Cost and Financing

NYU's CIO allocated $106,400 for Fiscal Year 2024 (September 1, 2023–August 31, 2024) to hire student and consultant staff, acquire licenses for software infrastructure to support R and Python applications, and pay for GenAI LLM usage.

NYU IT is using these funds to develop GenAI infrastructure like the NYU Enterprise-Level GenAI Adoption Platform for testing GenAI LLMs using RAG[1], Posit Connect to host GenAI applications, and API[2] development. They are developing the domain-specific applications in partnership with NYU Shanghai Libraries. The NYU IT Research and Instructional Technology unit is absorbing additional costs for the infrastructure to accommodate Posit Connect and on-prem LLM hosting.

For the domain-specific applications, the business unit that the application is being developed for provides ongoing testing and maintenance resources. NYU IT will provide resources for technical tasks like upgrading application components and infrastructure. The business unit will fund any LLM usage costs for the application. Any IT infrastructure costs related to the applications will be added to the annual NYU IT budget from central university funds.

New York University (cont.)
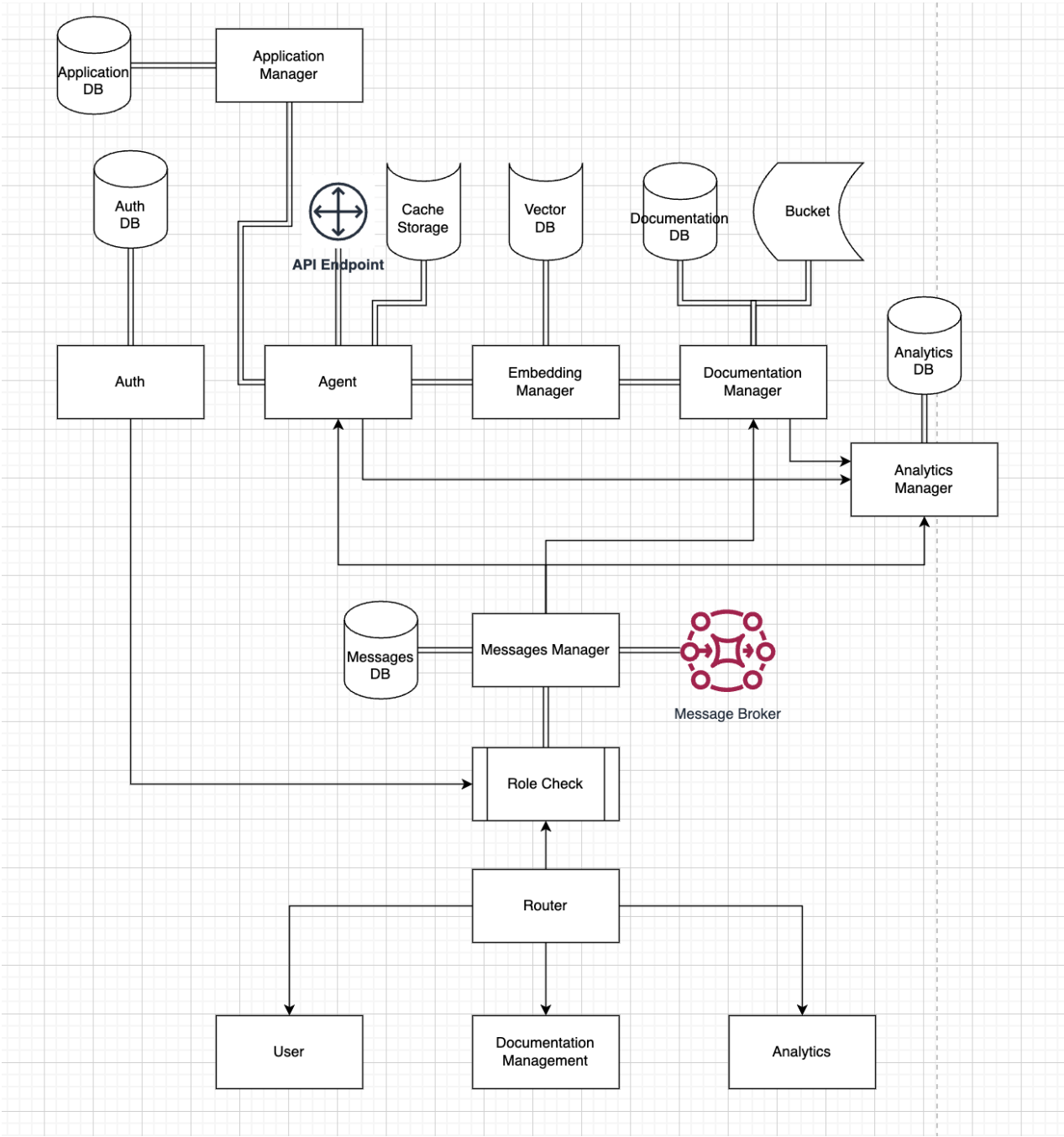
York University's AURA Platform

AURA Empowers Campus Users with Customizable AI Agents

Type of Institution: Large Research Institution

Location: Toronto, Ontario

Enrollment: 53,100 students

FY26 Operating Expenditures (In Thousands): CAD 1,006,500

Application Overview

In 2023, York University developed its **Automated University Response Assistant (YU AURA),** a platform that enables campus users to customize and deploy AI-powered agents developed by York University's IT (UIT). AURA currently supports the following key AI agents across campus, among many others:

- **SAVY**: Serves as a bilingual student virtual assistant
- **AI Course Assistant**: Connects students with course resources and tutors them on course content
- **Research GPT**: Supports York researchers literature review and knowledge sharing
- **OUR Assistant**: Responds to advising inquiries over email, reducing administrative workload for advising staff
- **Procurement Assistant**: Guides staff through procurement processes
- **Leonie**: Answers prospective student inquiries, supporting recruitment workload

The IT Innovation and Academic Technologies team developed AURA to exploit the potential of combining AI with York's data while strengthening IT and campus capacity for AI innovation. In the future York intends to complement AURA with other AI platforms, as they look to automate more complicated AI workflows.

Application Functionality

YU AURA is a no-code platform that enables York users to customize and deploy AI agents. UIT provisions instances of AURA to users, which also allows users to develop AI agents from ideation to deployment. AI agents on AURA use the latest GPT models and can be trained on York data and engineered to serve a particular purpose.

For UIT

When developing agents, AURA enables UIT to change out and experiment with LLMs, embeddings frameworks, reranking models, in a streamlined interface. AURA supports LLMs from a variety of providers (e.g., Azure OpenAI, Google, Amazon Web Services). UIT can also monitor and augment agents even after campus users have deployed them.

For Campus Users

Users who want to customize an AI agent can request access to AURA and its associated agents from UIT. Once granted approval, users can train models with their own data sources using a drag-and-drop feature. All front-end features like colour, logo, and placement on the website are customizable. Users can also control model behavior by setting their own system message (e.g., meta prompt or high-level behavioral instructions like response tone, voice, and format). These agents can then be added to York websites, systems, and applications. For example, UIT developed AI course assistants that faculty can customize and train with their own course content and deploy within their course learning management system.

Pages 2-3 outline the AI agents created on the AURA platform and deployed on campus.

York University's AURA Platform – AI Agents

AURA Agent: SAVY Student Virtual Assistant

SAVY is a bilingual AI student virtual assistant available to all York students. SAVY serves as students' first point of contact at York providing personalized support to help them navigate the academic and student experience. Students can ask SAVY questions on a a range of topics, including:

- How to enroll, grades, and course and program changes
- Fees, scholarships and financial aid, including OSAP[1]
- Mental health and well-being
- Academic advising referrals
- Campus life and events
- Career advising information
- Campus services

AURA Agent: AI Course Assistant

The AI Course Assistant can answer questions about course logistics, tutor students on course content, and even assist faculty as they design course materials.

1) Course logistics

Faculty can upload syllabi, Q&A files, and class notes to the course assistant so that it can answer course logistics questions like assignment value, assessment details, office location, and academic integrity.
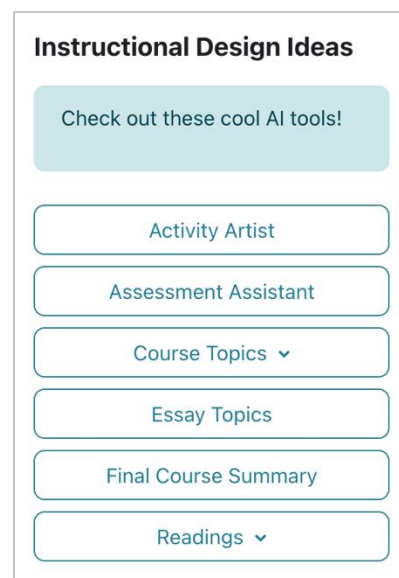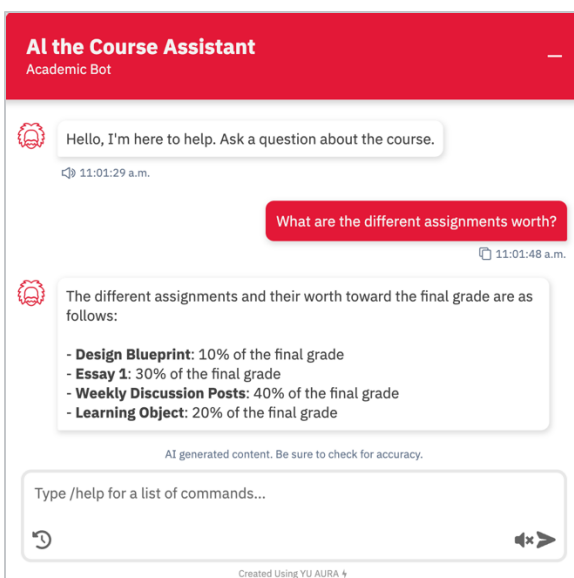
2) Tutor

The course assistant can create class notes based on uploaded documents and provide tutoring assistance. For example, students can select specific course materials and ask the tutor to coach them on topics or even generate practice quizzes. The assistant can even explain to students why answers they selected were wrong.

3) Instructional design (yet to be deployed)

The instructional design addition to the course assistant will help faculty build learning outcomes, design assessments, suggest readings and topics, and generate course summaries.

The AI course assistant was initially piloted in two courses, but now over 30 faculty have customized and deployed their own AI course assistants.

The course assistant is accessible on York's learning management system, Moodle. It also includes features such as text-to-speech and conversation export to PDF.

AI the Course Assistant
Academic Bot

Hello, I'm here to help. Ask a question about the course.
11:01:29 a.m.

What are the different assignments worth?
11:01:48 a.m.

The different assignments and their worth toward the final grade are as follows:

- **Design Blueprint**: 10% of the final grade
- **Essay 1**: 30% of the final grade
- **Weekly Discussion Posts**: 40% of the final grade
- **Learning Object**: 20% of the final grade

AI generated content. Be sure to check for accuracy.

Type /help for a list of commands...

Created Using YU AURA

Instructional Design Ideas

Check out these cool AI tools!

Activity Artist

Assessment Assistant

Course Topics

Essay Topics

Final Course Summary

Readings

1) Ontario Student Assistance Program.

AURA Agent: **Research GPT**

Faculty can upload research (e.g., documents, journal articles, and research notes) to create custom agents that serve as expert knowledge sources on specific topics, as all AURA data is securely stored on York servers. Faculty can use Research GPTs as learning and thought partners. For example, Research GPTs can more accurately assist with literature reviews or with testing ideas, given their baseline knowledge on a topic.

Research GPTs can also facilitate collaboration with other researchers and faculty. Faculty can share their Research GPT with researchers in and outside of York, allowing others to interact with their research and even expand its knowledge base.

AURA Agent: **OUR Assistant**

The OUR assistant automatically answers email queries the Registrar's office receives, reducing the administrative workload for advising staff. OUR assistant helps the 14-person office answer the 55,000+ questions it receives over email annually.

The OUR chat assistant, which is built on AURA, is added to Microsoft Outlook as an authorized user. This allows OUR to receive and automatically answer emails through Outlook. OUR was trained on the Registrar's annually updated knowledge base.

AURA Agent: **Procurement Assistant**

The Procurement Assistant enables users to access accurate, real-time information about purchasing processes, vendor contracts, and ordering instructions. The assistant is trained on York-specific procurement data and integrates securely with internal systems, ensuring privacy and reliability.

AURA Agent: **Leonie**

Leonie is a virtual chatbot that can be deployed on individual college or campus websites that answers prospective student questions such as what programs are available or how to apply for that college or school.

One college and one campus currently host customized versions of Leonie. For example, the Glendon campus customized Leonie to be bilingual and respond to prospective students with campus-specific answers.
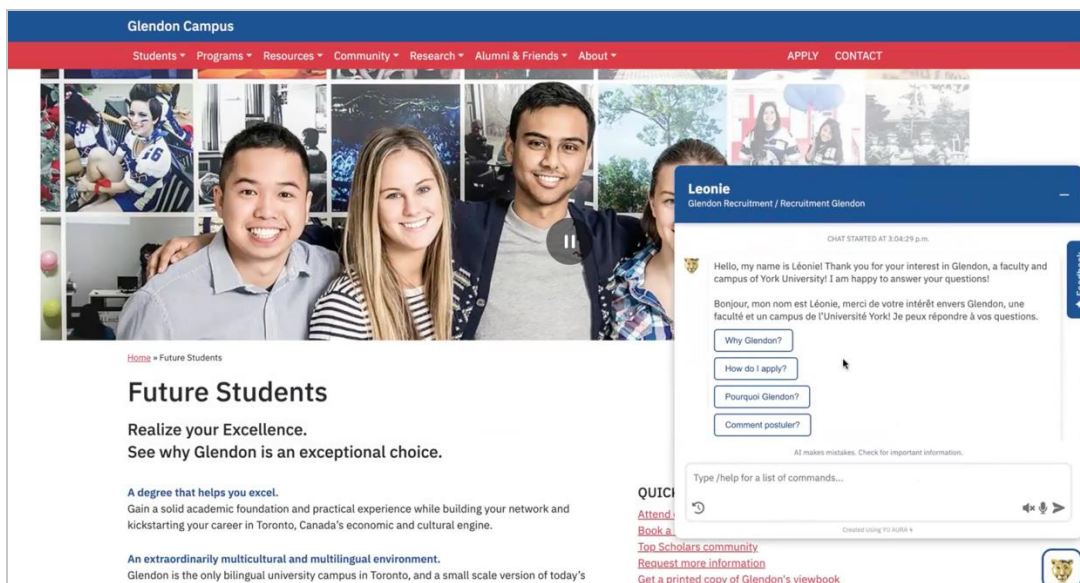
How York University Developed AURA

Brad Strom
York University Chief Information Officer

Patrick Thibaudeau
Director, IT Innovation and Academic Technologies

Vidur Kalive
AI Architecture Lead

Isaac Kogan
Student, IT Innovation Team

Development and Implementation Details

April 2023: Starting to Experiment

- York's CIO and three-person IT Innovation Team begins experimenting with AI tools to bring to campus to showcase AI's utility and potential.
  - York had previously worked with IBM Watson to build SAVY, a virtual student chatbot assistant, dating back to 2018. The team wanted to add generative AI capabilities to SAVY, as IBM Watson incurred high costs.

July 2023: Forming a Vision for AURA

- York receives access to a variety of large language models (LLMs) like GPT-4 through the University of South California's (USC) Azure OpenAI service.
- With this access, the IT Innovation Team's initials goals are to:
  - Develop a low-code/no-code platform that anyone can use to create a custom AI agent for themselves
  - Create an AI course assistant to support instructors
  - Transfer over Savy from IBM Watson to Azure's OpenAI service
  - Explore other types of AI-powered agents to create for campus consumption

November 2023: Developing the AURA Platform

- The IT Innovation Team builds the prototype AURA platform in three months, primarily employing open-source frameworks.
  - IT develops AURA within the Moodle LMS environment. IT uses Moodle as the primary platform for interface design; it also provides built-in authentication.
  - The team creates a custom RAG[1] system able to index institutional data, so that LLMs provide grounded and York-specific responses.
  - AURA uses LlamaIndex for indexing and vectorizing data. Quadrant and MySQL are employed as vector databases. The team also builds a custom document parser that uses headers and nodes to ensure tables are properly linked and translated into well-structured JSON data.
  - All AURA data is stored securely on Microsoft Azure, through USC's deployment of OpenAI services.
  - The team spends this entire period iteratively comparing their process for uploading documents into AURA against vendor tools (e.g., Microsoft Copilot) to see which produces the best results. They continuously modify AURA so it performs better than vendors at this point.

- *Agent-specific development on the AURA platform*:
  - *SAVY*: IT transfers SAVY from IBM importing all data into AURA. IT also works with student services to gather documents to further train SAVY with.
  - *AI Course Assistant*: IT prioritizes the AI Course Assistant because of York's strength and emphasis on teaching and learning. The team is also experienced in partnering with teaching and learning and working in the LMS (Moodle). They first present a proof-of-concept to the Vice Provost of Teaching and Learning and then the Teaching Commons, garnering support before showcasing it widely at a campus teaching and learning conference. IT initially tests the assistant with two faculty, then slowly expands testing to 30 or so other instructors that are granted access.

Development and Implementation Details (Cont.)

   – *Miscellaneous agents*: IT develops a variety of other agents (e.g., Research GPT, OUR Assistant, Procurement Optimization, and Leonie), which go through varying levels of development and testing (but none were as labor-intensive as SAVY and the AI Course Assistant).

August 2024: Launching the AURA Platform

- AURA enters into production; campus members can now widely request access
- SAVY and the AI Course Assistant launch

September 2025: Supporting AURA and Evaluating Next Steps

- IT continues to build and scale AI agents, supporting campus experimentation
  - *Agent-specific development on the AURA platform*:
    - The team is evaluating features to add to the AI Course Assistant, informed by instructor and student feedback, and in close collaboration with the Vice Provost of Teaching and Learning.
- The CIO is not tied to indefinitely supporting AURA and is evaluating future options. However, he wants to scale up the teaching and learning-focused AURA agents at a baseline.

Project Cost and Financing

- AURU initially cost $400,000 to develop; now, it only costs 2,500 operate annually. Responding to 20,000+ questions asked over the course of AURA's testing period only cost York $100.
- UIT currently covers usage with their budget and hopes to continue to provide AURA for free. However, UIT does utilize an internal tracking system for token usage and each custom AI agent deployed is assigned to a cost center. Therefore, if a department employs an agent that consumes considerable tokens, IT can request the department cover additional costs.

University of Birmingham and StudyStash Partnership

Birmingham Pilots StudyStash, an AI-Powered Personalized Learning Platform

Type of Institution: Large Research Institution

Location: Birmingham, United Kingdom

Enrollment: 38,000 Students

FY26 Operating Expenditures (In Thousands): GBP 584,971

Case-in-Brief

Two undergraduate students from the University of Birmingham developed an AI-powered personalized student learning and faculty insights platform called StudyStash, kickstarted by Birmingham's Elevate incubator programme.

Birmingham piloted StudyStash within the Computer Science department in 2023 and eventually institution-wide. The pilot found higher student achievement and course completion. Today, StudyStash is an incorporated company deployed at eight higher education institutions (including Birmingham).

AI Application Overview

StudyStash is an edtech platform that uses generative AI and neuroscience-based learning algorithms to convert course content into learning aids and assessments such as flashcards, short videos, virtual tutors, and personalized tests. StudyStash personalizes content based on real student performance to prioritize the concepts students uniquely need help with. StudyStash also reduces faculty workload and provides them with insights into students' understanding of course content, generating reports that detail where students are struggling and recommend targeted questions for future lectures or assignments.

University of Birmingham's StudyStash Timeline

2023-2024
- Two Computer Science undergraduate students develop StudyStash with support from Birmingham's Elevate programme, a student entrepreneurship incubator. Elevate provides initial funding and networking contacts for the StudyStash founders.
- The StudyStash founders pitch the platform to Birmingham leadership and launch a pilot in Computer Science with 730 students.
- StudyStash hosts demonstration sessions with faculty and students to introduce the platform and encourage adoption.

2024-2025:
- Birmingham's Pro-Vice Chancellor for Education and Deputy Pro-Vice Chancellor for Education Innovation connect the StudyStash founders with contacts in other faculties, eventually expanding the pilot university-wide. The pilot grows to over 5,000 students across 25 courses and all five colleges.
- The Deputy Pro Vice Chancellor for Education undertakes an early-stage evaluation of pilot efforts, conducting a mixed-methods evaluation (i.e., focus groups, surveys, usage and assessment data). The pilot and initial evaluation assess the impact of StudyStash on student attainment, satisfaction, confidence, and stress reduction.

2025:
- The pilot concludes at the end of academic year, but Birmingham students retain access to StudyStash. They are now in negotiations with Birmingham to continue and expand StudyStash use on campus.

Source: EAB interviews and analysis.

StudyStash Pilot Results and Platform Functionality

StudyStash One-Year Pilot Results

StudyStash surveyed faculty and students throughout the duration of the pilot. The results:

Adoption
- 2,192 students voluntarily signed up for StudyStash
- 34% of students regularly used[1] StudyStash outside of lecture time (avg. 2 hours per week per active student)
- 5,207 documents uploaded to StudyStash

Teaching and Learning Outcomes
- 58.8% of students cited that StudyStash improved their academic performance
- 78% of students reported increased confidence managing study workload
- 54.5% of students felt more organized with their studies

Application Functionality

This section outlines the core features of the StudyStash platform for both students and faculty.

How StudyStash Helps Students

StudyStash converts course content into personalized, interactive notes by breaking content into smaller chunks, highlighting key concepts, and creating visual concept maps. To help prepare students for assessments, it generates flashcards, mini-games, and practice tests. The platform adapts to individual student needs by targeting areas of difficulty, optimizing content delivery based on exam schedules, and sending reminders to stay on track. Students can also interrogate course content with StudyChat AI, the platform's virtual chat assistant.

StudyStash Learning Dashboard

---

1) Regular use is defined as a user who has logged in on at least five unique days during the
   module's semester and created at least two unique resources (flashcards, notes, or tests).

Source: EAB interviews and analysis.

StudyStash Platform Functionality (cont.)

Application Functionality (cont.)

How StudyStash Helps Faculty

StudyStash generates weekly insight reports to help faculty understand where students need additional support. StudyChat AI also takes pressure off faculty by answering an average of 25+ student questions per week and directs students to additional resources (e.g., office hours) as appropriate.

StudyStash Report: What are Students Struggling With this Week?

Report highlights the topics and questions where students struggled, along with the number and percentage of students affected.
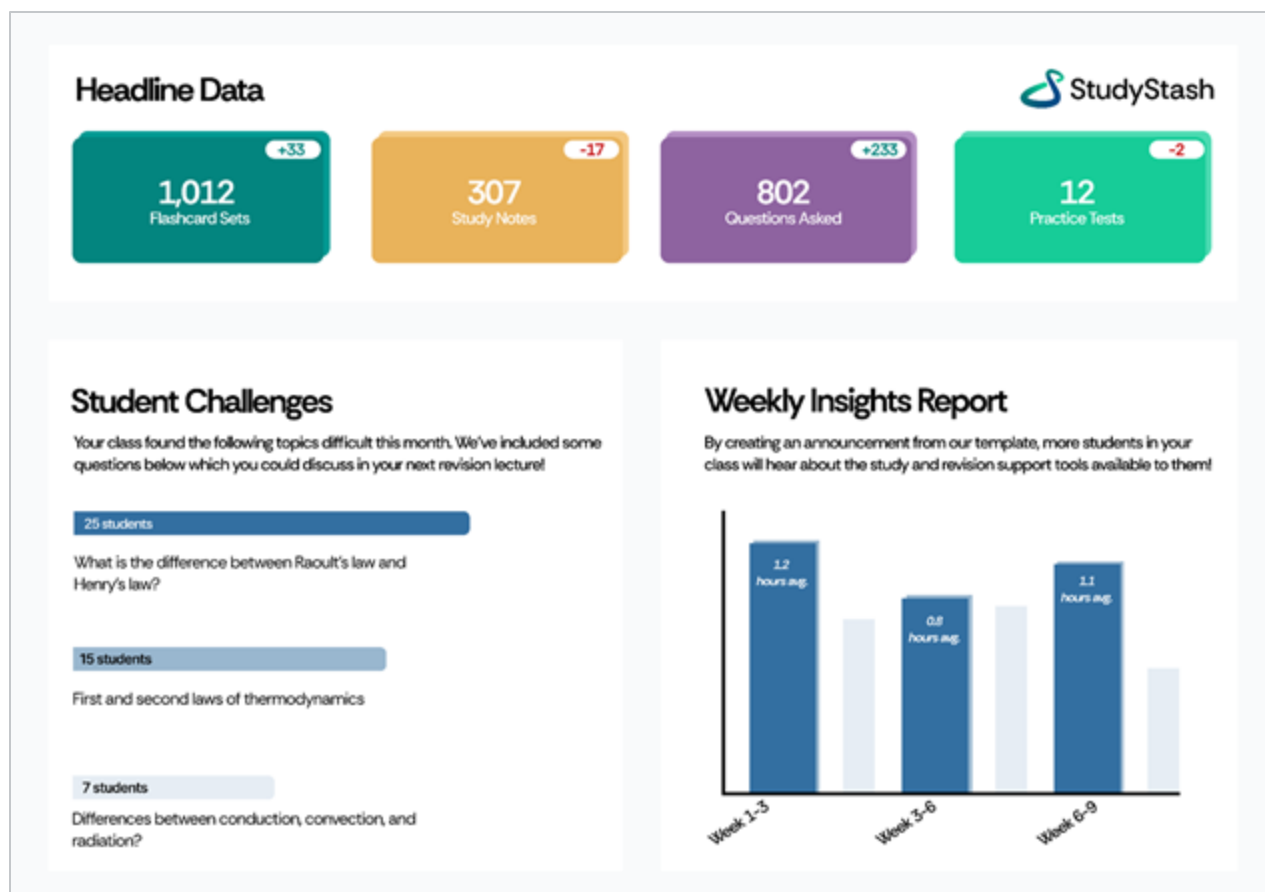
Developing the StudyStash Platform

Key Participants

Ben Ward
StudyStash CEO, Co-Founder
University of Birmingham Alumni

Jonathan Graham
StudyStash CTO, Co-Founder
University of Birmingham Alumni

AI Application Architecture

The StudyStash application consists of two main components:

1. **Learning algorithms** that dictate when and how course content is delivered based on research-based learning principles.
2. **Generative AI models** to automate content generation from course content from the university's learning management system (LMS).

Learning algorithms: StudyStash developed algorithms based on research-based learning principles that decide when, how often, and what type of content is shown. A selection of these underlying learning principles include:

- *Spaced repetition*: Tracking student mastery of materials to surface topics and schedule review sessions at the best intervals to drive retention.
- *Confidence-aware testing*: Adjusting assessment difficulty to target a student success rate of 60–80%, balancing need to motivate while continuously pushing students. This is called balancing test error rates.
- *Testing pedagogy*: Varying the order of concepts and question types based on pre-assessment data and course exam schedules.

Generative AI models: StudyStash uses LLMs[1] like GPT-4o via OpenAI's API[2] via enterprise agreement to generate content and chat features. StudyStash uses other models (like Gemini 2.5 Flash or GPT-4) depending on the task. LLM applications include:

- Creating flashcards, visual concept maps, tests, learning mnemonics, etc. from course content
- Generating student responses in chat (StudyChat AI), grounded in course content
- Turning any file (PDFs, PNGs) types into editable not

StudyStash Platform Process Explained: Turning Course Materials in Personalized Learning Aids

Step 1: Processing & Indexing Course Content

Course materials can be uploaded from the user's file system or through LMS deep-linking. Once a file (e.g., lecture slides, notes, PDFs) is uploaded, StudyStash:
- Chunks content into segments and generates embeddings for semantic understanding
- Stores information in a vector database to enable context-aware search

Step 2: Generating Responses Grounded in Course Content

StudyStash uses Retrieval-Augmented Generation (RAG) architecture enhanced by semantic search and grounding techniques to ensure LLM outputs are rooted in uploaded course content. The system retrieves contextually relevant excerpts from content (e.g., lecture notes, annotated slides, and diagrams) to construct grounded prompts.

Beyond content grounding, prompts are assembled based on a student's profile to personalize LLM outputs. This includes learning analytics such as topic-level retention curves, historical performance patterns across question types (e.g., multiple choice vs. open-ended reasoning), learning modality preferences, and metadata from tutor interactions (e.g., uncertainty expressions, hesitation patterns).

---

1) *Large language models.*
2) *Application programming interface.*

64   **eab.com**

Developing the StudyStash Platform

Step 3: Employing Learning Algorithms to Optimize Content Delivery

Learning algorithms operate independently of the LLM to decide the timing, format, and complexity of content delivery. For example, when the system detects that a student's retrieval strength for a topic has dropped below a threshold, it can issue an instruction to the LLM such as: "Generate 15 flashcards targeting weak points in Topic Y, emphasizing retrieval practice and cross-topic integration with Topic Z."

Integration and Accessibility

- Faculty and students can access StudyStash directly through the LMS, such as Canvas, Moodle, Blackboard and D2L Brightspace.
- The platform also includes accessibility features like screen reader support and keyboard navigation. For example, students can highlight any portion of a PDF directly within the platform and ask related questions in the same chat, eliminating the need to copy and paste large blocks of text.

Application Development and Investment

- £600K-£900K ($800K-$1.2M) for first application development cycle
  - Five types of developers (frontend/backend developers, DevOps, QA engineers, and machine learning developers)
  - Two years of product development and iteration with thousands of students from ten K-12 school districts.
  - Cloud infrastructure setup (e.g., Microsoft Azure), DevOps pipelines, and security provisioning
- Partnered with [EdCortex](#), EdTech consulting firm with specialist PhD neuroscientists, to develop algorithms to ground application in best practice learning techniques
- £370K ($500K) allocated for development over the next year (FY26)
  - New enhancements planned include adding video conversion and transcription, an improved mobile application, and extending the *LTI 1.3 integration* to include data collection for module and course completion outcomes.
- Tooling/licensing costs: £37K ($50K) budgeted over the next year (FY26)
- Ongoing compliance: £45K ($60K) yearly (*ISO27001*, *SOC 2*, *Cyber Essentials*, *AI compliance*) + additional for legal compliance
- The above costs are for the web application; there is a separate set of costs for developing and maintaining their mobile app on the App Store and Google Play.

Source: EAB interviews and analysis.

ABOUT EAB

At EAB, our mission is to make education smarter and our communities stronger. We work with thousands of institutions to drive transformative change through data-driven insights and best-in-class capabilities. From kindergarten to college to career, EAB partners with leaders and practitioners to accelerate progress and drive results across five major areas: enrollment, student success, institutional strategy, data analytics, and diversity, equity, and inclusion (DEI). We work with each partner differently, tailoring our portfolio of research, technology, and marketing and enrollment solutions to meet the unique needs of every leadership team, as well as the students and employees they serve. Learn more at eab.com.

39206