

Jill Watson: Scaling and Deploying an AI Conversational Agent in Online Classrooms*

Sandeep Kakar, Pratyusha Maiti, Karan Taneja, Alekhya Nandula, Gina Nguyen, Aiden Zhao, Vrinda Nandan, and Ashok Goel

Georgia Institute of Technology, Atlanta GA, USA
{skakar6, pmaiti6, ktaneja6, anandula3, gnguyen37, yzhao750, vrinda, ag25}@gatech.edu

Abstract. We present a highly performant and scalable conversational virtual teaching assistant, Jill Watson, to enhance teaching presence in online classrooms. Jill Watson leverages the generative AI capabilities of ChatGPT and the underlying OpenAI’s GPT large language models, along with dense passage retrieval and retrieval-augmented text generation to answer student questions about instructor-approved courseware anytime and anywhere. This courseware may include natural language text from textbooks, user guides, video lesson transcripts, course websites, and class syllabi. This version of Jill Watson has been deployed in several classes with close to 1300 students at the Georgia Institute of Technology as well as in two community colleges in the Technical College System of Georgia. We have found that Jill Watson enhances the positives of conversational courseware (such as answering questions and engaging in conversations anytime and anyplace) and suppresses the negatives of large language models (such as biases and hallucinations).

Keywords: Virtual Teaching Assistants · Intelligent Tutors · Conversational AI Agents · Conversational Courseware · Intelligent Textbooks.

1 Introduction

The growing popularity of online programs in higher and continuing education [1] has lent much urgency to reducing the gap between the learning experience of online and in-person learners. Online classes may have hundreds, if not thousands, of students, and they typically deliver educational content asynchronously. While this dramatically expands the reach of education delivery, it also makes it challenging for instructors to adequately attend to the learning needs of their students. Many student questions in an online classroom may go unanswered or not be answered in a timely fashion. This gap between the online and in-person modes of instruction can impede effective learning by online learners [2, 3]; lead to low student engagement [4]; and adversely impact retention in the program

* This research has been supported by NSF Grants #2112532 and #2247790 to the National AI Institute for Adult Learning and Online Education. We thank members of the Design Intelligence Laboratory for their contributions to this work.

[5]. Unfortunately, there are no cost-effective means of having enough human teachers who can answer all questions posed by online students anytime and from anywhere to alleviate this gap.

Artificial intelligence-based virtual teaching agents (VTAs) can help alleviate the lack of student-teacher interactions in online classrooms if they are easily scalable and cost-effective [6]. To this end, we have developed a new version of Jill Watson, a virtual teaching assistant (VTA) that leverages generative AI. With OpenAI’s ChatGPT at its core, Jill Watson can answer student questions and engage in extended conversations anytime and anywhere. By itself, ChatGPT is a conversational agent based on the GPT family of large language models (LLMs). While ChatGPT has an uncanny ability to answer questions and converse with users with a human-like text, it suffers from two major drawbacks. First, LLMs underlying ChatGPT are trained on a massive corpus of text data on the internet, and they implicitly store their knowledge in model parameters. ChatGPT relies on this knowledge to answer user queries. As such, if used in a classroom, ChatGPT’s answers may not reflect the instructor’s expectations based on the class materials, thus failing to support the desired learning outcomes. Second, LLMs are statistical models that answer a simple probability question – given a training data set and a string of words, what is the most probable next word? As a result, LLMs, and hence, ChatGPT, are prone to hallucinations, i.e., they can generate factually incorrect answers devoid of reality (hallucinations) [7], or even undesirable speech [8]. This in part is where Jill Watson comes in. Though having ChatGPT at its core, Jill Watson mitigates hallucinations by using retrieval-augmented text generation (RAG) [9] with dense passage retrieval (DPR) [10] to constrain ChatGPT to answer student questions using only the class materials provided by the instructor. Further, Jill Watson also employs moderation to block irrelevant questions from being processed and to block answers containing harmful speech.

Our work makes several contributions. First, Jill Watson is highly scalable and optimal for learning at scale. It has been deployed across several classes with up to several hundred students at multiple institutions. It can answer questions based on natural language text from instructor-approved documents of any length, including textbooks, user guides, transcripts from video lessons, text from presentation slides, course websites, and class syllabi, schedules, and policies. In theory, given enough computing infrastructure, there is no limit to the number of classes, students, or institutions that can benefit from Jill Watson deployments.

Second, Jill Watson is conversational. It allows students to use a private chat window enabled as a Learning Tools Interoperability (LTI) [11] tool in class Learning Management Systems (LMSs) to ask follow-up questions and converse about course material within the confines of its knowledge base. This interactivity can motivate students to explore content-related questions; answer questions related to course-related activities; and prompt reflection on course content. Thus, Jill Watson may enhance student’s cognitive engagement with the courseware.

Third, Jill Watson is appropriate for a classroom setting due to grounding in the approved instructional material, thereby increasing the trust of the VTA by both students and instructors. The provenance of its answers is not only checked for grounding, but they are also filtered through built-in guard rails against undesirable speech and irrelevant questions.

Fourth, Jill Watson is cost-effective as it relies on pre-trained models along with RAG to provide relevant answers to student questions while minimizing hallucinations. The alternative to this is computationally intensive and expensive fine-tuning of underlying LLMs using hard-to-obtain training datasets with fidelity to courseware. This eliminates the cost of creating proprietary models while making it easy to incorporate ongoing technological enhancements in the fast-paced space of LLMs and AI agents.

2 Related Work

The Community of Inquiry model [2, 3] points to the centrality of teaching presence, i.e., the presence of teacher-student interactions in the classroom, to achieve effective learning outcomes through individual reflection (cognitive presence) and collaborative construction of knowledge (social presence). Given that provisioning adequate teaching presence in online classrooms is often challenging, their work admits technical intervention through VTAs. VTAs can answer questions about course logistics and content anytime and anywhere, thereby amplifying interaction between human instructors and students and hence, increasing teaching presence.

In a first for VTAs, Goel and Polepeddi introduced the original version of Jill Watson built on top of IBM’s Watson platform in 2018 [12]; hence the name. This version of Jill Watson answered students’ questions on class discussion forums of online classes based on a digital library of previously asked questions and human answers to those questions. However, this version of Jill Watson showed demographic bias because the data of the previously asked questions skewed toward majority groups [13]. As a result, the next version of Jill Watson was constrained to answer questions related only to course syllabi, schedules, and policies [6]. This Jill Watson not only helped the students by providing them with a 24/7 question-answering service, but they also supported teachers by offloading routine question-answering and thereby saving precious time. However, the construction of a Jill Watson agent for a new class was prohibitively expensive because the need for training it anew on data for the new class required several tens of hours of manual effort. Further, Jill Watson could only answer a limited type of questions, and that too standalone questions; it could not engage in extended conversation. Goel et al. extended this work with a question-answering AI agent embedded in a natural language user guide to answer questions about another AI agent [14].

In contrast to versions of Jill Watson grounded in a well-defined corpus of relevant but limited knowledge discussed above, open-ended question-answering models such as ChatGPT based on LLMs derive their knowledge from large

non-specific textual training datasets sourced from across the internet and other natural language sources. As noted in the previous section, such models may give answers that are either irrelevant or factually incorrect. These issues are mitigated by using retrieval augmented generation (RAG) with LLM-based models [9]. Very recently, Wang et al. report an implementation of a retrieval-based chatbot for use in education [15]. While they validate this approach with a set of 20 questions, they do not report having deployed their VTA in live classes. In the same vein, Taneja et al. present a conceptual design of the Jill Watson version presented here and evaluate its performance using standard datasets while comparing it to a chatbot constructed using the OpenAI Assistant API [16].

3 Overview of Jill Watson’s Architecture

Unlike previous versions, the new version of Jill Watson is deployed as an LTI tool in the class LMS and students interact with it through a private chat window. This affords a private channel of communication between a learner and Jill Watson. In addition, LTI-based deployments offer easy scalability across several alternative LMSs such as Canvas and Blackboard.

A Jill Watson agent for a class has a modular architecture with functionality enabled by three modules – the agent knowledge base, the agent memory, and the question-answer pipeline.

3.1 Knowledge Base

The first step in generating a Jill Watson agent for a new class is to create its knowledge base by preprocessing natural language courseware provided by the instructor as PDF documents or text parsed from the course website. Figure 1 illustrates the document preprocessing step, which is performed before agent runtime and hence, it does not affect runtime latency. We use the Adobe PDF Extract API to extract text from the PDF documents. All text is sectioned into pages which are further divided into paragraphs. The paragraphs are then grouped into passages of at least 500 characters (about 100 words). There is a 50% overlap between consecutive passages for redundancy and continuity of context across passages. The raw text in these passages often has special characters, unwanted spaces, and formatting anomalies. We prompt ChatGPT with these passages to generate passages with clean text along with a 2–3-word long header. ChatGPT is also used to generate a summary of the clean text to make implicit information in passages more explicit. The clean passages along with their headers and summaries are then served to a text encoder to generate two sets of embeddings that are stored in the agent knowledge base for later use in dense passage retrieval.

3.2 Agent Memory

A MongoDB database stores information about agent-student interactions to serve as a source of student-specific question history for conversation. This elim-

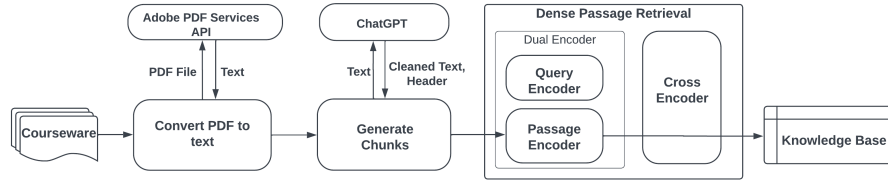


Fig. 1. Jill Watson document preprocessing

inates reliance on LMS to provide this information. This is important as LMS often do not store the question history.

3.3 Question Answering

Figure 2 shows Jill Watson’s runtime architecture. The runtime performs the following functions - Jill-Watson-LMS interaction; co-reference resolution; skill classification for relevant document selection; question-answering; and moderation. A REST server enables Jill-Watson-LMS interaction by exposing Jill Watson as an API. This allows for integration with any LTI-compatible LMS.

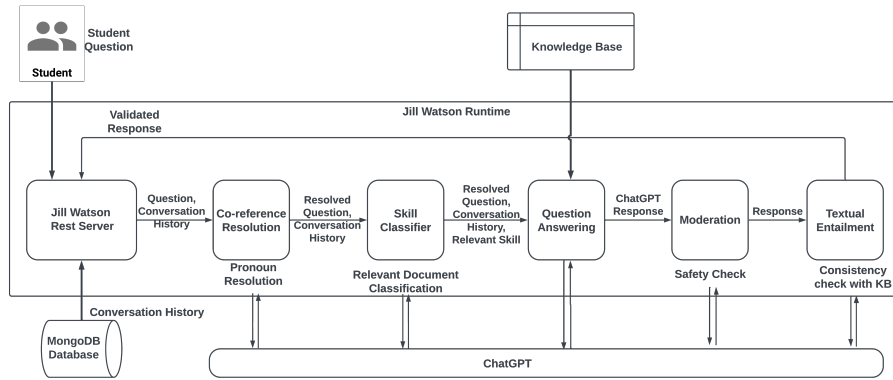


Fig. 2. Jill Watson runtime architecture

Co-reference resolution replaces indirectly referenced entities (pronouns) in a query with explicit nouns or noun phrases based on the conversation history [17]. This ensures the continuity of dialog. The resolved query is then input into a skill classifier to determine the skill required to answer it. Queries that require content retrieval are forwarded to the Contextual Answering Skill; irrelevant queries and greetings are trivially answered, respectively, by the Irrelevant Skill and Greeting Skill.

Figure 3 illustrates the implementation of Jill Watson’s contextual answering pipeline. Contextual answering involves answering questions based on relevant information retrieved from the knowledge base. Jill Watson uses dense passage retrieval for this task [10]. First, based on the required skill, we retrieve relevant passages with clean text and their summaries along with their embeddings from the knowledge base. Second, we compute query embeddings of the co-reference-resolved queries. Third, we compute the similarity of the query embedding with context embeddings for each passage as well as its summary text. The similarity score of a passage is the maximum of the two numbers. Fourth, we sort the passages in decreasing order of similarity and select k passages with similarity above a threshold. To improve retrieval accuracy, we introduce a cross-encoder in our design to re-rank the selected k passages, which are then grouped into batches of five for prompt construction.

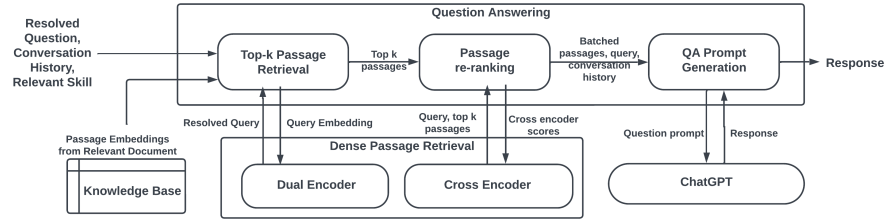


Fig. 3. Contextual answering pipeline

Prompt construction for the question-answering constrains ChatGPT to only answer student queries if the context has the requisite information, or else, refuse to answer. Further, it should be able to respond to follow-up questions to converse. For the latter, we use the conversation history for the current user extracted from the agent memory. We construct a prompt with the first batch of passages as context, conversation history, and resolved user query. If ChatGPT fails to answer the question using the first batch of passages, we use the second batch of passages, and so on, until ChatGPT generates a valid answer. The validity of the answer is checked by prompting ChatGPT. If a valid answer is not found from any of the batches, Jill Watson lets the student know that it cannot answer that question based on the provided context.

A textual entailment module checks the response for consistency against the context retrieved from the knowledge base. The task assesses the logical relationship between a context (C) and a hypothesis (H), where C entails H , contradicts H , or is neutral to H . We incorporate these checks to validate whether context C fully supports the generated answer A . If not, the answer is either rejected or flagged with a low-confidence warning, prompting users to verify the answer independently. Our method utilizes ChatGPT for textual entailment,

with instruction-based prompts yielding the highest recall for identifying non-entailed answers.

Finally, to ensure safe and relevant responses, Jill Watson checks input queries and responses for harmful text using the OpenAI Moderation API. The API allows Jill Watson to detect different categories of harmful text and discard such text with an appropriate message to the user.

4 Results and Discussion

The current version of Jill Watson has been deployed since Summer 2023 in several classes at the Georgia Institute of Technology (Gatech) as well as classes at two community colleges - Wiregrass College and Columbus College - in the Technical College System of Georgia. Table 1 lists some salient information for these deployments along with natural language documents that Jill Watson uses in each class to converse with students. These include instructor handouts, textbooks, lecture transcripts, and text parsed from course websites. Figure 4 shows the accuracy of Jill Watson’s responses to a large number of synthetically created questions for a wide variety of documents used in deployment in alternative courses. These responses were graded by humans. The accuracy ranges from 75% to 97%. Lower accuracy is mostly traced to parsing challenges posed by tables commonly found in syllabi.

Tables 2, 3, and 4 show demographic data for the Fall 2023 Knowledge-based AI class at the Georgia Institute of Technology and the English class at Wiregrass College, and they illustrate that the two student bodies have starkly different gender, age, and ethnicity distributions.

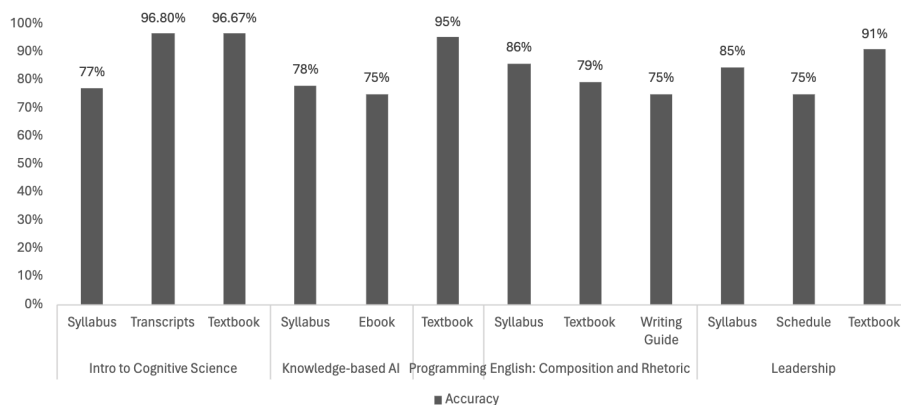


Fig. 4. Accuracy of Jill Watson responses to synthetically created questions for various documents.

Table 1. Jill Watson deployments. In addition to the knowledge base components listed, Jill Watson also answers questions about class syllabus, schedule, and policies.

Semester	Institution	Course	Knowledgebase components	# Students
Summer 2023	Gatech	Human-Computer Interaction	Course website	385
Summer 2023	Gatech	Knowledge-based AI	E-textbook	216
Summer 2023	Gatech	Intro to Cognitive Science	- Textbook - Video transcripts	115
Fall 2023	Gatech	Knowledge-based AI	E-textbook	198
Fall 2023	Wiregrass College	English: Composition and Rhetoric	- MLA Writing Guide - E-textbook - Instructor handouts	100
Spring 2024	Gatech	Knowledge-based AI	- E-textbook - Course website	223
Spring 2024	Gatech	Intro to Cognitive Science	- Textbook - Video lecture transcripts	58
Spring 2024	Wiregrass College	English: Composition and Rhetoric	- MLA writing guide - E-textbook - Instructor handouts	60
Spring 2024	Columbus College	Leadership	Course readings	15

Table 2. Student Gender Distribution

Course (Fall 2023)	Female	Male
Knowledge-based AI	28.54%	71.46%
English: Composition and Rhetoric	78.00%	22.00%

Table 3. Student Age Distribution

Course (Fall 2023)	17-24	25-34	35-44	45-54	55-64
Knowledge-based AI	12.50%	64.79%	17.08%	4.16%	1.45%
English: Composition and Rhetoric	67.50%	23.50%	6.00%	2.00%	1.00%

Table 4. Student Race/Ethnicity Distribution

Course (Fall 2023)	American Indian /Alaskan Native	Asian	Black/ African American	Hispanic /Latino	Two or more	Unk	White
Knowledge-based AI	0	51.25%	3.54%	6.45%	2.71%	2.50%	33.54%
English: Composition and Rhetoric	1.00%	1.00%	40.00%	11.50%	4.5%	0	42.00%

Table 5. Example of a real student conversation with Jill Watson from the Knowledge-based AI class in Spring 2024 (conversation has been truncated for brevity). Note the depth of the conversation between the student and Jill Watson

Turn	Student Question	Jill Watson's response
1	Do you have sample questions to quiz about means end analysis?	Certainly! Here are some sample questions to quiz your understanding of means-end analysis: 1. What is means-end analysis in problem-solving? 2. How does means-end analysis differ from generate and test? 3. What is the role of a heuristic in means end analysis? 4. Explain the concept of state spaces [...] 10. Can means-end analysis guarantee computational efficiency or optimality of the solution? Explain. If you have any specific questions or need further explanation on any topic related to means-end analysis, feel free to ask!
2	Means ends analysis is a universal problem solving technique. It is not guaranteed to provide a solution. It attempts to solve a problem by traversing through a potential state space [...] Is this correct for the first one?	Yes, your understanding of means-end analysis is correct for the first question. Means-end analysis is a universal problem-solving technique that aims to solve a problem by reducing the differences or distance between the current state and the goal state. It involves traversing through a potential state space and applying operators or moves to bring the current state closer to the goal state. [...]
3	Are the operators the potential moves being generated from a given state?	Yes, that's correct. In means-end analysis, the operators refer to the potential moves or actions that can be applied to a given state in order to progress towards the goal state. These operators represent the available actions or transformations that can be performed on the current state to generate new states. [...]
4	For question number 2, Means-End-analysis differs from generate and test because it is reducing the distance between the current state and the goal state. To get the states for means end analysis, the agent will generate[...]	Yes, you are correct in stating that means-end analysis differs from generate and test in terms of reducing the distance between the current state and the goal state. However, there are a few additional differences between the two problem-solving approaches: [...]

Table 5 shows a real conversation between Jill Watson and a student in the Knowledge-based AI class. The student is using Jill Watson to test her skills and receive Jill Watson’s feedback. This is an example of a student’s deep engagement with the courseware. What is interesting is that though the student asks specifically about questions related to means-end analysis, Jill Watson generates questions for related topics taught in the class. ChatGPT also generates a good quality quiz to the same question, but it focuses solely on means-end analysis. This means that Jill Watson, by having access to course content, generates questions that encourage students to think more deeply about related concepts in the course. The ability of Jill Watson to support deeper understanding and critical thinking however requires more investigation.

Table 6. Example where ChatGPT hallucinates an entire response but Jill Watson does not. The entirety of the ChatGPT response is a hallucination.

Question: How does Thagard explain the concept of 'free will' within the frameworks he discusses in his book "MIND: Introduction to Cognitive Science"?
Jill Watson: Thagard does not explicitly discuss the concept of "free will" within the frameworks he discusses in his book "MIND: Introduction to Cognitive Science." The book primarily focuses on explaining various aspects of intelligent behavior and the different approaches to modeling the mind. Thagard's emphasis is on understanding the nature of the mind and the cognitive processes that underlie intelligent behavior, rather than delving into philosophical concepts such as free will.
ChatGPT 3.5: In his book "MIND: Introduction to Cognitive Science," Thagard addresses the concept of free will within the context of cognitive science, exploring how cognitive processes give rise to our sense of agency and decision-making. Thagard's approach to free will in "MIND" is consistent with his broader views on consciousness and cognition, emphasizing the role of neural mechanisms and computational processes in shaping human behavior. While "MIND" covers a wide range of topics in cognitive science, including perception, language, and reasoning, Thagard's discussion of free will typically focuses on the following key points: [...]

Table 6 illustrates Jill Watson’s ability to mitigate the hallucinatory behavior of ChatGPT. When asked about the concept of 'free will' in the context of the recommended textbook in the Intro to Cognitive Science class, Jill Watson maintains fidelity to the source material, responding that there is no explicit mention of 'free will' while ChatGPT’s elaborate response consists entirely of a hallucination and is not grounded in the actual content of the textbook.

Table 7 presents sample student comments gathered from the class discussion forums and surveys to gauge student perceptions regarding Jill Watson. The comments are reproduced without any editing, and they illustrate a positive student experience on the whole.

We close this section with two preliminary observations from the Knowledge-based AI class in 2023 indicating that access to Jill Watson may positively impact student performance. All students in this class were asked to consent to participate in an A/B experiment to measure the impact of Jill Watson on teaching

Table 7. Student comments on Jill Watson from class discussion forums and surveys.

Student A: Does Jill provide any course content related answers? Student B: Yes, it does. It can answer questions like “What’s common sense reasoning?”
Student A: Can you ask Jill how to do assignments and projects then? Student B: Yes, it will give you step by step general instructions on how to approach the assignment. It won’t provide any extra information that’s not already present in the assignment description.
Student C: The Jill Watson upgrade is a leap forward. With persistent prompting I managed to coax it from explicit knowledge to tacit knowledge. That’s a different league right there, moving beyond merely gossip (saying what it has been told) to giving a thought-through answer after analysis. I didn’t take it through a comprehensive battery of tests to probe the limits of its capability, but it’s definitely promising. Kudos to the team.
Student D: +1 really like the upgrade - I am using it to actually learn the lecture material. One feature request would be to persist the conversation, I had closed the tab and lost my prompts. Again, kudos!
Student E: Jill helped me find out where I needed to go when I was lost on blackboard.
Student F: I often times need assistance with finding certain assignments and Jill Watson is very helpful with directing me towards the right place.

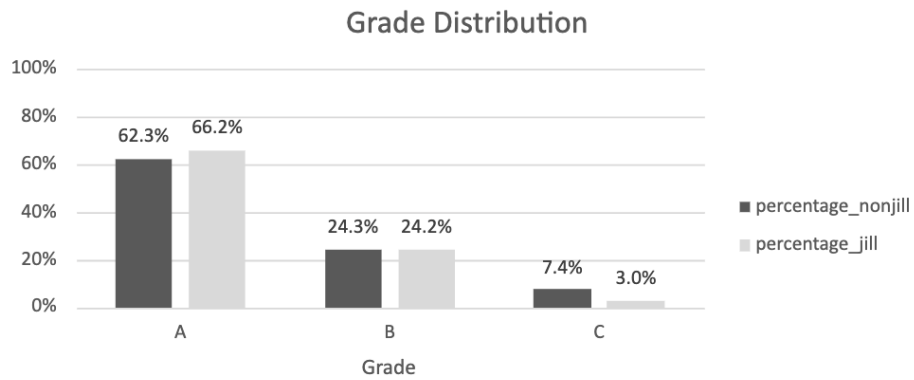


Fig. 5. Grade distribution in Knowledge-based AI in Fall 2023 class

presence using the Community of Inquiry instrument [18]. Those who consented to participate were randomly assigned to treatment and control groups. The treatment group was given access to Jill Watson, while the control group was not. The impact on teaching presence will be published elsewhere; we report here other important results. Figure 5 shows the proportion of students receiving final grades of A and C in the two groups. The proportion of students with an A in the treatment group is 3.9% higher than that in the control group; the proportion with a C in the treatment group is 4.4% lower than that in the control group. Note that while this is a promising result, the statistical significance of this one data point and Jill Watson’s causal relationship with it is undetermined.

We also compare the mean final raw score for those students in the treatment group who interacted with Jill Watson at least three times with that for students in the same group who did not. The mean scores are 91.4% and 86.6%, respectively, for the two populations. The probability that the null hypothesis that the two populations have identical mean values is true is $p = 0.007$, which is a statistically significant difference at the $p < 0.05$ level.

5 Conclusion

We presented a highly performant and scalable conversational virtual teaching assistant, Jill Watson, based on the GPT family of LLMs. Jill Watson is resistant to hallucinations. It answers student questions and engages in extended conversations based on instructor-approved courseware anytime and anywhere using safe language, thus supporting both students and instructors. It has been deployed in several classes at the Georgia Institute of Technology and in two community colleges with about 1300 students. Qualitative feedback by students shows an encouraging level of engagement with the AI agent. We see preliminary evidence that Jill Watson may support deeper understanding of the subject matter but this needs to be investigated further. We also see preliminary indication that Jill Watson positively impacts student performance but this too needs to be explored further.

References

1. Shah, D.: By The Numbers: MOOCs in 2021. <https://www.classcentral.com/report/mooc-stats-2021>. Last accessed 14 Feb 2024
2. Garrison, D., Anderson, T., Archer, W.: Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*. **2**. 87–105 (1999). [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
3. Anderson, T., Rouke, L., Garrison, D. R., Archer, W.: Assessing Teaching Presence in a Computer Conferencing Context. *Journal of Asynchronous Learning Networks*, **5**(2), 1–17 (2001)
4. Estrada-Molina, O., and Fuentes-Cancell, D.-R.: Engagement and desertion in MOOCs: Systematic review. *Comunicar* **30**(70), 111–124 (2022)

5. Badali, M., Hatami, J., Banihashem, S. K., Rahimi, E., Noroozi, O., Eslami, Z.: The role of motivation in MOOCs' retention rates: a systematic literature review. *Research and Practice in Technology Enhanced Learning* **17**(5) (2022). <https://doi.org/10.1186/s41039-022-00181-3>
6. Goel, A.: AI-Powered Learning: Making Education Accessible, Affordable, and Achievable, arXiv:2006.01908v1 [cs.CY]. <https://doi.org/10.48550/arXiv.2006.01908>
7. Ji, Z. et al.: Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, **55**(12), 1—38, (2023). <https://doi.org/10.1145/3571730>
8. Zhang, B. et al.: Comprehensive Assessment of Toxicity in ChatGPT, arXiv:2311.14685 [cs.CY]. <https://doi.org/10.48550/arXiv.2311.14685>
9. Huayang, Li. et al.: A Survey on Retrieval-Augmented Text Generation, arXiv:2202.01110 [cs.CL]. <https://doi.org/10.48550/arXiv.2202.01110>
10. Karpukhin, V., Oğuz, B., Min, S. et al.: Dense Passage Retrieval for Open-Domain Question Answering, arXiv:2004.04906v3 [cs.CL]. <https://doi.org/10.48550/arXiv.2004.04906>
11. 1Edtech's LTI Standards, <https://www.ledtech.org/standards/lti>. Last accessed 22 Apr 2024
12. Goel, A. K., Polepeddi, L.: Jill Watson: A Virtual Teaching Assistant for Online Education. In *Learning Engineering for Online Education: Theoretical Contexts and Design-Based examples*. 1st edn. Routledge, New York (2018)
13. Eicher, B., Polepeddi, L., Goel, A. K.: Jill Watson Doesn't Care if You're Pregnant: Grounding AI Ethics in Empirical Studies. In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp 88–94. ACM, New Orleans, LA, USA (2018)
14. Goel, A., Nandan, V., Gregori, E., An S., Rugaber, S: Explanation as Question Answering Based on User Guides. In *Explainable Agency in AI: Research and Practice*, CRC Press, Boca Raton (2024). <http://dx.doi.org/10.1201/9781003355281-4>
15. Wang, K., Ramos, J., Lawrence, R.: ChatEd: A Chatbot Leveraging ChatGPT for an Enhanced Learning Experience in Higher Education. arXiv:2401.00052 [cs.CY]. <https://doi.org/10.48550/arXiv.2401.00052>
16. Taneja, K., Maiti, P., Kakar, S., Guruprasad, P., Rao, S., Goel A.: A Virtual Teaching Assistant powered by ChatGPT. Accepted for publication. In: *Proceedings of Artificial Intelligence in Education: 25th International Conference, AIED 2024*, Springer, Recife, Brazil (2024)
17. Sukthanker, R., Poria, S., Cambria, E., Thirunavukarasu, R.: Anaphora and Coreference Resolution: A Review. arXiv:1805.11824 [cs.CL]. <https://doi.org/10.48550/arXiv.1805.11824>
18. Arbaugh, J.B., Cleavelan-Innes, M., Diaz, S. R. et al.: Developing a community of inquiry instrument: Testing a measure of the Community of Inquiry framework using a multi-institutional sample. *The Internet and Higher Education* **11**(3), 133—136 (2008). <https://doi.org/10.1016/j.iheduc.2008.06.003>