Jill Watson Doesn't Care if You're Pregnant: Grounding AI Ethics in Empirical Studies

Bobbie Eicher, Lalith Polepeddi, Ashok Goel
Design & Intelligence Laboratory, School of Interactive Computing
Georgia Institute of Technology, Atlanta, GA 30332
bobbie.eicher@cc.gatech.edu, lpolepeddi@gatech.edu, goel@cc.gatech.edu

Abstract

Jill Watson is our name for a virtual teaching assistant for a Georgia Tech course on artificial intelligence: Jill answers routine, frequently asked questions on the class discussion forum. In this paper, we outline some of the ethical issues that arose in the development and deployment of the virtual teaching assistant. We posit that experiments such as Jill Watson are critical for deeply understanding AI ethics.

Introduction

Herbert Simon, one of the founders of the discipline of Artificial Intelligence (AI), asserted that AI is an empirical science (Simon 1995). He stated that the design of each AI system is an experiment. He suggested that a productive methodology for making progress in AI is to conduct experiments that explore the space of AI system designs for various tasks, compare the performances of various AI systems at different tasks, and abstract general principles of intelligence that provide a mapping between the AI system design and the system performance. He argued against AI research based on mere conjecture and speculation; he wanted conjectures and arguments about AI to be grounded in experimental data, not in philosophy or theology. It is remarkable that much of AI research follows Simon's methodology, even if only implicitly. It is also noteworthy that much of the progress in the field has come from this paradigm.

We propose that, like AI itself, AI Ethics (AIE from now onwards) too is an empirical science. Now that we have a reached a stage in history that AI agents are actually influencing human society (as well as other AI agents), the design of each AI agent is not only an exploration in the space of AI agent designs but also an experiment in the space of AI agent ethics. In fact, if it's our goal to build only ethical AI agents, then the two design spaces are indistinguishable: the design of AI agents logically implies the design of ethical AI agents. We propose that a productive methodology for making progress in AIE is to conduct experiments that explore the space of AI agent designs for various tasks, compare the performances of various AI agents including their behaviors from an ethical perspective, and abstract general principles of intelligence that provide a mapping between the AI agent design and the agent performance including its behaviors from the viewpoint of AIE.

In this paper, we describe one such experiment in AIE. Our work has revolved around the use of AI techniques to create interactive agents capable of acting as teaching assistants for courses in computer science at the university level. One such virtual teaching assistant is called Jill Watson: Jill autonomously answers routine, frequently asked questions on the discussion forum of a Georgia Tech class on AI. As we worked to develop and deploy Jill Watson, we have wrestled with questions pertaining to student needs, deception, impact on human teaching assistants, our own responsibilities as teachers, and how we might be reshaping the educational experience of thousands of students. Our goal here is to share some of our reasoning as we have made difficult decisions about how to proceed with the project, and what we have empirically learned about AIE along the way.

Background and Motivation

In January 2014, Georgia Institute of Technology created an online edition of it's Master of Science in Computer Science degree called OMSCS (Belkin 2013). Although the OMSCS program began with just a few hundred students, by the Spring 2017 term it had expanded to more than 4,000 (Georgia Tech 2017). As the number of students in the program has grown, there has been a corresponding growth in the amount of communication, coordination, and teaching assistance need to support the student needs.

To get a sense of the scale of the required teaching assistance, consider the online class CS7637: Knowledge-Based Artificial Intelligence (KBAI) that we developed in 2014 and have regularly taught as part of the OMSCS program since Fall 2014 (Goel and Joyner 2016). By the Fall 2016 term, the KBAI typically included 300 or more students with a total of more than 12,000 contributions to the class discussion forum (Piazza). This can be compared to the less than 2,000 contributions in the digital discussion forum of the traditional residential section we offered in parallel. Table 1 gives a deeper sense of the differences in the participation rates on the discussion forums of the residential and online sections of the KBAI class in Fall 2016. Each item in Posts is an individual thread, while Contributions include all responses and followup discussions associated with the thread. Instructor responses and Student responses are answers to questions in

	Residential	Online	Change
Posts	455	1201	+3x
Contributions	1838	12190	+5x
Instructor Responses	195	640	+3x
Student Responses	75	356	+5x

Table 1: Comparison of forum participation between the residential and online sections of the KBAI class in Fall 2016.

an existing conversation, rather than prompts for new conversations.

In addition to sheer numbers, students in the OMSCS program may be in any time zone in the world, which means that posts requiring attention from the teaching assistants are likely to arrive around the clock. Finally, online students do not have several meetings a week with the instructor where they have an opportunity to ask questions, and therefore their course experience overall is more dependent on how responsive the teaching team can be on the forums, regardless of the time of day or night when someone might seek help.

Development Context

We developed the virtual teaching assistant Jill Watson in the context of the aforementioned OMSCS course on KBAI (Goel and Polepeddi 2017), memory, reasoning and learning. All students have access to a set of lecture videos, and complete a mixture of writing and programming assignments over the course of the semester. In developing the online KBAI course, we anticipated some of the scaling needs of the OMSCS program and thus we used AI techniques to help with the learning right from the start. For example, we built 150 exercises and 100 cognitive tutors into the video lessons, which comes to about 1 exercise for every 6 minutes of video on average (Goel and Joyner 2017).

However, we did not fully anticipate the traffic on the discussion forum of the online KBAI class. Thus, we built Jill Watson in the fall of 2015, more than a year after the development of the KBAI class. Each semester, the class includes a teaching team including the course instructor (typically Goel), several (human) teaching assistants, and Jill Watson. The instructor and the teaching assistants communicate with the students through the class discussion forum that acts like a virtual classroom. The (human) teaching assistants are all graduate students: in recent offerings of the KBAI class, many of the teaching assistants had previously interacted with Jill Watson as students.

Although our initial motivation for developing Jill Watson was to manage the load on the (human) teaching assistants in the online KBAI class, we now use Jill Watson in the discussion forum of the residential section of the KBAI as well. We estimate that by now more than 1000 students and more than 40 teaching assistants have interacted with some version of Jill Watson.

Outline of the Jill Watson Experiment

We initially developed Jill Watson for use in the KBAI class in the Spring 2016 term (Maderer 2016). Figures 1 and 2 illustrate student interactions with Jill Watson (Goel and Polepeddi 2017).

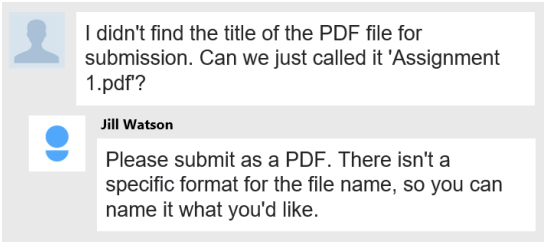

Figure 1: Example interaction between student and the virtual teaching assistant

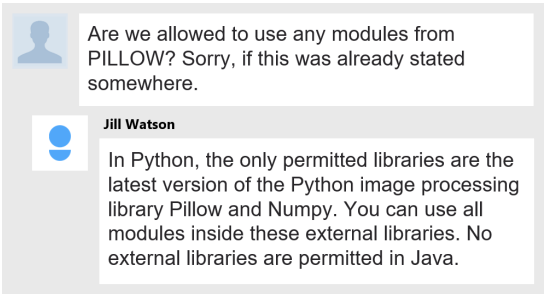

Figure 2: Example interaction between student and the virtual teaching assistant

In the initial deployment, we introduced Jill Watson as a member of the teaching staff, giving no indication to the students that an AI experiment was in progress. This is because we wanted to determine if we could develop an AI teaching assistant that could answer simple routine questions on the class discussion forum with enough authenticity that the students would not suspect her true identity. We were well aware of our ethical responsibilities both as teachers - the experiment could go awry, disrupting the learning in the class - and as AI researchers. Given that this was an experiment engaging human subjects, we formally asked for and received permission from Georgia Tech's Institutional Review Board so that we could conduct the study with the element of deception. Until we were confident of her ability to post accurate answers, we also had a human review everything she wanted to post.

It was not until the end of the Spring 2016 term that we told the students that one of their teaching assistants secretly was an AI agent, a product of the same concepts and methods they were learning in the KBAI class. As Figure 3 illustrates, the student reactions were very positive (Goel and Polepeddi 2017). In fact, after using Jill Watson in multiple semesters, we have never had a negative reaction from students.

Redesign of Jill Watson

Jill Watson has remained active in various forms through all of the offerings of KBAI class since its initial deployment

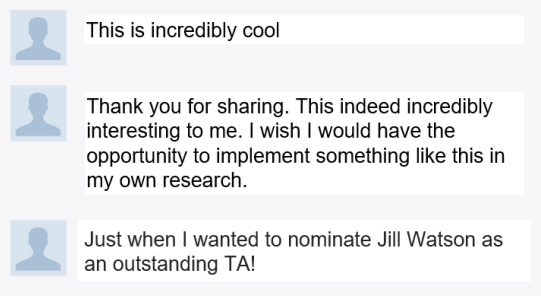

Figure 3: Several real student reactions to the end-of-semester announcement

in January 2016. (In fact, it is active even now, as we write this paper.) creation. To maintain its anonymity in each new semester, the virtual teaching assistant is given a new name (masculine or feminine). So that students cannot easily identify the virtual teaching assistant through the school directory or social media, human teaching assistants also operate under pseudonyms for the duration of the semester.

At the end of each term, the students are polled and asked to discuss which members of the teaching team they believe to have been the AI agent(s). We also invite them to share their thoughts on why they picked out particular individuals. The students themselves open conversations to discuss this question over the course of the semester as well. However, it is only after the official poll at the end of the semester is complete that the instructors join the conversation and an announcement is made with the correct answer(s). It is also at this point that we open a conversation specifically sharing how Jill Watson works and how she uses the various AI techniques they learned in the course.

Over the two years since we first developed Jill Watson, we have extended and expanded its capabilities with additional features, such as the ability to respond to student introductions and to open new threads of its own. All of its features continue to operate autonomously. However, human teaching assistants do see what she posts and may offer a correction or clarification if needed.

Expanding to New Courses

Most recently, we have begun work on a new version of Jill Watson intended for use in an introductory computing course that teaches programming in Python. This effort has involved both changes to the technology and a fresh start with gathering and organizing data in the new class. Thus, it has proven to be a good opportunity to reexamine some of the design decisions we made in developing the original Jill Watson and their results. It also provides an opportunity to review some of the ethical considerations in developing and deploying virtual teaching assistants more broadly.

There are Questions, and Then There are More Questions

Once the original architecture of Jill Watson was functioning, we were faced with the issue of how to decide what kinds of questions she should first try to answer. Our initial decision was that we should only attempt to answer those questions that we had already encountered previously in the discussion forums, including postings from previous semesters, rather than attempting to predict what kinds of things students might want to ask or how they might phrase the question.

Some of the results of this decision were easy to anticipate. It was obvious that the questions answered would tend to be logistical in nature; although there are infinitely many things students might be curious about, questions like ”When is this assignment due?” or ”How many pages are we required to write?” tend to be of interest to almost all students in almost all classes.

What we did not initially consider was the differences in the types of questions and comments we could potentially see. For example, every student is asked to post an introduction detailing some basic information like their background, why they’re participating in the class, what their particular interests are, and any other information they’d like to share about themselves. This is aimed at starting the course on a friendly and welcoming note, as well as giving the students opportunities to make connections among themselves and to give us insight into any areas of particular interest we might want to focus on for a given semester.

Each semester, therefore, the teaching team makes a point of creating an individual and personal response to every student introduction. When we decided to give the virtual teaching assistant the ability to participate in this activity, we followed the common practice of relying entirely on building the ability to respond to the types of posts we were actually seeing in our pool of posts from previous semesters, rather than trying to speculate about what we might see someone post. When we had finished the first round of that project, we looked back and realized that the system’s training made it capable of responding to a phrase like ”will be father for the first time” as in figure 4, but would not react specifically to something like ”I’m pregnant” as in figure 5.

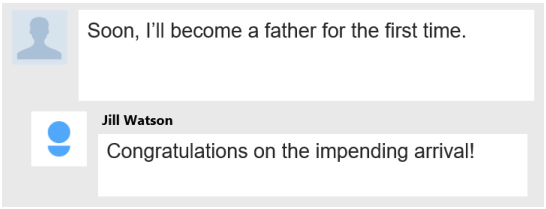

Figure 4: Example of a reaction the virtual teaching assistant might give.

We stopped to examine why this had happened, and realized that it was a consequence of the demographics of the program. Women are a minority group in Computer Science, making up only 32% of students in introductory programming courses and 16% of CS degree earners (Sax). This demographic imbalance is even more pronounced in the OMSCS program itself, with less than 14 percent of the students in the program identifying as female (Georgia Tech 2017).

While we don’t have exact numbers for parents and ex-

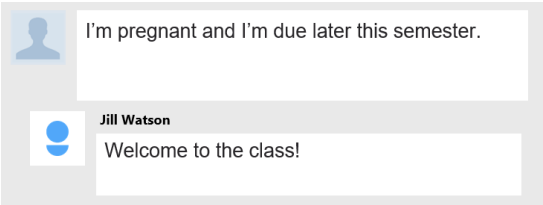

Figure 5: In contrast, it would not specifically comment here.

pecting parents in the OMSCS program, the information we do have makes it reasonable to expect that there are far more fathers than mothers in the program. Furthermore, the history of gender-based discrimination in STEM fields may make women more hesitant to call attention to this kind of personal detail than men (Teague 2002).

This has led us to add two concerns to the list of things that we look at when making design decisions:

- Might different demographic groups express the same question or need in different ways that could make it easier for one group to receive help from an automated system than another?
- What is the potential emotional impact of having a system that knows how to respond to concerns or interests that are relevant to a majority group, but potentially overlooks similar issues concerning minorities?

It's difficult to get around the fact that we can only train on the data we have, and that the greatest impact will always come from focusing on the concerns affecting the largest number of people. We also don't have any solid data on how much of an impact this issue would have on the students.

Still, we are now making a conscious effort to look at how the system behaves and to ensure that it is as inclusive and welcoming as possible to as wide a range of people as possible. This is also an area where having humans working alongside the system is invaluable, because where it fails to handle cases that are uncommon, humans can do so instead. In fact, one of our hopes for the system has always been that students with unusual, or even unique, questions and concerns would be able to get more attention than would be possible in a setting where the common situations weren't handled automatically.

Student Theory of Mind and AI

In previous work, we have argued for the importance of Theory of Mind (TOM) in interactions, and particularly collaborative efforts, between humans and artificially intelligent agents (Eicher et al. 2017). Theory of Mind is, essentially, the ability to build a mental model of what someone else is thinking and how they might react. In groups of humans, it has been shown to be a stronger predictor of group performance on tasks than any other known measurement of the skills or abilities of the individual participants. (Engel et al. 2014) This is true even when the interaction is virtual (Baron-Cohen et al. 2001).

Given that education involves a variety of social interactions and collaborative efforts, we believe that this issue applies specifically, and especially, in the context of constructing an AI that is intended to act as a teacher. To be effective, the virtual teaching assistant must be able to draw conclusions about what the student means by a particular question and why they would be asking it. It can use this information to select among the available answers for that topic. It should also provide answers in such a way that students are guided toward understanding what they need to do or say if they want to be more effective in seeking help.

The Tension Between Authenticity and Deception

For the purposes of our work, we've taken to using the word *authenticity* to capture the idea that we want to build something that feels like interacting with a human. This was a research goal in itself, but we also have a working hypothesis that it's a beneficial approach for our students. They've spent a lifetime interacting with people and turning to teachers for help, and we want them to feel as if they can approach this in exactly the same way that they always have. They shouldn't need to learn any special commands or spend more time thinking about how to phrase a question or read a response to account for the fact that they're interacting with one of our agents.

It's notable that there are indications that humans will tend to interact with a computer as if it's human, even when they know the truth. This phenomena has been termed *ethopoeia*, defined as "a direct response to an entity as human while knowing that the entity does not warrant human treatment or attribution" (Nass and Moon 2000). What we don't yet know, however, is how reliable or complete this effect is. People who used early versions of search engines learned to shape their queries to the capability of the tool, while Google has been working more recently to make the tool capable of responding to questions one might pose to a human (Giles 2012). In a sense, what we're doing is attempting to discourage students from expending energy on thinking about the virtual teaching assistants as something that requires a distinct approach to questions. That way they can fully focus on their question and post something once that's suitable for understanding by classmates and human course team.

Through the lens of TOM, it could be said that the goal of authenticity is to build an agent that does a compelling job of behaving according to the expectations a student would have of the behaviors and responses that would come from a human instructor. One of the key early goals of these experiments was to work on building an experience that met this goal of authenticity as well as possible, so that we would be serving the needs of the students with a minimum of disruption or additional effort on their part. This makes the deceptive element of the experiment key, because it's the best way to set up a situation where the students who are seeking information will ask questions in the same natural way that they would pose them to any human, rather than deliberately trying to account for whatever other expectations they might have of how an artificial agent will interpret a question or behave when providing an answer.

The Threat of Deceptive Bots

The most common focus of curiosity, and sometimes objection, to the virtual teaching assistant project has centered around the fact that we tried to make Jill seem as human as possible (Miller 2016), and it has been likened to bots in other situations that try to conceal their identity for much more nefarious purposes. It has even been cited as one example of this situation by people advocating for regulation requiring that all such bots reveal their nature and source (Etzioni 2017). This concern about misleading bots attempting to deceive or influence humans about real and important issues is both valid and timely (Schultz 2017; Markoff 2017), but we feel that it's actually an additional reason that our project is important.

Our students spend an entire semester interacting with bots that aren't announcing themselves. Since the fall of 2016, we have also had many students working on an optional alternative project track where they work to build their own bots that can hold discussions or answer questions. This is a rare and meaningful opportunity for them to actively engage with the question of what it's like to experience these interactions, what it ought to be like, and what kind of obligations designers have.

At the end of each semester, we also hold a poll where we explicitly ask them to tell us who they believe were actually artificial agents and which members of the team they believe were human. This sparks discussion among the students (in addition to what happened organically throughout the term) about what kinds of qualities signal that something is probably the work of artificial intelligence along with how one can be identified. In essence, a side effect of our experiment is that we're giving students training in how to spot whether they're talking to a human or not.

Even if regulation happens that attempts to control whether bots announce themselves or not, it is unlikely that it will be completely effective in eliminating deceptive bots. Given that, we feel that it's invaluable that we're both indirectly training our students to recognize other bots and giving them the opportunity to have the experience of being the person who may unknowingly be interacting with them. We hope that they'll use this knowledge to become more informed citizens, and that those who go into the field of artificial intelligence professionally will use this experience to make more thoughtful decisions about how to go about their own work and what kind of experiences it should offer to users.

Students as Current and Future Builders of AI

We are also working under the assumption that at least some of the students of KBAI will later go on to build their own bots. In fact, there has been so much interest in this area that we have added an alternative project track to the course where students are asked to build their own chatbot whose sole purpose is to answer questions about the class based on the content of the course syllabus.

Therefore, we have taken steps to consciously treat our students as future producers of technologies similar to what they're experiencing in class. We actively encourage discussion about the experience of interacting with the virtual teaching assistants, and a variety of related ethical issues in artificial intelligence such as whether it's important to disclose the identity and origin of every agent or how much concern we should place on the issue of impacting human jobs. At every step of their course projects and in all of their written assignments, we also ask them to think about the parallels between what they are doing and human cognition. While there's not time in class to go into these issues as deeply as other work has recommended (Goldsmith and Burton 2017), we're at least ensuring that every student is challenged to think about these issues before they begin developing projects that may have an impact on the larger world.

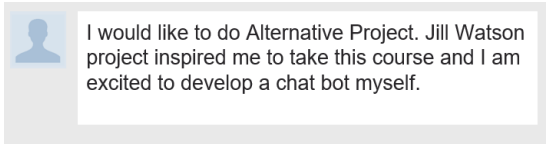

Figure 6: Excerpt from a student post.

While we don't know how many of these students will do further work on technologies involving artificial intelligence, we hope that these steps will help to turn this into an experience that guides them to do so in a more conscious and ethical way. In particular, we believe that having the personal experience of interacting with agents that don't identify themselves as separate from humans will help our students to develop a greater sense of empathy for what that experience could be like for other people, and how they should shape their own agents to account for it. Work at Harvard University has projected that the OMSCS program will soon be graduating as much as seven percent of the MS degrees in CS, we believe it's important to the field that we make a point of producing graduates who can thoughtfully address the ethical issues around their work (Goodman, Melkers, and Pallais 2016).

Considering the Intended Audience

Student Background and Interests

One of the chief reasons that we've always felt that this project is ethical in the context where it's being done is the nature of its audience. Our project has operated in a course that consists primarily of graduate students, with some upper-level undergraduates. All of them are students in computer science, and all of them have expressed personal interest in the topic of AI by signing up for a course about it.

This means that we are dealing with students who:

- A background in Computer Science to understand what kinds of things a computer might do.
- Specific interest in the topic that led them to sign up for the course.
- A possible future in making similar interactive agents in the future.

These conditions are key to our decision that it's ethical to continue running the experiment each semester without

telling the students in advance which of the teaching assistants are actually human and which are agents that we created. We feel that the potential negatives of this approach are very limited, particularly considering that it's now well known to incoming students that this is a practice in this particular course (and anyone who doesn't know will soon see posts from their classmates speculating about the teachers for that semester).

In contrast, we have been working to expand the project to include an introductory computing course at Georgia Tech which is also offered freely to the general public through edX, and we have decided to have the new versions of the agent that run on that platform freely announce themselves. The chief reason that we've done this is that those students are typically new to computing and may not have any interest in artificial intelligence. We'll have further discussion later in the paper about the background work involved with expanding to an additional course, and what kind of considerations we took into account before deciding how to do it.

In contrast, we've recently been experimenting with applying similar technology to one of Georgia Tech's introductory computer science courses that focuses on teaching Python. The student population here is very different, particularly because the space where we're operating is the edX edition of the course that's open to the public. These students

- Likely do not understand computers or programming well at the beginning of the course.
- Have much more varied backgrounds. (The course is open to the public, rather than only those students who have been admitted to Georgia Tech.)
- Have not expressed any specific interest in artificial intelligence.
- Are at the beginning of learning to program, and therefore there's no specific reason to believe they will go on to create interactive agents of their own.
- Are participating in a course where, like many free edX courses, there is limited-to-no dedicated human support.

Based on the differences between the students we're dealing with for the KBAI class and those in introductory computing class, we decided that it wasn't appropriate to run a deceptive study. The new agent, which we've dubbed Noelle King, actively announces herself as a product of the same programming language that the students are learning. Due to the limited human monitoring of the edX forums, she also volunteers information in her posts about where the students should send feedback if they feel she's said something that's incorrect or confusing.

Conclusions

The Jill Watson experiment illustrates the empirical nature of the science of AI Ethics (AIE). While we surely can (and will) speculate about Jill Watson's ethics, ultimately it is our observations of her behaviors and interactions with the intended users as well as our experiments with her evolving design that provide insight into her ethics. As we illustrated in this article, given that Jill Watson's replies to new questions are based on her repertoire of previously encountered questions and answers. Most of the past questions came from the demographics that dominate the class, with the most prominent imbalance being the number of men compared to women. In this way, Jill Watson exhibits a bias towards the demographics majorities in her subject population and may underserve minorities, which is especially troubling in a STEM field where educational and corporate institutions are working to increase diversity. This is an issue we're actively monitoring and attempting to limit.

It follows that AI research must take ethics deeply into account when designing AI agents. This especially true of AI agents that directly interact with humans because their behaviors are likely to have a more explicit on humans. However, ethics now is a matter of concern in design of all AI agents; AI is transforming into AIE. Nevertheless, while we must take ethics into account when designing AI agents, ultimately there is no substitute for experimentation with agent designs to deeply understand AIE.

In conducting experiments with AIE, deception is sometimes acceptable, for example, when we want to test the authenticity of an AI teaching assistant's interactions with students compared to the interactions of human teaching assistants. If we were to tell the students about Jill Watson's identity as an AI teaching assistant, we would have necessarily biased students' perceptions of her authenticity.

Finally, it is noteworthy that there are two sets of ethics here. On one hand, we want AI agents to behave ethically; on the other the humans involved in experiments with AIE have their own ethical requirements. For example, as teachers we have responsibilities toward the students in the class where we're running the Jill Watson experiments, such as being truthful and transparent. Finding the right balance between our ethical responsibilities and that of the AI agents we create and experiment with is not always easy. In fact, finding this balance is itself a part of the experiments we conduct in AIE.

Acknowledgments

We are grateful to Georgia Institute of Technology for multiple GT-Fire educational technology seed grants (2016, 2017). We are also grateful to IBM for multiple Faculty Awards (2014, 2015, 2016). In addition, we thank IBM for giving us access to the IBM Watson Manager starting in Fall 2014 and to IBM Bluemix starting in Fall 2015. The original version of Jill Watson made extensive use of IBM Bluemix (though the more recent versions have been based mostly on our own homegrown technology).

We are grateful to Spencer Rugaber for many broad discussions on AIE. We are thankful to other contributors to this project, including Parul Awasthy, Ananya Roy, David Joyner, Angela Smiley, and Sydni Peterson. Finally, we thank the more than 40 (human) teaching assistants and more than 1000 online students who have participated in the Jill Watson experiment.

References

Baron-Cohen, S.; Wheelwright, S.; Hill, J.; Raste, Y.; and Plumb, I. 2001. The reading the mind in the eyes test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry* 42(2):241–251.

Belkin, D. 2013. College offers online master's degree for $6,000. Accessed: 2017-11-14.

Eicher, B.; Cunningham, K.; Marissa Gonzales, S. P.; and Goel, A. 2017. Toward mutual theory of mind as a foundation for co-creation. Presented to the International Conference on Computational Creativity, Co-Creation Workshop, June 2017.

Engel, D.; Woolley, A. W.; Jing, L. X.; Chabris, C. F.; and Malone, T. W. 2014. Reading the mind in the eyes or reading between the lines? theory of mind predicts collective intelligence equally well online and face-to-face. *PLoS ONE* 9(12):1 – 16.

Etzioni, O. 2017. How to regulate artificial intelligence. https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html. Accessed: 2017-10-15.

2017. The numbers. https://www.omscs.gatech.edu/prospective-students/numbers. Accessed: 2017-10-27.

Giles, J. 2012. Why Google will soon answer your questions directly. https://www.newscientist.com/article/mg21428676-400-why-google-will-soon-answer-your-questions-directly/. Accessed: 2017-10-29.

Goel, A., and Joyner, D. 2016. An experiment in teaching artificial intelligence online. *Journal for Scholarship of Technology-Enhanced Learning* 1(1).

Goel, A., and Joyner, D. 2017. Using AI to teach AI. *AI Magazine* 38(2).

Goel, A., and Polepeddi, L. 2017. Jill Watson: A virtual teaching assisant for online education. Presented to the Learning Engineering for Online Learning Workshop, Harvard University, June 2017. To appear as a chapter in Dede, C., Richards, J., & Saxberg, B., Editors (in preparation), Education at scale: Engineering online teaching and learning. NY: Routledge.

Goldsmith, J., and Burton, E. 2017. Why teaching ethics to AI practitioners is important.

Goodman, J.; Melkers, J.; and Pallais, A. 2016. Can Online Delivery Increase Access to Education? Technical report, Harvard University.

Maderer, J. 2016. Artificial intelligence course creates AI teaching assistant. Accessed: 2017-11-15.

Markoff, J. 2017. Automated pro-Trump bots overwhelmed pro-Clinton messages, researchers say. https://www.nytimes.com/2016/11/18/technology/automated-pro-trump-bots-overwhelmed-pro-clinton-messages-researchers-say.html. Accessed: 2017-11-14.

Miller, P. 2016. Professor Pranksman fools his students with a TA powered by IBM's watson. https://www.theverge.com/2016/5/6/11612520/ta-powered-by-ibm-watson. (Accessed 2017-10-27).

Nass, C., and Moon, Y. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues* 56(1):81–103.

Sax, L. J. Cracking the code: Making the CS degree accessible for all students.

Schultz, M. 2017. Twitter uncovered more than 200 Russia-linked bots. http://nypost.com/2017/09/28/twitter-uncovered-more-than-200-russia-linked-bots/. Accessed: 2017-10-29.

Simon, H. A. 1995. Artificial intelligence: an empirical science. *Artificial Intelligence* 77(1):95–127.

Teague, J. 2002. Women in computing: what brings them to it, what keeps them in it? *ACM SIGCSE Bulletin* 34(2):147–158.