JORDAN BOYD-GRABER

QUESTIONING ARTIFICIAL INTELLIGENCE

UNIVERSITY OF MARYLAND

tufte-latex.github.io/tufte-latex/

2022

1
Introduction: The Unanswered Questions

> Since this is a book about question answering, I should try to answer some questions to start things out. I'm going to try to answer: "what's this book about", "who am I", and "who can use this book".

AI has been defined by asking questions even before it was "a thing": Alan Turing designed a question answering proceduring to determine if machine is intelligent. In the years since, AI has continued to be defined by questions: answering questions on *Jeopardy!* and on our phones.

This book looks at not just the *how* of computers answers questions but also *why* answering questions is so important for AI. Answering that "why" questions requires us to go back to mythological, civil service exams, and game shows and connect them to AI question answering.

We call someone smart if they can answer questions: the pass a test, win on *Jeopardy!*, or offer a witty retort in a debate. This is not just a recent phenomenon. Question answering is deeply embedded in our culture: in myth Œdipus answered the riddle of the sphynx; getting into a school depends on passing a standardized test; good governance depends on civil service exams; a trivia subculture has grown around answering all manner of questions.

AI has been defined by asking questions even before it was "a thing": Alan Turing designed a question answering proceduring to determine if machine is intelligent. In the years since, AI has continued to be defined by questions: answering questions on *Jeopardy!* and on our phones. From IBM Watson on *Jeopardy!* to Siri and Alexa in every home, the ability of computers to answer is a common yardstick of how smart computers are.

As we'll see this is important in both contemporary civilization and is fundamental to artificial intelligence. The goal of this book is to help explain how computers answer questions on your phone, on game shows, and to fight fake news. Importantly, from my perspective, this won't just be about computers answering questions: humans and computers have a lot to learn from each other, and I hope we'll be able to discover new things about those interactions.

This book looks at not just the *how* of computers answers questions but also *why* answering questions is so important for AI. Answering that "why" questions requires us to go back to mythological, civil service exams, and game shows and connect them to AI question answering.

To be transparent, this book has an agenda: we can better understand artificial intelligence by looking at its ability to answer questions through this historical lens. If we want to call a computer smart—indeed, smarter than a human—we should make sure the competition is fair. Moreover, writing and asking questions is an art that has been refined over decades.

Part of the story I want to tell is how we got here. And why little things like the decisions made at a tiny British library in Bedfordshire fifty years ago have locked the artificial intelligence community into a particular way of evaluating whether a computer is smart. These historical decisions have shaped the systems that we have today; as a consequence, we are stuck with the crappy answers from our smartphone.

But it didn't have to be this way! The very definition of computing and artificial intelligence was based a thought experiment that is closer to how computers and humans should interact with each other. However, that definition—from a boffin in Manchester called Alan Turing—was too fanciful to really be implemented in the lab in the twentieth century. I'll argue that we could do something closer to Alan Turing's vision, and you, the reader, can judge whether it's (still) too fanciful.

This is not just a narrow question for those building question answering systems. I think that the way we interact with our smartphones and computers will define the future of human–computer interaction and thus the shape of the modern, AI-infused economy.

1.1 Who am I?

I'm excited to be writing this because it's something that I'm passionate about: answering questions is central not just to my research program but also to my day job as a professor at the University of Maryland: teaching and guiding students. And it's also important to me as a hobby: even before this became part of my research program, I spent weekends writing and answering silly trivia questions. I'm hoping that this will help more people appreciate the joy and beauty of posing and answering questions.

In my day job, I have built question answering systems that used the structure of Wikipedia to better answer questions, was one of the first to apply deep learning to the QA task, and played a lot of exhibition games against trivia whizzes like Yogesh Raut, Ken Jennings, and Roger Craig.

I'm also a professor of computer science, and one of the courses that I teach is a course on question answering (and this book is the textbook for it).

As a hobby, I'm also a trivia enthusiast. I played QB for Caltech (where I got BS degrees in computer science and history), started Princeton's pub quiz (where I went to grad school), and appeared on *Jeopardy!*. To be clear, I'm not very good at trivia: our QB teams at best cracked top ten nationally, and I came in second place in the one game of *Jeopardy!* I taped (my students are particularly fond of mocking my poor performance on the video game category). But this has brought me a familiarity with the trivia community and how it works, one that I think gives me a unique (or deranged) look at how computers answer questions.

In contrast, I've won more awards as a researcher. Probably the one I'm most proud of is the Karen Spärk Jones award (we'll learn more about her in Chapter 5) from the British Computing Society, but I've also gotten an NSF CAREER award and "best of" awards from Intelligent User Interfaces, Neural Information Processing Systems, the North American Association for Computational Linguistics, Empirical Methods in Natural Language Processing, and the Conference on Natural Language Learning. I've published over a hundred peer-reviewed publications on interactive machine learning, question answering, and exploring document collections; and this isn't my first book, with David Mimno and Yuening Hu, I wrote *Applications of Topic Models*. Although I should say that because I'm a professor, this is mostly the work of the students that I've worked with over the years (I thank them by name at the end of the book).

But the biggest prize that I've gotten is to have a career (and permissive bosses) that allow me to combine my hobby with these intellectual pursuits. And I'm even more lucky that that career allows me to now share it with you.

1.2 How the Book is Structured

While this book is obstensibly about how *computers* answer questions, that is not where the book begins. After a brief introduction laying out the organization of the book, we go to the *history* of humans asking each other questions. This history is important because it explains why humans value answering questions. This provides historical context on why humans equate answering questions correctly with intelligence but also because key elements for artificial intelligence—item response theory and preference models—were developed from looking at how humans answered questions.

The next stop is looking at *today's* question answering, begining with

Alan Turing, arguably the father of computing and artificial intelligence. Two wartime efforts—one at air and one at sea—began a rivalry in question answering, which we call the Cranfield and Manchester based on the college towns home to Alan Turing's theorizing on intelligent computing and Cyril Cleverdon's practical evaluation of computing systems. Having set the stage of this rivalry, we then see how these ideas built the foundations of modern AI from the methods that let *Watson* win on *Jeopardy!* to datasets like Google's Natural Questions that set the stage for smart chatbots and the models like GPT.

Finally, the book closes with the *future* of question answering, focusing on what it means for an AI to be intelligent: how to turn Alan Turing's thought experiment into reality and how to make sure that computers don't trick us into thinking they're intelligent (but just cheating on the test). And if computers are becoming intelligent the question is how to *align* them with users, which sees the return of item response theory and preference models: this time built up so we can better understand why AI says the things it does. Finally, we close with a discussion of the "big questions" that computers cannot answer on their own: how can the models developed from answering questions from humans best work alongside humans for work and play.