Purdue University

## Purdue e-Pubs

# All the Robots Are Coming! The Promise and the Peril of AI

Ian Mulvany
*SAGE*, ian@mulvany.net

Peter Brantley
*University of California, Davis*, naypinya@gmail.com

Ruth Pickering
*Yewno*, ruth@yewno.com

Elizabeth Caley
*Chan Zuckerberg*, ec@chanzuckerberg.com

# All the Robots Are Coming! The Promise and the Peril of AI

*Presented by Ian Mulvaney, SAGE; Peter Brantley, University of California, Davis; Ruth Pickering, Yewno;
Elizabeth Caley, Chan Zuckerberg Initiative*
*Moderated by Heather Staines, Hypothes.is*

*The following is a transcript of a live presentation at the 2017 Charleston Conference.*

**Heather Staines:** Good morning. I'm Heather Staines, the director of partnerships for Hypothes.is, an open source nonprofit collaboration technology. And I'm happy to bring you today four speakers who will be talking about the promise and the perils of artificial intelligence. Our plan for the session is to keep the presentations brief because we do want the bulk of the time to focus on conversations with you. I'll just introduce the speakers briefly. We have Ian Mulvaney, he's the head of product innovation for SAGE. He's going to do a brief intro on "What is AI?" I think these days, it's sort of like everything is AI so nothing is. Ian's going to help clarify that for us. He'll then be followed by Peter Brantley, who is the director of online strategy for the University of California, Davis, and I think well known at this conference. Our third speaker is Ruth Pickering, she's the managing director for Yewno. And then our fourth speaker is Elizabeth Caley. She is the director at the Chan Zuckerberg Initiative working on the Meta project, and she'll be joining us remotely. Without further ado, I'm going to hand it over to Ian.

**Ian Mulvaney:** Okay. Hello everybody. I've promised to teach you what AI is in five minutes, so I'm going to do that, and then I'm going to have a couple of slides about some personal reflections on where we are with AI.

What is AI? So, just some context. We have a lot of data and AI is just one of a number of different methodologies for dealing with data at scale. And I've put four classes of how today we could deal with data at scale. I'm not going to go into any detail, but what I think is really interesting about these different four classes of how we deal with data at scale these days is there are open source tools that are available for all of those types of classes of ways of dealing with that data, and I think it's really fascinating. So, when we think about AI, one of the really key models is machine learning and deep neural networks, and they're pretty simple conceptually. So, you start with some training data, my letter "A" here, and I've got a model and the model actually has no correlation whatsoever to the training data, and you start by asking "What's the difference?" And then you get a difference, a delta between your model and your training data, and then that feeds back into the model, but you train your model with a huge number of examples of that training set and the differences can pile up, and that model gets richer and richer and richer, and eventually you can ask questions like, oh, that was supposed to be a picture of a cat at the top, and you can ask questions of like "Is my test data correlated to my model?" And "Can I tell from my test data whether this thing is an 'A' or whether it's a cat or whether it's not an 'A'?" And so we build up these models by just training our system with this test data, and that is what machine learning is in a nutshell.

Now, one of the key changes that's happened over the last couple of years and is over time we're seeing a huge amount of data becoming available and the cost of computation coming down radically, and so we can train these models with large data sets and run those computations very cheaply, and that allows us to get to the place where these models are becoming very powerful. And so with these models we can do things like explore data, so the models are very good at doing things like identifying objects and images. We can do face detection or we can do prediction. These models have been very good at beating humans at tasks like playing Go because they can predict ahead what plays that the opponent will be able to make, and they can even be used to generate new kinds of information and data, so at the bottom is an example of an image that's been generated from a Bayesian model.

Now sometimes these machines don't get things quite right. So, on the left we have an example of image identification, which was identifying people as guerrillas. On the right-hand side we have an example of a prediction algorithm, which was predicting that only white people could win at beauty contests, so we have to be careful with how we train these models. Also, if you know how the model is built, you can create data that inputs into the model that confuses the hell out of it. So, on the left hand we have a picture of a cat, but the machine thinks it's an avocado, and on the right-hand side we have a 3-D model of a turtle, but the machine thinks it is a

gun. And so you have to be careful about how we construct these models, but it is really just as simple as training with lots of data examples to create this model, which is then used to do these kinds of tasks, so now I hope you all know what AI is.

Now, I'm going to talk about some of my own reflections on what I'm doing and my thoughts on where we are with AI. Very brief. So, I am head of product innovation at SAGE. Some of those problems that I identified earlier come from a lack of an understanding of how to train data well. So, we think at SAGE that is very important to get social scientists involved in this mix, and so we're trying to support the tools that help get access to data and skills, and so we launched this year an online learning platform to teach social scientists how to program.

Now where are we with respect to AI at the moment? This guy on the left is a guy called Henry Maudsley; he's one of my heroes. He invented a machine that allowed people to create nuts and bolts that were interoperable. Before Henry Maudsley, every single nut and bolt in the world could only work with the one that it was built for. They were not interoperable. That kind of led to the creation of the Industrial Revolution. Yeah. That's where we are today with AIs. AIs are handcrafted. They are extremely delicate. There is a dark art, but we are about to approach an era where they are becoming productionized, and so I think we're right on the cusp of seeing them filter into many, many areas of our lives. And then my last slide is I, however, remain an optimist. I think that AIs are coming but they will be friendly and they will help us deal with data at scale. Although I've been reading a few articles this week that has shaken my conviction on that a little bit, but I'm going to remain optimistic and I'm going to hand it over to the next person on the panel. Thank you.

**Peter Brantley:** Hi, everyone. I'm Peter Brantley and thanks for coming. In this talk I'm going to actually take the voice of doom and the dark side as a partial counterpart to Ian's more optimistic take. Because both sides are present in AI and both sides are really important and we have to head into this world very much with our eyes open. So, I want to raise a couple of things that are maybe not immediately present in our concerns but are potentially looming very soon; I'll say also that I'm typically a technology advocate, particularly a network advocate. I got my start in networking with a pre-Internet system called the CDC Plato, which unless you're at least of my age you won't know anything about, and I think you'll either

have to look it up in Wikipedia or consult Brewster Kahle's Open Library to find some of the manuals for it. There is actually an emulator.

So, as Ian mentioned, currently AI works by trying to derive patterns out of large masses of data and when it does that, if it is trained well enough, it can make inferences out of that data, such as the famous inference of, "Oh, this looks like a cat." But if you train an AI well enough over more and more sophisticated problems, then those kinds of associations start to feel like insight, something that heretofore has been the province of humans, and that is what is unique and special about us. But, fundamentally, and I don't think any serious AI practitioner would argue otherwise, what an AI is doing is making an observation, drawing inferences out of correlations of data, but those data may not be causal. Even if something happens in a temporal timeline, so an AI sees one thing happening and then it sees another thing happening, you can't infer that the first thing causes the second thing. So, that's actually a very significant thing to keep in mind.

Now, AI can be classified in two broad ways. Right now we're working mostly with what practitioners call Narrow AI. A Narrow AI is something that focuses on a particular kind of application, like a certain kind of image recognition, that kind of thing. Many of these kinds of AIs, particularly the ones that are now fairly sophisticated, are noninteractive. They pull data together and then they draw an inference. I'll give a couple of examples. One is you can string an airfoil with a lot of sensors for electromagnetic resistance, temperature, and so forth, and you can begin to draw inferences with these data. For example, "Maybe there is a potential crack forming in the airfoil. I, the AI, need to make a signal that the wing should be examined." Another example would be as we start aging more in place, we are wearing more and more personal sensors, and our homes are becoming full with Internet of Things devices that monitor ambient conditions. It is certainly in the works that an AI will begin to synthesize those data and pull them together, allowing the AI to start making inferences like, "Hmmm, maybe this person might be nearing a cardiac event. I need to signal a doctor or a caregiver to examine this person." So, these are Narrow AIs, and fortunately if we make a mistake developing those AIs, typically those mistakes are fairly limited either in scope or in time. Worst case, God forbid, the AI screws up in the aging home example, but we would learn from that and correct that very easily and do better next time without too much wider consequence.

But, AI is increasingly invading areas of social interaction, and here is where I would start to raise a caution. As AIs start trending more toward what we call General AIs or Broad AIs that attempt to synthesize data across a much wider array of inputs and try to make a much broader set of inferences available to us, then there must be some elements of caution. To give a couple of examples here, by analyzing a wide swath of video and textual data, can an AI suggest from its monitoring of conservative political or religious groups that an act of terror might be near? That's a much more difficult and more fraught kind of prediction. Similarly, by looking at food pricing indicators and satellite terrain photography or imaging, can an AI suggest that a population is moving because of ecological stress, or is it because of racial genocide? So these are much more profound discussions and the impacts potentially are much more persistent than figuring out whether I am nearing a cardiac stress episode in my home.

So, I want to suggest, and this is my last slide, that bias is something more than just mere misclassification in AI, and it is really important for us to grapple with that. AI is informed by the societies in which it emerges, but it also informs society in turn. Recently one of the founders of Facebook made the observation that—unknowingly perhaps, or unpredictably to him as a founder—that the growth of Facebook as a phenomenon has changed how we interact with each other. So, in a similar way, the emergence of AI changes how we make assumptions about our world. It changes how we make assumptions about how our machines, our computers, and how the network will talk back to us. As a consequence, AIs pose the threat of supporting hegemony, of violence, or exclusion of populations or segments of society, if we get them wrong. And generally in AI there is a trade-off between an accuracy of an inference and the intelligibility of that AI itself. We are already entering a phase where AIs are in contest with other AIs, and on the cusp of where AIs are capable of building other AIs.

So, I leave with a reference to the article that Ian mentioned in passing, which many of us have read recently. It's an article by the British artist and critic James Bridle, who observed that there is an emerging number of videos targeted to very young children on YouTube that are very disturbing, often with sexual characterizations and violence, and that these videos seem to be a combination of human-designed and machine-made. In other words, they're algorithmically constructed and then thrown into YouTube for discovery and consumption. So, this kind of dialectic between us and the network is becoming more and more pervasive, and it enables what Bridle described as "infrastructural violence against our own society." These are dangers that we can surmount, but we must bear them in mind as we put our hands together to try to build new intelligent systems that can serve us and move us forward, rather than threaten us. Thank you.

**Ruth Pickering:** Good morning, everybody. So, I am delighted to be here this morning talking to you. I work in technology and we specialize in artificial intelligence, and last month I was in London and I was catching this taxi to the airport, and I was chatting to my driver and he said to me, "Where do you live?" And I said, "Actually, I live in California. Obviously, I sound very English but I live in California." He said, "What do you do?" I said, "I work in technology and AI." And he said, "What?" I said, "Artificial Intelligence," and he said, "What?" And I said, "Algorithms." And it just made me really realize that actually AI is far, far from being mainstream. So, when you're talking about AI, it's a bit like saying "medicine." If you work in medicine you could be a doctor, you could be a nurse, you could be a gynecologist, you could be a neurosurgeon, very, very different things, so what we specifically do is a blend of computational linguistics, graph theory, and machine learning. And what we're trying to do and we often refer to ourselves rather than artificial intelligence is augmented intelligence, because what we're really trying to do is help overcome some of the problems that people face every day. So, one of the real challenges is the sheer volume of information that exists and that is created across a very, very wide range of sources that are often fragmented and dispersed.

So, how we use AI is we ingest huge quantities of high-quality information and our algorithms read them all in a way that isn't possible for humans to read that quantity of information both in terms of what already exists, what has been created historically, and what continues to be created on a day-by-day basis. Then we use AI to read the full text and extract meaning in the form of concepts, and then what we're trying to do is to re-create a neural network model, so we knit the concepts together in a big graph network, and so what you see is on the right this enormous multidimensional graph and within the graph it's entirely adaptive, so as new information is ingested, if a concept becomes related to another concept in terms of something that was published that day, that week, the graph will move, so the graph is a fantastically flexible mathematical

instrument, and what's key is that it can be interrogated by different users for different purposes in many different ways, and so however you want to search for information or to find knowledge, you can put in a different query and a different element of that graph network will be surfaced.

So, I wanted to talk to you about a couple of different things and how we hope that they help people overcome this big problem around finding what they're looking for today. So, in terms of how you can create products using artificial intelligence, starting off you can take this huge quantity of data and we take data across all kinds of different sources, it could be anything from the newsfeed to books to journal articles, reference sources, and we also ingest different languages. Then, we run a series of algorithms across that data and that creates this huge network of relationships, and what is really, really important is that part, so those relationships between all those different concept sources across all those different domains of information, and once you've built those relationships into this enormous graph network, you can use that in different ways. So a financial analyst can use it to search for trending use in their sector. A researcher can use it to understand connections in a particular field of research, and one of the things that we are particularly keen on trying to help people do is to uncover inferences across different domains of information. So, if you go to a search bar today and if you put in a query, and if you go to the right place and you put in a good query, hopefully you'll find the answer to your questions, but you probably won't find something else. So, we know statistically that 93% of people stay on the page one of their results list, and we also know that 63% of people take the first three entries only, but are they really the best sources of information that are available and that people could find?

So, using a graph visualization, what we're trying to do is give a completely different perspective to information to help people understand in a more intuitive way concepts, their relationships, the context, the broader perspective, all of that area, and so the way in which we share information and you can see in this visualization on the right is really designed for people to understand the broader context. So, you could put in any query and you'll get a single node in the center and then you'll get a concentric circle of nodes around that, and those are all related concepts. So as you look around you can see if you find anything interesting to you, and anything you see is interactive and clickable. So, you can click, whether it's a concept or whether it is a relationship between

concepts and really try and understand, and then the research process is not just knowing that two things are connected. It's understanding the reason that they are connected and ultimately being able to get back to the source information and understand the kind of key documents behind that.

The other thing that you can do algorithmically is as you are reading the full text of documents—I think I'm missing a slide, sorry, Peter. I'm going to give up going back—but what I was going to say is as you're reading the full text of documents, you can also categorize any document. So, you can say "this book is 63% about engineering of which developmental robotics is 84%," etc., so you can provide this fantastically accurate picture, and one of the things I really like about AI and I like about the technology is that it is unbiased. So it is reading this huge quantity of information and it is able to accurately and consistently apply this kind of categorization and extract information in a consistent way, so really hopefully what we are going to be able to do in this space is to help people spend less time being frustrated, trying to look for the right website, trying to look for the right source of information so that they can find what they're looking for more quickly and spend more time on their research and actually on the argument and thinking about it. So, that is where we're trying to come at this from. And I will now hand it back over to Heather.

**Heather Staines:** Thanks. We'll try to find the missing slide so that it can be included in the presentations later. And next up we have, I hope, Elizabeth?

**Elizabeth Caley:** Hello!

**Heather Staines:** Take it away, Elizabeth.

**Elizabeth Caley:** Good morning from San Francisco. Thank you so much for allowing me to participate remotely.

We at Chan Zuckerberg have been focused on how can we use different technologies including artificial intelligence to help scientists make new discoveries. So, our organization, Meta, is a company that was based in Toronto that has been around for seven years, focused on the application of artificial intelligence on scientific literature in order to make solutions for scientists. We were acquired by the Chan Zuckerberg Initiative, a philanthropy based in San Francisco, earlier this year. You'll hear similar patterns and themes between all of the presenters thus

far, but to start off—a little bit about why we've been doing this work as Meta and why Chan Zuckerberg is now doing this work. We believe it is important to accelerate the impact of science by enabling awareness and understanding of scientific knowledge for free to all scientists and all consumers of scientific literature around the world. So, how can we make the fastest progress in biomedical sciences by facilitating better understanding of the knowledge that is encapsulated within scientists' brains and within multiple sources including scientific papers?

So when Meta was started several years ago, the underlying premise was how to take the information within scientific literature and transform it to meaningful, easy to navigate set of connections, and then take that further to do then predictions about where science is headed or where a particular field is headed or which brand-new papers that have been published are going to be really impactful in your field.

We have been focused on biomedicine to start with. How do you take the 27 million papers that have been published in biomedicine over the last 200 years and transform them into a knowledge graph, that you have heard a little bit about so far this morning, and I'm sure many of you are familiar with. How do you take all of that information and harness it, recognize everything that is going on and tease out the relationships and then use that knowledge in order to predict where things are going into the future? We have built a knowledge graph. Every time a brand-new scientific paper gets published, our systems, using artificial intelligence, using a series of machine learning algorithms, will process that paper within a minute or two and basically add that contribution to scientific knowledge into the knowledge graph. And that works just as well for the 90th paper on breast cancer that will be published this week as it will be for the first paper that has ever been published on a brand-new concept. For example, say the first paper that was published on breast cancer.

The only way to actually do this, and do this at scale, is something that we realized several years ago is to use machine learning and very specifically supervised machine learning, which is a technique or a set of techniques that we've been mostly talking about this morning. That involves the ability to take large sets of data and apply them and use them to train a model and then use them to test how well that model has performed.

Just to give you a flavor of the types of problems that are particularly well suited to the kind of work that we do, we use artificial intelligence to recognize and map entities. Again, we heard a little bit about this earlier, so when a term or technique or a gene or protein is mentioned within a piece of text, how do you recognize what that is, even though it might be referred to in many, many different ways across many different authors over time? Then another set of problems is how do you disambiguate authors, affiliations, papers, and again there's lots of examples inside our field. How do we know this John Smith who published this paper is actually the same or different from that John Smith who has exactly the same name, and works in similar fields, but may or may not be the actual same person? That is a problem that machine learning and algorithms can be applied to.

Third: recommendations. I think that's one of the most familiar applications of machine learning—familiar as recommendations on Amazon's website. How do you use artificial intelligence in order to generate recommendations for users? And then how do you predict on this information? How do you take an individual technique, say a new emerging technique in biomedicine, and predict how much impact it might have in the future? Well, you do that by modeling based on past events and similar entities and how they have evolved over time. Then those predictions can then hopefully help humans make decisions and plan for what they might want to do based on probabilities of what might happen in the future. This is related to the last category as well. How do you forecast what might be really impactful within a field given how fast science is moving, particularly how fast biomedical science is moving? That's been a focus for us as well.

In summary, what we've been trying to do is take machine learning and apply it to scientific literature. Here's one example of how that's being used. This is Meta, which is an application that is available for free for researchers in order to stay on top of their scientific research. We are bringing people back to reading the papers by really helping them sort through that overwhelming amount of new information that has been generated in a single day within the sciences. This is an example of an end result that can come out of a machine learning program with the help of a machine, learning experts, and a whole bunch of data. So, with that, I'll hand it back over to Heather. Thank you very much.