

Jay Shaw on AI

Camille Castelyn and Jay Shaw

Keywords: Narrow AI, Machine Learning, AI and Healthcare, Artificial Intelligence, Public Health, Community Integrated, Technologies, Bioethics, Interdisciplinary Social Sciences, Model, Clinical-Patient Interface, Technology Corporations, Ethical Design, Policy, Data, Population, Health Data Regulations.

Camille Castelyn 0:04

Hello, and welcome to Voices in Bioethics Podcast. I'm Camille Castelyn. And today it is my great pleasure to welcome Jay Shaw. He's an assistant professor in the Department of Physical Therapy at the University of Toronto, with a cross appointment to the Institute of Health Policy Management and Evaluation. He serves as a Research Director of Artificial Intelligence, Ethics and Health at the University of Toronto Joint Center for Bioethics. And he's the adjunct scientist at the Women's College Hospital Institute for Health Systems Solutions and Virtual care.

Welcome, Jay.

Jay Shaw 0:41

Thank you Camille. Thank you so much for having me.

Camille Castelyn 0:44

So as our listeners might know, technologies, and specifically AI in healthcare is developing at an extremely fast pace, while the ethical and social implications of these technologies really lag behind. So maybe just first, which kind of AI systems are we talking about today in health care with your research, Jay?

Jay Shaw 1:05

Yeah, thanks, Camille. I think the field of artificial intelligence is an interesting one. And it's developed quite quickly, particularly the sort of popular understanding of what AI is and what it means. In the early phases of the discussions of AI ethics, there was a preoccupation with the sort of existential threat that's posed by what's called artificial general intelligence or AGI. And that's a version of artificial intelligence that I think is viewed as sort of science fiction by many of the people actually building AI and machine learning models, particularly with applications in healthcare. And that sort of funnels the conversation toward what the community refers to as narrow AI. And in fact, many of the computer scientists (data scientists) working in the AI field prefer not to even use artificial intelligence because of how that phrase has become so hyped up and simply refer to the much more

specific analytic method that they're using, like the variant of machine learning that they're employing when building a particular model.

So narrow AI, what is a narrow AI, it simply refers to the application of the analytic techniques, building algorithms, building models, to perform a highly specific function. And with artificial intelligence, one of the best ways to understand that is in terms of making a prediction. So, I'll give a couple of examples that I think are illustrative. There's prediction in a sort of intuitive sense, where, for example, one could build a machine learning model that brings in different kinds of data to predict the volume of patients that will arrive at an emergency room. And there's a hospital in Toronto that had an early really positive use case around artificial intelligence, where they brought weather data and data from sporting events, like the existence of sporting events and the effects that would have on traffic and people coming together in a closed space. They brought that data together with their hospital throughput data and enhance the accuracy with which they predicted how many patients would arrive in their emergency room that enhance their ability to appropriately staff the emergency room and think about the sort of throughput of patients through the hospital. So that's about predicting something very concrete. But I want to give another example of prediction that I think helps to kind of clarify the wide variety of use cases that can be made using prediction as a primary tool. So, if you think about a chat bot, and I don't know if any listeners have had the opportunity to interact with a chat bot, but it's quite an interesting experience. I was on a telecommunications sort of cell phone provider website recently and was typing into the chat for help. And after a few minutes, I asked, "Is this a chatbot? Am I talking to a chatbot?" And it responded, "You're speaking to a virtual assistant". It was in fact a chatbot. But what a chat bot does is it assesses the text that you're sending it, and it predicts the best possible collection of texts that should be sent in response. So that's also a prediction. So those are just two examples of used cases of narrow AI, where it's the machine learning model is performing a very specific prediction to accomplish a very specific task. So a bit of a lengthy overview, but I would say that the kind of artificial intelligence, I'm interested in, and that's most sort of demanding of ethical, social, legal analysis is these applications of narrow AI in healthcare and public health.

Camille Castelyn 4:53

Okay, great. That's interesting. And I have also chatted with a chatbot. And it was an interesting experience. I think ours was a bit more limited, but definitely interesting, but interesting that you couldn't tell whether it was one or not. Yeah, they're getting better and better. So when we talk about these narrow AI systems, your research, then also looked at what the ethics and the values are that underlie the design of these technologies. And you also suggested an approach which focuses more on community integrated systems, rather than a capitalist approach, if I can call it that. And it seems that more often than not, these systems are designed with a main goal of being profitable. And that values such as equitable access, sustainability, justice, and otherwise good societal values have a much lower priority. So maybe we can delve into that now for a bit. What would you say, are the current design ethics and values, broadly speaking, in many of these systems?

Jay Shaw 6:00

Yeah, thanks, Camille. That's a good entryway into kind of setting the context for how, in my work, I go about thinking about the most important ethical and social issues to consider. And I would say, as a bit of preamble that this is where I draw heavily on the interdisciplinary social sciences in my thinking about bioethics, because understanding the ethical issues presented by the development of AI for health, I think requires an understanding

of the broader political economy in which those AI systems are being developed. And when I say political economy, I mean, the broader collection of institutions and stakeholders or actors who have various different incentives, and in fact, conflicting incentives in that broader ecosystem that's responsible for the development, deployment and scale of AI systems in healthcare. So, you know, you can imagine that a clinician scientist working in an academic health sciences center, building a model to help inform clinical decisions at the clinical patient interface, has a very different set of incentives that govern what they're trying to accomplish, than does a big technology company trying to amass data from health systems, as we've heard about in the UK and in the United States, most prominently, in order to generate models that will ultimately help to enhance the search of electronic health records, for example. So social science, to me helps to understand who are the different actors in those kinds of different AI design scenarios? What are their incentives? How much power do they hold? How do they influence the AI design process? And with what consequences? So when we ask the question, I would say you've posed a question about values. I would say that that conversation is closely linked to a question about harms, what are the potential harms of AI systems that are being built for health care and for public health? And I think that the answer to those two questions, you know, first, about values and second about harms is very different for those different design scenarios. So then we can step back and say, for example, I'm doing some work with the Public Health Agency of Canada, with a colleague, Angela Power and a student, Heather Decker, to think through some ethical guidance for the development of AI technologies for uses in public health. And we've had to have quite a bit of conversation about who was the guidance for? Is it for university-based researchers who get publicly funded grants and are working in the public interest? Or alternatively, is it for technology companies with a profit motive? Because the kinds of guidance and the kinds of governance strategies we would suggest are very different in those kinds of scenarios? So maybe just to kind of carry this one step forward. I'll talk about the role of technology corporations, and what that implies for values and harms. Would that be okay, Camille?

Camille Castelyn 9:23

Yeah, please go ahead.

Jay Shaw 9:25

Sure. So you know, and the reason for kind of taking the conversation in that direction, relates to a paper that I recently published, led by a PhD student at University of Toronto, Joseph Donia. It's a paper we're both really excited about on frameworks for ethics and values in design of technologies, such as artificial intelligence. And in the case of thinking through the value of these frameworks, we confronted the fact that so much of the design of these technologies happens in the context of technology companies. And the fundamental incentive of technology companies is to generate, as you've said, technologies that can be profitable. And that can scale widely. So that context has important implications for whether and how a designer can build client centeredness or community centeredness or racial justice, or inclusivity into the design of that technology, and thinking through that particular challenge, and the role and influence of that, I'll use the word neoliberal corporate context is really crucial for framing the possibilities of ethical design. So acknowledging the context in which design is happening is one really crucial feature for understanding the ethical and social implications of the design of those technologies and their consequences. I want to just address the other point you raised Camille, which was around community engagement and community engaged design. I think, you know, for several years now, bioethics has been tinkering around the edges of a racial reckoning and a reckoning of the role of social justice in bioethics practice and research. And kudos to the recent efforts since the murder of George Floyd to really much more deeply engage

with this discourse and to contemplate a future for bioethics that is thoroughly anti oppressive, I think that's, you know, crucial and really important. One of the most important ways, I think, to carry that, there are many, but one of the most important ways to carry that ethos forward is to partner in meaningful ways to co-lead initiatives with affected community members. So if we're talking about a public health application of artificial intelligence that will predict which particular communities or sub populations ought to be screened for a particular disease, then developing that technology with those community members is going to be a crucial aspect of an ethically designed socially just version of that technology. So where I refer to community engagement as a fundamental feature of the design of artificial intelligence technologies, I'm referring to that as a strategy to partner with and build the experiences, expectations and value systems of those communities into those technologies. Now, that's very difficult to do even for a university-based kind of publicly funded public values-oriented researcher, let alone in a context that's dominated by a drive to make money.

Camille Castelyn 12:52

Yeah. Now, I think that's extremely important. I really love what you say about focusing on the local context. But I wanted to ask for how easy is that to implement? But the fact of the matter is that it doesn't always have to be easy. And for justice to be served, then definitely, yeah.

Jay Shaw 13:10

Yeah, and I appreciate that question. And I think that's exactly the right question. How easy or difficult is it? And how do we do it? Right? Over the past few years, this for the team that I work with, we've been engaging in that sort of challenge quite deeply. And particularly given that everything is happening, virtually, we found it to be a real challenge. It's required finding a way to get devices to community members who can't afford them, thinking through internet access, thinking through digital literacy, or a phrase I heard yesterday from a colleague in Pennsylvania in Philadelphia digital preparedness, digital readiness of people to engage with those technologies. But ultimately, it requires time and commitment and relationship building. That's the number one piece of feedback from community partners in the literature and in our experience is that it takes time, and you need to build relationships. And once those are in place, then you have trust, and then you can build meaningful community engaged research approaches.

Camille Castelyn 14:16

I think that's so interesting, because so many times like if you say, how do we develop these systems, as you said, like social sciences, and bioethics has a big role to play in the development of these systems. And that's often contrary to how we think of it or because we're caught up in this thing where only the big players get to develop the AI or the systems you know.

Jay Shaw 14:41

And maybe I would just add at the Joint Center for Bioethics at University of Toronto, we have really invested in building a network and a research community where we can partner with people building technologies to think

through the values that underlie the design processes. So we have several projects ongoing that are focused exactly on that. And I also think there's a role for thinking critically about bioethics how bioethics approaches its work in this instance, because I suspect that, you know, the history of bioethics has not been a history of community engagement. It's been a history of relying on philosophical ideals that are brought to bear on particular decisions. And shifting toward community engagement really changes that logic of bioethics. And I think that's an important shift that's underway and is one that has rightfully garnered a lot of attention, both supportive and critical.

Camille Castelyn 15:42

Yeah, no, definitely. And making all the stakeholders heard, who at the end of the day will be using these systems are the people who are on the ground doing the work. So definitely important to include all of them. And it's really great as all the work that you guys are doing, it sounds really valuable to have that resource as well, that you're developing in terms of the guidance of how these systems should be developed in public health that you're referred to at the beginning. And I wonder how many other countries actually have that in place?

Jay Shaw 16:17

I can comment on that one of the projects I'm working on is a design ethics project with some researchers and clinicians and machine learning scientists to build an ethically designed technology and a healthcare program in which it's embedded, that uses motion data to predict particular kinds of outcomes. And that includes, you know, so this is a home-based physiotherapy application that uses a wristwatch to sense motion data and identify whether and how people are doing their exercises at home. And it's been this really rich, wonderful experience collaborating with this diverse team who we put in a grant together to get this. So ethics was there from the very beginning of this project. And I really commend the project team for having that degree of foresight. Early on, in that experience, I put out a tweet on Twitter, asking the broad community, "Who has experience working through the necessary challenges that arise when doing this kind of work when you have different worldviews and different incentives coming to bear on the design of a particular kind of technology?" and I was just thrilled with the response we got. And it turns out, just based on our review of literature and the sort of social networking that came through that experience, we identified people doing exactly this kind of work in Boston, in San Francisco, in Munich, in Paris. There's quite a community of scholars who are doing work on design ethics. It's an emerging field. I would say there are important lessons that continue to be learned about the right ways, the practical ways to engage a diverse group, particularly given the kind of well recognized differences in worldview between those who are schooled in mathematics and science kind of mindset versus those who are schooled in a social science and humanities kind of mindset. I'm oversimplifying there but it's a nascent and growing body of work. And there are communities working on this scholarly communities around the world.

Camille Castelyn 18:30

Well, that's great news. And I can only imagine the challenges but also the interesting challenges that arise while doing that kind of work. And I think we live in a world where diversity is often shunned or but I think it's amazing if you can actually work on a global team or a global level like that and just learn from each other as well. So that's really great. Yeah.

Jay Shaw 18:53

The other thing, I might say here, if it's okay, is the World Health Organization published a set of guidelines on ethical governance of AI for health last summer and that's been an important piece. It reviewed, in the process of putting that piece together, a large group of international experts reviewed a huge body of literature synthesize that into a set of principles and recommendations. And it's largely a summary of the work that existed prior but it's very important that a group like the World Health Organization puts that kind of guidance out. They have important moral suasion and to be able to hold those guidelines up, where organizations of any kind are building AI to be used in health. That's a very important tool for ethicists who are approaching this work.

Camille Castelyn 19:44

Yeah, definitely. And I can also just think how that must connect to the policies and how if we have at least a benchmark like the WHO guidance, then that's a good starting point as well. And as we finish up, if we have these guidelines, what role does the legal implications play? Like is it that these systems are kind of free to develop the way that they are? Are there any legal binding regulations at the moment or not really?

Jay Shaw 20:16

Yeah, really important point. So you know, I'm not a legal scholar, but I've done some work on the policy and regulatory frameworks around AI for health. And there are two distinct but related issues, at least that occupy my attention in the policy landscape. One is around strategies to regulate software as a medical device. This is really about apps, apps and other software systems that interface with healthcare. Like you could imagine add-ons to an electronic health record that informed decision support systems for making healthcare decisions as an example.

But other mobile apps that you download onto your phone, that relate to, you know, a wide variety of health-related uses, from something as benign as kind of tracking your running route through to something much more, I would say, serious and health related terms with respect to something like seizure detection. So those apps the risk they pose needs to be assessed, and they need to be regulated, approved, or taken off the market in various ways.

And the Food and Drug Administration in the United States has a particular model, Health Canada's model is very similar. European regulation has a similar approach, as well, how these apps are regulated is an important topic to attend to. The second domain of policy is around data governance. What data count is health data and what are the policies and regulations that govern how those data can be used. There's an important example with respect to motion data, which I already raised. Because motion data when collected by a healthcare provider in Canada, or the United States, is considered health data. But collected outside of a healthcare context by for example, you know, a fitness app, is not considered health data. And the frameworks governing the data in those two instances are very different. Now, even when something is considered health data, existing policy frameworks allow for it to be de-identified to a certain degree and then put to secondary uses. And that is a major topic of discussion by governments and ministries of health around the world, at the moment around how to think about secondary uses of health data, particularly because healthcare systems often have sort of treasure troves of huge amounts of health-related data from defined populations. So, there are these sorts of distinct policy and regulatory issues. And I would say that there's a huge unregulated space around data that can be used to infer health status and other health related phenomenon. So, the way that people shop online, the apps that are used, how long they spend using apps, many of these things can be used to predict mental and physical health realities. And yet, none of

those data are covered by health data regulations in countries around the world. So that massive unregulated space is going to be a topic of great interest and really important policy attention in the coming years.

Camille Castelyn 23:40

Yeah, that sounds fascinating, as well. And as you say, a lot of work that still needs to be done. But thanks very much for sharing your research expertise and the work that you've been doing as well with us today. It's been very interesting and great to hear that there are places where the stakeholders are all being accounted for and at least starting to think about these issues.

Jay Shaw 24:04

Yes, thanks so much for having me.

Camille Castelyn 24:06

Thank you, Jay.