

# AI and You

Transcript

Guest: [Fiona McEvoy](#)

Episode 201

First Aired: Monday, April 22, 2024

Hello, and this is episode 201. My guest today is Fiona McEvoy, author of the blog [YouTheData.com](#), with a specific focus on the intersection of technology and society. She was named as one of “30 Influential Women Advancing AI in San Francisco” by RE•WORK, and in 2020 was honored in the inaugural Brilliant Women in AI Ethics Hall of Fame, established to recognize “brilliant women who have made exceptional contributions to the space of AI Ethics and diversity.”

Fiona’s a really good role model for how a young person can carve out an important niche in the AI space, especially for people who aren’t inclined to the computer science side of the field. We’re going to be talking about that perspective, her journey to becoming an influential communicator and the ways she carries that out, what it’s like for young people in this social cauldron being heated by AI, and some of the key issues affecting them. Here we go!

Well, Fiona McEvoy, it’s a pleasure to have you on AI and You, and I wonder if you could describe for us your trajectory here, the path that led you into this business of AI and ethics, because I always find those stories interesting and they help our listeners understand vicariously how they might do the same.

Well, first of all, very happy to be here and to be invited, Peter, so thank you very much. Yeah, I can’t, I think like a lot of people who sort of write at the sort of intersection of technology and society, I arrived via a non-traditional route. I’m not sure there is a traditional one, but hopefully there will be, or there is that’s currently being established. So going right back to what I studied at undergraduate, I studied classical history and I also have an English literature degree as well. So I was interested in, I guess, civilisations and literature and all the things that are, or at least feel, very far removed from AI and technology. I worked in communications and you can probably tell from my accent, I’m from the UK originally, for around 10 years or so and then ended up moving to San Francisco and taking a bit of a career break. I did a Master’s degree in something that I’d always wanted to do, which was a Master’s in philosophy and I never imagined that I would end up writing about the ethics, effectively the ethics of science and technology, but that’s where I found an interest and being in San Francisco, it felt relevant to start thinking about how those ethical principles that I was being grounded in as part of my Master’s degree were relevant to the technologies that were being built sort of several miles down the road. I was at San Francisco State, I should say that, which has got a great Master’s programme for philosophy, I should give them a plug. But, the science and tech we were writing about was all older stuff. It was like CCTV and things like that. And so I was very conscious that technology had moved on somewhat since some of the syllabus had been written and I asked my thesis committee if I could perhaps write something about what at the time, and you don’t hear this term very often, but at the time people were talking about big data and how big data was

being used and I was interested in this idea that vast amounts of information was being collected about human beings in the world, people and their families, individuals, and analyzed, and in some ways being used to do two things, to make decisions about them, what they were able to acquire, in some cases thinking about a job, or credit. And on the other hand, also to be nudged and persuaded, maybe thinking about advertising and that kind of thing. And I just thought it was interesting and there was something in it. And so I wrote my thesis on big data and the future of autonomy. And so, it was quirky, there was no one really on my advisory committee who had dealt with this sort of thing before. And I ended up publishing it and got to speak at some conferences. And then, you finish a Master's degree and I'd already been in work 10 years before that. And I didn't really know what to do. And it was suggested to me, and I honestly couldn't tell you who suggested it, that I start a blog and just start writing about some of the stuff that I observe. And so I did. It was very much an outlet for me to think about things, to continue my philosophical career in some ways and with some portal for it. And to just keep abreast of what was moving in technology, which obviously at the time and I'm talking about 2017 now, 2016, 2017. And so I kept that going. And then, I found that if you write something with enough frequency and it's in any way connected to the current zeitgeist, people will start to read it and they'll start to comment and they'll start to share it. And that's exactly what happened. And I found myself, with no technical experience whatsoever, no grounding in computer science to speak of beyond real layman's knowledge, being asked to appear on panels and give keynotes and things like this, which - talk about imposter syndrome, it was off the scale at that point. But I started to settle into it. And I thought, it's right that ordinary people should be reflecting upon this stuff. I want to make this stuff accessible if I can. I want to try and understand it. And maybe other people can come along with me and can help me or maybe it will help them learning alongside me. And so I've sort of seen myself since then as a little bit of an everyman figure reading stuff, trying to understand it, trying to understand what it means, mostly really for ordinary people. And so that's how I came about. I came to where I am now. And I mean, ethics, I'm not an ethicist. Ethics is a component of it, but a lot of it is just a reflection and analysis and adding my two cents where it's wanted or needed or possibly not in some cases.

How did the move from the UK to San Francisco change your path here? Did that have the effect of landing you smack dab in the middle of ground zero of the development of AI at the right time? Was that a catalyst for you?

I think that's absolutely right. It was it was a sort of convergence of, I was moving, it was the right time for me to do something else, which is why I went and took that career break and did that. There's two years doing a Master's degree. And at the same time, there was obviously, unbeknownst to me, this ramping that was occurring of, technology, this second wave - another wave, I should say - of innovation that was happening at the same time. And it was just so obvious to me that that is where my focus should be. And I guess I brought to it my job. My last job in the UK. I was working with the financial sector. I was working in the city, as we call it, London's equivalent of Wall Street as a communicator, which is what I've done for many, many years now for the British Bankers Association, which no longer exists, but is effectively the lobby group for the banks that are headquartered in in London, or in Europe even. And it was interesting because what I learned there, which was a lot to do with trying to understand how to

cultivate trust among consumers and build a foundation of trust and a better relationship with the end user, seemed to me to be increasingly applicable to an environment where you have the Facebooks of this world and Googles, where again, the end user tends to be on the wrong side of an information imbalance. And a big entity is telling them to trust. And so I think there's an interesting communication dynamic there. And I think that I could see some parallels from what I'd done previously, and that was certainly helpful.

The philosophy background is interesting because I've talked with people in the past who were doing - some time ago - philosophy degrees and PhDs who never expected it to develop into this sort of thing. They were doing Heidegger and Kant, and thinking, this is fun, but I'll never get a job out of this except teaching philosophy to other people who need to get the degree. And then suddenly they find themselves in the middle of the AI ethics explosion. To what extent is the field of AI ethics now deriving some of its principles, some of its fuel, from fields in philosophy that had not been turned over much for commercial purposes before?

I think "hugely" is the answer to that. And it's a really interesting dynamic. I mean, there is precedent, right? Like medical ethics is a prominent field and you kind of see it there, but it is, it's incredible how quickly it's grown and how, and I was the same, by the way, the reason I went to do this sabbatical was I'd been studying in the UK, I was at Durham University and I was studying classical history and I ended up studying Aristotle and Plato and I loved it and I never forgot about it. But it was always a hunch in the back. I fancy going back and studying a bit more. So my absolute, my absolute reason for going back was to sort of get more into the classics, to learn more about these sort of ancient philosophers and ended up doing something that superficially is extremely different to that. But actually, to your point, I mean, we're dealing, when we talk about AI ethics, we're dealing with principles that have their roots in history, very, very long history, thousands and thousands of years back. And we are dusting those off. And I know, I don't know if you're familiar with Shannon Ballor, who's currently the University of Edinburgh. She was at the University of Santa Clara. But she talks about virtue ethics, which is this very specific component of the ethical spectrum. And she's bringing that to technology. And it's functionally extremely helpful as a lens through which to think about how technologies should be used, how they should be built, how they should be built and how they should be received. And so it's interesting now. I mean, on one hand, I mean, the, as you said that, the cynic in me kicked in, I think, because I think about when I first started writing about this stuff, I was concerned about how, if we're all shepherded by algorithms, to what degree do the algorithms begin to sort of determine us? If our choices are on some level the result of some algorithmic nudge, at what point do we stop determining the algorithms and they start determining us? And now thinking about the field of AI ethics, I wonder now whether there is a whole field of ethics and a whole bunch of people that are now dancing to the tune of the Silicon Valley drum because they want, they see careers there. And I don't blame them for that. But I wonder to what degree it's now shaping a whole generation of ethicists and whether that whether there's something in that actually just as a thought as I'm having live.

Wow, you just went right to free will and determinism.

I did, right?

I mean, that is fascinating, again, because when we have things like recommender algorithms changing people's psychology, then it certainly does intersect with that. And I think some of the, I think another example of what I was thinking about was 1967, Philippa Foote comes up with the trolley problem. No way is she thinking about autonomous vehicles. She's just thinking about how can we gain insights into people's value systems? And now people are all like, how do we embed this in code in a self-driving car? So what about this field most drives your evident excitement and passion for it?

Because I think that it's extremely important that conversations that affect, or decisions that affect everybody - and when I say decisions, I mean the decision to build, create and innovate or the decisions that drive those things - it's extremely important that those are not had behind closed doors in Silicon Valley or equivalent areas or in government offices or, at VC firms that they're out in the open and that we can all have a view on how our universe is going to change, how our world is going to change and how we live in it. And so for me, I'm driven by this idea of opening it up and making it accessible and starting with simple ideas, looking at specific technologies or specific benefits or specific impacts and analyzing how they are going to touch people's lives for good or for bad. Understanding whether a harm's creeping in, but also where this could be a step change that makes the world a better place when you think about developments in health care and technology that applies to environmental sustainability and such like. It's not all bad. It's not all naysaying, but it's just an analysis that I think allows people to at least understand what is going on around them. Because I think at the moment, there's a lot of people who hear the term AI a lot and they don't necessarily understand what it means, aside from apocalyptic ideas of, sort of AI gone mad and taking over the world, if that makes sense.

So for a long time, virtually everything that was being done in this field constituted creating awareness of these issues for people who would otherwise be - AI, what? What are you talking about? Now, of course, there's a lot more awareness and has the balance shifted into more of, "We know about the problems, but what do we do about them?"

Look, I think the balance has shifted. I'm not a policy person and I'm not a regulator. And so I can have an opinion on how we should address the problems, but I'm interested. So here's how I'm thinking about it at the moment. When we leave formal education, we come out with a whole bunch of stuff that we've learned, right? And I'm thinking when we graduate high school, that age, and, you know what a glacier is, and you can do a quadratic equation, and you might know, some of the history of the country that you were raised in, and there's a whole ton of stuff. And most of it you might not ever use again. I have never had a cause in my adult life to use a quadratic equation. I can sort of list the planets in order from Earth, but I've never had any real reason to deploy that information. But nevertheless, it's like this spectrum of latent information that sits in my head. And I'm glad that I have it because when I see something, I can use that. I can access that part of my knowledge, my baked in information, and it helps me understand things. So when I see a glacier, I know it's a glacier, and I have that information. I use it very

rarely, but every now and then you see a glacier. And I think there is still, even though there's more awareness of AI, I still think there are vast swathes of people that do not have even the most basic ideas about what AI is baked into their latent information reserves or whatever you want to call them. And I think that is worrying. And those people - I'm not talking necessarily about people in middle age or older - there are people who graduated college within the last five, six years who are necessarily going to be living in a world working, operating, having children, families in an environment that is, living within the machine almost, right? Because as I said, it touches every component of our lives. And they never learned about this stuff at school. They did not learn about it at university and everything else they get is mediated through anecdotes that friends tell them, stuff in the media, which we all know can be completely hyperbolic. And so I think there is still a lot of work to do, to try, and I'm doing that in a very small way, reaching very limited numbers of people, but I still think there is a baseline layer of knowledge and information that still needs to be, still needs to permeate the general public and it isn't happening. So yeah, we should be trying to solve these problems. And I think that's one component, but I definitely think there is a communications problem and an information problem and an education problem when it comes to whether people understand what it is that they're dealing with when they click a cookie button and accept all the cookies; or whatever it might be, or when they agree to have a job interview that uses a video that then is going to be used for analysis of their facial expressions in order to determine their emotions and whether or not they are a driven person. People don't realise that this is going on and they might not care and that's fine too, but I think just having that information available out there and easily accessible is important.

**What's the constituency or demographic that you feel has the most at stake here and why?**

I mean, it's easy to say young people, so that's where I'll start because, usually that's, those are the guys that have the most at stake in any given generation. I do think that I was thinking to those, that, I guess that 24, 25+ range as well, they're already in the working environment, they're already having to deal with these terms and they don't necessarily have anything beyond what they've read in the media or maybe some of them have picked up a book, but I do think there are issues there as well because if you're in your early 20s you've got, conceptually at least 40 years of your career left and you're self-educating on this stuff, I think. And I do understand because I speak with, at and with colleges a fair amount and sometimes with high schools as well, that there is work being done to give the grounding that I'm referring to at the moment with informal education and that's great, but honestly I feel like, and I do understand that there is a lot of work being done on this front, but there's still work to do universally. I still think that, older people as well, there are lots of reasons why they might be deeply vulnerable in a society, particularly where you think about technology like deep fakes, synthetic media, where they're quite literally for the first time, the same as all of us, but with certain vulnerabilities, inhabiting a world where you quite literally cannot believe your eyes and that is a change-up for the human condition.

It reminds me, the paradox that I've observed in my own interactions with people of all stripes is that the younger ones have more at stake, but paradoxically it's the older ones that have the more anxiety.

Yeah, I think that's absolutely right.

And, speaking of young people, we were talking about you having some plans for writing a book before we started recording. You want to say something about where that might go, what that might cover?

Yeah, absolutely. So, when you've written however many as 80-odd blogs, you sort of start to think that you've got a book there anyway, which I can tell you in no uncertain terms that you do not. You have a collection of disparate ideas, most of which were thrashed out in the moment with a little bit of thought, but not too much, because there are other things going on and it's just, an instantaneous reaction to something in the media very often. But that said, I do think that I've been writing, and I've been writing [youthedata.com](http://youthedata.com), my blog, since 2017 now. And it has very naturally settled into a few different categories where the reason I look at these categories is because I think that there's cause for concern there for whatever good might come there are reasons for hesitation. So, like surveillance is one of the big ones, synthetic media, as I've mentioned before, nudge, algorithmic nudge, as somebody termed it, and various other categories that affect the human condition. And so, the idea really was to pull those together in a way that lays a groundwork that might be useful to younger people and younger people, because I think that they're important, but also, in terms of like, as you say, like who's got the most at stake, but also because, my natural writing style, I like to try and be a little bit funny and sassy, and I try as much as possible to write in a way that isn't in the total abstract, as I think often these things are. And so, yeah, as I say, I'd like to play a role in helping create something that introduces these ideas in a funny, helpful, irreverent way, but at the same time, drawing very clear parallels with certain parts of history, and trying to emphasise where the individual might have, at least reason to pause or read the small print.

What's it like trying to help young people, that demographic, when they are at ground zero of so many conflicts? You've got social media trying to influence their decisions, change their psychology; you've got generative AI coming into the mix, and schools going, Don't use this thing, because we don't know how to assess you, you might be cheating. And they're going, well, if I can use this to write the answer anyway, why are you training me to do something that a computer can already do? And a lot of these issues are being worked out live in real time with them as the experimental subjects.

Yeah, I think that's an extremely good question. And there are so many forces playing down. I mean, I was lucky enough, I taught for a couple of semesters at SF State. And one thing I learned - and they were undergrads, is, young people are a lot smarter than you maybe think they are when you haven't been around them for a while. And some of the observations that were coming at me were extremely erudite. And I think making points similar to the one that you've made that, there are lots of contradictions going on. And a lot of conflict at the heart of their existence right now. I think, for me, I mean, I'm often, and hopefully, for good reason, put in the AI ethics

bucket. But for me, one of the reasons I sometimes resist it is that AI ethics is quite prescriptive. Very often, as recipients of other people's AI ethics, we are subject to quite forceful prescriptions of what is right and what is wrong. And I don't want to do that. I don't want to be telling people what is right and what is wrong. I don't want to be enforcing an ethical framework upon anyone, any demographic or anything else. For me, I think it's about forearming young people with the right questions and not necessarily having all of the answers. So that's my approach is to say rather than saying, for example, all government surveillance that involves facial recognition is bad. the question is, like, is it always bad? Like, is the system in China that, is that oppressive? If it is, then why is it? like, why do we understand that constant surveillance to be something that is a negative and unpacking those ideas, unpacking those ideas of privacy, so people can make their own decisions about whether these things are permissible or not. And I think, honestly, in writing something, and it is difficult in a point, with generative AI really being upon us now, I'm very conscious that I'd like to write something that lasts and it's very difficult in extremely fast moving environment. But I think the only way to do that is to just to introduce questions rather than try to be smart about answers.

Exactly. I couldn't agree more. What do you find are some of the most powerful questions that you asked them to ask?

Well, yeah, I mean, like I say, like the book I'm writing is really quite across the board in terms of different components of our technological lives. So there are a ton of questions. I think one of the ones that's really interesting to me at the moment, and they've had all the conversations about is this idea of AI and creativity, right? So not necessarily the IP problem and the copyright problem, that's a whole bag of worms in and of itself. And we all know that. But, can AI be creative? Is it okay if we if we outsource new large components, creativity and film and music and art to a system, to an unthinking system? Or is there something baked into the notion of creativity that requires human consciousness in order to make it valid? I mean, I like questions like that, not least because they're philosophical, and it takes me to those ideas, but also because I think they're important. And there's no right answer explicitly. But I think in having those conversations, we can start to at least determine things that we don't like. And then, that's where I would hand off to those that have that's where I would hand off to those that have stronger ideas on how to regulate, how to contain, and how to police this stuff.

So in essence, it sounds like you're training them to create the next generation of AI ethical frameworks.

Yeah, maybe that's what it is. Maybe it's trying to have a little hand in forearming a new generation of people in the how to think about this stuff and how to draw conclusions that might be helpful.

Right. And it just boggles my mind to think about what it's like for people growing up where they will always have known generative AI, large language models, and somewhere not far off in that progression will be artificial general intelligence.

Well, I don't know if I agree. That it's that far off or---

That it's that close, you mean?

I don't think it's that close. But that's just my opinion, right? I think, I've been in an autonomous vehicle in San Francisco, and it's okay. Like it can't necessarily park that well, sometimes like we're okay, we're doing like AI is functionally doing some things that are great right now. They certainly don't want to disparage it. And I'm not a Luddite when it comes to this stuff. But do I really think that an artificial superintelligence is around the corner? I don't know. I suppose it depends how you define it, right. And, and then there's that just because something is super intelligent doesn't mean it has some emergent consciousness. So I certainly don't believe in the latter, whether you could have something that is sort of generalised, but complete AGI, I'm just very sceptical. But we should be thinking about it, we should be talking about it, we should be preparing for it, because I'm not the expert.

And more importantly, we should be doing what you're doing, which is preparing especially younger people to think about and tackle those issues, because they are also parts of a giant, unsupervised experiment that's being done on the human race by itself.

Yes, right now. Absolutely.

I believe you've got some events coming up soon, or you are going to a conference in Montreal. Tell us about that.

Yeah, I do. I am very fortunate. This is my, maybe my third or fourth year chairing the World Summit AI Americas, which is in Montreal, there is another one people might be familiar with in Amsterdam, which is the European one, but the one in Montreal is a phenomenal event. We'll have Joshua Bengio, I believe is going to be there and a few others I should know, but I don't, so check out the website. But it's, it's an incredible couple of days where, we hear use cases. And at the moment, that's really interesting, because generative AI use cases, I mean, that's really the hot topic. But also, there is a lot dedicated to responsible AI, AI for good as well, which is often the two are conflated, but they are separate. And it's great for me, because I am a learner, right? So I get to listen, direct and ask questions and be in the mix with some of the absolute smartest people in the world who are thinking about these problems and do something that I'm not doing, which is solving for problems and innovating and creating incredible systems that will do brilliant things in most cases, and I get to listen, try to assimilate. I mean, obviously, I try to write afterwards, and express my understanding of what's going on right now. What are the hot issues? What are the hot buttons? So yeah, it's great. And I'm always honored to be involved.

When you're at an event like that, which has got a large number of technologists, people who are driving the AI engine at the moment, designing it, what is your relationship with that role? How do you see the jigsaw puzzle of how you fit into what they're doing?

Yeah, that's a really good question. And it's definitely changed. When I first started, I think I did my first panel, which I ended up moderating, I should say, was 2018. And there was definitely, like this resistance to people like me, whatever that means, who had no real knowledge of the technology, no computer science background. I mean, absolutely not. Right. And that we didn't



know, what we were criticizing, worries about algorithms and what they're being used for, and it turns out, I think that there was good reason to be worried about those things. And that evolved, and it's evolved hugely. And there was definitely a turning point, and I couldn't tell you exactly what or when it was. And now I see there is a huge receptivity, especially among younger people in the field to understand how they can build technology responsibly, what they should have at the front of their mind, what end users should they be envisaging when they're building, to make sure they're inclusive, and those kinds of things. And so I don't know that it's my role to educate them in any way. But I'm certainly happy in many respects to be the conduit to, I know a lot of people who work in this consultant AI ethicist, and they're able to give very practical instruction that I'm not able to give to technology startups and those working in the field. And so I'm happy to be that conduit in many cases. But it feels, and maybe I'm wrong, maybe the door slams closed, and there are, and everybody thinks it's all fluff and nonsense still, but I'm something of an optimist. And it feels to me as though the field and those within it are much more receptive and open to, the idea that they have to keep in mind some of the potential harms. And also, honestly, it's like for many of them, I think the greatest case for that is that it just makes good business sense. Having a biased system, especially if it's a consumer facing system is a terrible idea as a business, right? You want to be able to have your technology appeal to and understand the demography of the environment, the population in which it serves. And so, yeah, I mean, there are lots of motivators. And increasingly, obviously, we're seeing, especially in Europe, there's regulations now as well. So it's carrot and stick a little bit as well.

I think you put it very well there. I think that they realized once they started intersecting more with the lives of non-technical people that they needed non-technical people to help understand that audience, that customer base, and to act as you say, that conduit to them. It's been a fascinating discussion. What would you like to leave our listeners with as final thoughts and how to find out more about what you're doing?

Yeah, sure. So you can always find me increasingly now. I had a bit of a pause, but Generative AI has very much brought me out of my hiatus. So you can find me at [youthedata.com](https://youthedata.com) where I'll be sort of giving my own two cents on the order of the day, whatever that might be. And yeah, and on there, you can also find details of conferences that I'll be speaking at and are often moderating or hosting. And yeah, I try to keep that updated and updated. And also people should feel free to contact me if anyone wants to. My blog is predominantly me, but I have guest bloggers as well. So if anybody feels that the inspiration, then I'm always open to having contributors. So yeah, you can find me at various conferences and that's all detailed on my website. So do have a look and feel free. Should anybody want to blog on YouTheData, it's predominantly me, but I do have guest contributors from time to time. So I'm more than happy to host blogs by anyone who has any thoughts on that intersection of technology and society.

Great. Well, thank you, Fiona McEvoy for coming on AI and You.

Awesome. Thank you.

That's the end of the interview. There's a link in the transcript to the [World Summit AI Americas](#) event.

In today's news ripped from the headlines about AI, an unidentified multinational company lost over \$25 million in a scam after employees at its Hong Kong branch were fooled by deepfake technology, with one incident involving a digitally recreated version of its chief financial officer ordering money transfers in a video conference call. Everyone present on the video calls except the victim was a fake representation of a real person. The scammers applied deepfake technology to turn publicly available video and other footage into convincing versions of the meeting's participants. The company employees in the call looked and sounded like real people the targeted employee recognized. Acting senior superintendent Baron Chan Shun-ching of the Cyber Security and Technology and Crime Bureau said the employee followed instructions given during the meeting and made 15 transfers totaling HK\$200 million to five bank accounts. Police are still investigating and no arrests have been made.

My guest next week will be Eleanor Drage, Senior Research Fellow at The Leverhulme Centre for the Future of Intelligence at the University of Cambridge, co-host of the Good Robot podcast and co-editor of the book "The Good Robot: Why Technology Needs Feminism." That's next week, on *AI and You*.

Until then, remember: no matter how much computers learn how to do, it's how we come together as *humans* that matters.

<http://aiandyou.net>

Get the book: <http://humancusp.com/book2>