AI and You

Transcript

Welcome to episode 22. It is an exquisite pleasure to introduce this week's guest, Pamela McCorduck. She is the historian of artificial intelligence, having written it down in two landmark books, first *Machines Who Think*, which was published in 2004, and the sequel, *This Could Be Important*, which came out last year.

Now, maybe you're thinking, hold on, AI hasn't been around long enough to start thinking about history, surely? These people are still alive, aren't they? Well, not all of them are, and it was a close brush with death by AI pioneer John McCarthy in 1974 that prompted Pamela to think that maybe someone should get these people to tell their stories before the Grim Reaper claimed too many of them. That someone turned out to be her, and she inserted herself into the community and thoroughly documented where AI had come from.

A quick note: This interview was recorded remotely for the obvious reason in 2020 and the audio quality isn't as good as we would like it to be, but you can still hear everything Pamela has to say.

I don't know about you, but I hated history in school. It was my least favorite subject. It was one set of meaningless actions and dates after another. It wasn't until much later when I read the book *Generations* by William Strauss and Neil Howe that I found an appreciation for history as a set of interlocking stories that were connected by the motivations of people in the context of the groups they belonged to and shaped.

In a similar way, Pamela McCorduck's books are so much more than the history I came to hate. She wasn't really a historian at the time, she was a novelist, for which we can be profoundly grateful, because it meant that her histories are rich with engaging and wonderfully humorous personal stories about the characters in them.

Just to give you a flavor for that I'm going to read the first couple of paragraphs from Chapter 1 of *Machines Who Think:*

> "Can a machine think?" This question is in a class with those snappy vaudeville comebacks: Does a chicken have lips? And like them, it ought to end the discussion at once by its self-evident nonsense. After all, we agree, our one essential, identifying property is thinking. Don't we call ourselves *Homo Sapiens;* declare that we think, therefore we are; and consider ourselves lifted above the rest of the earthly beasts by our capacity for symbol making? And if we are lifted above the beasts, no use even to talk about machines. Can a machine think? Does a chicken have lips?

> Yet for all its absurdity we find the idea irresistible. Our history is full of attempts— nutty, eerie, comical, earnest, legendary, and real—to make artificial and tell to make

artificial intelligences, to reproduce what is the essential us, bypassing the ordinary means. Back-and-forth between myth and reality, our imaginations supplying what our workshops couldn't, we have engaged for a long time in this odd form of self-reproduction.

I hope that gives you an idea of why I am so enamored of Pamela's books.

Some of Pamela's many other achievements are that she is the author or co-author of eleven published books, four of which are novels, and seven focused on the intellectual impact of computing especially AI. Other books that she's written include *The Universal Machine* and *The Fifth Generation. S*he has written works appearing in *Cosmopolitan* and *OMNI* and the *New York Times* and she is a contributing editor to *Wired*. She received a BA from the University of California at Berkeley, which you may know from some of the other things we have said in the show and my book is a hotbed of AI research and philosophy, and MFA from Columbia University. For three years she was also a vice president of the Authors Freedom of Expression organization, PEN American Center in New York City.

So without further ado let's get to part one of the interview with Pamela McCorduck.

Hi, I am here with Pamela McCorduck, who is the historian of AI; not just the historian, but the *raconteuse*, if you will. Her book *Machines Who Think* just humanizes the giants of the field. I would say it relentlessly pricks the balloon of academic reserve in the way it just gleefully describes so many fascinating details. The irresistible image I get is of you merrily lifting the skirt on their professional distance, while they desperately try and push it back down, sort of like Marilyn Monroe standing over a grating. Well, at least it's irresistible to me. And it's not just a history, it's high drama, it's soap opera, it's poetry. And I wonder if when the people in that field were encouraging you to proceed with this, whether they had an idea of what they were going to get when they were talking to a novelist as opposed to a, what we usually think of as a historian.

That's an interesting question. Nobody ever said, "Don't print that." Nobody ever did. They were very open with me, and of course, as I describe in a subsequent book, *This Could Be Important*, I was really friends and social friends with a lot of them, most of them, especially the founding fathers. So the subsequent book gets to tell a little bit more of the personalities, although it's true, *Machines Who Think* did have something about the personalities in the book.

Well, it's irresistible again in that book. And in fact, both of them, the amount of personal detail that you tell, I think that it is a gift to future historians who want to know, above all about the people in this field because anyone can read the textbooks.

True.

But where else are they going to find out about the people behind them, which is what ultimately, history should be more interested in? And you mentioned your subsequent book, which I've already described, and when you compare and contrast the people at the beginning of this era of AI with the people that we have now, the giants in the field, some of which you

get into describing in *This Could Be Important*. How would you say they've shifted along any dimension that occurs to you?

That's an interesting question. I think one of the things that's most noticeable to me, is the financial incentives in the field now. There were no financial incentives and AI in the early days. People like Alan Newell and Herb Simon and other people like Marvin Minsky, they would go and beg, essentially, for money. Now, I was just reading that the number of academics that have been hired away into industry is phenomenal. And this has some very bad downsides in the sense that who's going to teach the next generation?

Right, in fact, you remind me of a line from Rana el Kaliouby's book *Girl Decoded*, where she and her mentor Rosalind Picard, go and ask for I think more grants, more funding in the department and the answer is, "You don't need more funding. You need to spin off", which probably wasn't a thing back in the '50s and '60s,

Oh, no, no, not at all. None of this was ready for any kind of development, really.

Do you think of people that you would like to have interviewed but didn't get the chance because, like, say, Alan Turing, they were either too far away or had shuffled off this mortal coil, in this case, both.

Yes. Alan Turing is the first one who comes to mind because he was dead by the time I took up my interest. But on the, I don't know, hundredth anniversary of his death, I wrote a little piece for *The Atlantic* saying, "Alan Turing saved my life", which he did because it was thanks to him that little kids like me could eat in the United Kingdom during World War II. The convoys were under terrible pressure, and it was thanks to Bletchley Park that they were protected much more than they would have been otherwise.

And you were born in the United Kingdom?

Yes, I was.

Right, and moved from there when you were how old?

I had just turned six.

And so, did you feel any pull towards the homeland there in any sense of some of the people in that field? Or some of the work that was being done that you have any, perhaps feel a stronger connection to than you might if you had grown up in San Rafael?

That's an interesting question. Yes and no. Yes, because such wonderful things were happening then and the brains were being allowed to do whatever they want it to do, and I loved that part of it. And I'm sure I'm an Anglophile. I mean, when you said Victoria, where you're based, I thought, "Oh, afternoon tea. Yes."

Yes, our motto is, "We put the British in British Columbia."

That's right.

You've spanned the gamut of AI history, not just the part that you're personally around for but you write about it from prehistory.

Right.

And so, booms and busts are a big thing in AI, to the extent that they have a name, the AI winter, the period where AI is the name that must not be spoken. If you want to get any funding, you call it something else. It has gone through those periods.

Peter, may I say something about that?

Yes.

You're right that funding is up and down but research is very steady. The first so-called AI winter was in the '80s, and I thought, "Gee, that's funny. Okay, I'll take a look." And I looked at the papers being published during that decade. They were fundamental to machine learning, and so on and so forth. "AI winter" means the money dried up, but the money dries up for capricious reasons.

Okay, so help us out here. In response to things like the Lighthill Report and perceptrons, there was, apparently, as we're saying here, [a] decrease in funding. When I started computing in my undergraduate days at the computer lab in Cambridge, there was work being done on AI there, but I'm pretty sure they didn't use the words "artificial intelligence" around it.

Well, it was invented by an American in 1955. So, you would have been around, but they probably could have called it something else.

Right. Right now we're in whatever the opposite of an AI winter is, some kind of heatwave. From your perspective of the hype cycles being burst, are we on the brink of another downturn like any of those? Just speaking from a historical analysis.

Yes, well, probably there will be some diminution. But you've got to remember, we've got another set of circumstances now that we never have had, and that is, we have a massive set of deep-pocketed funding in China. The Chinese adore AI. They are doing everything they can to develop it very wisely, as far as I can see. Anyway, they're not letting anything hold them back. This makes the notion of an AI winter or even a diminution in funding, a little harder to predict because we know that those wheels are grinding over there. And if you believe that keeping up, never mind being dominant in AI is important, you got to keep going.

So are you seeing an AI arms race in the offing?

Yes. Sorry to say.

Wow, what a great time as if 2020 couldn't be conflicting enough. You say that computer science on the whole is regrettably ahistorical. It kind of reminds me of one of my perceptions of being an immigrant to the United States that one of the things that drives American innovation is that they're not on some level aware of the existence of other countries. So, they

feel compelled to do whatever invention is required because it just couldn't happen anywhere else. Does computer science in this relentlessly looking forward and not backwards? What are the strengths and weaknesses of that worldview?

I see it as a weakness because, boy, you better pay attention to what all the smart people are doing, not just the smart people around you. I think the emergence of the British firm DeepMind, that woke people up. Yes, they got bought by Google and so they're "American", but this was a kind of wake-up call, and it will happen again and again. It will happen with Japan. It will certainly happen with China. I was just looking at the list of countries that have national AI strategies. I mean, there are places you'd think don't even have a computer, never mind AI. Anyway, yeah, you've simply got to pay attention to what's going on elsewhere.

You mentioned 1956 and the dawn of the term artificial intelligence. And that was from the Dartmouth conference?

That's right.

And when I look at the things that they said, the famous prediction there, they said, "We're going to do these things in this summer", and it was basically "Figure out how to make machines think."

That's right.

And we look back on that now and go "Haha", and it seems like overweening hubris, or ignorance, except these were the smartest people on the planet at the time. That cannot be that simplistic an analysis of that position. Are we either not interpreting their words correctly, did they have a different standard of what they said meant at the time, or what kind of blind spot was going on that these people would make that statement?

Well, I think the fairest thing to say is, we didn't have a clue how complicated intelligence was, is. We thought, "Oh, we can figure this out. Yeah, take a summer to do it." Now it's 60, 70 years later and we're still working.

Were they talking at any point to psychologists, cognitive scientists, if that field existed then, or anyone that looked at what we've got up on top of our necks?

Oh, well, Newell and Simon themselves were cognitive scientists, cognitive psychologists, I guess they would have called themselves. But the instrumentation didn't exist. The only way you could examine what was going on in a human mind was to interview them. Say, "Okay, let's have this little experiment where you predict what number is going to come up next in the sequence and think aloud about why you're reasoning this way." And so, you'd have some, surely, student getting paid $5 an hour or something. "Okay, I think the next number in the sequence will be three because I haven't seen a three for a while", and that kind of thing. There was no instrumentation. None. And it was only in the late '70s and the '80s that we began to be able to scan brains, have PET scans, that kind of thing. Oh, this made a big difference - MRIs.

Right. Earlier in this series, I talked to a neuroscientist who actually works with brains and does MRIs so we have some idea already of that. Listeners might want to look up the previous shows with Ryan D'Arcy. And that reminds me how there are estimates now of the cognitive computing power of the brain, trying to translate it into computational equivalents that differ by dozens of orders of magnitude. On one end of the scale, according to Ray Kurzweil, we are already there. On another end of the scale, we could be two decades away of Moore's law from getting that. What was your reaction to that level of uncertainty?

It tells you that the science is fundamentally undeveloped - got a long way to go. Here's slightly peripheral to this. I went to hear the great primatologist Frans de Waal give a talk in New York City. I just wanted to know what was new with the primates. He's one of the leading primatologists and he was talking about primate intelligence. And he said just offhandedly, "You know, we really didn't know how to measure this until AI came along. AI showed us what questions to ask." I thought that was so interesting. Gosh. And then you see things like the intelligence of slime molds. What? And yet, if you set up the questions correctly, you can see intelligence being exhibited in a slime mold. Now, this would never have happened in the '50s, the '60s, the '70s. Never. Intelligence was something that humans had, period, end of story.

But, well, intelligent slime mold can always go on to have a rich career in politics… Not taking any sides there.

Well, I understand.

And it makes me think that every time we ask the question, "What is AI?", we have to ask, "Who are we?"

Yeah.

Was that level of introspection on display at the beginning of this field?

Oh, yes because that was the whole point of that part of AI was "What is thinking? Who are we? How do we do this?" And so, these pioneers conceived of AI as a way of modeling cognitive behavior. It wasn't to build a killer machine that would go out and do killer things. I mean, they admitted, yeah, that would probably happen many years from now, after I'm dead kind of thing. But the real scientific interest was, "Can you model human intelligence in a way that is better than the ways that they had been doing it before?"

The title of your book, though, *Machines Who Think* always provoked something in me, and I feel like I'm walking into an obvious trap here, but I haven't found it answered anywhere in the other book, so I'll bite. Why is it *Machines Who Think* and not *Machines That Think*?

That was kind of a subtle reference to Marvin Minsky's quote of the 1970s, "You know, humans are just meat machines." And it was a moment where we were beginning to see that if humans were different from other species on the planet, they had a lot of things in common with all the other species. And so, Marvin could say very casually, "Yeah, we're meat machines. Machines can think. We're the obvious example." So, to call the book *Machines Who Think* was to

reference both that and possibly the emergence of personality from what we still considered machines.

So, it wasn't a double entendre about any of the people that you were profiling?

Oh, no, no. Look, anybody who reads that book could tell that I adored them all. I really did. I just had such admiration for them.

Even the ones that there were conflicts with? I mean, some of the most amusing lines are around Joseph Weizenbaum when you describe "professional detumescence" - which has got to one of my favorite words, I may have it engraved on something - and that he had lost confidence in his scientific manhood. I mean, this is page-turning stuff.

Yeah, well, I probably wouldn't write like that now. I'm kinder than I was then.

But that actually reminds me of the first person I interviewed on this show Audrey Tang is the Information Minister for Taiwan.

Oh, yeah.

And she said that their strategy in dealing with disinformation, particularly around the coronavirus, was to fight rumor with humor. And it occurred to me that you may have snuck in a great deal of opinion and information that might otherwise have been viewed as more contentious by surrounding it with such entertaining prose.

I hope so. I hope so. You should have seen what my editor made me take out.

Well, what else are editors for? What do you think of AI when you look at now GPT-3 which has come out since your last book, [and] Google Duplex and its conversationality. And she's so perky with the California upbeat accent and "huh" in response to the conversations that Duplex - I think she may have another name now - has with people on the phone. It seems like we're a stone's throw away from at least cracking the original Turing's description of the Turing test criteria. What do you think the pioneers in this field would say if they saw that?

I think they'd be thrilled, frankly because it is the result of enormous amounts of human brainpower figuring out how to do this, how to do this well. It's incredible. But we all know the downsides of this. And those downsides are one of the things that worries me about AI now, that the possibility of that kind of faking is-- Oh, it leads to all kinds of things. And we are not catching up with detecting those fakes as fast as we can make those fakes. That could do a lot of damage to the human society.

From your perspective of integrating the views and experience and expertise of so many people who have been in this field, what advice do you have on how we should govern the development of those tools?

If I had good advice, I'd bottle it and sell it. I don't. I notice that the number of references to ethics in AI has skyrocketed. Well, yay, that's a very good thing. I mean, this wasn't even in

anybody's mind back in the old days. So I think those people who are studying the ethics and the implications are in a much better position to say anything about this than I can. But I congratulate them for taking on the task.

Your writing is so - I keep returning to this thing, but it's so entertaining, so gripping that I wonder whether you found yourself in describing AI, anthropomorphizing it, because that's what writers do, is anthropomorphize a lot of things for the purpose of getting attention. But did you ever find yourself anthropomorphizing it to the extent that you felt, "Oh, I need to walk this back or people will start thinking that this is more capable than it really is"?

No, I really didn't. And the reason I think I didn't is because I could see how flawed these programs were. I could see the big gaps in their performance. I was a real cheerleader for what was accomplished but I didn't kid myself that it was anything near human performance. Now that changed. In some aspects of AI, it is as good as if not better than human performance, but that wasn't the case when I was writing my histories. And when we see a program performing better than a human, we also have to think, "Yeah, but you know, it couldn't get itself to the airport and get on a plane. It can only be a fabulous player of Go or a fabulous player of chess."

Although it is interesting how many things we thought were the province of humans turn out to be remarkably imitable by a sufficiently large network, like composing music, and at least to some extent, creating art, and now, conversing - holding conversation. There are some conversations that I've seen with GPT-3 that were this close to the Turing test five minute barrier. No one would possibly think that they weren't talking with a person. It's easy to get it to give itself away if you know what to do but--

No, for the ordinary, naive user, you could go on for a long time.

Right. And so now, well, let's talk about the left brain, right brain because what we now call good old-fashioned AI, which obviously wasn't called that, to begin with, was it symbolic logic then? Or was it--

Well, it was symbolic—Yeah, it used symbolic logic in a real sense but it was—These researchers were more interested in symbolic thinking, as you and I think about it. That is, we have a word called honor. What does honor mean? So, the meaning of these very abstract concepts was fascinating to them. They wanted to get to the essence of that, and we haven't yet.

And with referencing Daniel Kahneman's System 1 and System 2 thinking, they were focused on System 2, this higher-order reasoning and symbol manipulation at that time. Now, it seemed as though that was the thing where AI would make it its mark first and that the hardest thing to do would be the System 1 thinking, which is this intuition, pattern processing. And yet, things have flipped. Now that's the easier thing for AI to do. We don't know how it does it, but if we train it on enough pictures of hippopotami, then it will know how to recognize one. It can't tell us how it does it, but it can do it now better than us if we have enough hippopotami pictures. Are we perhaps in our - again, focused on the future mindset, are we not sufficiently aware of just how seismic a change that is to shift between those two thinking patterns?

We're not aware, no. And it would be very, very salutory to be more sensitive to those things. But again, it also has to do with the very fuzzy ideas of what thinking was considered to be for years and years. And we're slowly getting to the point where we think we know what thinking is, but we associate it with intelligence. And we can talk about intelligence-- biologists can talk about intelligence from the cell level to the human cranium and for all I know beyond that, but we are really in a pioneering age when it comes to making science out of intelligence, science of thinking. Each of these things you've mentioned is a great step forward but man, we got a long way to go.

Intelligence… reminds me I heard on a previous show about eliminativism, the philosophy that—I think of it as *The Princess Bride* philosophy: "You keep using that word, I do not think you know what it means." It was applied to creativity and they said, "Whatever this means, it may not have any meaning, so we should stop talking about that." And I started to think that intelligence is one of those things that we have one word for it and so we think we can have one number for it. And that seems so dangerously simplistic that we ought to stop thinking of it in those terms and maybe just put the word aside.

Yes. Minsky used the word *suitcase*. "This is a suitcase word. You have to unpack it and find out that it means 10 or 12, or 105 different things." And we don't do that; we say, "Oh, well, that was intelligent."

Yes, we're going to have to come back again next week to finish the interview because we got through just so much juicy conversation that we couldn't fit it all in one episode. I know we don't often think about history in computer science because we're too busy doing things that have never been done before – after all, just about every program ever written had never been written before – or if it had, somebody made a mistake. But the history gives us a foundation for asking important questions about the directions we're taking in the future.

In the latest AI headlines, the hybrid swarm intelligence UNU correctly predicted the winner of the World Series before the season started. "Hybrid Swarm Intelligence" sounds extremely cool and slightly threatening, but it is referring to what amounts to an AI-assisted real-time voting system. It was observed some time ago that in contests at fairgrounds like the ones to guess the weight of the cow or how many jellybeans in the jar, that no matter how hopeless some of the answers might be, the average answer was always quite accurate. So if you could figure out a way to get the average answer from a lot of humans, you'd get a good answer to the right kind of question. Louis Rosenberg devised UNU, which is a system for finding the answer to a question as a group of people who are logged in influence each other's decision by pulling a little puck around the screen between the possible answers. The idea is that as you see how other people are voting, you might modify your input in real time, and collectively the decision is an optimal one.

It's called a hybrid *swarm* intelligence because of Louis Rosenberg's observations about the behavior of honey bees, which collectively – and here comes the *swarm* part – make optimal decisions about where to build their nests or hives based upon aggregating the input of lots of scout bees. And UNU has made some extraordinarily good picks, like the Kentucky Derby superfecta. This time, back in July, UNU predicted that the Dodgers would win the World Series (for our listeners outside the US, the "World

Series" is what Americans call a nationwide baseball tournament), even though the Dodgers hadn't won it in decades. And it was right.

In next week's episode we'll conclude the interview with Pamela, and you'll hear more that shows what I was saying about history informing our perspective and judgement of where we're going. We'll talk about the changes in attitudes towards women in computing. We'll also talk about C.P. Snow's "Two Cultures" a landmark essay about the divergence between the worlds of science and the humanities. And I'll ask Pamela outright what we should learn from the history of AI. 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