Beyond and Within: AI Talks

Episode 2 | Decoding the Code

As an essential part of *Ether's Bloom: A Programme on Artificial Intelligence,* the new podcast Beyond and Within: AI Talks invites experts, artists and researchers to dive deeper into the world of Artificial Intelligence (AI) and some of the vital questions it raises. It takes the listeners on a journey of questions, analyses and knowledge, while exploring the many sides of the fast-growing technologies around AI.

In the second episode, *"Decoding the Code"* podcast host Eliane Eid talks to historian and professor Orit Halpern about the power and presence of AI in our daily lives through technological and historical changes. How is AI affecting the fundamentals of what it is like to be human? And how can intersectionality and Artificial Intelligence exist in the presence of categorised data?

(As this transcript is based on a conversation, it contains colloquial phrases. It has been edited for better readability.)

Eliane Eid: Hello everyone, and welcome to our podcast Beyond and Within: AI Talks. I am Eliane Eid and I will be moderating the episodes of this podcast that was developed by the team here at the Gropius Bau. What is Beyond and Within? This limited series of five episodes will take you to a journey of reflection, curiosity and conversations regarding AI with a focus on a more artistic perspective. Within this format, we are discussing points of accessibility, fears and possibilities that come within this technology. We wanted to create this podcast as we had a lot of questions, concerns and ideas about how AI is changing our daily and artistic life. So, we decided to jump on this ride and open up the conversation with experts, artists and researchers who have been studying and working with AI for quite some time and asked their opinion and knowledge about all those questions that are emerging. This podcast is part of *Ether's Bloom: A Programme on Artificial Intelligence*, where we are hosting different artistic presentations and diving into the topic of AI on different levels, through a Writer in Residence, workshops, and discussions. The thoughts of this programme will inform the development of an app which we will launch in 2024.

I would like to welcome Orit Halpern, who is a professor, scientist and historian. Her work bridges the histories of science, computing and cybernetics with a focus on the relationship between reason, race, bias and democracy. Hello Orit and thank you for joining us.

Orit Halpern: Hello Eliane, thank you for inviting me.

Eid: I'm going to jump right in and ask how we can define in a simple way 'cybernetics' or what it means, and how it's also related to intelligence because I also think it's a heavy term for the public.

Halpern: Yeah, thanks for that question. I think it really helps to just historically situate this term. Cybernetics emerged out of Second World War research on anti-aircraft defense in the United States, where people started realising how human beings act under stress. For example, an aeroplane pilot would act very repetitively that certain aeroplane pilots, when shot at, would veer left and certain which veer right, and that there was a statistical regularity to that behaviour. So people started thinking: Oh well, human beings under stress seem to be acting mechanically or predictably, so maybe we can predict the actions of human beings using statistical methods. This opened up a way of thinking that started to say that maybe there is a way to predict and act on the world and maybe control it. If we start treating humans not different than machines, but the same, instead of making separations between what is a machine, and

what is human, instead, we should just look at how the particular object, agent, or person behaves. This might allow us new methods or new ways of thinking about both what is the human but also what is a machine. This is why this term got very popular because people were very, very concerned in wartime, in predicting the future behaviour of armies, of groups of people and trying to use computation or statistics to do that. That necessitated the production of new types of machines. But as I've mentioned, also ideas about people. This grew into a broader idea that began to think about the world, not in terms of all of these varied things that live in it, but rather in terms of behaviours and numbers. And so the idea started to change;: maybe we need to rethink everything we know, not in terms of static objects or by looking at whether that's a flower or a tree or a person, but instead, thinking of it in terms of communication and information. Now, this may all sound really strange, but of course, today we almost assume it is a given, right? It's a natural assumption that we live in something called an 'information economy', that information and communication are things that we can build, that we can model, that we can sell, right? We sell people's data and interactions between people, and between people and things, for example, social networks are something that we can actually engineer, build, or model and sell. So cybernetics in many ways is at the heart of reimagining our world as one comprised of information exchanges, but also a world where the difference between how we understand what the human is and what per se a machine is, is increasingly degraded. That isn't really the question, is it human or is it a machine? But instead, what does that agent do? Is it a consumer? Is it a creator? And we see those debates unfolding today right around ChatGPT 4 and questions about whether machines can be conscious if they can emulate or do what humans can do. In order to even ask or think that, we had to have gone through an amazing moment where we actually rethought what humans are and what machines are. And of course, the term that interests us in this podcast: what intelligence is, what behaviour is. All those things had to be reimagined in order to make machines and human beings compatible to work together. Cybernetics in many ways is exactly a term that encapsulates that form of thinking. The actual term was coined by an MIT mathematician, Norbert Wiener, who was doing this work on anti-aircraft defense, and the original definition was the sciences of communication and control in the animal and the machine. It's not even a separation between humans and animals. These are very seemingly complex ideas, but they're actually ingrained in our everyday assumptions about how machine intelligence, machine learning, and Artificial Intelligence work, but also a kind of assumption that we get up every day in the morning, probably use our cell phones or get on a network. You don't think twice about the fact that you live in a world that's computer-mediated or digitally mediated and that you're sharing information. And you consider writing all those memos and emails, blogs and "X's", I guess they're no longer "Twitter", but you probably don't even think that's a productive and good thing to do with your day. We assume that information is vital to our economy, to our lives, and to who we are, so in many ways, cybernetics is a place where people first began thinking about that.

Eid: You mentioned history, which I think is a huge part of how we deal with technology right now, but how can history help us understand how technology changes over time?

Halpern: One of the questions is always: What stays the same and what changes in the field of history? We asked that question for a number of reasons. The first is: if nothing ever changes, then it's almost like the future's already arrived, right? Why bother trying to change anything? Why bother trying to build more ethical or equitable technologies if, after all, everything's always the same? So the first thing history does is teach us that things change. And not only that, it helps us rethink another really important term I'd like to discuss, which is the idea of 'determinism'. Very often, we have ideas of technological determinism. For example, AI just has to increase, these corporations will always get bigger and greater, we have to give our data over, we have to build systems in a particular direction because economies have to grow and they have to grow in one particular way, and so on. These kinds of beliefs are so heavily ingrained for example, that AI in its current incarnation is unstoppable. We believe that we can't not have ChatGPT 4, it has to be more and more ChatGPT 4 or some other version. These kinds of ideas often create many problems in

terms of homogenising the type of technology we have, and of course, failing to deal with the injustices and racism or sexism that's in the technology. The first thing history does is that it shows us that no technology is inevitable. It is the result of particular social historical conditions amounting to each other. Secondly, history lets us see things like structural injustice and the way they're ingrained into particular technologies. Now that's very abstract, so let me just give you a particular example: A lot of people are concerned in contemporary AI and big data systems that they pull data, but they make certain assumptions about their users, and those assumptions come from longer histories. For example, throughout the 19th century, there was an effort to quantify intelligence. Maybe you've heard of an IQ test, the IQ test was originally introduced in the First World War to actually discriminate against Black soldiers, and to demonstrate that certain groups were less intelligent than other groups. So that this test would supposedly show and give us a quantifiable, verifiable, and objective measurement of human intelligence, that could show that groups like Blacks or women were less intelligent than groups like men. Of course, class also played an important role in the development of these psychological tests. That history of attempting to quantify intelligence has great resonance with contemporary statistical methods that are always trying to quantify and create scales that separate and figure out what type of user you are. As that becomes the sort of measure of intelligence, these concerns that older ideas about who's smarter, who's dumber, who's better, might actually be encoded into contemporary scoring systems used to train Artificial Intelligence agents. History lets us show that there's no such thing as pure data because data doesn't just lie out there. It actually has to be made and we have to collect it, we only collect certain types of data according to what we need and what we want to do. And the question is of course: Who's "we"? For example, my credit card company is not interested in what I'm thinking or feeling or my dreams at night unless they can directly turn them into my credit card and introduce some shopping experience. They have very targeted interests, they want to get me to click on something and buy it. The data they're going to collect is going to be targeted towards that activity in the same way that people who are interested. Whether it's real estate or whether it's targeting political interests, they are going to be collecting data in particular ways that shape the way we understand the world, that fits the agenda that they would like. History helps us understand the history of data collection practices and also the history of how we've understood who's human and what's human and who gets to be treated as a citizen and who doesn't, leading to how those things shape our contemporary statistical and data collection systems.

Eid: It's almost like society became this huge data factory, because every day, everything we're using, whether it's social media or any other format we are taking part of, our data is being kind of predicted and taken while being used in a way or another. But my question would then be: How is data being categorised and is it also essential to be categorised in certain ways and in certain formats?

Halpern: That's a big question! People like to create meaning, produce narratives, and organise the world. What we do now is question how we do those changes over time. And of course, I'm a historian researching mostly from within the United States, so I'm going to be speaking from a particular Western history of science perspective. I'm sure we could find many other examples from other places. For example, in mediaeval Europe, if you look at beast theories, people would classify certain types of rocks as "living". People would have a totally different organisation of nature that we would never imagine today. We understand rocks as being inorganic and bacteria as being organic. We have clear separations and understandings between humans and animals. We don't have a world full of unicorns and other strange beasts. When you look at these beast theories, barnacle geese were imagined to grow as barnacles and then change as geese. Now to our minds, that seems insane. These people must just not have had science, they must have been crazy, but actually, they weren't. Because in their worldview that made perfectly good sense, they didn't have the same ideas about life, death, biology, or that the body perhaps is a mechanism, in the same way that we do. They occupied a kind of theological world where these transitional states between organic, inorganic, and different types of life forms were possible and even seemed probable. And the same thing is true for us today. We're constantly changing our categories.

For example, since Linnaeus (Carl von Linné), most people knew that we're Homo sapiens. There are species, there are genesis and we've created a very elaborate system of separating different living things and putting them in boxes. And of course, that's constantly shifting. As evolution became an idea in the nineteenth century, people began to look for ways to figure out whether race was a real entity, so racists and nineteenth century eugenicists, often tried to make kind of explanations about why, for example, Black people were racially different, that race was actually an existing category biologically. But increasingly in the contemporary twentieth century, for example, the fact that genetics are pretty similar across the species puts into question some of our centralising ideas about race. Suddenly those categories that made things very clear, change, as do changes in technology. Gender and sex for example have radically changed in the course of the last few decades, as technology makes it possible for people to embrace a whole range of sexualities and genders so that I can not only be binary, male or female. Those are categories that are changing. They're changing because of technology, and they're changing also because we're understanding the world differently. In fact, even some of these older categories, such as species, are changing because there's an entire new branch of evolutionary biology that is showing that there are tons of forms of reproduction that aren't even through sexual reproduction, that viruses have inserted themselves into our genomes. There are a lot of new forms and ideas of carrying interspecies or post-species , if you will, evolutionary biology. So those are very specific examples, but they demonstrate that what we categorise, and how we categorise, is constantly changing. That's something we thought was a really solid category, male or female, or Black and White, which over time has been demonstrated to be unstable and has had to be re-categorised and changed. Whether it's a necessary thing to put things in boxes, I can't say. I do know that obviously we often do it as a matter of heuristics and bureaucratic necessity maybe, or desire. But I would also say that it's important when considering classification to realise that there's always variation and that classificatory categories change, and that one of the challenges with Artificial Intelligence and all big data infrastructures, of course, is trying to understand how we keep changing the categories and how we maintain diversity in our classification orders. We have more than one system in place and this is where history again matters, which classification system we're using, how we're gathering data, who's doing it, and what the assumptions are. These concepts are about the world that is organising those data sets. So am I working from an assumption that male and female are a binary, or do I have another set of assumptions about how sexuality and gender work? Those things are going to fundamentally shape whatever database I'm going to create.

Eid: Exactly. Because at the end of the day we live in a diverse society and we're all complex human beings. How does intersectionality play a role in Artificial Intelligence and what does it mean in a format like AI?

Halpern: One thing I'd like to do is to maybe situate the term 'intersectionality', and then I think we should have a conversation about whether we want Artificial Intelligence to be intersectional or not and what that would mean. At least within an American context, intersectionality emerged out of the field of legal studies and Kimberlé Crenshaw kind of coined it. I don't want to create origins and I'm sure there are other intersectionalities, but for purposes of my discussion, intersectionality was a particular term that was linked to 'civil rights'. And 'civil rights' I really mark as an important term because it's not just human rights, it's not just recognising that people are human, it's actually about political enfranchisement, it's the question of what laws are necessary to give people power, actual political power to exercise their power as citizens to vote, to be able to participate in society, to be educated, and so forth. It's fundamentally a political idea and it's fundamentally the idea of intersectionality that the law deliberately ignores intersections between race, class, and gender in order to essentially create weak spots for places where we actually need to take action. Which will create the capacity to say that we're giving everybody rights, but essentially to exclude large numbers of people from accessing power by essentially deliberately obscuring the fact that they're not impacted just by one category, but by multiple categories. For example, as we have current debates about affirmative action, for example, in the American universities, there's a

question of whether we should take more people from Black or Hispanic/Latinx communities or other groups because they've been historically and structurally discriminated against. But these laws have been under attack, and so people are saying instead of using that, we'll use class, but people in return are saying: that if you just separate one of these categories from the other, you're essentially missing the big picture. You're essentially missing the facts, so you're allowing a lot of people to be excluded, because in theory, you're only looking into making it fair, and it looks fair for one category. But essentially, people who are under multiple categories of administration or multiple categories in their lives can't and won't be assisted by this particular law. And for her (Crenshaw), it was particularly true around sex and gender, where essentially she suggests that even if I have a law that says: I want to allow poor students into particular universities, for example, if you don't look at the fact that being female and being Black and being poor constitute a particular structural inequity, that makes accessing that university education particularly difficult. You're going to essentially exclude these people and you're also going to exclude the ability to understand why certain groups are being held behind or being excluded from social institutions, because you're fundamentally not understanding that, for example, race can never be thought outside of sex, because after all, it's a category that's fundamentally linked to things like reproduction. That's why there were so many laws in the United States against miscegenation until the seventies. Like, different groups, for example Black and White people, could not marry, because fundamentally, racism is deeply tied to ideas of sexuality and reproduction that are contingent on certain normative concepts of sex. For Crenshaw, this is a really critical politically active question. So the question whether AI is intersectional is really interesting, because on one hand, one could say these new big data sets and the kind of fantasy, at least of AI, is that they're hyper-intersectional. I can combine so much data, I can find new correlations between my shopping, my age, my sex, many things I label myself on the Internet, etc. relatively rapidly, and I can create all sorts of new types of correlations and new types of people and new types of subjects. We have a million generations and subsets of categories and things like that. On the other hand, these forms of categorisation are fundamentally instrumental and at least as contemporary systems are mostly organised, are not really in the interest of empowering me in any kind of genuine, legal, or institutional manner. Therefore, the fact that I'm correlating data sets doesn't necessarily mean I'm correlating them with any kind of actual mode of civil power. Then the issue is, is it intersectional in any sense that actually allows us to understand structural violence, or does it merely recombine data and complex data analytics that have nothing to do with the fundamental question of attempting to understand structural violence and enfranchising groups that have been marginalised legally?

Eid: How can we look at intersectionality within those technologies while hoping that they can affect us in a positive way or in a more fundamental way?

Halpern: I'm basically a technophile, I know that doesn't sound like this from this conversation so far, as we are kind of focussed on what we already know around it. Especially when everybody mentions norms, race, sex and discrimination when it comes to Artificial Intelligence systems. But of course, they can do a lot of good, whether it's enhancing e public health or just helping us in our daily lives, accelerating tasks we dislike or in general augmenting our ability to communicate with others, it certainly augments our senses and our analytic abilities. There are lots of things that AI does that are great or could be great. I think that part of that is about how we treat technology. And I'm going to make a poll here that we need to think about the technology not in isolation. So we can't really talk only about AI, we have to talk about ecology, but also terms like 'epistemology'. That sounds a little weird, but epistemology is how we understand the world, how we represent it, and how we approach it. It's really important to understand that Artificial Intelligence is not just a single technology, it's not just Large Language Models, it's a way of approaching the world where we try to turn things into data, including things that weren't previously datafied. We have a very comprehensive sensor infrastructure and we want to approach the world through statistics and pattern-seeking, and that's a particular approach and it shapes how we build organisations, it shapes how we build technologies and the technologies are constantly going to change within that. And it's also the

world, as I mentioned, that we fundamentally understand, is comprised of communication, information, and data. Once you see AI as that, you start to say: Is it just about focusing on making the parameters or the data sets better? – which of course we should. A lot of people are working on doing that. But is it also about the broader social, institutional, and regulatory infrastructures within which Artificial Intelligence is located? For example, we can worry about whether the particular set of models is discriminatory and what parameters we need to change. For example, in the current issues involving social networks and democracy, and particularly with systems like X, there's a clear predilection towards the alt-right. Not only is it because pattern-seeking algorithms of course seize upon simple repetitions of slogans. Obviously certain things are easier to say than others and easier to be picked up by the AI, because fundamentally, these are owned by people who are pretty supportive of these groups and are intentionally setting the parameters to assist them. In those contexts, the question then is: is it all social networks or is it also broader regulatory and infrastructures that guarantee free speech for everyone, not just the alt-right? People are very concerned about discrimination against people with different capacities as well as different groups due to the capacity of these systems to generate particular ideas/stereotypes or to discriminate in things like job testing. There are all sorts of places, for example China, where they're introducing social scoring, and things like this really impact access to jobs, access to services, and social services in the United States. There's a lot of concern about automation, for example, systems for the poor, because not only is it already a group that's vulnerable and uninformed and relatively undereducated in terms of the use of these technologies, but of course, it's increasing surveillance on particular target groups, which means that obviously, people are going to show up as committing more crimes if they're being watched more often. There's a kind of concern with predictive policing and other systems that the AI will end up working that way. But the question is: Are we going to stop using AI or are we going to create a more robust civil rights protection regime that would allow you to prosecute, not technology by technology, but actually create frameworks where you have a ground for saying this is discriminatory. For example, very strong affirmative action or equal opportunity laws across the board, not just for Artificial Intelligence, for employment, and other things may very well work better than trying to regulate AI independently. And a great example of that is in the European Union, where the recent laws involving AI have basically scaled AI from most dangerous, never used, to, you know, you can go ahead and develop AI. Of course systems like migration have been left totally off, so you can do anything you want at your border with any artificial intelligence you want. This would directly affect refugees, immigrants and people attempting to enter Europe as their rights are limited, especially if governments start using Artificial Intelligence. So the question is: Do we need a more broad, robust civil rights regime for more types of people, or do we need to regulate each technology in its little cubicle? My argument is that we're probably going to need a combination of both, but we need to be thinking ecologically and sometimes that means having much more robust structural systems in place that can take account of discrimination not only through AI but through any mechanism, but they may allow people to have the institutional, the legal and regulatory tools to be able to even pursue discrimination when it does happen in Artificial Intelligence systems and also to make possible things like free speech. But again, free speech issues and democracy, part of that isn't just about one system, it's about whether we're investing in creating an ecology that's diverse and where people are getting their information from more than one channel.

Eid: Looking at the time now and talking about learning, I have one final question. If we go back to learning and a bit to close it with history, there is a saying that history always repeats itself. But I had a history teacher who told me that history never repeats itself because we're never located in the same exact situation twice. This stuck with me ever since I was 14 years old. I think a lot about it now with AI because AI is learning from our histories. And it made me realise: Is AI also learning from our mistakes? Because I still see those mistakes, in biases and non-diverse languages. So I just want to have this as an open question and hear your thoughts about this.

Halpern: I have to agree with your history teacher. History doesn't repeat, but that doesn't mean you can't learn from it. The problem I think that we're concerned with is that we're in a stage where for the first time we have machines that can both create narratives that we find seductive and interesting and that we buy into. I can kind of tell when they give me a ChatGPT essay, but it's compelling. At the same time, it's the first moment where maybe we could repeat history because our machines are only recording and regurgitating the past. So you're bringing in a really important question, which is, if machines are writing - and writing stories we believe - and stories that inform our imagination and how we think about the future, but they're only using the past to write those stories. This is a fundamental question that hits on ideas of innovation, creativity, what makes us human and not mechanical. We're really under pressure right now to find ways to think with our machines, not against them, but differently from them. And that's really the challenge. I guess that's where institutions like your own, and arts and culture, and the humanities really have a place to augment our imaginaries by trying to use histories and use our life experiences and the richness that encompasses the multitude of human cultures to really try to imagine things that haven't been recorded and haven't been done. To tell stories and make worlds that we haven't seen before. I think that takes imagination and cultivation, like I said, always with and through and around our machines, but also in supplemental and ways that machines can't do, and to produce new types of data and new types of stories that simply can't be automated.

Eid: Thank you Orit, it was very nice to dive a bit into the different topics and it was lovely to have you here today.

Halpern: Thanks.

Eid: I have to thank you all for joining us today on this episode and for being part of this journey of discovery and conversation. I also need to thank everyone who helped develop and produce this podcast. Madeleine Köberlein, our co-producer. Luis Kürschner, our sound designer and editor, Çağla Erdemir for all the assistance, and of course Clara Meister, our supervisor and program lead of the AI project here at the Gropius Bau, along with the institution for providing us with this space to take the audience into this journey.

The Gropius Bau and the programme are funded by the Ministry for Culture and Media. As for our next episodes, we will be diving deeper into topics related to hallucination and different artistic practices. So stay tuned.