

Crossing Channels

Interdisciplinary answers to today's challenging questions

A Podcast series hosted by **Richard Westcott**

Presented by

Bennett Institute for Public Policy Cambridge

Institute for Advanced Study in Toulouse

With support from

INVESTIR EN FAVENIR

anr

UNIVERSITÉ TOULOUSE CAPITOLE

Season 4 - Episode 4

HOW ARE DATA AND ALGORITHMS IMPACTING OUR LIVES?

With

Jean-François Bonnefon, IAST

Gina Neff, University of Cambridge

Jeni Tennison, Bennett Institute

S04E04: How are data and algorithms impacting our lives?

HOST

Richard Westcott (University of Cambridge)

GUEST SPEAKERS

Gina Neff (University of Cambridge), Jeni Tennison (Bennett Institute for Public Policy and Connected by Data), Jean-François Bonnefon (IAST)

Richard Westcott 00:02

Hello and welcome to Crossing Channels. I'm Richard Westcott. How are algorithms and data impacting our lives? That's the subject of the latest in our podcast collaboration between Cambridge University's Bennett Institute for Public Policy and the Institute for Advanced Study in Toulouse, as ever, we're going to use the interdisciplinary strengths of both institutions to explore a complex issue, how do these algorithms and data shape the way we work, the way we live and the way we interact? What ethical challenges do they raise, and how can governance frameworks address questions of power and accountability? Finally, we'll look at what role they play in shaping the future of our society, and how can we ensure it works for the public good? And I'm sure we'll be talking about AI a bit today as well.

Richard Westcott 00:55

To explore these issues. Today, we have Gina Neff from Cambridge University and Jenni Tennison from the Bennett Institute. Gina, start us off. What does your research focus on?

Gina Neff 01:04

I run the Minderoo Centre for Technology and Democracy, and so it does what it says on the tin. We think about how digital technologies change everyday lives, and how we can make sure that those technologies work for people, communities and our planet.

Richard Westcott 01:21

And Jenni, one sentence introduction of your main research, please.

Jeni Tennison 01:25

So I run a campaign called Connected by Data, which campaigns for communities to have a powerful say in data and AI. So I work with affected communities and how they can participate in decisions about data and AI.

Richard Westcott 01:40

And joining us from the last, we have Jean-François Bonnefon. Jean-François, remind us of your main research interests.

Jean-François Bonnefon 01:47

I'm psychologist, so I run experiments with people to see how they want algorithms to behave and the kind of thing they do with them.

Richard Westcott 01:55

Okay, so this is a huge question. Algorithms and data have quietly become the backbone of modern society, influencing how we work, how we communicate, how we access services and how we make decisions. The UK's new Industrial Strategy was optimistic about the transformative role of data in driving economic growth and boosting innovation. Recently, the Data Bill was also introduced to enable the secure and effective use of data for the public good. So, let's explore how data and algorithms are shaping our lives, their governance, their ethical implications, and what the future holds for society. First of all, though, Gina, I want you to just very quickly define what we mean by an algorithm and what they do.

Gina Neff 02:37

Well, everyone listening to this podcast probably has many algorithms in their kitchen. So an algorithm is simply a set of instructions like a recipe. Recipes are great examples of algorithms. If you have the ingredients and you follow the recipe, then you might come up with an output at the end, a tasty dish or a complete flop or failure. That's what an algorithm does. So simple as that, it's a recipe for getting something done.

Richard Westcott 03:12

First of all, Gina, let's go with you. Your research explores how the algorithms and digital tools are reshaping professional environments. So what do you see as the most significant changes in how algorithms affect the way we live and work.

Gina Neff 03:24

Well, just like those recipes in the kitchen, we are surrounded now by rules and processes that are managing the data about us and data that we interact with in many situations. So recommendation algorithms, for example, one of the common things that people interact with, right? How are our music apps and our video apps giving us things that delight us? Are they recommending things that I want to see, or they just cycling one bluey episode after another for me? The challenge is when those choices around what steps should be followed, around the decisions made in making those algorithms don't fit with our notions of fairness. So for example, one of the ones that I think is kind of a terrible example all the way around, which is the NHS liver donation matching algorithm. Now, this algorithm was made with the best of intentions. How do we get the most deserving people, the neediest people, who are the most likely to benefit? How do we get the most good out of liver donations. And yet, a simple choice that was made in the designing of that algorithm meant that people under the under 40 were hugely discriminated in this algorithm. They were aware of it. Their doctors were aware of it. They were aware that younger people were not able to rise up to the levels of getting on the list for getting those liver donations. And those choices were simply a choice of which measure of success were they going to use to build that program. We see those choices in so many cases where algorithms go wrong. For example, the city of Boston turned over their school bussing program to a bunch of, to two really bright, MIT PhD students, really great mathematicians, who made a set of choices that infuriated parents across the entire city, because parents have made decisions around how they organise their busy mornings, and these two said, Oh, well, you know, it won't matter much if the pickup time for children from their houses change by an hour and a half with less than two weeks notice. So people don't understand what choices were made in these algorithmic systems. They don't understand what they can do about it, and then they're told simply, well, computer says no, they have very little recourse in that.

Richard Westcott 06:15

Jean-François, let's bring you in here. So your work examines the intersection of algorithms and human decision making exactly what we've just been talking about with algorithms influencing everyday choices like we've just heard. How do you think it's reshaping our own autonomy and the decisions we make ourselves?

Jean-François Bonnefon 06:34

Well, what is clear is that the machines, the algorithms, are changing our culture in many, many ways. Like they create content that competes with the content created by humans. They also steer our attention. For example, recommendation algorithms steer our attention to the content created by some humans or some machines. And also the algorithms are kind of training us by rewarding some behavior, if you think of social media, for example, and the algorithm that boosts some of your posts or mute others of your posts. So gradually, the machines are training you to produce the kind of content they like or the kind of content that the people will program the machines like. So all these impacts can be seen as decreasing our autonomy. But at the

same time, generative AI has given people autonomy to explore, express their creativity in a way that they could not before, and the recommendation algorithms, they can also help you explore a niche, very personal interest that you would not have discovered without machine intervention. If I can actually tell you a personal like those, I realized that I was kind of into romance novels, which I would never have discovered before. The machine told me, hey, you know, I think you should try this stuff. You might not think it's for you, but you should try it. And I've discovered a personal interest into some new genres, like like this. So the net effect on autonomy? That's difficult to say, because autonomy is rather qualitative concepts. Machines are doing things to our autonomy in both direction. But as per the net effect, I'm not sure how we decide that.

Richard Westcott 08:30

Yeah, it's really interesting, isn't it? Are we teaching them, or are they teaching us, or a little bit of both? Where is that line? Jenni, so let's, let's look at how they affect public services, things like health care and welfare and so on. So, how do you think these systems are changing people's day to day experiences with the essential services?

Jeni Tennison 08:48

So yeah, we're seeing a lot of algorithms and AI being rolled out across the public sector, along with digital transformation more generally. And you can really understand why that's happening. We've got public services that are under immense strain, that need to find ways of doing things more effectively and efficiently, and that means that behind the scenes, there are a lot of algorithms that are used to prioritise, to triage requests or complaints that are coming in that are used for predictive purposes. For example, you know, used in order to determine where to send police patrols so that they're going to be most effective. What we're also seeing more of is take up of generative AI within the public sector. So, my personal anecdote, I had a call very recently with a GP, and as well as the normal, 'this call will be recorded for training purposes' message that you hear all the time, they also said, and 'we'll be feeding it into an AI in order to create notes' and transcribing text using AI is usually very, very accurate, but I also read studies that show that where that transcript is used and then AI is used to create recommendations they can be inaccurate, and it's very difficult for doctors to detect when those inaccuracies are happening. There are issues to do with bias, to do with the kinds of things that Gina was talking about, where historic data is used in order to create an algorithm, but that historic data is biased, and so the the algorithm that comes out is biased, but also I'm interested in how it affects kind of the more human aspects of it. What does it mean for me as a patient to know that my doctor is using AI in order to record things, maybe isn't paying attention as much as they might do if they had to make the notes themselves.

Richard Westcott 10:50

Do you think there is an issue with public perception of this technology?

Jeni Tennison 10:54

I think that there's a real issue with trust in technology. We know that because of newspaper stories, and what we see of the way in which algorithms are used, because of the way in which we see data breaches happening in public, and those are the kinds of stories that rise to the surface. It means that people are at a baseline distrustful of how algorithms and data might get used, and that leads to, I think, some mistrust about where they see public sector services using those algorithms, that means that it can lead to distrust in those organizations.

Richard Westcott 11:34

Yeah. I mean, we're talking about data here. These massive data sets fuel all of this stuff, and how that data is collected, it's so critical to the result that comes out, like you've just been saying. So let's look at how it's governed. How can policy makers, for example, Jenni and organisations as well, design the frameworks that empower communities to have a stronger voice in how their data is managed and used cause when I think about data, I think about it as a valuable thing that organizations and companies want to own and make money out of.

Jeni Tennison 12:03

Yeah, and I think that we we see the companies want to own and make money out of all the time, right? That's the kind of narrative that we see from the way in which social media companies use data. But the narrative for public sector is use and deliver public value from right and use in the public interest. It is kind of undeniable that that data can be used for, you know, medical research, developing new forms of diagnosis, developing new forms of treatment, perhaps improving education, having more personalized services, there's a bunch of things that are really could be really good. And when you go out to the public, and the kinds of methods you can use, like citizens juries to to really give both sides of the story to the public, and you ask them to deliberate about what the pros and cons are and what kind of path to take, they really see those public interest benefits in using in using data, and want to see it used like that. Their concerns are about privacy. Their concerns are about also commercial exploitation of data and what that might mean for future commercial decisions about them, but also the unfairness of people making profit out of what is essentially a public resource. And so I think that it is really essential to bring those voices into the decision making about data and AI, a lot of the decisions that we need to make are around these very difficult to unpick ethical lines about what's good and what's bad and what should be done, what shouldn't be done, what kinds of constraints and controls need to be in place, and where those can be looser, and you need to have very context based conversations about individual applications in order to really get to the bottom of what the right course of action is.

Richard Westcott 14:01

I'm gonna bring you in here Gina. So basically, how do you design it so that you can combat perhaps misinformation online, for example, but also protect individual rights and freedoms? Where do you see the governance of the data and the ownership going?

Gina Neff 14:15

One of the projects that we're working on at the Minderoo Centre for Technology and Democracy is a European wide project. We have 17 partners in 11 different countries working on a project called AI for trust, and there we are using the latest AI tools to try to combat mis and disinformation, basically understanding that governance of our social media platforms shouldn't solely be in the hands of social media platform companies, and that people should be able to get on the front foot, as it were, with understanding what emerging kinds of mis and disinformation are. So we're using state of the art algorithms for recognizing. A video, audio and text based mis and disinformation using state of the art parsing algorithms and social media network analysis to understand how messages are moving in order to create a better picture of what kinds of things are happening. We're starting first with misinformation around health, climate and migration, and we're doing that in multiple languages. So what we know about how content moderation works for social media platforms is that English is pretty good, right? But Romanian, not so much. Greek not so much. So we have Polish, Romanian, Greek, Spanish, English and Italian on this project. And the hope is to be able to develop something that, as Jenny said, is for the public good. We're funded by the European Commission. We're hoping to deliver this for the EU. What we're hoping to do with this project is to be able to give science journalists and policy makers and others civil society organizations a little bit of a head start of when new kinds of themes and memes emerge. So you can do what's called pre bunking, where you say, you know, there will be some things that you will see that look like this, and that's not what where the science is, or that's not where the evidence is, and that we know works much better in a consistent way than just doing labeling work.

Richard Westcott 16:32

Okay, let's talk about the ethics and AI. Because we've been talking about AI a lot, they're all sort of intrinsically linked. Let's talk about how it connects to technology shape in the future. So this is especially true for AI, which raises unique ethical questions, particularly when it comes to making autonomous decisions or operating on massive data sets. So Jean-François, your research explores the moral dilemmas in autonomous systems, particularly in AI. What frameworks or approaches do you think are most effective for programming AI systems to navigate ethical trade offs in high stakes scenarios? Such as healthcare or transportation.

Jean-François Bonnefon 17:11

Yes, okay, so for example, my work is focused on the ethical trade off in the way we program autonomous vehicles self driving cars, because in the system, you might find yourself in a critical situation where you cannot maximize the safety of everyone on the road, maybe you're going to try to protect the passenger of the car in case of the collision, but that means you're going to increase the risk or endanger the life of a cyclist. So in those situations, you cannot win on every front. So you have to tell the machine how to balance the ethicality of the situation. The key problem here is that these problems do not have a universally accepted solution, so in that situation, well, that means that you have to balance a set of values in a way that humans find acceptable. And that's part of what we call the alignment problem, making sure that the AI makes decision in a way that reflects the way humans would balance values.

But then you're faced with the issue of who you're going to ask. It's horrible, but you still have to ask some people what they would want the car to do. And the question is, you have different sets of people you can ask. You can ask the companies that develop the cars, you can ask the ethicists who advise the regulators, you can ask also the road users, the people who are going to walk the streets where these cars are driving, and those people might have different preferences or different way to balance their values these different groups, but also within the groups, you might find subgroups that don't agree on the way to proceed, including among the experts and the ethicist. The problem is, what do you do with this data once you realise that no one agrees, or that even within a group people disagree. So you're not going to be able to list everything that the machine should do, which means that at some point you're going to have to give high level directives to the machines and let the machine autonomously decide how best to align with those high level directives.

Richard Westcott 19:25

Okay, Gina, I'm going to bring you in here. So what steps can organisations take to build more inclusive and ethical algorithms into the entire life cycle of AI development? So from the design all the way through to the deployment?

Gina Neff 19:38

You know, when I think about all of the great work that Jenni and Connected by Data are doing in public services. It's so important because the conversation we're having now about AI and the public sector is that it's somehow sprinkling this very dust of efficiency onto organizations that are quite complex, rather than thinking what's the co evolution going to be as people learn to work with these tools, and how are they going to change the tools themselves? So I've spent 16 years understanding automation and large scale construction, and the short story there is the roll out of those automating tools did not go anywhere near what the rhetoric at the beginning of the adoption said it would be right. It was going to revolutionise the workplace. It was going to completely change the nature of high end construction. It was going to change different kinds of jobs. It was going to do away with old fashioned bureaucracies, remap contracts, change professions, get rid of some types of work, and none of that happened. But what did happen is when people were empowered to make new kinds of choices, they used those tools to get work done, and sometimes they had to create new ways of working to do that, but it was only because the people with the expertise on the ground understood what needed to happen that we got to where we needed with that innovation. And that would be the one lesson that I would take away. If you want safe, trustworthy, responsible, AI, you absolutely need to be working with people in their jobs to build that capacity. Otherwise it's not going to work.

Richard Westcott 21:27

It's really interesting, very human driven.

Gina Neff 21:29

Ultimately incredibly human centered. That's the success story.

Richard Westcott 21:34

Geni, I want to bring you in here because well you heard what Gina just said there, about everyone thinking, AI is like the this great silver bullet to solve efficiency in the NHS and in all of our in all of our different systems and public services, what are you finding of the expectations, say with governments and policy makers, with AI, and how do you think that you can make the AI align with what the public actually wants and maintain that trust we were talking about earlier?

Jeni Tennison 22:00

I think one of the things to just recognise is that any algorithm can go wrong, and what matters is what we do when it does go wrong, and how we respond to failures. And a really nice example of this is the Horizon scandal with the Post Office, where there was an algorithm detecting where there were discrepancies between what the post office was saying that the sub postmasters were saying and what the central systems were saying. And the assumption was that that meant that the sub postmasters were committing fraud and they were prosecuted. What happened there was that the bureaucracy around the decision making systems was such that it assumed that the decision making was right, that the algorithms were right. What we need to be assuming is that algorithms can go wrong. They can be faulty. There are places where they will discriminate against people, where the result that they give for a particular circumstances is wrong, and therefore we need to have the plate things in place to allow complaint, to allow redress and to allow monitoring of the system that lets us see when there's a systemic problem with a particular algorithm or a particular AI. So that's why we argue for it's not just individual people being able to see how the system is operating. It's empowered and powerful civil society organisations who can have people's backs and can be really holding government to account for the way in which, say, benefit systems or health care systems or ed tech systems are operating, that there's transparency that enables them to see when algorithms are being deployed, but also the impact that they're having on people.

Richard Westcott 24:04

Let me ask you an honest question, though, do you think there's the bandwidth of the money to have these checks and balances in place these organizations that can monitor?

Jeni Tennison 24:12

I mean, it's definitely the case that civil society organizations are underfunded and it's hard for them to operate and to do that proactively, but actually, one of the big boundaries and big costs that they encounter is chasing information from the public sector that they could be publishing it proactively, that we could, we could. I mean, I'm a technologist, I know that we could build into systems that we're building the kinds of transparency that will enable them to be monitored much more effectively and much more efficiently in order to enable those feedback

cycles to happen. So yes, of course, we know that public sector finances are strapped and civil society organizations are strapped, but I think we actually remove costs from the system by building it in, building in this, these kinds of facilities really early on.

Richard Westcott 25:04

Okay, let's finish up by talking about the hopes and fears, the opportunities and the risks for algorithms and data. And we'll go around the table, and you can all say what you think the biggest opportunity is, perhaps, and what you think the biggest risk could be. So Jean-François, why don't we start with you?

Jean-François Bonnefon 25:20

For me, the biggest opportunity is that these algorithms know. They force us to be explicit as a society if we want the algorithms to behave well, well, we all have to agree on the ground rules of what is good behavior, what is acceptable and unacceptable, what constitutes bias and what constitutes unbiased decisions. And so these questions maybe were not explicitly addressed at level, at the level of society, and at the level of clarity that the algorithms require. So I think this is one of the biggest social opportunities that the algorithms force us to have an explicit conversation. And I guess the risk is that, well, we're not ready, or that we fail to actually have this conversation, and we end up with a system that is even worse than this sort of implicit equilibrium that we had before. But, you know, I'm an optimist. I think we're gonna make it good.

Richard Westcott 26:20

Good news! Gina, what's your future gazing kind of for us.

Gina Neff 26:25

Well, I think that when you use new tools to empower people at work, you can have really incredible things. So I'm excited about the work that we're doing with trade unions around the world to both build their capacity. Around AI safety, but also help them understand and navigate and negotiate for these changes, and help put workforce in the driver's seat of how these tools and technologies are used. I'm also really excited about what data and research can happen at scale, so I do a lot of qualitative research, and I think that we have a real capacity to hear from more people than ever. It's one thing to interview one one person. It's great. What if we could use these tools to help us gather rich insights from 501,000 people in at a scale that we just don't do in that so I think there, there are ways that we can empower people and empower their insights. The risks, I really think are, and I think the risk is, is real. What may be happening is that we are entrenching today's powerful into a situation where they become unreachable, and if we don't act now, we will have a society where the inequities we face are greater and I think, perhaps insurmountable. So I do think that we can't lose sight that today's AI and algorithmic economy is dominated by some of the most powerful, largest companies that

have ever existed in human history, and that those companies are acting in extreme self interest, and we have work to do to hold them to account.

Richard Westcott 28:16

Jenni?

Jeni Tennison 28:16

On the optimist side, I think that we could get to a state where the people who are affected by public sector uses of data, and AI are active co developers and co designers of those systems, and feel that they are working in their interests and that they are working alongside government to make public services work for everyone. And as Jean-François Bonnefon said, you know, in some places where we know that human decision making has been really flawed, there is the prospect of making those systems fairer and better for for people who have been historically marginalized and discriminated against. The risk, on the other hand, is that that doesn't happen, and by not involving the people who are affected by these systems, they feel even more like they are just subjects to the machine, right that is operating on them, that they have no power in that system, and and that that, you know, creates even further kinds of distrust in public services, in the public sector, in government, and leads to, you know, really, I mean, Gina, Gina talked about the power of companies. There's also the power of governments. And the kind of, the kind of power that they could have with algorithms if we had a bad government is quite scary. So like the really bad kind of situation is the authoritarian government empowered by algorithms acting on people who no longer have power in the system.

Richard Westcott 29:58

Well, that's all we've got time for. Thank you so much. Three fantastic guests today. Really, really enjoyed this conversation. So thanks to Gina Neff from Cambridge University, Jenni Tennison from the Bennett Institute, and Jean-François Bonnefon from the Institute for Advanced Study in Toulouse. Now let us know what you think of this latest episode of season four of Crossing Channels. If you enjoyed it, then do leave us a review. It helps us shape future episodes and helps people find us. So we do appreciate it, and do please listen to other crossing channels episodes, including our last one, where we discussed whether the world is becoming less democratic, interesting, bearing in mind what you've all just been saying. Please join us next month for the next edition, where we'll be talking about green finance.