# Accelerating AI Ethics

## AI and Human Rights: Professor Yuval Shany on AI, Law, and Global Accountability

## Transcript

August 2025

*www. afp.oxford-aiethics.ox.ac.uk*

**00:00:10 Dr Caroline Green**

Hello and welcome to the accelerating AI ethics podcast of the University of Oxford. I'm doctor Caroline Green, and in each episode, we explore bold ideas, innovative thinking, and creative responses to the ethical challenges posed by artificial intelligence.

Today's guest is someone who reshapes our understanding of how technology intersects with law and human rights.

Professor Yuval Shani is an internationally renowned expert on international law, human rights and digital governance. Formerly chair of the UN Human Rights Committee, Yuval has spent decades exploring how law can respond to the challenges posed by emerging technologies, ensuring that human rights remain central in an increasingly digital world.

Now, as a fellow of the Accelerator Fellowship Programme here at Oxford, he's pioneering research into how AI can strengthen democratic governance and accountability. His work asks profound questions about how societies can leverage technology ethically and legally to enhance, rather than diminish, human rights protection.

Yuval, welcome. It's a real pleasure to have you with us today.

**00:01:21 Professor Yuval Shany**

Thank you for having me. It's nice to be here.

**00:01:23 Green**

Yuval, before we get started and diving into AI and human rights, tell me a little bit about yourself.

**00:01:30 Shany**

So, I'm a law professor. I teach. I taught for many years international law, Hebrew youth, Jerusalem, and I've been teaching for the last three years or so what is called Transnational Law here at King's College, and in the last two years have been affiliated with the Ethics and AI Institute here in Oxford. Last year, as a visiting fellow, and this year as an accelerator fellow like you mentioned before, I mean, I did spend 8 years of my career working for the UN in the Human Rights Committee, which is the body of experts. We are 18 experts, the committee still comprises of 18 experts that monitor the application of a treaty, the International Covenant on Civil and Political Rights, to 173 states that have joined that treaty and in that capacity, I both reviewed state reports for explaining their level of compliance and the measures they've taken to comply with their legal obligations, but also sitting in in in a quasi-judicial capacity, hearing individual cases by persons who believe that their rights were infringed, and also formulating what are called general comments, which are basically normative statements about what states really should do in order to better enhance the protection of human rights.

And like you said in the intro, Caroline, I've been involved in the last years quite heavily in this interplay between human rights, law and new technologies, which is quite a challenging field. Both of research, but also of practise.

**00:03:02 Green**

So, what strikes me about the work that you have been doing within the UN, you know, it's above what usually professors would be doing, who are researching and teaching and so on. I would love to hear more; what made you decide to focus on human rights in your career?

**00:03:22 Shany**

Well, I mean, I grew up in a place where many human rights issues, of course, come to the front. But I think human rights is really the place where international law, this general body of norms that apply in traditionally in the relations between states. This is where the rubber hits the road; where international law actually meets regular persons, and domestic legal systems and requires us to think not only about our domestic laws and our domestic rights, but also about how we can sometimes leverage some global perceptions of what we are entitled to obtain, and what our government is required to provide us.

And I thought that this is actually a very interesting intersection between domestic law and international law, between public law and private law. And, this is also where you can utilise your legal tools. Like you said, in order to make a difference in the lives of real people, because they do come to you either as a consultant or as a lawyer, or as an academic, but also as a member of a UN body, to actually help them with real life problems, which is not always something that academics do, certainly not academics who are are focusing on international law so. Over the years I I found that within international law, the area that I felt is most consequential and where I can make the most significant impact on people's lives is this interplay with human rights lawyers.

**00:04:58 Green**

I have got lots of questions in my head now. One thing I'm really also interested in is, you know, the United Nations as an organisation. I see it as something that is actually a place of hope really. It brings nation states together, talking about how you know, we envision our societies.

What role do you think the United Nations is playing in today's world? Because I also think that even though it is a place of hope, I think a lot of people have lost confidence in the United Nations and see it more as like an administrative heavy institution rather than something that can make meaningful change.

**00:05:42 Shany**

Yeah, I guess it's it's true that it's it's, it's all of the above. I mean, and I think you're right in first framing it as as a place of hope, in the sense that many of the problems that we are facing and we'll we'll talk soon about technology, can only be resolved through international cooperation. These are not the kind of problems that governments can resolve on their own because of the magnitude of the problem, because of economic interdependence, because of questions relating to political capacity, etc. But also, what we talked about up until now. I mean, human rights are often exactly the kind of problems that governments don't have the will and sometimes the capacity to resolve, sometimes the protection is needed exactly vis-à-vis one’s own government, I mean, and historically, human rights were developed in that context. And the only framework that we have for dealing with global problems in a structured and principled way is the United Nations and the Treaties that, I mean, I mentioned the Covenant on Civil and Political Rights. I should say that there are eight other treaties of that nature that protect our basic human rights. And there are other, of course, many treaties in other areas of international law, they represent a certain hope for the application of appropriate, just and fair standards to everyone around the world, through mechanisms that are again created under the auspices of this body, in order to make a real difference. At the same time, there is a famous book in that deals with actually not with the UN, but rather with the US Supreme Court, which is called the Hollow Hope. That hope is somewhat hollow because this organisation has been facing over the years very significant problems that have to do, not only with bureaucracy but even more fundamentally, with lack of funding and lack of resources to undertake serious work. But also, lack of political will by the stakeholders, who are ultimately the states to, confer on this body the necessary legal powers and material resources to actually make a difference. So, when you work for the UN, it's always this tension between what you what you think that you should be doing within an international framework and the frustration of what you can actually deliver within that framework. And for many people this is a disappointing experience.

I actually experienced more directly the half full part of the glass. I thought that we were making, for a lot of people, we were making a difference, and the very fact that you know, people had someone that they could look up to, and make their voice heard, that was in and of itself a very important feature of what we were doing, so we were not able to help many of the people who approached us, we were able to to help some, but I think we were able to give hope to many more and and that is in and of itself is a very important task that at this point in time at a global level, only the UN, with all its many deficiencies, can fulfil.

**00:08:59 Green**

If we now come to the topic of artificial intelligence and human rights.

I think we now live in this brave new world, right? You also just highlighting the role of government and how they are, you know, the main they are mainly responsible for the protection and respect of of human rights, but AI is something that is developed by big private corporations who've now got this very big role in our lives. How do you see human rights play out in that field? How do human rights apply to private corporations? These big AI tech companies?

**00:09:38 Shany**

Right. So, we're dealing here with with a multifaceted challenge, which would have at least three dimensions that could be referred to. And I will deal with the one you raised the question of applicability to companies last just to because even before you get to that very serious challenge, you still have two previous challenges that you need to deal with.

First is the question of do you have the laws to deal with this sort of technology? I mean, human rights, the Treaty that I mentioned was drafted in 66 in 1966 on the basis of of a UN resolution, the Universal Declaration on Human Rights from 1948. These standards were developed 70, 80 years ago, of course they didn't conceive of the kind of practical challenges that we are now confronting. There was no Internet, there was no social media. We didn't have computers that have the capacities that the computers that we now have and this whole idea of artificial intelligence was something akin to science fiction. What we need when we are thinking about applying human rights law to artificial intelligence, we first have to ask ourselves, do we have the right toolbox, do we have the necessary norms to even go there when there is obviously a temporal distance between the the time in which the norms were created and the time they are now applied.

Secondly, you have states that actually see this technology as (and they are regulators), they see this technology as something that gives them quite a lot of power. Artificial intelligence in some areas, such as the security or military industry, is something that states are quite keen to utilise and they see this as something that they would not be too keen to give up. So, any attempt to actually regulate this in ways that would deprive these new candies, these new sources of power from state, is something that would be politically very difficult to pull off.

And then you get to the third dimension and that is that, unlike many of the classic human rights challenges that we were thinking about in the mid-20th century, where a government was limiting freedom of expression, or, where a judge in a domestic court was applying the law in a biassed manner ‒ for that we did have some sort of a framework, because in many countries we had constitutional norms that could actually be then adopted at the international level ‒ we don't really have the same kind of frameworks, conceptual and practical frameworks ‒ when those who are exercising much of the authority that we are now experiencing as potentially oppressive or discriminatory or otherwise unfair, are private entities, private companies that are not democratically accountable. They're not typically subject to constitutional law.

We need in a way to reinvent the wheel, at least both with regard to the substantive norms that would apply, but also, how do you go about intervening in the decision of a company that applies an algorithm for content moderation online? Should we treat it as a government or should it treat it as a private business? If it's a private business, they can do everything that is not prohibited by law. If it's a government, it's exactly the opposite; they can only do what the law authorises them to do. So, the dominance in this sphere of private companies, the size of these private companies, the global field of their activities, the difficulty sometimes in even pinpointing the activity and actually identifying in which country, in which jurisdiction it is taking place. Where is the content moderation occurring? Where is the cloud? Where is the the generative AI software moving or the application? Where is it physically located? [This] really creates an impossibly long list of practical and theoretical questions which makes the applicability of these 1940s 1960s treaties almost impossible.

**00:14:09 Green**

Hmm, so we need new solutions, don't we? And within your project, you've been looking at potentially a new AI bill of human rights. You've been doing a lot of research. Can you share some of your insights? What do you think? Do we need a new AI bill of human rights?

**00:14:27 Shany**

Yes, we need a new AI bill of human rights, and the short reason is everything that I just said before. I mean the the tools that we have are not fit for purpose. They're not fit for purpose to apply to a technology, to AI system. They're not fit for purpose to bind governments on the basis of existing laws, and they're not fit for purpose for applying directly for corporations. And yes, in principle we would need to go in that direction. I think there is a general ‒ I mean there is no controversy around the basic proposition that human rights should apply into space, and that we, as human beings, do not lose our human rights and our human rights claims just because there is a new technology around the corner or a new technology in the rooms in which we engage in interactions with other human beings or with other computers these days.

The question is, how do you get from point A to point B? And so far, what we have seen is some attempts ‒ and much of my research here at Oxford has been to try to dissect and to understand what has been attempted so far; what are the shortcomings? But also, maybe the promising aspects of what has been attempted, and what more needs to be done. We have seen some attempts, some of them, I would say, more impressive than others at the international level, but also sometimes at the national level to broadly speaking, regulate AI systems in ways which are compatible with human rights values with human rights.

The main claim that I'm making in my research, or that I'm advancing through my research, is that so far, most of these efforts that we have seen either tend to be too general in nature, or too specific in nature. We have seen on the one side of the spectrum, the Council of Europe adopt recently, last year's Convention, what they call a framework AI Convent, that identifies very broad principles like transparency or equality or human dignity as principles that should apply to AI. But the problem is this is a bit like motherhood and apple pie. Everyone would agree on the principle.

The trick is how do you act? Really implement it, how? How do you actually break it down to concrete obligations of conduct that would apply both to states but also to companies? And this is something that the Convention doesn't go far enough. On the other side of the spectrum, you have the European Union introducing again last year the AI Act, which is a very detailed regulatory instrument which lists and categorises different AI systems on the basis of the risks that they that they pose, and requires actually those who disseminate those systems to undertake, inter alia, among other things, what they call a fundamental rights compatibility review, but it doesn't really define what are these fundamental rights that are at stake. It assumes that we know it, that when we see it, that we know, we can identify human rights violation when we see it, and I think here one of the one of the difficult issues is conceptualising ‒ is giving a name to a human right entitlement that we would actually want to protect. The regulation is actually quite good in generating a technical inspection, a technical evaluation framework, but it is missing the the principles that they do not really identify, in ways which are implementable. And this is exactly this middle level regulatory space, which I think something like an AI Bill of Rights, could satisfy.

In closing, I would say that there is one attempt that had been made in that direction and that is an attempt that was made under the previous Biden administration in the US. This is not something that the existing administration has kept on, but they did try to actually formulate a blueprint for an AI Bill of Rights. This was not a legally binding document. This was more in the form of a guidance document that would advise other federal agencies in the US as to how would they implement human rights to AI systems and this is actually quite a good model of this mid-level range of regulation which identifies rights, but not just in a very abstract sense, but rather in a concrete conduct guiding sense, which could then become implementable by administrative action and maybe in other contexts also applicable directly to companies.

**00:19:43 Green**

Thank you so much for that overview of where we are at the moment when it comes to AI and human rights, I think it's really fascinating to hear. I'd love to hear a little bit more. What do you think could be the content of an AI Bill of Rights. Do you have any idea of what kind of articles you'd like to see, which rights could be protected and specifically?

**00:20:04 Shany**

Right. So, I think there are five or six basic trajectories. And, of course there could be some hair splitting as is in every case of articulating a Bill of Rights and I don't feel wedded to a specific number or a specific formulation, but I would say that any Bill of Rights should be dealing with questions relating to access; access to AI systems, I mean, we often think of AI systems as a threat to human rights, but it's important to to emphasise, and actually AI systems (like you said in the introductory remarks) could be a very important component of enjoying human rights.

If you're thinking about the right to health ‒ having access to AI systems that can improve the quality and accessibility of health that we receive, that could actually be a very important component in a new health environment, and it's important to ensure if we are thinking about access to health as a human right (and of course, it is a human right) and we're thinking about access to AI systems in certain contexts, as a human right, then you can see how these two notions align with one another. So questions of access, and some were actually introduce within access also, questions relating to safety and reliability, so that we are not only accessing AI systems that exist, but may miserably fail. We need to ensure that these AI systems are sufficiently robust, sufficiently accurate so as to protect our human rights.

Then there is another cluster that deals with equality. We know that there are very serious concerns about the impact of AI systems on equality, which have to do both because of concerns of bias being baked into AI systems, but also because of lack of transparency of this black box situation, which is allegedly, arguably making it easier to actually mask instances of discrimination. So that will be, I would say the second cluster.

Then there will be a cluster dealing with data protection. There are a lot ‒ I mean AI systems are very our systems that feed on private data on personal data. And there are very significant issues regarding how this data is collected, how it is protected from being traced back to the individual in question. There are also sometimes questions, and I know that you've worked on this regarding intellectual property rights that are pertaining some of the ‒ whether it's training data or data that is being used later on in the life cycle of an AI system, so this will be another cluster of issues.

Then are questions regarding transparency *per se*, which is not something we normally think about as a basic human right in an offline world. But here actually the question of transparency and explainability is actually quite important to mitigate the perceived arbitrariness of AI systems. … It can be used effectively as a tool of authority, of a tool of power, that is exercised against us and if we have no basis of understanding why certain decisions are being taken, or what is the worldview that is informing the outcomes of AI systems that are becoming more and more important to our lives, than our ability both to protect our core interests but also to understand the world, the world around us; to have this access to a basic epistemology, this is significantly compromised. So there are calls and, I would certainly support them for introducing a right to transparency, or explain and/or explain ability in this.

And finally, the final cluster that an AI bill of of human rights … has to do with questions regarding certain facets of life, certain interplays that we do not feel comfortable relegating to AI systems. We already see, in some spaces, the right not to be subject to automated decision making. Within the EU context we know that already in 2016, the GDPR, the General Data Protection Regulation, has limited the ability of private and public entities to use automated decision making in certain contexts. For instance, it may kick in when you ask for a loan from the bank for mortgage, or when you ask for a benefit from a government agency. And in these contexts, individuals may be entitled to obtain some human inputs either to have a human take the decision or to opt out to a human decision maker, or have a chance to appeal a decision that was taken by an algorithm before a human. But at least in my understanding, and this is part of the work that I've done here, I think maybe we should be thinking even more ambitiously and more broadly, not only about decisions that are taken by algorithms, but also about interactions with algorithms and there could be certain areas where we as human beings are not comfortable with having only an algorithm on the other side of the interaction. So that could be in a patient, physician kind of interaction ‒ we may want to actually insist on seeing a human physician and not an algorithmic physician that would diagnose us.

It could be in a in a caregiving context, where we may want to be treated by a human nurse and not by a robotic nurse. It could be in an academic context where we may want to be actually taught by a human professor in a university and not by an algorithmic professor, and we could, it could also take place in a in a, in a service context where we would want to be served in a very a consequential context, for instance, relating to our most important benefits by human being and not by a machine. So, in all of these cases we can of course disagree about what should be the precise contours of such a, right. And maybe whether there should be a right, but I would submit that these are areas where very fundamental human values, such as dignity, such as liberty, such as equality, solidarity, as well come into play and could justify the introduction of a human rights framework to supplement the more safety or reliability type of regulation that we see around.

**00:27:24 Green**

I guess one of the main challenges though, I'd love to hear what you think about this is that AI is developing so fast, we're seeing new applications all the time. And I wonder, I try and think of the future of what might be possible, what AI can do for us. And here I assume that that must be a big challenge when we're also thinking what could a new AI Bill of Human rights look like so that it's actually adaptable to the speed at which AI develops.

**00:27:54 Shany**

Yeah. I mean, one of the arguments you often hear against regulating AI is that “Well, it's moving so fast. So, if we regulate it now, there will be in, in two years’ time a very different type of a system that would fall between the cracks”. But this is a bit I find it a bit disingenuous in the sense that this is (not you), but in the sense that not regulating is also regulatory choice. Because it means that for the time being, and the time being is now, and people are already affected now by the technology, it's not something that would only kick in down the road. We need to actually if we think that AI has the potential both to advance human rights, but also to dilute human rights. We need to address these problems now. Now, human rights law, I would say two things. First, regarding regulation we've heard these mantras, I would say when the Internet was introduced, there was quite a reluctance to regulate this and we've also heard these mantras when social media was introduced, and you know, we we've heard that it's time to move fast and break things. And the end result is that the technology is then being entrenched into economic, social, political structures in ways that make it almost impossible to then change things back from where they were. So, the decision not to regulate is not only a decision about regulation ‒ now it's also decision about the potential for regulating in the future, and if we decide not to regulate now, we may not be in a position to regulate even in the future.

But with regard to human rights, I would say that these arguments are not really operat[ing] with the same kind of strength, because like I said before, human rights were always designed to be future oriented. They were formulated in the 1940s in international law, in some constitutions, the US Constitution was, you know, the basic Bill of Rights was formulated in the 18th century ‒ so this idea that you have to formulate rights that would be effective now, but also would be sufficiently flexible to also govern and address future challenges ‒ this is something which is very intrinsic, and I think formulations such as the ones that we have seen with the Biden White House that talks about an anti-discrimination component to AI systems, this is not something that creates significant problems in terms of stifling the future development of technology. So there are ways to introduce human rights, human rights oriented regulation or introducing human rights entitlement in ways which would be sufficiently flexible to accommodate change and not to stifle change, which is of course important. We don't want to do that either.

**00:31:01 Green**

Yeah. I want to pick you up on that point a little bit more because we hear a lot here along the lines of innovation and human rights that doesn't work right. So human rights and other values for thinking about that kind of rights oriented regulatory system that will stifle innovation. But what I really liked about what you know you've said now several times is that actually AI and technology can help us in human rights protection, it can make people's lives better. Is there a way that we can really challenge this narrative, that innovation and human rights protection doesn't work with each other, that there's this stifling of innovation if we're also thinking of human rights risks and harms, how can we work against that narrative?

**00:31:51 Shany**

Exactly in the way you articulated this. I mean, we don't see these values as opposed to one another. We don't see human rights as an enemy of change. We don't see human rights as an enemy of innovation, quite the contrary.

There is a right to development under human rights. There is a right to there are positive obligations in in fields like health or welfare or education and to the extent that technology could actually support this, yes, we are very much in favour of further development ‒ and even expedited development. So, there is no inherent tension. At the same at the same time, I don't think we have any example in any other sector where we have decided as a matter of policy that we don't want guardrails that, you know, factories can work without health and labour standards, and without the work, you know, without protecting workers. And cars could travel on roads without any safety rules. This is not a choice that we have made in any sector of our economy in any other form of industrial or development, or something of that nature. And it is quite mysterious to basically to accept the proposition; “this is the one singular exception to the rule where we are where we have to think about development in ways which are divorced from human welfare”, and I think there is a general sense, at least again, in these grand international declarations such as the global Digital Compact, which is a UN statement from from last September, that the development of AI should be human centric, … we don't know always what that means. But at least as a starting point, I think this is about right ‒ that we want this technology to serve us, to help us, to make our lives better, and therefore we should evaluate this technology and identify those aspects that are in line with our rights. Aspects that actually could further advance our rights, and aspects which are in tension with our rights. And then we have to find work arounds and for instance, I mentioned data protection ‒ this is one area where technologies such as encryption, or creating some other ways of anonymising data could provide us with technological solutions for human rights problems. I don't see this plea for data protection as necessarily something which goes against the idea of innovation. It's simply another, you could say, technical requirement that we want innovators to work on, right?

The same thing with transparency. The idea of XAI, this idea of explainable AI is again, not something that is goes against innovation. It actually fits in a whole stream of Innovation, of how to render these machines more predictable, more, intelligible to human beings that interact with them, and this actually is also important for quality control, because then those hallucinations, or mistakes that could happen would be less plausible if we know better what the outcomes of these machines is expected to be.

**00:35:27 Green**

You know, I think you're, opening up a whole new perspective on looking at human rights and AI because as I said earlier, the common narrative is human rights and AI are ‒ it's all about risks and harms. And it's about, tuning down AI and how we're not to roll it out across society. But what you're saying is that actually, let's be excited about what AI can do for us, for our rights, and how we can use it in a way that actually, makes our lives much better. So it’s a really positive way of looking at human rights and AI.

**00:36:00 Shany**

Yeah, but I don't want to sound too positive and I and I will say that one of the implications of what I said, and I also alluded to that in the context of the AI Act, we have to think about what we're doing. So, this ideology of moving fast and breaking things, this is extremely problematic. And I realise that much of the industry has basically trained itself to to believe “well, we'll first release the first prototype and then we will improve the product as we go along”. A human rights analysis would say “no, you have to engage in some reasonable process of risk assessment, and you have to do an ethical, but also a legal review of what are the risks.

I would say it's, somewhat of a speed bump in terms of the launch process. You have to be a bit more careful when you are launching new products ‒ to think also about their implications, like other safety concerns that you would have to factor in. That's another…. I don't want to sound too sanguine about, you know this is one happy marriage. There are certain trade-offs. There are certain prices will be paid but, I don't think there is a fundamental tension, right? I mean, it's not that this, even if we slow this by 10% because you would have to undertake more examinations before launching new products, these products would be in the long run better for society. So we do want, from a human rights point of view, the innovation to kick in and to kick in fast and to bring as many people as we can into this world of AI. But we have to do this maybe a bit more prudently and smartly than we have done up until now.

**00:37:46 Green**

Yeah, I also thinking of the UK Government and I know that also other governments around the world are also looking at AI in that way. But there's huge enthusiasm for AI for example, within the health service, right? We are facing big problems in terms of resources and so on. Recently we've had a new health plan for England published and here AI is also [presented as] one of the big lights at the end of the tunnel that will help everybody with their AI companion in their pockets and that's going to replace some of the GPS services and so on, so I guess that is also one of those areas where you would say, you know, we need to be mindful on how do these systems actually work, what are their impacts on people's rights.

**00:38:36 Shany**

Yeah. No, I would say I would say two things. First, yes, I mean we have to be a bit more careful and to do a proper analysis of what it means also from a human rights point of view and whether in terms of trade off, whether greater access and you know without … having to wait two weeks to see a GP, this is also a serious problem from a human rights point of view. So, if you can cut, you can cut the two weeks waiting period to two days. This is a significant gain from a human rights point of view, but you should also be mindful of some other problems which we mentioned; discrimination and transparency, right to interact with the human being in certain circumstances, etc. That's one thing I would say.

The other thing is that I don't think human rights give you most of the answers to these questions. Human rights is a framework, and it's a framework which highlights trade-offs and some of times it is trade-offs between competing rights; sometimes it's trade-offs between competing rights of different people. And sometimes these are trade-offs between rights and efficiency, or other societal economic considerations, they are not foreign to human rights law.

So, the right to health is an excellent example. The right to health is not an absolute right. It's a relative right. And its implementation does depend on the availability of resources. Ultimately, these are decisions that democratic societies have to undertake through democratic processes.

And I think that this is another area where we have not seen these problems ‘problematised’, ([that’s] a bit mouthful) as human rights problems that require some democratic input. If you only think about them as manage economic issues that you have to move numbers on a budget sheet from one column to the other, then this is something that will often be decided at a technocratic level. But if you actually conceptualise these issues as as serious issues of principle, that have significant trade-offs, this is an area where you would want actually elected politicians, elected representatives, and sometimes even broader constituencies to have a significant input. I mean, one thing I can say in favour of the AI Act, which I was a bit a bit, I was a bit critical of before, is that the AI act did actually adopt very specific arrangements regarding some very contested policy issues. So, for instance, the use of biometric data in surveillance in CCTV surveillance, or the use of subliminal influence on consumers ‒ and these are decisions that, some of them, are actually hotly contested in the European Parliament, and that's a good thing. So, you could basically criticise the outcome whether they have gone too far or have not gone far enough. But at least you had a a robust, democratic deliberation around these issues. I'm not sure that this is true across the board regarding the introduction of AI into public services, and certainly not the introduction of AI by private companies into markets. We're seeing this now actually taking place in in a slightly different context regarding social media and minors and now we are seeing some states trying to walk back decisions that they haven't really consciously made 10 or 15 years ago, regarding access of minors to social media and what it does for their mental health and welfare, and what have you. And now, but of course, walking back something is much more difficult than addressing the problem when it is still under manageable proportions.

**00:42:43 Green**

I want to go back to that deliberation and also in democracy and engaging the public with AI, and those decisions that are being made, and I would say, probably quite cynically, but are are being made across public services, for example, without citizens actually really knowing that this is happening? Right? And AI is part of our lives now ‒ what do you think are now the steps that we, as citizens, could take to get more control over what is actually happening within government in our lives and so on ‒ so we've got a say.

**00:43:24 Shany**

Yeah, that's of course increasingly difficult in a very polarised societies where accountability of office holders towards broad constituencies is being challenged on a number of levels and you have a technology that, we we've discussed this before by private very large private companies that may be too global and too large to regulate; so we do have serious problems regarding empowerment and disempowerment of citizens *visa vie* those in power both in government and and in private companies. And I think the one of the appeals of this Bill of Rights approach, which again I should clarify that I'm not advocating that the UN would publish a treaty tomorrow. It's more of an agenda which would we would have to find, you know, practical ways of slowly advancing. But one of the steps which I would want to encourage through this discourse is [to] encourage individuals and civil society groups and other change actors to actually conceptualise and actually name those issues, as human rights issues and use the vocabulary of human rights both to raise awareness to individuals, to the entitlements that we should have, also to the risks to those entitlements that the new technology could potentially constitute. And also encourage them to try to use the different mechanisms that we have domestically and internationally to challenge these applications of power, these applications of authority against us, without our consent, without a proper process of democratic deliberation, and hopefully the even the mere threat of more lawsuits, more public campaigns, more what is sometimes called the UN, public shaming debate, when the state is being taken to task for failing to engage with the public about a sensitive issue, maybe this would give some politicians, but also some companies, room for pause in terms of before being too quick to launch a new approach to health services or education or welfare, what have you, try to think of what would be, what could be potentially the rights pushback that this could generate.

**00:45:57 Green**

I'm coming to my last question, and that is about the project that you've been doing as an accelerator fellow with us at the Institute for Ethics and AI. I think we can look forward to having a White Paper published by you quite soon. Can you tell us a little bit more about that?

**00:46:14 Shany**

Yeah, my project here was really to try to think about this idea of coming up with Ian AI Bill of Rights or a blueprint for an A Bill of Rights, both in terms of the substance, and we've discussed this, and the mapping exercise that I have done looking at domestic and international instruments was basically quite helpful in fleshing out these what we call these clusters on which we are starting to see the emergence of human rights thinking and even some human rights norms.

But also, to think about what could be within, as a practical matter, how such an agenda could be advanced. For this I have engaged with the help of the, with the help of the wonderful staff of the programme here, in a process of public consultations with experts and stakeholders, from different parts of the world, who are working on similar issues or are concerned about similar concerns. We've held a meeting here at Oxford, but we've held a meeting in Geneva with people from the UN human rights ecosystem. We've held a meeting at Harvard and we will be soon, very soon, having a meeting in Pretoria in South Africa, so as to get both global, regional, Western, Southern perspectives to these issues. And those consultations are actually very useful for me, both to access the kind of knowledge and ideas that are out there, but also to get a sense of the field of what is actually, what could be used for, what could be plausible at this point in time, and I'm hoping that the the White Paper would in and of itself, would not be a treaty to be adopted, but it could generate further discussions among domestic, regional and global bodies, as to what steps could and should be taken in order to advance this sort of agenda. I would see this project as an attempt to trigger or to accelerate, (and that's a very good title for our Accelerator Programme), accelerate deliberation of the need to fill existing gaps in the human rights discourse and the human rights catalogue of norms in this area.

**00:48:44 Green**

Wonderful. Yuval, thank you so much for this conversation. Listeners can learn more about your housework and then also read the White Paper and the other fellows’ projects by visiting the Fellowship website at afp.oxfordminusaifx.ox.ac.uk.

This has been accelerating AI ethics, a podcast from Oxford's Institute for Ethics and AI. If you enjoyed this episode, please subscribe and share. Until next time, thanks for listening.