HOUSES OF PARLIAMENT

Joint Committee on Human Rights

Uncorrected oral evidence: Human rights and the regulation of AI (HC 1262)

Wednesday 29 October 2025

4.10 pm

Watch the meeting

Members present: Lord Alton of Liverpool (The Chair); Lord Dholakia; Tom Gordon; Baroness Kennedy of The Shaws; Afzal Khan; Lord Murray of Blidworth; Lord Sewell of Sanderstead; Alex Sobel; Peter Swallow; Sir Desmond Swayne.

Heard in Public Questions 28 - 33

Witnesses

I: tèmítópé lasade-anderson, Executive Director, Glitch; Javier Ruiz Diaz, Technology and Human Rights Lead, Amnesty International UK; Alex Pirlot de Corbion, Director of Strategy, Privacy International; Silkie Carlo, Director, Big Brother Watch.

USE OF THE TRANSCRIPT

1. This is an uncorrected transcript of evidence taken in public and webcast on www.parliamentlive.tv.

2. Any public use of, or reference to, the contents should make clear that neither Members nor witnesses have had the opportunity to correct the record. If in doubt as to the propriety of using the transcript, please contact the Clerk of the Committee.

3. Members and witnesses are asked to send corrections to the Clerk of the Committee within 14 days of receipt.

Examination of witnesses

tèmítópé lasade-anderson, Javier Ruiz Diaz, Alex Pirlot de Corbion and Silkie Carlo.

Q28 **The Chair:** Welcome back to the proceedings of the Joint Committee on Human Rights. At the present time, we are looking in detail into AI and its impact on human rights. We have a second panel with us today. I am going to introduce them before we turn to the questions.

We have Silkie Carlo, who is the director of Big Brother Watch and no stranger to the Joint Committee on Human Rights. Silkie works to further human rights and equality, particularly in the fields of state surveillance, policing tech, big data, internet regulations and AI. She and Big Brother Watch do this through parliamentary lobbying, strategic litigation, investigations and public campaigns to successfully change policies and law. She has spearheaded several national campaigns, including one against live facial recognition surveillance; her work on this was featured in the Sundance-nominated Netflix documentary "Coded Bias". She is regularly invited to give expert evidence on civil liberties matters to the UK Parliament and has also given oral evidence on technology and human rights issues to the European Parliament and the Bundestag.

Next, we have Javier Ruiz Diaz, who is Amnesty International's UK technology and human rights lead. Javier is a UK-based advocate working on human rights and digital policy. Prior to his role with Amnesty UK, he worked for a number of civil society and consumer organisations, both in the UK and internationally. He is an associate at the Sussex Centre for Law and Technology and the Centre for Inclusive Trade Policy, where he focuses on the rights impacts of digital trade. He is also a member of the intellectual property policy insight forum at the Department for Business and Trade. Javier is the former policy director of the Open Rights Group.

We are also going to hear from tèmítópé lasade-anderson, who is the executive director of Glitch, a charity that works to ensure that tech does not facilitate discrimination towards black women. tèmítópé is a Nigerian-British-Canadian writer and researcher, and a PhD candidate researching digital intimacy at King's College in London. tèmítópé is currently a 2025 affiliate at DISCO Network's Black Communication and Technology Lab; the acronym stands for Digital Inquiry, Speculation, Collaboration and Optimism. She was a fellow at the Centre of Advanced Internet Studies in 2024. tèmítópé's work has been published in academic journals and independent media.

Our final panellist will be Alex Pirlot de Corbion, the director of strategy at Privacy International. Alex is the convenor of Privacy International's strategy team and oversees its network of partners across the world, global advocacy, and public engagement. In addition, she works on an array of issues at the intersection of human rights and technology.

We look forward to hearing from you all individually. My colleagues will come in on individual questions later but, once again, let me kick off our proceedings by asking an opening question on the management of risks,

data, privacy, bias and discrimination. I know that some of you were able to sit through the earlier session and heard our discussions around those subjects, but can you give us any specific examples of AI breaching people's human rights in the UK today?

Alex Pirlot de Corbion: Thank you very much for giving us the opportunity to share the work we have been doing in this domain, both in the UK and worldwide. By way of introduction to this first question, I wanted to offer three precedent-setting examples, showing how the use of AI is triggering human rights concerns, particularly in the UK.

The first is the practice of industry and government making a mad dash to hoover loads of data into AI systems—including models—before we even noticed and without restricting their conduct. As a result, it is really hard to know how much of our personal data is being processed by companies and Governments. When we have been able to hold them to account, such as in our latest case against FRT firm Clearview AI, which was scraping people's photos from social media into its system and selling its services to police across the world, we were able to get an order that it must delete that data, including that of people in the UK.

The second example that I wanted to share is that many UK government departments are very keen to deploy algorithmic decision-making. For example, in 2019, Privacy International uncovered how the DWP used an algorithm to identify and investigate individuals as part of its fraud investigations. Other examples have been raised by other civil society organisations around the vetting of universal credit claims on the basis of the use of algorithms. As you may know, plans are currently under way to use algorithms to scan millions of bank accounts as well.

Another example I want to share is that—as we have documented—the Home Office has seemingly been using algorithms for immigration purposes, including deciding whether somebody should be detained or deported, or whether they should be tagged 24/7 with GPS tagging. This is an issue we have recently brought to the attention of the Information Commissioner's Office.

The third point is that, unfortunately, we have very little clarity on how government agencies are using AI tools, particularly agencies that hold vast amounts of data and use third-party services. By way of example, it was disclosed in August of this year that the Israel Ministry of Defence has stored vast amounts of data on Palestinians—including communications data—on big tech cloud services and is using AI on this data. We know that the UK intelligence agencies are using the same big tech providers, but what we do not know and do not have clarity about is how that data is being processed.

I close by emphasising the importance of considering the breadth of human rights being affected by the use of AI. While at the onset we are talking about data and how the way it is being exploited is triggering concerns for the right to privacy, when the UK Government consider the use of AI in health, for example—as was recently mentioned by the

Health Secretary—we are also talking about the right to health. When we talk about FRT in public spaces, we are talking about the right to freedom of assembly.

The Chair: Thank you very much. I have to call the meeting to order because a vote has just been called in the House of Lords. I do not know whether my colleagues intend to vote, so I will suspend the meeting while I find out.

Sitting suspended.

The Chair: Welcome back to the meeting of the Joint Committee on Human Rights which was interrupted by a vote in the House of Lords. That has been completed and my colleagues are back in their places.

We have heard from the first of our witnesses. I turn now to Silkie Carlo, who is going to give her opening statement on AI and the threats and challenges that it poses. The floor is yours.

Silkie Carlo: Thank you, chair, and thank you to the committee for giving Big Brother Watch the opportunity to share our research and analysis. Responding directly to the question about examples of artificial intelligence involved in human rights breaches today, I echo the evidence given that it is incredibly hard to get information in this area. Although we know that such harms are happening, finding detailed information about them takes a lot of hard work. The examples that I want to give are mostly from Big Brother Watch investigations spanning welfare, education, policing, justice, healthcare and national security.

I will talk further in this session about the use of facial recognition in policing, which now operates nationally on a large scale. Already this year, our estimate from police figures is that around 3 million people have had their faces scanned by police facial recognition technology. Less than a 10th of a percentage of those scans result in a positive identification and an arrest. So AI has clearly contributed to a massive expansion of mass surveillance in this country.

We have also seen examples in the criminal justice system. Our investigation into Durham Constabulary a few years ago found that it had developed its own AI tool called HART—the harm assessment risk tool—which is essentially a recidivism scoring tool. Before a decision was made about whether to charge an individual and prosecute them, they were put through this AI system, which basically gave a risk assessment of reoffending. A decision was then made whether to prosecute that person or put them forward for a rehabilitation programme. AI is being used by a British police force to decide whether to prosecute somebody, so I think we are getting an idea that the consequences of AI use are really significant now. In that particular case, we found that postcode and location were significantly weighted variables in the system. Individuals' postcodes and locations were having a significant bearing on whether they were prosecuted.

I will refer to some automated decision-making examples as well as AI, because they are a hair's breadth away. You will all be familiar with the A-level exam grading scandal, where around 40% of students received grades lower than they were anticipating. That also related to their postcode and location, so disadvantaged students were unfairly impacted.

Our investigation into the use of algorithms in welfare found that a DWP algorithm wrongly flagged 200,000 people for benefits reviews. A significant proportion of the people who had been flagged were wrongly flagged, which obviously causes issues for both the DWP and those individuals.

I want to reference national security as well, because this is one of the most opaque areas which we know very little about. We know that AI is being used by our national security agencies. For example, there was a GCHQ paper in 2021 called *Pioneering a New National Security*. It contained very vague copy about AI and cyber ethics, and no real detail about how it was using AI, but I noted that the agency said it would use AI against disinformation, including "machine-assisted fact-checking", and that AI would be used to mount online operations.

The risks of this will not be lost on anyone. It is an incredibly sensitive area, especially if we have intelligence agencies involved in fact-checking information on the open internet, where it is very hard to geo-fence. It could well be that British people's social media posts, for example, could be impacted by our own agencies using AI for fact-checking. There is a very significant risk that we will have a future Snowden who will suddenly reveal the use of AI in the agencies, which may be eye-watering. If it is anything like the previous revelations we have seen, it could be completely out of kilter with what the average citizen would expect.

Finally, one of the most chilling examples of AI being used in opaquely in this country comes in the healthcare setting: the liver transplant benefit score. This is an example of AI being deployed in healthcare to determine a life or death decision of whether an individual is offered a liver transplant or not. No humans were involved in overseeing the score and there was no appeals process. People affected by this were not fully informed about what happened.

I would recommend a fantastic long read about this in the *Financial Times*, looking at the case of a young woman called Sarah Meredith. She required a liver transplant and was not being offered one. Through her and her family's research, she realised that this AI tool was being used and that even if different data points changed she would not get a different outcome from the tool. She was not, under any circumstances, going to be offered a liver transplant. She then realised that it was her age that was the obstructive factor, because she was a young woman and this tool had been set up in a way that favoured transplants for older people, even though it would be life-saving for her. She had to equip herself with knowledge and attempt to challenge that system.

The point is that it is very hard to get these examples but they are out there. AI is already being used with very significant effects and a high degree of opacity. We are not talking just about future risks; the harms of AI are happening today.

The Chair: Thank you, Silkie. That was a very clear exposition for the committee. If other examples come to light that you are aware of, do share them with us as well. tèmítópé lasade-anderson, would you like to address this issue of what we know about the existing risks of AI?

tèmítópé lasade-anderson: At Glitch, our focal population as a charity, and who we advocate on behalf of, is black women and black gender expansive people as a specific racialised and gendered demographic who face historic and continuing marginalisation. The issue area we look at is racial and gender injustice as it pertains to the use of technology and power. That will be the framing of my comments today.

Harms are breaches of rights but, to give some context, we think about these in three different ways: as individual breaches of human rights, collective breaches of human rights and societal breaches. Individual breaches are when the interests of a person are wrongfully thwarted or contravened. Collective breaches are when a demographic or specific group of people are disproportionately impacted. Societal breaches, as we frame them, are a contravention of societal interests which can impact democracy. In this case, the use of AI in the public sector is a particular threat.

Right now, I want to focus on an individual case which is very pervasive and predominantly around individual and private sector use of AI. There are numerous generative AI tools which allow for the proliferation of non-consensual intimate-image-based abuse, what is known as deepfake intimate-image-based abuse or so-called revenge porn. These AI tools and apps allow someone to use a tool to denudify a photo of someone or create sexually explicit imagery of them. This can then be shared online without consent; in relation to intimate partner violence, it can be used to threaten or coerce a victim or survivor.

As you may know, there have been several high-profile cases of this. The most recent was last January, when sexually explicit AI-generated deepfake images of the musician Taylor Swift were created and shared online. Research in 2024 from #MyImageMyChoice, an advocacy group, found that, in the previous year, 98% of all deepfakes found online were sexually explicit and 99% depicted women. That was a 17,000% increase in new deepfake images that had been posted online from 2019 to 2023.

Concerningly for us, deepfake intimate-image-based abuse exploits not only gender vulnerabilities in terms of who is disproportionately impacted but racial stereotypes, which leads to a unique and severe form of victimisation for black and other racialised women. This means that deepfakes also result in the compounding of existing structural and systemic racism and sexism that impacts victims' personal and professional lives.

Glitch and others, including professors and legal scholars, argue that deepfake intimate image abuse fundamentally constitutes a violation of human rights for victim-survivors as per the Human Rights Act. The impacts of these violations can be far-reaching and long-lasting. As Professor Clare McGlynn and Professor Erika Rackley have found, there are individual harms of physical and mental illness, together with the loss of dignity, privacy and sexual autonomy, which combine to constitute a form of cultural harm that directly impacts individuals, as well as society as a whole in terms of misogyny and gender-based violence.

As described by the Victims' Minister, Alex Davies-Jones, intimate image abuse is a deeply degrading and misogynistic form of abuse. Victim-survivors describe it as a shattering social rupture or a violation, and it can devastate all aspects of their lives. We say that this degrading abuse constitutes a violation of Article 3 of the Human Rights Act, which secures the right not to be subjected to torture or to inhumane or degrading treatment or punishment. Intimate-image-based abuse also constitutes a violation of Article 8 of the Human Rights Act, which is around the right to privacy. Existing case law—Volodina v Russia in 2021—shows that the European Court of Human Rights accepted that cyber violence, including image-based sexual abuse such as non-consensual intimate-image-based abuse, constitutes a violation of Article 8.

It is important to say that, while there is existing legislation in the Online Safety Act that criminalises the sharing of this abuse, this specific type of AI-facilitated abuse does not have any form of redress or remedy. Earlier this year, the Women and Equalities Committee put forward a clear recommendation in its fourth report of the Session that, to tackle this abuse, we need to amend the eligibility criteria of the criminal justice compensation scheme to bring claims from victims of sexual offences perpetrated online within its scope, which the Government rejected. The Government also rejected the recommendation to introduce a swift, inexpensive statutory civil process, as has been established in other regions, such as British Columbia in Canada, on the basis that this function is available through the Online Safety Act, but, as the previous panel mentioned, individual redress through these legislative options is very costly, very difficult to access and often does not provide an individual route.

The Chair: Thank you very much, tèmítópé. My colleague Sir Desmond Swayne will now come in about bias and discrimination.

Sir Desmond Swayne: You have answered my question comprehensively already.

The Chair: Thank you for that. We will hear from Mr Ruiz next and then we will turn to a question on enhanced surveillance and facial recognition.

Javier Ruiz Diaz: Thank you very much. I am representing Amnesty International UK; we are the UK section of our global movement. We want to make clear that the examples we have found in the UK are only a small part of the evidence that the whole of Amnesty International has

collected. There is a lot of work in the Netherlands, for example, and in many other countries, where we have identified digital algorithmic artificial intelligence harms.

Many of our findings are quite similar or potentially overlap, so I am not going to repeat what my colleagues on the panel have said, but we are finding that it is very hard to identify artificial intelligence harm when we look at harms around algorithms and technology. One of the problems is the inevitability and exceptionalism in the way this technology is framed: we think that it is driving a sense that we have somehow to go over human rights: this force is unstoppable, so we have to move aside. That is a particular problem in itself.

We are not saying that this is just normal technology, although to a point it is a question of scale. There is a question of higher degrees of autonomy and control, but there are also new things that extend and maybe transform. For example, there are questions around the impact of geoguessing, which is where you can show a picture to an LLM and, in some cases, find out where it was taken. The harm is that this could lead to the identification of a person reporting anonymised evidence of human rights abuses. However, it could also lead to identification of where the abuse took place, so it works both ways.

We have submitted our report on two areas that we work on in the UK: predictive policing and the automation of social security. I am not going to go through the report in detail and reproduce what we put in writing, but we are seeing some things that are fundamentally new and which could potentially amplify the problems with AI. We would like to know how these things work but we are not able to because of the lack of information.

Something that is potentially new is the statistical nature of AI. It does not create certainty and there is an acceptance of risk. Under this system, some people will be wrongly identified by facial recognition. We know that this is going to happen, so the question is, "Do we accept this?" We can try to hide it, or go one way or another, but there is a fundamental issue about accepting a certain level of risk while, at the same time, trying to present these systems as determinative.

For instance, let us say you are a criminal: we have identified what you are doing and we are taking you to a police station to interrogate you. We have a degree of confidence, but as we tweak the systems to provide more confidence we will increase the number of false positives or negatives. This statistical way of thinking is something that we have not really caught up with and it is not reflected in the way that policy is made. Eligibility criteria and risk profiling has an effect on people when they get their benefits cut or they are stopped by the police, but it does not say "we think this is happening"; it says "we know this is happening". That is problem number one.

The other point is that the human interface referred to by the previous panel—the anthropomorphism of these technologies—is creating other

problems. The replacement of human interaction in the benefits system is creating situations where vulnerable people or those with complex cases are unable to deal with chatbots or avatars. The intrinsic limitations of these systems in dealing with certain complexities or processes is not being acknowledged.

Finally, these technologies are not neutral. With predictive policing, we are reproducing the bias of racist policing and building it into an automated system. This is something that we need to deal with, but we find it hard. It is a chicken/egg situation: unless we change the practice and generate new data we are not going to have an automated system that is not racist, because the reality is that you are feeding it racist data. Those things are very hard to solve through proposals that are seen elsewhere as potential solutions.

When it comes to analysing people as data, one overall concern that we have is the concept of social scoring, which is to understand someone on the basis of their profile rather than evidence. This is being challenged in other places like the European Union, and we think it is at risk of happening in the UK.

Overall, we are worried that the UK may become a dumping ground for technologies that are not allowed in nearby jurisdictions because other places are putting in red lines. We really need to start taking action.

The Chair: That is very helpful Mr Ruiz, and thank you to all four panellists for raising the curtain for us. The committee is very concerned about one thing you mentioned: the use of AI for enhanced facial recognition. My colleague Lord Dholakia has a question about this for Silkie Carlo.

Q29 **Lord Dholakia:** Recent evidence we have gathered on this is the fact that the sample of some minority communities in this country is so small that the type of evidence you can obtain will not be very relevant. What is being done about that particular matter?

Silkie Carlo: Do you mean about the bias in the facial recognition algorithm itself?

Lord Dholakia: Yes.

Silkie Carlo: That is something we have repeatedly raised concerns about and, as a result, the Metropolitan Police undertook a report with the National Physical Laboratory. That report claimed that when it used the algorithm at a certain threshold, discrimination against people of colour was now not statistically significant. The Metropolitan Police would probably give its own account of that report, but our view is that it is not satisfactory to say, "If we tweak the settings, the discrimination will not reach a statistically significant level". It also relates to scale. Our concern is that the testing was carried out in a very small trial operated by the Metropolitan Police. If you use this nationwide and constantly then even a small error margin quite quickly becomes tens of thousands of people. If

this is happening in a discriminatory way, then there will be further human rights and legal concerns.

It is a very well-known and very well-documented problem with facial recognition algorithms. For sure, they are getting more accurate. I am afraid to say that the lack of regulatory action in this area means that some people have paid the price for almost 10 years now since the police started using this. It was absolutely awful when they first started using it. I have monitored police using live facial recognition on the ground for many years. I have witnessed women being flagged as wanted men, and people of different ethnicities. It was really, really poor. There have been lots of misidentifications over the years. For example, in the early use at Notting Hill Carnival, there was one year where there were 100 or 101 flags and 99 of those were misidentifications.

The Chair: That is an important point about where we are in terms of regulation and law. Perhaps I can bring tèmítópé in here as well, because I am interested about existing protections against discrimination and bias, whether at the Notting Hill Carnival or anywhere else for that matter, and how those biases can be addressed effectively. Do existing laws work at the moment? If not, how do we enhance them? Perhaps Silkie could continue for a moment on that, then we will go to tèmítópé and see if anybody else wants to chip in.

Silkie Carlo: The law as it stands has not meaningfully restricted the use of facial recognition technology in this country. As a result, NGOs operating on small budgets, like Big Brother Watch, are afraid they will become the regulators of last resort. Big Brother Watch is now backing a judicial review of the Metropolitan Police's use of live facial recognition technology. That is being brought by an anti-knife crime campaigner called Shaun Thompson, who I witnessed being misidentified at London Bridge. He was held by police for about half an hour, treated quite badly and very nearly arrested, so he and I are now undertaking a joint judicial review.

The Chair: I am just being advised that we cannot discuss live criminal cases and I am slightly anxious that we should not stray into that territory.

Silkie Carlo: It is not a criminal case; it is a judicial review. Do you mean any litigation whatsoever?

The Chair: Any live case, whether it is a judicial review or a case before the courts. We cannot discuss those details in open committee in this way.

Silkie Carlo: That is understood. I will also refer to Liberty supporting the challenge by Ed Bridges. A pattern has emerged where NGOs have to try to uphold convention rights in the area of AI, in particular around facial recognition, by taking up very lengthy and costly judicial reviews. As was mentioned by the witnesses you heard from before, the Bridges challenge was five years ago, and only now are the Government talking

about consulting on a legal framework for facial recognition. The Government have been asleep at the wheel, to be honest, and we are a bit behind the times. There is no meaningful restriction in law on the use of live facial recognition.

Another thing which is desperately underdiscussed is the fact that the passport database and the immigration database have been turned into mass facial recognition databases, in a way that I do not think anyone would expect should be permitted given our Article 8 privacy rights. There are other uses with facial recognition as well. NGOs are having to undertake litigation in order to make human rights matter.

The Chair: I was struck by your phrase that you have become the regulators of last resort. That is placing an enormous burden on non-governmental organisations. tèmítópé, do you want to come in on this point? We will then perhaps bring in the other two members of the panel.

tèmítópé lasade-anderson: Looking at it from the perspective of an individual, the Manchester police force recently deployed facial recognition without giving prior information or seeking consent. There is no real way an individual can opt out if they are walking down the street and there is a police van. I am sure my colleagues here will know that there have been cases where people have said, "Hey, I want to opt out" and the police in the van then question them and suggest that they cannot opt out. Individuals are unable to choose not to have their faces scanned in this way.

As the panel mentioned, there is accidental legislation around AI use but there is immense fragmentation and insufficiency around specific issues of discrimination and bias. There are existing rights-based protections, which come through the Equality Act, the Data Protection Act, the GDPR, the Human Rights Act, and the public sector equality duty, but these were not designed to govern AI, whether on issues of opacity, scale or autonomy issues or whether it is upstream or downstream. They also rely on individuals and civil society, or regulators that are underresourced and ill-equipped with the skills to challenge any of these harms ex post—we are not even talking about ex ante or preventative measures. That is a fairly succinct explanation of where we are today.

The Chair: That is very helpful, thank you. So our laws are laws by accident, they are fragmented, and there are huge gaps. Two of my colleagues have been so provoked by what you have said that they want to ask a supplementary before I come back to the other two panellists.

Q30 **Afzal Khan:** I just want to ask a question on facial recognition technology. Could it be argued that the technology is progressing and improving, and therefore that some errors you may be getting now will be ironed out with improvements? The supplementary question I want to ask is that I have a passport which is recognised by AI. What I find interesting is that when I have gone to other places there has been no problem, but when I come back to Britain the AI never recognises me, so I have to go through the manual system. Is that an error, or has it been

loaded in such a way that picks up certain types of individuals?

The Chair: Let us take Lord Sewell as well, because that is a really interesting question.

Lord Sewell of Sanderstead: I was just thinking about being positive about AI, as none of you is giving it to us as something that could help. We know a proportional demographic involved in knife crime, particularly in London, and the police already use profiling to try to find perpetrators. The human side of it is already involved, using human intelligence rather than artificial intelligence. Would it not be better if we had technology with facial recognition that was sharp enough to recognise criminals? In other words, why not argue for more AI? Why argue against it? AI would be more likely to find the perpetrator accurately, rather than someone at the behest of using human instinct. I want to push you to the other side to say, on a human rights level, would it not be better to argue for using the tool around racial bias?

The Chair: We have some excellent questions there. I want to make sure all our panellists get in, so Ms Pirlot de Corbion, you start then we will go to Mr Ruiz, and then to the other two.

Alex Pirlot de Corbion: Thank you very much for giving us the chance to weigh in on this discussion. Coming back to your question, we are veering towards a tool that would rely on accuracy also, and that the tool would work well. Currently, given the example that Silkie shared and other examples that we are giving, we know that it is not working properly and, on the scale that is currently being deployed, it risks becoming—if it has not already—a tool of mass surveillance.

There are two points that are particularly concerning about the use of facial recognition. First is the scale of surveillance that is possible. It is possible not only to identify but to authenticate individuals on a mass scale, which means that it is no longer targeted to specific individuals but is directed at groups of people—for example, at protests.

Secondly, with regard to the scope of facial recognition technology, Silkie mentioned the live use of FRT but there is also retrospective FRT, which means non-live or static. It is no longer just about being seen on a CCTV camera; there is the ability to identify and verify a person's identity in real time, which renders this technology even more invasive than the current deployment of CCTV cameras.

The third point I wanted to make before handing over is that, in terms of the harms that are experienced by people with the deployment of such technologies, there is a clear distinction between being monitored anonymously versus being identified. The protections that come with that anonymity are really important in our democratic society. Otherwise, there is a chilling effect on a person's behaviour: whether they choose to attend a protest, or to commute on a certain path where there might be more facial recognition. The aspect of protection that comes with anonymity is really important and linked to our democratic values.

Javier Ruiz Diaz: Following on from that, the question of data to improve the quality of systems was debated quite a lot in the European Union at the time of the passing of the AI Act. Of course, we are solving the discrimination problem by adding more data, and that can be a good thing in some contexts. For Amnesty International as a whole, the formal position is that facial recognition as a tool of mass surveillance is fundamentally incompatible with human rights and just cannot happen. That is the bottom line. We can then question: if it is not mass surveillance, when is it not? Then we can look at discrimination. Of course, you can compound human rights problems: you can have a tool for mass surveillance that is also discriminatory against certain people.

In that context, one problematic issue is indirect discrimination. AI is very good at making connections and inferences between things that are not immediately obvious. The problem we find is that you are not targeted because you are X; you are targeted because the machine has found something else. An example is the young lady who was unable to find out why she was being refused a liver transplant. It could be something like your age, or it could be something else as a proxy for your age. That is something that we find particularly problematic. Again, some technologies and certain intentions—not even just the technology itself—seem to be fundamentally incompatible with human rights.

In terms of other problems, we find that current protections around data protection, access and transparency do not seem to work at all. It is pretty much impossible to find out in detail how your data is being processed. We think we definitely need new safeguards to build on the basics we already have.

The Chair: So it is fundamentally incompatible with human rights?

Javier Ruiz Diaz: For us, mass facial recognition is fundamentally incompatible. When it becomes mass surveillance, we can start working back from that to see if there are some circumstances where it is not, but that is the principle and the common position. I will be happy to share the documents from the international organisation.

The Chair: Thank you. Mr Khan will want an answer to his question about passports, but I would also like to ask tèmítópé if bias in AI systems can ever be addressed effectively. If it can, what should we be doing to try to ameliorate the worst effects of bias?

tèmítópé lasade-anderson: That is a tricky one. I am not a technologist but there is this idea that you can debias a dataset. Looking at facial recognition technology, which is sometimes used on mass databases, the gangs matrix for example is one database which has an overrepresentation of people of colour, in particular black men. There is this argument that if you have a database that has a proportionate number of images of all different people and you use that data to train the algorithm then you have an algorithm that is debiased. That does not necessarily eradicate the issue, because if you still have a small percentage of people who are misidentified or have error rates that can

still be a large number of people overall. Similarly to Amnesty International, Glitch would say that any mass surveillance tool—for example facial recognition or digital ID—would be fundamentally incompatible with human rights.

The issue with discrimination and bias is also around existing structural and systemic issues. Silkie mentioned postcodes, which are proxies for different kinds of demographics. When you have a postcode that is for a lower socioeconomic area, for example, which tends to have more engagement with policing, then you already have a notion that there is an issue with this area. When you apply that database into an automated tool, it is likely to suggest that people from this area need to be surveilled in a predictive policing sense. In the case of recidivism, it will be someone who has a number of previous engagements with the criminal justice system.

It is not just a matter of having a debiased dataset. That is simply not good enough. The fundamental issue with AI is that it can result in increased discriminatory outcomes, so it is really important to ensure that there are guardrails or red lines. Although the risk-based approach of the EU Act was rightly criticised, there need to be some red lines as to where AI must not be involved. In terms of generative AI in particular, there is a difference between individual issues and the private sector's use of AI; individuals have some choice about which tools they may wish to use or engage with. When it comes to the public sector, there is no choice: you do not get to decide how you engage with healthcare, the criminal justice system or the education system, where automated decision-making tools are being used today in the UK. So, because of the public sector—

The Chair: I am sorry to interrupt, but should there be different regulations for the public and private sectors?

tèmítópé lasade-anderson: Yes, you would need that. As our previous colleague mentioned, it depends on the sector, the issue area and how that regulation would apply, but there cannot be one flattened approach to AI regulation across the board.

Silkie Carlo: I appreciate these follow-up questions, and I do not want to sound as though I am arguing for or against AI in general. If we can use technology to improve policing and the criminal justice system then that is fantastic. Obviously, as a group that is interested in protecting and promoting human rights, we focus on the risks because they are so prevalent at the moment. It is a fairly consensus position that most human rights groups view the police's use of live facial recognition as a breach of individuals' rights. You have heard Amnesty's position on that. The Equality and Human Rights Commission and 130 human rights groups have called for a stop to live facial recognition being used in the UK based on the information that we have seen.

I do not think that accuracy is the central issue with technology like live facial recognition; in many cases with AI, I do not think the key issue is accuracy whatsoever. It has been an issue and it means that lots of

people suffered real human rights harms and discrimination as a result in the early years of deployment, but you are absolutely right: the Metropolitan Police's algorithm has got a lot better. It is absolutely true that it is making far fewer misidentifications than it used to, but there are a host of other issues.

The first is the lack of regulation. Even the police are saying that they are making up their own policies as there are no clear guidelines on how to use this. Essentially, they are having to make it up as they go along. There is very little accountability and transparency, not only with facial recognition but other types of AI. In many of these systems, if you are flagged, you may not know why. That contradicts very basic British democratic values about transparency and fairness. It also means there is a chilling effect. If you feel that surveillance systems or other AI systems used by government departments could be watching you, judging you, flagging you or making decisions about you, and you do not know why and you cannot challenge it, that cuts through our very long-standing values about fairness and what we expect from public authorities.

The other risk inherent with something like live facial recognition and similar AI systems is the scale and speed at which they work. You will hear from other people that this is a great benefit of AI and of facial recognition, but actually it means that it lends itself to mass surveillance and to changing the fabric of the country that we live in. You cannot imagine a situation at, say, Westminster station in which a police officer took a fingerprint from every single person who came through that station just to check, and said, "We'll throw it away afterwards, don't worry". The fact that data can be processed at scale and very quickly does not make it any less intrusive; it makes it more intrusive because it is happening at an enormous scale.

Earlier, the witnesses were asking, "Have we seen this film before?" I think we have. Thinking about dystopian mass surveillance that can be used in that way, and given that we are on the brink of a national mandatory ID system, which is a facial biometric system, we have to think about those risks very carefully. Live facial recognition and AI mass surveillance being used in this way can spiral out of control and be used in many other ways by public authorities in the blink of an eye.

The Chair: Thank you very much. I know that Mr Sobel wants to ask about mitigation, so that is a very good point to come in.

Q31 **Alex Sobel:** I was just down the corridor yesterday with the Secretaries of State for DCMS and for DSIT, discussing AI in the creative industry sector. We focused on the ingestion of data by creators and—less so—the ability of individual users to opt in or opt out of AI materials on a streaming platform.

I am a graduate in information systems; the whole-systems life cycle in AI or any information product is much bigger than just data in, data out; there is the design stage, the development stage, modelling, then the data stages. From your experience, do you think that threats to data

security and privacy occur all the way through that life cycle, or are there particular stages that we should look? If we are going to legislate and regulate this, we have to know which end of the telescope we are looking from. Are we looking right through the telescope from the beginning until the point at which it sees? That was a terrible analogy, but you get what I mean. What measures could we undertake to regulate AI to prevent or mitigate threats from it in terms of the development life cycle?

Alex Pirlot de Corbion: I will respond to your point on the data life cycle. As you have seen already from the examples that were given by the first panel, and the second panel as well, there is a wide range of AI systems. They all function in different ways in different contexts, but one common denominator of any AI system is the data. The way we see it at PI is that AI systems are exploiting that data. The short answer to your question is that there need to be safeguards at every stage of the life cycle, as Dr Kiden mentioned earlier, as there are potential threats to privacy and data security because it is always about the data.

To give you a longer answer that you might appreciate today, there are three moments in the data life cycle that are key. The first is data collection and data gathering—the input data, as you mentioned. How much is there? Where does it come from? What kind of data is it, and about whom? This is a key concern that has been raised, for example, in the development of large language models built by companies like OpenAI, Meta, xAI, and Mistral. What data is being processed by these companies and where is it coming from?

Similarly, in the example I gave earlier, Clearview AI was scraping photos from social media and from the internet. Again, we should be questioning where that data comes from. That is why it is essential to scrutinise the sources of data and how it was obtained in the first place. Is it being repurposed for a purpose for which it was never intended? Is that new purpose proportionate, necessary and lawful? The failure to get this right at the onset triggers privacy and data security concerns even before an AI system has been deployed.

Secondly, the stage in the data life cycle that I wanted to elaborate on is the processing activities to which the data that has been collected will be subject to. That really depends on the design of the AI system. It might have a set of parameters or it might not. That really depends on the design level. At this stage, we are particularly concerned about how the data that was collected will be exploited and analysed with a particular purpose in mind. To give some examples we have seen here in the UK, take Palantir's role in the NHS: what processing activities was the data given to Palantir subjected to? Similarly, we have been working very closely with unions and workers for many years now. We have seen how data and algorithms from workers in the gig economy are being used by employers to rate them, to decide how much they should be paid and even whether they should be given a job. Decisions are being made at that level.

We have talked a lot about transparency around AI systems. You might have heard of the black box of an AI system. What is really concerning at that particular stage of the life cycle is that even the designers or users of an AI system might not be able to explain what those processing activities are. It is a real challenge to know what is being done to that data and to make sure that it is not being used to undermine the integrity of the data or to undermine people's rights.

The last point around the data life cycle is when it comes to its application and use. I gave a few examples about the Home Office appearing to use algorithmic tools for immigration purposes, including deciding whether a person can remain in the UK, whether they face deportation, and whether they will be subject to GPS tagging or have their freedoms restricted through detention. We have heard similar examples from the DWP about deciding who deserves to have access to benefit and welfare support. These tools are really impacting people's dignity and autonomy.

When we look at this level of the data life cycle, we need to ask whether the AI system is being used to solve a problem that was identified and whether, in solving that problem, it creates more harms and risks for people. It is possible to mitigate all the threats I just mentioned, but at every stage where decisions are being made the starting point should be: what is the data? Is it being protected and how is it impacting people? What mitigation measures can be implemented along the way? I will stop there, but I would be happy to elaborate.

The Chair: It is very helpful to give us those guardrails. If people want to add to that subsequently then please feel free to write to us.

I know that Mr Ruiz wants to come in, but we have a question for you now, so maybe you can add it to your answer. We are going to go back to Mr Khan; I do not know whether you are able to tell him what remedy he will get, or what happened to his passport, but what can we do about things going wrong? How can we put right any of these injustices that you have been describing? That is your point is it not, Mr Khan?

Q32 **Afzal Khan:** That is right. How easy is it for people affected by AI to seek redress? Does the law offer appropriate remedies for affected individuals? The supplementary to that is: is there a case for a dedicated regulatory body or ombudsman to address AI harms?

The Chair: Mr Ruiz, this one is for you. In a moment we are going to come to Tom Gordon MP, who is going to ask you about your recommendations, which is the trickiest question of them all. That will be a chance for you to summarise any points you want to make to us about practical things we can do.

Javier Ruiz Diaz: I will combine the two responses. Something we need to remedy, as we have said before, is a lack of a full understanding of the problem. Picking up on the arguments that have been made, if we are deploying systems in the public sector, a useful remedy would be to

discuss where the resources are coming from: is this substituting some other functions that are already being provided? We should also look at the cost—even the opportunity cost—of going down this path.

One fundamental pre-remedy would be participation and including participatory processes in the design when deciding who is going to benefit: is the organisation or individual people receiving the benefit? You asked for international examples before. We are seeing good examples in Taiwan right now, including a good participatory process for developing legislation.

On remedies, we have found that the situation is not good at all. Starting from transparency and the lack of information from a data protection point of view, as we said before, it is very hard to find out how decisions are made or even what data was used to make them. From the point of view of equalities, there is the idea, "Is this fair?" The result may be fair in one sense, but from the point of view of the individual it is not. When it comes to indirect characteristics, unpacking the levels of potential discrimination is proving very hard, so we need additional safeguards and to create specific rights to obtain more information about decisions. It is not just a matter of the human in the loop; we need to take apart the decision process and not just have the human as a rubber stamp. That is something that we need to be very careful about.

In terms of a particular regulator, I was in this room a few years ago during an inquiry with people from Oxford and Cambridge universities looking at a similar question about whether we should have one regulator. At the time, the consensus from many people in civil society was that it was better to get the regulators to do their job properly. I am not 100% sure about our position—we have not got to that point—but we have to be careful about putting all our eggs in regulators. Looking at the role of Ofcom in the Online Safety Act, it has to employ a huge number of people. It has been given the role of making decisions that probably should be made by someone else, but it is forced into juggling questions such as, "Is this a freedom of expression problem?", and trying to create rules. It may be unfair to expect regulators to take on the issue of AI and the question "Is this a fair decision?"

The Chair: Mr Khan's question is specifically about whether there should be a dedicated, specific regulator that does nothing else other than this.

Javier Ruiz Diaz: Look at how Ofcom has to deal with these problems: what kind of regulator can deal with everything that we have talked about, from facial recognition to benefits? As I said, we are not sure, but if there was to be a regulator it would require huge resources. Also, I do not think we can talk about regulation without looking at the state of water and rivers. I am a keen water sports person but I dread to paddleboard in rivers because it is dangerous. Water regulation has been an absolute failure. We have to be careful about our expectations. Whether there is one regulator or many, we need to be very careful about resources, not giving the regulator too much responsibility and providing very clear guidelines so we are not passing the ball on to the

regulator. We must also make sure that there are other elements of accountability for the regulators.

The Chair: Thank you, that is very helpful. The last question is going to Tom Gordon.

Q33 **Tom Gordon:** Thank you, chair. At the end of this inquiry, we will put together a report to the Government and there will be recommendations in that. We have heard a breadth of different issues and potential solutions, but what would be your top one or two recommendations to the Government to tackle this issue?

Javier Ruiz Diaz: For us, as well as what we said before, one would definitely be to create new specific protections. Then we need clear red lines. As we said, there are some things that we think should not happen—maybe facial recognition and social scoring. This is not to say that there should not be research on technologies, but there should be some very clear red lines in the application of those technologies about things that should not happen.

There are some other areas that are more about the detail of how you go about it. Something that is particularly problematic in the accountability of technical systems is that we encounter issues around commercial confidentiality. This should be tackled. I work on digital trade and, unfortunately, we are signing up to trade agreements that specifically ban access to source codes in the name of protecting the intellectual property of companies. Removing any barriers to accessing technology that makes the black box a little less opaque would be useful. There are obviously many more details that we could process, but those are definitely the most important.

If the framework convention of the Council of Europe goes ahead and does not get sabotaged along the way by greater powers, it will force the UK to take some form of action. It would be good to take the opportunity to expand rather than do the minimum, particularly when it comes to trying to add something on the role of private companies that may provide systems. In supply chains, it is very hard to decide, "Is this providing a service or is the company buying a product that is providing the service?" Trying to tackle commercial confidentiality would be quite important.

tèmítópé lasade-anderson: Not all AI tools carry equal risk when implemented; the risk is relational. Biometric and surveillance tech present severe direct threats. For Glitch, the main thing is really just to plus one what Javier said: there need to be red lines and hard lines as to where we should not see AI being used. For us, that would be in the public sector in particular, as it is very difficult to see under what conditions of transparency AI could continue. We would also agree that there needs to be a new regulatory and legislative framework that particularly looks at AI, and these need to be rights compliant. In particular, there need to be options to challenge a decision. That is not only about transparency but about redress. I will leave it there.

Silkie Carlo: It would be encouraging to see sufficient weight being put on AI risks within Governments so that, if they continue with this very optimistic and very commercially focused outlook, it has to be paired with regulations. Otherwise, we will continue to see litigation and high-profile mistakes being made.

It is a concern that so far in this Parliament we have seen the Data (Use and Access) Act, which has basically decimated individuals' rights and protections against automated decision-making. Whereas we had an assumed prohibition on automated decisions being made about people where the effects would be significant—you have heard already about the harms caused, so that was not working very well—the Government have basically liberalised automated decision-making for all purposes. It is really serious and has not been given enough attention. I am afraid we are going to see many more harms as a result. I do not think that was actually a Labour policy or a Conservative policy; it was a DSIT policy. It came from civil servants and everyone has basically been too asleep at the wheel to do anything about it.

At this stage, we need something like an EU-style AI Act or a digital Bill of rights: a specific piece of legislation that takes a risk-based approach to artificial intelligence and can deal with many of the different areas that this committee has heard about, that can put red lines in, as has already happened in Europe, and give people the confidence to use AI in the places where it will benefit people. But, at this stage, we certainly need overarching legislation to deal with the most serious risks. My understanding is that AI legislation is coming, but it is growth focused and not addressing the risks.

The Chair: That is really helpful. Thank you for pointing us to some guardrails and restrictions that might be introduced, and models that we can have a look at from elsewhere. The last word goes to Alex Pirlot de Corbion. You have given us great evidence throughout the whole of the hearing. Please give us those two things that you would want us to take away.

Alex Pirlot de Corbion: At the risk of repeating what my colleagues have said, the first one is around accountability. As we have heard today, the current regulatory framework is fragmented and not working well. We need to see the UK Government taking steps to ensure that a strong, legally binding instrument is in place, because the ethical guidelines we have seen so far are not sufficient. White Papers are not putting on a lot of pressure or giving incentives. Self-regulation by companies is ineffective. An important element around that accountability piece is that it needs to be founded on the UK's existing national and international human rights obligation, rather than a focus on innovation and growth, as Silkie just mentioned.

Importantly, as has come up a lot in today's hearings, it needs to apply to both the private sector and the public sector, and it should apply to both personal and non-personal data. We have not touched on that issue very much today, but a lot of the data that is going into AI systems is not

personal data but can reveal personal data. That is an important element to bring in here.

The second one is around transparency. We have talked about this a lot today, but the UK public cannot be kept in the dark when it comes to their data and how it is being processed through AI systems as they go about their everyday lives. The UK Government, their partners and the private sector need to be forthcoming about how they are designing and testing AI systems. We have tried freedom of information requests and we are not getting that information. Unless we know what is currently going wrong, we will not be able to remedy it.

My final point—if I may take the liberty of adding a third little cheeky one on the end—is that we are really concerned that we are creating a dependency on big tech companies. Unless this comes in very quickly we might set ourselves up for an irreversible dependency, and we will be dependent on big tech companies to make life-changing decisions about people and about their rights.

The Chair: Anyone who has watched our proceedings today will know that these are awesome issues that you have been touching on, which are phenomenally important. We are extremely grateful to you all for your generosity in sharing your expertise with us. As I mentioned earlier, if there are points you want to follow up on, please feel free to be in touch with our committee specialist and let us know what those are. We will certainly consider them when we come to our recommendations.

At the end of the previous panel's remarks, I quoted Stephen Hawking and Geoffrey Hinton. Given what Silkie Carlo said earlier about dystopia and the worries that people have around these issues, I remind those who are watching our proceedings today of the seriousness of what is at stake in terms of their human rights. Yuval Noah Harari has said that, throughout history, humans have been tempted to pursue powers they cannot handle; with AI, we are summoning a power you cannot control. The dangers are self-evident. What we do about them is a matter to which this committee is now going to give a lot of attention, and I hope that when we produce our report, it is something we can be proud of and that you will feel was worth your time in giving us your evidence today. Thank you very much for being here. With those words, I will now close our proceedings.