HOUSE OF LORDS

Justice and Home Affairs Committee

Corrected oral evidence: New technologies and the application of the law

Tuesday 12 October 2021

10.30 am

Watch the meeting

Members present: Baroness Hamwee (The Chair); Baroness Chakrabarti; Lord Dholakia; Baroness Hallett; Baroness Pidding; Baroness Primarolo; Lord Ricketts; Baroness Sanderson of Welton; Baroness Shackleton of Belgravia.

Evidence Session No. 4 Heard in Public Questions 52 - 67

Witnesses

I: Professor Karen Yeung, Interdisciplinary Professorial Fellow in Law, Ethics and Informatics, Birmingham Law School; Peter Dawson, Director, Prison Reform Trust; Silkie Carlo, Director, Big Brother Watch.

USE OF THE TRANSCRIPT

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

Examination of witnesses

Professor Karen Yeung, Peter Dawson and Silkie Carlo.

Q52 **The Chair:** Good morning, everyone, and thanks particularly to our witnesses for coming today—Professor Yeung, Silkie Carlo and Peter Dawson.

We all have a number of questions to ask you. I hope that there will be time at the end to ask some unscheduled questions, as it were, from what comes out of the evidence. Please respond to one another if you feel that that is appropriate. We are formal, but that does not mean to say we are stodgy.

Welcome to Professor Karen Yeung, who is the Interdisciplinary Professorial Fellow in Law, Ethics and Informatics at Birmingham Law School. I am baffled and impressed. Peter Dawson is director of the Prison Reform Trust and Silkie Carlo is the director of Big Brother Watch.

Mr Dawson and Ms Carlo, forgive me for being formal but I think I probably should be. You have both given us written evidence, each focusing on a particular tool and how each is used. Could I ask you both, because we have all seen the written evidence, to give us a brief introduction? We will then, I hope, go on to more searching questions. Mr Dawson, first, on the security categorisation framework.

Peter Dawson: Thank you very much for the opportunity to give evidence.

The Prison Service is not known for being at the cutting edge of anything very much, so I would not claim that the tool we have described is particularly sophisticated in artificial intelligence, but it is an interesting example. There are people round the table who will know this very well, but for those who do not, prisons have security categories that relate to how difficult it should be to escape from them and how difficult it is to cause trouble within them. Prisoners are categorised at the start of their sentence, and periodically thereafter, to make sure that they are not held in conditions that are more secure than they need to be. The gap between the most secure and the least secure is huge. At one end, escape should be virtually impossible in a prison such as Belmarsh. In an open prison, escape is not an issue at all, and people can, essentially, leave if they choose to, although they suffer consequences for doing so.

The sort of prison they are held in makes a big difference to the prisoner's way of life. If they are serving an indeterminate sentence, and their release depends on the Parole Board—so that a decision is taken about their future risk—the prisons in which they are held during their sentence can have a huge impact on their chances of being released at the earliest opportunity. The stakes are very high.

Historically, categorising prisoners has been a paper-based process done in individual prisons, with a great deal of human judgment applied to it. Because it has been paper-based and purely prison based, it is very difficult to interrogate, and actually very difficult for the Prison Service corporately to manage, which creates all sorts of difficulties for it at a strategic level. It is also difficult for prisoners to challenge.

The change that has been introduced, simply, is that the process has been digitised at the first stage. The first categorisation that a prisoner gets is informed by a digital process that has an algorithm within it. Someone in the prison is then allowed to overrule that algorithm.

We put that in as an example, in some ways because it is an example of good practice. We feel we have been well consulted. There was a pilot scheme. The digitisation process means that a great deal more information has been extracted about its impact, but we think it also shows some of the risks that attend the process of making a process digital and some of the responsibilities that probably now fall to the Prison Service in seeing it through and assessing its impact.

The Chair: Is it kept up to date with further information being fed in? Secondly, you framed it in terms of risk of escape. Is it also risk of criminal activity within prison and so on? Would it extend across behaviours?

Peter Dawson: Yes, it is updated periodically. It is an interesting question as to the extent to which the first categorisation stays with someone and can be altered. It is about more than escape. It is about risk of disorder. It is certainly about risk of criminal activity within prisons. One of the motivating factors in updating the policy was to try to give greater weight to the risk of people committing crimes within prison and rather less weight to the offence that they originally committed and the length of their sentence.

The Chair: Ms Carlo, facial recognition.

Silkie Carlo: Yes. Thank you very much for inviting me to be here.

I submitted written evidence focusing on facial recognition: in particular, live facial recognition that is a biometric surveillance tool whereby a camera is put in a public place by the police—at least that is the context in which I am giving evidence—and members of the public are biometrically scanned as they walk past the camera and checked against the database, a watch list, in real time, producing alerts that the police can then follow up. The technology is very rapidly doing a biometric assessment of the individual’s face, which is sensitive personal data in data protection law. It is a bit like a border security check, except that it is in real time and in a public space.

We might also want to talk about retrospective facial recognition. Very recently, we have found that some police forces have started trialling that as well. It is more like a forensic capability, whereby someone’s image, which may be taken from a surveillance camera or public footage of some type, can then be checked, looking back at potentially massive police databases, social media databases—or anything. The possibilities are endless.

There are severe rights and legal issues with that. There are significant accuracy issues, particularly with sex and race, which I am sure we will come on to. In particular, South Wales Police and the Metropolitan Police have used live facial recognition. This was started with a great deal of secrecy. In fact, as you know, there has not really been any parliamentary discussion about it. It is something that has gone under the radar. There is really very little specific regulation of it at all.

Even more under the radar than that has been that a number of police forces are working with private companies to use facial recognition. I can talk a bit more about some of those partnerships.

The Chair: Thank you.

Q53 **Baroness Chakrabarti:** I should say at the outset that in the past I have had the pleasure of working with both Peter Dawson and Silkie Carlo, many years ago with Peter in the Home Office and some years ago with Silkie at Liberty.

I want to ask about the proportionality of the tools you have been discussing in tackling the needs they are addressed towards. How effective are they, and how else might the objectives be achieved? If it is your view that there is disproportionate interference, what would be a more proportionate method?

I would like to begin with Silkie Carlo, if I may. My understanding is that Big Brother Watch recommends prohibiting both live and forensic, retrospective, uses of facial recognition. I want to probe as to whether there are any circumstances at all in which you think that their use would be acceptable.

Silkie Carlo: First, on proportionality, I think live facial recognition is one of the most disproportionate and extreme surveillance technologies we have ever seen in Britain, certainly as an overt capability. It is already the case that without any parliamentary involvement, and in near-total absence of a regulatory regime, tens of millions of Britons have been subjected to facial recognition scans and will not even know about it. That seems to me a wholly unacceptable position. There are lots of people who have been wrongly stopped by the police and, in some cases, have had some quite traumatic incidences as a result.

To judge proportionality, you would have to look at what the police have gained by doing that. The Metropolitan Police, whom I have followed most closely, have achieved eleven positive matches in four or five years of operation. In that time, they have scanned tens, if not hundreds, of thousands of people on the streets of London and created a lot of distrust among communities, particularly by repeatedly deploying at the Notting Hill carnival—there is inevitably a racialised element to that—and in the borough of Newham, which is the most diverse borough of London.

Not only have they only had eleven true matches, but they have generated an awful lot of false positive matches. Their current rate over the entirety of their deployment is 93% false positive matches. I struggle to see a world in which that could be classed as anything near proportionate and it is why, unfortunately, in the light of all that evidence, the independent review by the University of Essex, which the Met commissioned, produced a damning report saying that it was quite likely that it could be found unlawful and that there were severe accuracy issues.

Unfortunately, in the light of all that evidence, the Met has committed to continuing to deploy.

In my view, the most prudent thing for Parliament to do to protect citizens' rights would be to put an immediate legislative ban on authorities deploying this technology, even if there are people who believe that it could have a future. I do not, but I think the case would need to be made. Some of the issues would need to be worked through, particularly the demographic accuracy biases. They are well documented internationally, and even in the Metropolitan Police's own evaluation of their technology, using their own police officers, they found that there was a statistically significant issue with facial matching and sex. Women were statistically less likely to have a true positive match with the technology.

On retrospective facial recognition, there is a complete lacuna in regulatory safeguards yet again. In terms of proportionality, at the moment it seems to be up to officers how they use it. Really serious rights are engaged, of course. It could be the right to a fair trial or the right to privacy. So many serious rights are engaged there. It seems bizarre that they are choosing to trial and deploy now. You could use it with body-worn cameras or CCTV. The possibilities are significant and endless. It could be social media databases. It goes as far as the imagination stretches.

However, you can see that it could theoretically have a forensic use, in the same way that DNA does. We would not accept everyone collecting all our DNA all the time, but if you had a clip from a serious crime scene and the only way you thought that you could match a suspect was by using retrospective facial recognition, you could envisage a scenario in which that might be a useful tool. I think an urgent moratorium on its use right now is needed, because none of the controls and safeguards are in place.

Baroness Chakrabarti: Could I turn to Professor Yeung with the same question on proportionality? You are the lawyer among our witnesses. I was concerned in Ms Carlo's testimony that this is not just a question of proportionality. There is a question of whether any of it is currently in accordance with the law for the purposes of Article 8 of the European Convention, given that it is happening without statutory basis and on a sort of do-it-yourself experimental basis by individual officers and forces.

Professor Karen Yeung: Thank you for this opportunity to share my thoughts with the committee. To give you some background, over the past two years I have been undertaking considerable and sustained research into the pilot testing of live facial recognition in several European countries, including England and Wales. Together with a postdoc, we have done a study of the publicly available information of the trials by the London Metropolitan Police service and South Wales Police. The Essex report, as my colleague mentioned, is very informative in providing some data about how those experiments or trials were conducted. One of the things that was notable about the London trials, unlike in Germany and in France, was that they used real, live suspects.

In other western European countries, they have been using volunteers to test the accuracy of the data. Then they have a full database of the people passing in front of the cameras. This has not been the case in London. They have been doing operational trials. They call them trials. Although they have purported to comply with data protection legislation, the documents that have been published pursuant to the requirements of the Data Protection Act 2018 are seriously deficient, in my view, in the extent to which they declared the operational purposes and the question of impact evaluation and proportionality.

When we get to the trials themselves, and the scientific methodology, we see very unrigorous procedures. There are changes every time a trial is conducted to tweak the methods that have been used. We do not have a stable and rigorous set of data on the basis of those experiments. We have an evaluation by the National Physical Laboratory, the UK's measurement service, which purports to provide an evaluation of the accuracy of the ten to eleven live tests conducted in London, but of course as they were using live suspects, we actually do not have data on accuracy because we do not know how many got away. We only have data on false positives. We have no data on false negatives, yet there are statements made in that report that the experiments have been demonstrably successful, with completely rigorous methodology as to how that conclusion was sustained. It really is a remarkable exercise in "Where on earth did you get this reasoning from?"

On top of all that, in those eleven trials 500,000 faces were scanned to produce nine to ten arrests. Many of those were individuals who were wanted for very trivial offences. All that meant real-time location tracking of many, many hundreds of thousands of British people going about their lawful business and not bothering anyone. This is a serious reversal of the presumption that people are entitled to go about their business in a lawful manner without being disturbed by the state. I completely support Silkie's view that this should be subject to very stringent regulations, if not an outright ban.

Silkie Carlo: Can I add to that? You might have detected a discrepancy between the figure I gave for true positives and the figure for arrests. That might give you a taste of just how extreme policing can be when supported by live facial recognition.

One of the deployments was at Remembrance Sunday. The individuals who were on the watch list were not wanted for arrest. It was a watch list of people, essentially, who had mental health problems. It was from the Fixated Threat Assessment Centre, which you might otherwise be familiar with. These are not people who are suspected of committing any criminal activity. They are not wanted for arrest. No one had thought about the implications of creating a watch list of people with mental health problems.

We also saw that in south Wales, where protestors were put on watch lists who were not wanted for arrest. They were put there for intelligence reasons. The scope for abuse is extraordinary.

Baroness Chakrabarti: Goodness me. No doubt, at some point the police forces concerned will want to share their views on what you have said with the committee. Does Peter Dawson want to respond on proportionality?

Peter Dawson: I will be brief. I do not think proportionality is the issue in the prison context. There are concerns, but I do not think the use of the tool is disproportionate to the problem that it is trying to solve.

The Chair: Thank you. Can I check, for clarity? You distinguish obviously between live facial recognition and retroactive use. You mentioned regulation when you were talking about retroactive use, but your view is that live recognition should not be employed at all. Is that at all, or not at all unless and until there is regulation?

Silkie Carlo: I cannot envisage a way in which live facial recognition for overt use in public spaces could ever be compatible with the right to a private life. That is a view shared by lots of human rights experts around the world.

The Chair: Thank you.

Q54 **Baroness Hallett:** My background is as a judge. I specialised in criminal appeals, so I need to declare that. I am aware of a number of really grave crimes where the perpetrators have been caught on camera and evidence of facial recognition has been used in court to prove their guilt.

Surely, when you have crimes of that nature, that should be factored into proportionality. Proportionality is always a subjective concept. It is very difficult for a judge when faced with that problem. I think the Court of Appeal in the south Wales case said it was not disproportionate. It is a matter of subjective judgment. Surely, where you can capture a perpetrator of a grave crime on camera, the court should be allowed to use that evidence.

Silkie Carlo: My understanding of the Bridges case is that the court did not make a judgment on proportionality, and that the issues were whether it was in accordance with the law and the lack of safeguards on who deploys live facial recognition and where. There may be more of this to be tested in the courts. I am sure that if the police go ahead we will see that.

Of course, in the serious cases you talk about, everyone wants criminals to be brought to justice. There is no disagreement about that. I am aware, of course, that facial recognition has been used on the custody image database for many years. I expect that is what you are referring to.

Baroness Hallett: I am talking about criminals caught on surveillance cameras.

Silkie Carlo: That is where an identified image is then compared against the custody image database, which has existed for some time. That has 20 million images on it, several hundred thousand of which the Biometrics Commissioner says are unlawfully held. If you have a proportionate custody image database and an isolated image of a suspect, that is the kind of scenario where facial matching has clear uses. That should be legislated and regulated to make sure that it is done in a safe and proportionate way. It is light years away from having a camera on the street biometrically scanning everyone as they walk past all the time.

Baroness Hallett: I was talking about the forensic use, not the live facial recognition. You talked about a moratorium on that too, until it is properly regulated.

Silkie Carlo: I think there should be a moratorium on the retrospective facial recognition technology that police forces are acquiring now, which allows them not just to compare one isolated image against the custody image database but, effectively, to do any sort of facial recognition matching with footage against potentially any type of database. That is a much more expansive use of the technology.

Q55 **Lord Ricketts:** May I briefly follow up Baroness Hallett on grave crimes? My background is more in national security than in the privacy and rights area. I am always uncomfortable with moratoriums and outright bans, because there are often circumstances that are very exceptional.

In the Skripal poisoning, it was possible to identify, I think using public camera footage, the arrival of those people at Gatwick Airport and their departure from Gatwick Airport, tracing them through to Salisbury and back again. In a case where chemical weapons have been used on the streets of the UK, surely there is a security issue to balance against the rights and proportionality issue.

Silkie Carlo: The focus of my evidence is on overt uses of facial recognition in public spaces. The national security application raises different issues. It might be that the security and intelligence agencies feel that they already have the right to use facial recognition lawfully in some circumstances under the Investigatory Powers Act. This is entirely unregulated with any specific regulation, other than just the backdrop of common law, the Human Rights Act and the Data Protection Act, which is what the police claim enables them to use live facial recognition. There is no specific legislation for overt facial recognition in public spaces. That is what I am specifically speaking about.

Q56 **Baroness Shackleton of Belgravia:** What deterrent aspect do you think there is? There are no statistics, because you cannot possibly have any, of how many people do not commit crimes because they know that this exists on our streets, and if they do they will get caught and the sanctions will be horrific. How do you factor that into the legalities of deploying this sort of surveillance?

Silkie Carlo: The flip-side of that is that there will be some people who will not go to protests, for example, if they know that facial recognition is going to be used. We saw some examples in south Wales, where there were individuals who had spent convictions but who were still on facial recognition databases, trying to move on with their lives. There was one gentleman who went out with his girlfriend to go to a concert. He was detained by police in a pub, which completely ruined his date, because he had been flagged by facial recognition. There are two sides of that.

What I can say from the way we have engaged with the police and the way their use has changed over the past four years is that it is now very overt in order to try to fit the system to comply with the law. When they first used it I went on an observation, and they had actually hidden the facial recognition van behind sheets of corrugated iron. It was comical. They were claiming that it was an overt deployment, but you would never have known that there was any camera there at all.

By the time it was last used, which was in February 2020, they had police officers handing out leaflets and signs saying, "Facial recognition cameras this way". If you think that you might be on a watch list or if you are a fugitive, you are not going to walk through that bottleneck straight under the gaze of the camera. Again, that is where I think that overt capability, more than anything else, is having a severe chilling effect. The statistics show that it is not catching criminals.

The Chair: Thank you.

Q57 **Baroness Sanderson of Welton:** Thank you all very much for coming. Moving on slightly, Professor Yeung, you have questioned claims of neutrality and objectivity in machine-learning systems. Could you tell us a bit more about that? To everybody, if data itself may not be objective, is there a way to responsibly use justice system data for risk assessments, algorithmic or not?

Professor Karen Yeung: Thank you very much. If you do not mind, I will look at my notes, because I have lots of reasons.

One of the great promises of machine-learning systems is that they provide us with scientific predictions. That is one of the greatest fallacies of the digital age, and I want to explain why. It is because these basically prediction machines are built by human programmers, and they are trained on historical data. That means that we need to consider the data and the choices that are made by developers in building the algorithmic model that is then turned into a prediction machine.

Let me start with a number of ways in which subjective judgment—to assume that is the opposite of objective and neutral decision-making—comes into the process. First, as I think you know, prediction models are trained on historical data, what has happened in the past. The basic idea is that if we can find patterns in the historical data, we can create a prediction model that will help us to make a prediction on unseen data if circumstances arise.

That will reflect existing historical patterns. Of course, that means that existing social practices that are reflected in the data will emerge in the predictions that you make, because that is how you design it. That is why Amazon is so good at predicting what kinds of products you are likely to buy. It has seen what you have bought before and what you have looked at. That is exactly the logic.

One lovely example of the problem I like to cite is quite an old study from back in 2016 by some Carnegie Mellon researchers. They had 1,000 simulated user profiles: 500 men and 500 women. They had them click on news sites. They wanted to see what kind of job recruitment ads were served up to the men and the women. It turned out that the male users were shown high-paying job ads six times more frequently than women, on the assumption that women have low-paid jobs—that women were not interested in high-paying jobs because historically they have not had high-paying jobs. That is a wonderfully vivid example of the way in which prediction machines will replicate the biases that are already built into your historical data. That is reflected in the data.

The second thing is that a lot of the tools that we have seen, particularly in the criminal justice context, are database tools that collect lots of data and put them into a single, very usable digital interface that can be accessed by front-line decision-makers. For example, the London Gangs Matrix is based on opinions from police and other sources of intelligence, including schools and local community officers, about whether or not someone has been seen associating with a known gang member.

The data that is used to decide whether to place people—these are children—on the London Gangs Matrix is opinion data. It is not hard evidence, but it goes into the Gangs Matrix and the person thereafter becomes a 'gang nominal': the database is given the sheen of objectivity because out comes a new smartphone and you get a little risk assessment, red, green or yellow, which is treated as if it is official objective evidence that this is a risky person or this is a member of a gang. That is also true at scale in the LAPD policing tool that is used, based on observations by police out on the street when they stop and search or when they even observe someone who is known to have had previous encounters with them.

The subjectivity that is built into the data and being stored on the databases then acquires a sheen of objectivity simply because it is stored in the database, and it is really easy to access. It gives you a nice colour, which is easy to interpret. That is the second way in which there is no objectivity in these tools.

The third one is more difficult to identify. It concerns what are called the abstraction choices made by computer scientists when they build prediction models. Crucially, they need to decide, "How do I understand this problem? What is the relevant input data? What kind of algorithms should I use to generate a particular kind of output?" In fact, that is not a science; it is an art. You make subjective choices about what is most appropriate to your organisational need, and there is no right or wrong way of doing it. You tune and you tweak, and you come up with a model that seems to do what you want it to do, but it is a highly subjective decision, just like any decision of fit for an organisational purpose. In particular, one of the decisions that is made is whether or not the data that is being used to train the models is an adequate representation of the phenomenon that the machine is trying to predict.

For example, in so-called recidivism risk predictors, the tools are supposedly predicting whether or not a particular individual is likely to commit a serious offence within a particular timeframe. In theory, it might be wonderful to have prediction machines that could do that, but if we were to do it accurately, we would need a comprehensive dataset of all the crimes that had been committed, the individuals responsible and their behavioural profiles to make a reasonably accurate prediction about whether or not a person fits a profile of someone likely to commit a future criminal offence. We do not have that data, of course. Lots of crimes that are committed are unreported, and we do not have convictions in all the crimes that are reported.

What happens? The computer scientists say, “Let’s use arrest data instead”. We all know that just because someone is arrested does not mean they will be charged, let alone convicted. There are all those crimes for which we have no arrests at all, and yet these tools are being used in Britain on the basis that they generate predictions about recidivism. We should at least be labelling them as re-arrest predictors and not call them recidivism risk predictors, because that is not what they are.

We see already that all those assumptions have been built into these tools, which the average punter will of course not be aware of, and certainly the front-line worker will not. These things really need to be subject to much more careful and rigorous scrutiny and evaluation before we roll them out thinking that they are going to massively enhance operational effectiveness. We still do not even have data that they do that.

Peter Dawson: Could I come in briefly on that, because of the relevance to prisons? It seems to me that there are three potential protections against the risks you describe, which are absolutely real. The first is the design and whether it is subject to external scrutiny. The second is the power of the individual to challenge information. In prisons, the systems for doing that are shaky at best. They are well described. There are good frameworks and they should work, but in practice they do not work well. The third protection is proper, transparent scrutiny of what actually emerges. On categorisation, the jury is out, but generally in relation to prison affairs, the willingness to really scrutinise whether a policy is having the impact you think it should have is lacking. It is particularly lacking when it comes to issues of discrimination.

We may come on to it, but specifically in relation to categorisation there was a very good, thorough and quite sophisticated analysis of its potential impact on race discrimination. We wait to see whether or not that follows through into evaluation, which is promised. We know that in other aspects of prison life where there appears to be discrimination, and there have been promises of evaluation, they have not been followed through. If you think about the David Lammy test, “Explain or Reform”, explain happens sometimes, reform rarely happens.

Baroness Sanderson of Welton: There is potentially a role, but there needs to be more transparency about what is going in and better scrutiny of what is coming out.

Peter Dawson: Exactly right.

Baroness Chakrabarti: Peter Dawson, to go back to the issue of design but also legal design and legal foundation, do the regimes within prison have a statutory foundation? Was there a source in law for these exercises in the prison regime?

Peter Dawson: Specifically on categorisation, I think there is existing practice of many years’ standing that they are seeking to improve. I am not sure that the application of the algorithm is any more than a redesign of the algorithm and its consistent use through technology. There has always been a calculation that informed the categorisation process, with a weighting of factors, but it has been done locally on paper and centrally invisibly.

Baroness Chakrabarti: Under administrative discretion.

Peter Dawson: Yes.

Baroness Chakrabarti: So now the algorithmic tools have been developed under that same, broad administrative discretion rather than with new regulation or new statutory authority.

Peter Dawson: There is certainly no specific new authority that I am aware of.

Baroness Sanderson of Welton: Ms Carlo, would you like to add anything to that?

Silkie Carlo: The key thing that always comes up is that we need to understand that in the datasets that are used to inform the development of new algorithms we have to accept that they represent histories of discrimination and try to understand that and in some way mitigate it. That is far more complex and challenging than it might sound, or that sometimes data scientists will admit.

To put that in the context of where we are right now, in the last eighteen months we have had some of the most significant public discussions about policing and racism and policing and sexism. We are nowhere near solving those. I think it is therefore incredibly optimistic to think that we can start developing algorithmic tools to do predictive policing. We have PredPol in the UK. It is incredibly optimistic to think that, while those conversations are still clearly unresolved, predictive algorithms will do anything other than just show you the areas that have been policed, the crimes that have been policed in the past and the types of people who are more frequently arrested or convicted.

It will be an enormous challenge to get beyond that incredible obstacle. Of course, there is also a trust issue. Trust in policing is, on many accounts, quite low at the moment.

Baroness Sanderson of Welton: It is going on what has gone before.

Silkie Carlo: Yes.

Baroness Sanderson of Welton: Thank you.

The Chair: We have heard information about arrests being fed in as if the arrest is an outcome. What about stop and search and suspicionless stops? Do those get fed in? Do we know?

Professor Karen Yeung: That is a very good question. One of the problems is that we just do not know because there is no comprehensive register that informs us about which tools are actually in use, how they are configured, what the input data is or how they can be challenged. This is a serious problem for academic researchers because it is really hard to find public data. It is an even greater problem for the British public and for democratic accountability itself. If there was one recommendation, it is the desperate urgent need we have for systematic accountability and transparency, which is currently lacking.

The Chair: I think we may come to more about that.

Silkie Carlo: The thing about stop and search is that, even if suspicion did not precede a stop, criminalisation often can happen, because of the use of offensive language, for example. As someone who has observed lots of policing deployments over the years, you see this quite often. When a suspicionless stop is conducted, sometimes a very thin justification for that stop will then be found. Inevitably, discriminatory patterns of stop and search will be reflected in crime data.

The Chair: Thank you. Baroness Primarolo, your questions follow on from that.

Q58 **Baroness Primarolo:** We are discussing this morning basically that we seem to have two sets of tools that we can use: human judgment, for which you have put a very strong case, is flawed but is the one we are using; and algorithmic capability. I would like to try to unpack this proposition: is there a way that we can gradually use algorithmic capabilities to assist, without the challenges that have been identified so far this morning? Not all human judgment is flawed, so how do we incorporate that human judgment into the use of algorithmic capabilities, given their clear benefit in huge amounts of data? Perhaps, Professor Yeung, you could open. I presume that the three of you are not saying, “Never, never use algorithmic capabilities to deal with volume”.

Professor Karen Yeung: My spam filter is one of my best friends. I am certainly not anti-computers. They can do some wonderful things, and definitely enhance operational effectiveness, convenience and all those things that are as good as motherhood and apple pie.

However, I think you are absolutely right. The challenge is how we can avail ourselves of the opportunities these tools offer in a way that augments human decision-making. That is accepted as a noble aspiration. The problem is that there has been so much excitement about the promise of big data that we have charged in and used tools just because we can. What we are really lacking is robust, sustained and careful research to identify the conditions under which the use of these tools genuinely augments human decision-making. Good systematic scientific research is hard, time-consuming and expensive. That is the kind of research we should be investing in, but unfortunately it is not. We are going for the big, showy tools that do biometric scanning in airports and so on and so forth.

I think there definitely is a role. At the moment, we do not have clear scientific systematic evidence about the conditions under which it is used, because it is very context specific depending upon the decision in question. It is fine for a spam filter, but it is much more problematic if you are using it, for example, to decide whether to take a child into care. Those decisions should not be subject to automation, but they are being subject to automation because of the claimed deficiencies and the reduction in resources devoted to front-line workers.

I think there is politically a tendency to go for the digital quick fix in the view that it will modernise and enhance operational effectiveness, without clear evidence that in fact it does that or has any effect in the capture of adverse impacts and downsides associated with its use. I think this is the problem; we are rushing in without careful, sustained research to really identify how we can optimise the use of these tools for the benefit of society without the unacceptable harms and abuses of power associated with them.

Baroness Primarolo: Thank you. If we could park the question about appropriateness—we might need to come back to that—and where it would be appropriate to use these technologies and where it would not, it seems to me that there are two central problems that the three of you have clearly identified this morning.

The first is how they are programmed and the assumptions that are made by the data scientists about the expectation of the outcome. The second is that human judgment and human understanding evolve, so we are able to recognise, perhaps more clearly now, that women are discriminated against and subjected to more violence because they are women, as a result of that.

Those seem to be the two crucial things about how we interact with an algorithmic capability. Could you unpack that at all? What are the steps? It is never going to be perfect. Then we add threat to it, or perceived threat, and it gets even more complex.

Professor Karen Yeung: I am just trying to get a little bit more clarity about your question. How would we think through identifying the circumstances in which it might be appropriate?

Baroness Primarolo: Yes. That is much clearer, thank you.

Professor Karen Yeung: Unfortunately, in the digital excitement, which I call the digital enchantment, there has been a tendency to grab the data that you have and do something with it. It is to turn the existing historical messy data that you have into a tool to predict something. Instead of being problem driven, we are being data driven.

There are lots of ways of thinking about how we understand a problem. In particular, we can understand it as a problem at the individual level—for example, that women are subject to violence: “If only she took more care, she would not be so subject to abuse". Or we can think of it as a problem in broader social and systemic terms and look at the social conditions under which these practices take place.

There is a preliminary question that I really urge policymakers and operational or organisational units to think about. How do we understand the nature of the problem? What are the different ways that we can think about what is going on before we reach for tools that necessarily assume that the problem is at the level of the individual? That is what individual risk assessments do. That might not necessarily be the most valuable, socially beneficial or inclusive way of thinking about the problem. It is to push back against the tool and think hard about the social problem first, and then think about the tool and the range of tools that we might use, digital or otherwise, and the need for research in understanding the nature of the problems before we start waving digital sticks at them.

Baroness Primarolo: Peter, can we pick up on the Digital Categorisation Service? The prediction is based on historical information. It seems to me that what you are describing, picking up from Professor Yeung, is that the diagnostic tool simply speeds up, with less human interaction, what would have happened anyway because of the historical data. Do you think that is fair?

Peter Dawson: It is a little more than that. It potentially improves it. No one in the Prison Service would say that the previous system of categorisation was likely to produce good or fair outcomes, and certainly not consistent outcomes. There is the possibility of consistency that comes from it. The thing that strikes me is the danger of the word "prediction" because people draw the inference that they are being given a prediction when clearly they are not.

The one thing I would add to your list is that these tools ought to prompt questions and thought about how you affect the outcome that might otherwise happen, not for the individual but in general. On categorisation, because it is the specific example we have given, the algorithm and the human judgment applied to it do not predict that a particular person is going to carry on committing crime within the prison environment. But they might say that it is possible that they will, and it is more possible that person A will than person B.

Of course, there is an immense number of things that the prison can do to test whether that is happening and to give the person the opportunity for it not to happen, and to reward them when it does not happen. It is a useful tool for guiding operational decisions that you will take in the future that can change the outcome.

The risks that Professor Yeung identifies are so common in the criminal justice system. Many people around the table will be familiar with the IPP sentence, where we decided 20 years ago that we could predict with reasonable confidence who was likely to commit very serious crimes in the future. We have held well over 10,000 people in prison on the basis of that false assumption. Perhaps the first conclusion I would draw is that we should have a massive dose of humility about what the tools can tell us, for all the reasons that have been explained.

Baroness Primarolo: Silkie, picking up that point about human or external challenge to the predictions, do you think there is a way to empower that challenge to prevent the clear problems that occur in these algorithmic capabilities—the ones that you have identified—or is challenging at the end fundamentally flawed?

Silkie Carlo: It is a very important point. It depends on the tool. This was something that was looked at in the Essex report that I mentioned on live facial recognition. When an officer receives a match, how often are they applying genuinely their own analysis and deciding whether to make an intervention? How often are they looking at what they have been presented with, assuming superiority from the machine's judgment and then going ahead and making an intervention by stopping someone, asking for their ID and so on?

It seems in that case that the matches the facial recognition system suggests are highly influential. They certainly can be. I once had the privilege of going inside a police facial recognition van and saw men being matched as women. Obviously, in those cases that is not very persuasive, but in some others it certainly has been.

On the other hand, if you look at a system like the harm assessment risk tool that was used by Durham police—another reoffending AI risk tool—it seems there that officers were using their own judgment. Again, you need to look under the bonnet. When we read through the literature on that system, we found that postcode data was being used as one of the variables to determine whether someone would reoffend. There is clear scope for discrimination with that.

We need to look at the systems. Often, when people talk about safeguards with algorithms, they talk about having a human in the loop. We need to think about having a machine in the loop of human decision-making where that can be proven to be beneficial, and then monitoring very closely to make sure that individuals are not unduly influenced by the predictions and recommendations of the machines such that they effectively abandon their own judgment.

Baroness Primarolo: I have asked quite a lot, but this is crucial. Does it fit? I think Silkie at the end touched on whether there is ever a place for it.

The Chair: Peter, in your written evidence you talked about equipping managers within prisons to use the system appropriately. Would you like to expand on that a little?

Peter Dawson: Yes, and it is not rocket science. The first thing is training people and giving time and space for people to understand what these things can and cannot do. A lot of the evidence you have heard today will never make it to the person who is given this tool to use. They will not understand its limitations.

There is something about the interpretation of data, which can be complex—understanding when a number is significant and when it is not. It does not particularly apply to the categorisation tool as it is applied to individuals, but the instruction to governors says that they should consider the outcome of the categorisation process specifically in relation to protected characteristics.

That is difficult. Even in a moderately sized prison the number of people with protected characteristics might be quite small. It will be difficult to tell from the statistics that are produced whether a difference is justifiable or not, whether it is statistically significant. There are ways in which the centre can equip people locally to understand whether what they are seeing is of concern or not.

The final element is confronting the biases that exist in people's decision-making. Some of that is about looking squarely at what has happened in the past and being prepared to look, with challenge, across the whole of your operation. I used to be a governor. I might look at my categorisation decisions and think, "Those are absolutely fine. They match my population and there is nothing to be concerned about". But am I also looking at disciplinary outcomes for prisoners? Am I also looking at the use of force against prisoners? Am I looking at who gets to education, who goes to the gym and what level of privileges people are on? It is about willingness to look squarely at all that and then say, "It is likely, given what I know about outcomes, that my decisions on categorisation may not be as pure as they appear to be and may not be lacking in discrimination".

It seems to me that we are at a very dangerous moment on that because there is a general inclination to say, "Maybe we have moved on from that, and maybe we are worrying too much about discrimination on the grounds of race and it can be explained away by any number of other factors". In the prison context, which is what I know about, that is not true. It is a really serious risk in relation to the specific issue of categorisation.

The Chair: On training, understanding what underlies what is being presented should include the confidence to challenge that, presumably.

Peter Dawson: Yes, and the confidence to take prisoners' views and challenges seriously. This shows up in micro ways. One of the advantages of digitisation, rather than the algorithm, is that evidence is held and that it can be accessed. Anyone familiar with prisons will know that we have decades of paper records that go missing. Lots of information is there.

One of the things the instruction did not do was to preserve any challenge that a prisoner makes against the decision. That is on paper, and it is lost. It does not follow the prisoner on to another prison. It is a tiny example, but I think it betrays an attitude that says that the prisoner's challenge is not quite as important as everything else. That is an easy thing to alter.

Q59 **Baroness Chakrabarti:** I think Peter Dawson's reference to humility is quite refreshing in any form of public life or decision-making. Certainly, transparency and legality are supposed to be a bedrock of democracy, particularly when we are talking about powers of compulsion over people, the criminal justice system and policing. These are things that can really change people's lives.

Given that both human decision-making and computerised decision-making are imperfect, if we think of the world pre-algorithms and we have public debates about, say, stop and search powers, hopefully we will have transparent public debate about whether there should be stop and search with suspicion or without suspicion. If it is to be stop and search without suspicion, how will that work? What will the checks and balances be? That is what we try to achieve by primary legislation, secondary legislation and even public guidance codes, *et cetera*.

If that is the case for traditional human decision-making, why can we not replicate that for the algorithms, which are in themselves lists of instructions/guidance but for computers rather than people? Am I right in that? Can you envisage a situation where these hidden decision-making chains could be more transparent and publicly accountable?

Professor Karen Yeung: Absolutely. I could not agree with you more on every single thing you said. These tools are being used to inform decisions. We have a very established constitutional administrative framework that recognises different forms of abuse of power, and we can draw on those principles to come up with a systematic oversight framework for ensuring that we have proper accountability, reviewability and transparency in the way we make these decisions, alongside and with the assistance of computer tools. It is just a lack of political will that we do not have them.

Silkie Carlo: If I might add to that very briefly, theoretically we have the right not to be subjected to purely automated decisions under GDPR, and if you are subjected to a purely automated decision in a law enforcement context, you have to be notified and you have the right to review.

One of the things I am concerned about is that that human review might be seen as such a safeguard that some of the other safeguards that are needed for algorithms are then minimised, and at the same time the accountability in human decision-making is also minimised, so accountability overall can be minimised and becomes obscure and abstract. I am not aware of cases where an individual has been told that they have been subjected to a purely automated decision.

The Chair: Thank you. We are about to come on to the safeguards, and I have just realised that we are only half way through what we knew at the start we wanted to ask, let alone what we have been thinking of since. I think we are going to have to speed up somewhat.

Q60 **Baroness Pidding:** I have a trilogy of questions for you. I am interested to hear each of your thoughts on them. What should be considered before a tool is deployed? How can impact assessment processes be most meaningful? We know that we must strike the right balance between using information to protect while avoiding infringements on our liberty. What safeguards should be in place?

Peter Dawson: I can say a little about the impact analysis process for categorisation, which we thought was surprisingly good. It was quite sophisticated analysis and it highlighted possible risks of future discrimination.

An element that I think was missing, specifically on race, was thinking about the subdivisions within race. Actually, it is different in prison if you are black, compared with being mixed race, having an Asian background or a Gypsy/Roma/Traveller background. That level of sophistication needs to be added, but it was a good analysis, in contrast to others that we have seen.

The additional safeguard that could come with it would be aggregating all those analyses and knowing whether or not they are common, what emerges from them and then what is done about it. We asked the Ministry of Justice what equality analyses had been done in relation to policy changes, and how many of those analyses had shown there was a risk of discrimination and a decision taken to proceed none the less. We asked how many produced recommendations for change and what had happened to those recommendations for change.

The Ministry does not have a central record of the equality assessments that it conducts, so it cannot answer any of those questions, and it does not know overall whether its processes are working. I can say for this specific example that it worked quite well, but I can also say that that struck us as unusual. Part of the safeguard is gathering all that data together and looking at it.

Professor Karen Yeung: If I may pick up on that, I have made some notes. I already mentioned the importance of thinking quite carefully about what the nature of the problem is before we jump in and assume that the problem is at the level of the individual and we should do some kind of individual risk assessment based on either a statistical or a machine-learning tool.

There is also the question of who is being made the subject of the algorithmic evaluation. The reality is that we have tended to use the historical data that we have. We have data in the masses, mostly about people from lower socioeconomic backgrounds. We are not building criminal risk assessment tools to identify insider trading or who is going to commit the next kind of corporate fraud because we are not looking for those kinds of crimes, and we do not have high-volume data. This is really pernicious. We are looking at high-volume data that is mostly about poor people, and we are turning it into prediction tools about poor people. We are leaving whole swathes of society untouched by those tools.

This is a serious systemic problem and we need to be asking those questions. Why are we not collecting data, which is perfectly possible now, about individual police behaviour? We might have tracked down rogue individuals who were prone to committing violence against women. We have the technology. We just do not have the political will to apply it to scrutinise the exercise of public authority in more systematic ways, in the way in which we do it towards poor people.

This is another problem. We are not looking at the way we use data in relation to the disparities in power that already exist in our social structures. I really urge that the whole of society should be thinking much more about the distribution, exercise and exploitation of power, because these are powerful tools and they are being used against the most vulnerable. We should always be asking, "Who is being assessed by these tools, and is it right that we should be assessing these people and not those people?" Why are we not thinking about those people, because actually we might strengthen accountability, transparency and democracy by looking at those people rather than these people where we have huge volumes of historical data about poor people?

Then we need to ask questions about the real value of the tool. What exactly is the social and organisational purpose? What is the potential for the abuse of power and for injustice? Is it lawful? That means to what extent does it serve legitimate and effective legal purposes, bearing in mind that we are no longer in a paper-based world?

This is one of my huge bugbears. There is a tendency to use argument from analogy, but we have always done it. We have always subjected people to scoring systems. We are just putting it in a digital tool, and if there is anything that illustrates the difference, it is the problem of the naked photo.

The problem is that once we have it in the digital tool, in data, it is scalable and I cannot erase it. It is there permanently. It can be disseminated in real time around the world.

Baroness Chakrabarti: The naked photo?

Professor Karen Yeung: Yes, like the naked photo. That is why it is not the same as a paper-based tool. A paper photo is not the same as a digital database. That completely changes the proportionality assessment. That means that we can scale injustice. It is exactly what happened in the Fujitsu case, with Horizon, the financial accounting software that was not even machine learning. That is what we have not grasped, and we really need to think about power and the scale of injustice that digital makes possible.

The Chair: Ms Carlo, you could say, "I agree".

Professor Karen Yeung: I am sorry. I got on my high horse again.

Silkie Carlo: I agree wholly. Really important points have been made. On a completely practical basis, once you have got through all those important points about why and what it is going to achieve, we could have more safeguards around procurement. There have been companies that have been very proactive in trying to—

The Chair: We have some questions on procurement.

Silkie Carlo: I will hold my horses.

The Chair: Shall we move on so that we get time to explore them?

Q61 **Baroness Shackleton of Belgravia:** I think Baroness Chakrabarti has already trespassed on this question, and you have answered it.

Baroness Chakrabarti: I am so sorry.

Baroness Shackleton of Belgravia: It is to do with the regulation of the technology. How would you characterise the legal framework around advanced technologies in the justice system? Would you like to see better compliance, a better framework and better regulation?

For example, Professor Yeung, I think you have given a very good example about arrest. If it was labelled on the packet that this is data and only compiled by arrest, but the arrests are flawed for the following six reasons, on a scale of one to ten you should discount this evidence and put it down to low because it is not very accurate. There is no compelling law to make somebody disclose for prediction, compared to what it is. The packets may not be labelled properly. If you were in medicine you would have to say what is on it, and it would have to comply.

Do you think there is scope for the judicial system to intervene to make these tools fit for purpose? If you do, would it have to be specific for each tool?

Professor Karen Yeung: I will try to be as brief as possible. First, I am asked this question a lot. I think our constitutional principles are completely up to the task. The challenge is translating them into what that means in safeguards in relation to specific tools. The problem is that the tools are sophisticated and hard for the average person, organisation or front-line worker to understand. That is why I think we need specific statutory regulation that translates and operationalises our constitutional principles into much more concrete, systematic safeguards.

I do not think it is necessary for every single tool to have its own bespoke scheme. It is perfectly possible to have domain-specific oversight, for example, in relation to policing and decisions about individuals that involve the deprivation of liberty or other kinds of rights. You could come up with a generic piece of legislation graded according to specific, contextual features. You might need some particular ones for specialised domains such as finance and so forth, but I think we have the constitutional principles that will equip us to come up with the kinds of legislative oversight that are systematic and ensure transparency, review evaluation and the opportunity to challenge. That would make these things much safer to use and would inspire public trust in a way that currently is not there.

Silkie Carlo: I agree. That term "domain specific" is very useful. In the case of facial recognition, it is very easy to isolate that. It would be relatively straightforward technology to legislate around, and I think it should be done in the way I set out earlier.

One of my concerns is that at the moment we rely on the legislative backdrop—in particular, GDPR, the Data Protection Act and the Human Rights Act. All those pieces of legislation are under attack by the current Government. My concern is that, if GDPR rights are scaled back, what does that mean about the right not to be subject to a purely automated decision? What does it mean about privacy rights? That could completely change the landscape again.

Baroness Shackleton of Belgravia: Thank you. I was not particularly confident that the appeal process, which Mr Dawson mentioned for prisons, was an entirely transparent and effective means, if you were caught by the computer, of wriggling yourself out and getting proper representation in order to to appeal against any judgment made. If you rely more heavily on initial arrest, you have to have a very effective appeal system, in my view.

Peter Dawson: That is exactly right. There are yards and yards of instruction about complaint systems in prisons, which, if they operated well, would be largely adequate, but they do not.

The only thing I would add is that for decades in prisons one of the ways in which policy has been improved has been through judicial review in individual cases. Of course, that is immensely more difficult than it used to be, and the policy will not be improved because of that.

Baroness Shackleton of Belgravia: Thank you.

Q62 **Lord Ricketts:** I want to tee up Ms Carlo's comments about procurement, and focus on the interaction between public bodies and private technology developers. I assume that often a company will come selling its wares and claiming that they are solutions to a particular problem. Since procurement is devolved, I wonder whether you think that police forces have the capability to understand what they are buying, to evaluate it and to be conscious of issues such as commercial confidentiality and the extent to which the private technology developer will have access to, and indeed even own, the data that is produced. Ms Carlo, would you like to start?

Silkie Carlo: Thank you. I will speak about it through the example of facial recognition because it is so instructive about how this can work. My understanding is that, essentially, NEC, which is a Japanese-based conglomerate, approached police forces for it to deploy live facial recognition, and that initially the technology was provided at little or no cost. Unfortunately, I think that then became the reason why the UK was an outlier in Europe in facial recognition deployment. It did not start with asking why. It did not start with the proper kind of assessment you would want to see when you are adopting a rights-altering technology.

I think there needs to be much stricter regulation on procurement and the adoption of these extreme new technologies. I cannot profess to have expertise about exactly how that should work. I know that there is active consideration about how it should work in local authorities. I am not sure how it should work in police forces, but there needs to be a proper process for that.

Another interesting interaction was when shopping centres started using live facial recognition, again on a trial basis. The police were brought in to support that deployment as well. You can see through the PR arms of big tech companies that there is creeping and informal deployment of very serious rights-altering technologies. I think a proper procurement system would help to clean it up a bit.

Lord Ricketts: Would more centralisation allow greater competence in judging what is being bought? By analogy, we have government IT procurement. When I used to be the Permanent Secretary in the Foreign Office, we all used to buy our own IT systems and we were all left face to face with technology companies. Frankly, we did not always have the expertise to buy appropriately. That might be one solution. Professor Yeung, do you have thoughts in this area?

Professor Karen Yeung: As you mentioned that, I was thinking of NICE evaluation of pharmaceuticals. That is extremely effective relative to the US, which is a total gong show in wasting resources on ineffective drugs. I think there is something to be learned from that experience.

One of my gravest concerns about procurement is the extent to which it allows public authorities deploying these systems to hide behind the protection of IP rights, which the private software developer typically invokes. For example, in the Bridges case there were claims that the facial recognition was discriminatory against protected groups, particularly women and ethnic minorities. The software provider was asked to give evidence. All he said was, “I can assure you that we have trained it on representative data”, but the court was not shown that data. The police had no access to the training data.

All we have are empty promises and no evidence to demonstrate or test whether the training data that was used to configure those models is in fact representative of protected groups at all. Fortunately, the Court of Appeal decided that simply advising the police, “Rest assured, we’ve checked that”, was not sufficient protection. They needed to be more proactive in ensuring compliance with the public sector equality duty.

The question becomes, “How much is enough?” My own view is that we should be able to have access, at least in a confined way, to protect the IP rights of the developer, so that there is independent evaluation of the underlying training set and the algorithmic model. It does not need to be made public, but I would like some assurance that it has been independently evaluated and that the data has been made accessible before it is used to make decisions about people in this country.

Lord Ricketts: In other words, some degree of accountability.

Professor Karen Yeung: Yes, that is right.

Lord Ricketts: Mr Dawson, do you have anything to add?

Peter Dawson: There is nothing I would add to that.

Lord Ricketts: Thank you very much. In the interests of time, I will stop there.

Q63 **The Chair:** Professor Yeung, can you think of any example where there is a terrible block of commercial confidentiality on the one hand but, as you say, IP rights on the other and where society has managed to get through that in any other sector? That can be your homework, if you like.

Professor Karen Yeung: I am thinking of the pharmaceutical example again, but then we are talking about big pharma and that is also problematic. No, I cannot come up with one off the top of my head.

The Chair: We seem to come up against commercial confidentiality quite a lot in this.

Q64 **Baroness Primarolo:** Professor Yeung, you used the example of NICE, which acts as that intermediary, taking on board efficacy as well as IP rights and technologies.

Professor Karen Yeung: Exactly.

Baroness Primarolo: Are you suggesting that there is room for a similar body on procurement and regulation with regard to the algorithmic—whether it works, is it appropriate and has it been properly assessed?

Professor Karen Yeung: I have to confess that I have not given that very serious consideration, but as a general model there seems to be a lot to commend that kind of approach.

Baroness Primarolo: Thank you.

Q65 **Lord Dholakia:** This committee will be making recommendations to the Government. It would be very helpful to know the top priorities that you would like us to address on this particular matter. If I may add a further question, have we missed out anything so far that we should consider?

There is one other point I want to raise with the professor. I was able to visit, with the Lord Chair, the Ethics Committee in the West Midlands. Does that have a role in the way that the police should be using such tools in the implementation of new technology?

Professor Karen Yeung: There are lots of questions there. On my top tips, if I was king for the day and what I would do, first, I would introduce legally mandated, systematic transparency for all uses of algorithmic tools in the public sector, particularly in the criminal justice system. There should be basic information requirements that should be disclosed. I will not go through them here. I have a paper that will account for those.

Sorry, I have lost the question. What was the third piece that you asked me?

Lord Dholakia: I was asking about the Ethics Committee in the West Midlands.

Professor Karen Yeung: Ethical judgment is really useful when you have a decision that falls in the grey area; when it falls on the side of legal acceptability but there may be questions about whether it is the right thing to do. For example, it is very common in relation to medical ethical decisions that are legal but where there may be an ethical quandary. I can imagine that similar kinds of questions arise in the criminal justice system.

My concern is that these kinds of oversight frameworks need a legal basis. They should not just be an optional extra that can be disregarded. Without systematic and legal foundations for the committees, I am really worried that they are not sufficient to guard against the kinds of systemic sustained abuses that can go on in a highly opaque way, because these are very opaque tools. It is a good start, but I do not think it is enough.

Peter Dawson: I would like a central register of equality assessments and a requirement to report against them on whether attention is paid to them and their recommendations, as well as the number of occasions on which they are overridden so that equality comes second.

On a very practical recommendation for prisons, Professor Yeung said that we are not in a paper-based world any more. We are in prisons; prisoners do not have access to information technology. In this particular respect, two things would be hugely helpful if they did. First, they could have access to information about them, and have it easily and in a language of their choosing. Secondly, they ought to have access to the organisations outside that are there to protect them, such as the Information Commissioner’s Office and external avenues of appeal. That is all possible, and in ten years’ time we will be astonished that it was not true now. It needs to be made to happen.

Silkie Carlo: I agree very strongly with those points. There is also a case for a central register or inventory of all public sector algorithms so that we can at least surface where they are being used. If they had equality impact assessments as well, that would be fantastic.

I strongly feel that the committee should recommend that the use of live facial recognition overtly in public spaces is prohibited. The European Parliament has just called for that. In the US, a number of facial recognition algorithms were paused for law enforcement use. This is now becoming an international standard. We have set a very bad example over the past few years, and we are grappling with major trust issues with policing.

We know that there are demographic accuracy biases as well. The thing about the public sector equality duty and the equality impact assessments is that you can surface a discriminatory impact, but you do not necessarily have to show that you can completely eliminate it in order to still deploy. In the context of the current lack of trust, as well as the serious rights abuses, that is a very serious thing.

I will give you one example of a case I saw that made me feel with greater urgency than ever that a ban is needed. On one of the deployments that I observed of live facial recognition, a 14 year-old black schoolboy in his school uniform was misidentified. He was walking down the street with his friends. Suddenly he was physically accosted by four officers in plain clothes. He would have been led to think that he was being attacked. He was dragged over to a side street and held up against the wall. He was asked for fingerprints and for his phone. He was asked for his ID, which of course he did not have because he was so young. Obviously, they established that they were police officers and that he had been misidentified.

This is a whole new thing for people to grapple with. Stop and search has its issues and discriminatory policing has its issues, but when you are being told that a machine is also against you, and that you can be stopped in this way and there is no law for it and no recourse, it is a whole new type of injustice. Used at scale, monitoring hundreds of thousands of people across the capital or elsewhere in the country, there is serious scope for injustice and discrimination. That is why an urgent legislative prohibition on the use of facial recognition, as we see being called for in Europe, is essential. Then, if people think there is a future for the technology, they can make the case, but right now all we see are very serious problems and risks.

Q66 **Baroness Shackleton of Belgravia:** The actual equipment to do the facial recognition is one thing, but deployment of the information when they have it is another. If you made facial recognition—collecting and harvesting the information—illegal, it could not be used, for instance, in the example that Lord Ricketts gave when we are dealing with serious crime. As to the policemen in plain clothes in the illustration you gave, the filter was obviously flawed because there was nothing to back up the information and that they had got it right. I am being slightly devil's advocate. Is not the problem arising from the use of it rather than the harvesting of it?

Silkie Carlo: Maybe I misunderstood the question, but I cannot see a positive use for harvesting mass biometric data from members of the public.

Baroness Shackleton of Belgravia: If they had the right checks in place before they went off—the four plain clothes policemen—to collect this 14 year-old—

Silkie Carlo: They believe that they do. They have tablets—

Baroness Shackleton of Belgravia: But what if it was regulated that you cannot do that unless you have, for example, applied to a judge?

Silkie Carlo: That is a completely different scenario from live facial recognition, where you are in real time. They receive a match alert on the tablet that shows a photo from a watch list and the photo of the individual who has just walked past the camera. Then they run to intervene.

Baroness Shackleton of Belgravia: But if the law said that you cannot do that and that you have to go back to base, go to the highest person in the police and trace the boy again as he is going along the road, or whatever, is that the problem, or do you want us to recommend that there should be no cameras and no facial recognition at all? I want to be clear what we are being asked to ask.

Silkie Carlo: First of all, to clarify again in relation to the example that was given earlier, I am not talking about national security. I am talking about overt use in public spaces. I am not talking about airports. I would like to, but I think that completely different kinds of rules need to apply.

For overt use, the scenario that you are suggesting—you see that you have a match and then apply to a judge to follow the person again—brings you back to square one. These are already people who are on watch lists. If they are wanted for arrest, that is precisely what the police should be doing. They should be proactively trying to find those individuals.

One of the things I am concerned about, especially with the funding cuts that there have been to policing, is that some of the cases we have been given refer to sex offenders. I am not reassured that there is a highly inaccurate catch net out there that might pick up a dangerous sex offender. I want the police to have the funding and the resources to find that individual before they roam the streets. If we are talking about wanted people who are suspected of serious crimes, it seems highly unsatisfactory to me that we would rely on facial recognition to do that.

The Chair: There may be a difference between reliance and use as one of the tools, of course. Do you put in the same category of being different, as terrorism, a watch list of missing people? I appreciate that there are a lot of issues with consent and so on about being found.

Silkie Carlo: I think different issues are raised there. I have not really seen anywhere that all the ethical issues and the impact on mental health have been properly evaluated. Obviously, that is not my expertise. From a privacy perspective, if you were to have citywide facial recognition tracking millions of people to try to identify people who do not want to be found, that is disproportionate on any account.

With these technologies, we are being confronted again with the question, "What does privacy mean in 2021?" If it means that we have cameras that can basically operate as anything between identity checkpoints and police line-ups in real time, that does not fit with any sensible conception of privacy to me. I think that would be incompatible.

Q67 **Lord Ricketts:** I have a very brief question, just to broaden the scope a little bit. Do any of the three witnesses have experience of any use of algorithmic technologies from elsewhere in the world to help law enforcement that have been positive? Every other industry uses algorithms all the time. Does any other country—we have had one or two references to the US—have examples, beyond the facial recognition area, of anything that we ought to be looking at as a committee for a positive experience of how modern technology can help law enforcement?

The Chair: Perhaps I could add to that. Are there good examples of registration and regulation from other jurisdictions? That is perhaps the other side.

Lord Ricketts: If necessary, perhaps we could have it in writing if you want to think about it. I am conscious of the time.

Professor Karen Yeung: I can think of one very good example. In the US it is being used by the Securities and Exchange Commission to prioritise how resources are distributed within the agency. It is being used for operational effectiveness, but not to identify high-risk offenders. There is an excellent report out by some Stanford researchers that looks at machine-learning applications by federal agencies, which I commend to you.

Peter Dawson: I do not think I can add anything.

Silkie Carlo: It would just be focusing on the negatives.

Lord Ricketts: Thank you very much.

The Chair: Do other Members want to follow up on anything? In that case, thank you all very much indeed. A transcript of this morning’s evidence will be sent to you in case you want to make any corrections. If there is anything that occurs to you at three o’clock on a Thursday morning, please write to us with anything that you would like to add. Thank you very much for your time. It has been a very lively morning.