

Science and Technology Committee

Oral evidence: The Right to Privacy: Digital Data, HC 97

Wednesday 8 June 2022

Ordered by the House of Commons to be published on 8 June 2022.

[Watch the meeting](#)

Members present: Greg Clark (Chair); Aaron Bell; Chris Clarkson; Katherine Fletcher; Graham Stringer.

Questions 273-380

Witnesses

I: Professor Sir Ian Diamond, National Statistician and Chief Executive at UK Statistics Authority; John Edwards, Information Commissioner.

II: Lord Kamall, Minister for Technology, Innovation and Life Sciences, Department for Health and Social Care; Dr Mark Toal, Deputy Director of Research Systems, National Institute for Health Research; Jennifer Boon, Deputy Director of Data Policy, DHSC.

III: Felicity Burch, Executive Director, Centre for Data Ethics and Innovation, Department for Digital, Culture, Media and Sport; Joanna Davinson, Executive Director, Central Digital and Data Office, Cabinet Office.

IV: Julia Lopez, Minister for Media, Data, and Digital Infrastructure, DCMS; Jenny Hall, Interim Director for Data Policy, DCMS.



Examination of Witnesses

Witnesses: Professor Sir Ian Diamond and John Edwards.

Q273 **Chair:** The Science and Technology Committee is in session. We continue our inquiry into digital data and questions of privacy. To help us with that this morning, I am very pleased to welcome John Edwards, the UK Information Commissioner, who is here in person. It might be relevant to say for people watching that before John Edwards was appointed in the UK, he served for seven years as New Zealand's privacy commissioner, and in that capacity, he chaired the international conference of data protection and privacy commissioners, which is now known as the Global Privacy Assembly.

I am also very pleased to welcome from Cardiff Professor Sir Ian Diamond, the national statistician and chief executive of the UK Statistics Authority. He has helped the Committee with our inquiries on multiple occasions. It is very good to have you back with us today, Sir Ian.

I will start with a couple of questions to both panellists, starting with John Edwards. In your role as Information Commissioner, you survey the whole data scene in the UK. What are the principal problems associated with data sharing, and what actions are you taking to address them?

John Edwards: Thank you, Chair, and thanks to the Committee for the invitation to assist with your inquiry. That is a big question that demands wide consideration. It is a multi-factorial problem. I think there is sometimes a rush to alight on the legislative framework, and to say that uncertainty has caused risk aversion. There are actually cultural problems that need to be overcome.

There are organisations that take great pride in their own mission and do not wish to share information with other organisations, even though, when joined up, they could provide a greater benefit for society. There are technological limitations to data sharing, some of which require considerable investment to overcome. There is also a lack of understanding about the facilitative and enabling nature of the legislative framework. That does, in some sectors, drive a risk-averse approach, whereby organisations are concerned about potential liability for sharing that is not supported by the law.

On the second part of your question about the role of our office in addressing those problems, we administer a principle-based set of legislation that has to deal with an infinite variety of data transactions throughout the economy every day. Because of that, it is set at a reasonably high level of principle. That is one of the great benefits: it can do all that work. The downside is that it creates a bit of uncertainty and organisations need to make a judgment. To address those concerns, we have done a few things. First, we have established what we call the sandbox. If organisations want to innovate in data use but are unsure about how the law might apply to it, whether it is allowed and what they need to do to mitigate risks, they can come to us, and we will work alongside them to workshop and work those things out.



Secondly, we have the data sharing code. We have a statutory ability to issue codes that provide more practical guidance, so that we go from that very high level of principle and demonstrate how the law can be applied in particular data sharing situations, such as data sharing for the welfare of children and the like.

Q274 **Chair:** Do you think we will see more data sharing in the years ahead? Is it your ambition to see more?

John Edwards: Yes, absolutely.

Q275 **Chair:** Do you think the constraints are principally caution on the part of data holders or public resistance to the idea of their data being shared more widely?

John Edwards: It is tempting but very difficult to speak in generalisations. I return to my answer to the first question—I think the obstacles to better data sharing are multi-factorial. There are these quite significant cultural obstacles within different organisations. There are pockets of public concern.

I had the benefit of seeing a preview of some of the questions you might ask. You will see in the evidence that I provide today a theme emerge of the importance of really clear information governance. That has the potential to provide the public with reassurance that a variety of different sorts of data sharing can occur in ways that are safe. One of the risks for an inquiry such as this is to look at this quite complex ecosystem and to assume that there might be simple or single answers. We need to be addressing this multi-factorial problem with multi-factorial responses, anchored in sound information governance and transparency with the public and the organisations that we are trying to support.

Q276 **Chair:** To take an example of that, this Committee obviously has an interest in scientific research and health and medical research. In this inquiry, Professor Ben Goldacre told us of the risks of using pseudonymised health data—where the identity is suppressed, but through increasing connections between databases and the use of AI, it may be possible to identify the person behind the data. Professor Goldacre advocated its eradication. What is your assessment of pseudonymisation?

John Edwards: My assessment would be that there are horses for courses. There may be examples of research, even medical research, that can be undertaken with pseudonymised datasets. It depends.

If you have sound information governance and a structure that enables a good risk assessment and research proposals can be put before that body and assessed and the risks identified and managed, there will be situations. I was trying to think of an example, and it is probably poorly informed, but you could imagine, for example, that you could have a piece of research that wanted to identify the relationship between the incidence of rheumatic fever amongst a cohort of children aged between five and 10, and the later development of cardiac disease. You might want to compare

that incidence to the incidence of children having rheumatic fever between 10 and 15 and later incidence. I think you could undertake that research quite easily with a pseudonymised database, without the risk of the data points being able to identify individuals.

There will be situations where pseudonymisation is an appropriate response to the risks that you have identified. There will be others where it will not be satisfactory.

Q277 Chair: If it is based on someone's medical records—I am sure it is possible to do the research and to answer the research question by using that information, but the concern is that researchers who are conducting that research could end up identifying the individual and then, through access to their medical records, could take an inappropriate interest in their wider health history. That is the concern. It is not a question that you couldn't do the research. It is a question of whether a by-product of that might be an insecurity about prurient or inappropriate interest in someone's medical history.

John Edwards: I imagine that Sir Ian is very well informed to assist the Committee with this. Again, there is no perfect answer. I don't think that we need to say that there is only one response. A trusted research environment is a very good way of enabling research on live datasets, but a pseudonymised dataset with synthetic data introduced, with differential privacy, with a number of emerging techniques that we identify in our report on anonymisation and pseudonymisation—our guidance from the Information Commissioner's Office—can assist. Again, the problem identified by Professor Goldacre is not entirely eliminated with the trusted research environment, because at some point you still have to trust the researchers not to abuse the identification of the individuals in those datasets.

Chair: We hope to put some of these questions to Professor Diamond, but we have lost the line to him in Cardiff, so we will stick with you, if we may. To that end, I turn to my colleagues Chris Clarkson and Graham Stringer, and we hope to get Sir Ian back.

Q278 Chris Clarkson: Sorry, John, it's just you now. What further steps, if any, are needed to give the public more confidence in data sharing? I am thinking particularly about how their data is shared with private companies. I think there is always an inherent assumption that when they are sharing it with the state, there will be an extra level of control, but obviously sharing among private companies will be an increasingly important part of this ecosystem. What additional steps can the Government take to ensure robustness and confidence in the way people's data is handled?

John Edwards: I think transparency, being open with the public about the kind of value proposition and the steps that are taken to ensure the safety of the data. This again takes us back to the concept of information governance. My office plays a role in that to reassure the community that we are part of that ecosystem and are looking out for them. Increasingly,



HOUSE OF COMMONS

there could be a role for us to provide ex-ante assurance. We often come along afterwards and assess whether a project has been undertaken in conformity with the law. Increasingly, we might play a role in assisting those proposing the data sharing projects that you suggest to design them in ways that can give the public confidence that the benefits can be derived without the risk of their data getting leaked, lost or abused.

Q279 Chris Clarkson: Is this the sandbox that you were talking about before?

John Edwards: I think the sandbox concept has been proven and has a really important role. There are other possibilities that I am exploring at the moment, including giving a commissioner's opinion up front. We put a lot of effort into chasing after things where they have gone wrong, investigating complaints, and then making a judgment about how the law applies to that scenario. If we front-load some of that, we might avoid the harms and provide confidence that the innovations we are trying to seek from responsible data use can be obtained in a way that is safe and provides assurance to the community.

Q280 Chris Clarkson: To what extent do you think the Government's current proposals strike the right balance between ensuring the data can be shared freely and the idea that the data should remain anonymous, or private and secure?

John Edwards: The national data strategy has at its heart a recognition that the legitimacy of those research projects, and getting value from a data economy, depends on a strong regulatory framework and a strong independent regulator presiding over that. I am confident that that remains a part of the Government's plans for the future.

I think you will be speaking to the Minister about data reform, and she will be able to give you a more up-to-date assessment of where those plans are. Certainly, in the consultation document that came out last year from DCMS, there were proposals to enable and facilitate research, but again within a framework of a strongly regulated and overseen regulatory framework.

Q281 Chris Clarkson: What input have you had into those proposals?

John Edwards: We have acted as a trusted adviser, so we work very closely with DCMS. The public role we have displayed has been one of advocacy for strong information rights and the protection of information, but also the safe use of data. We made a public submission. My predecessor signed that off just before she left post last year, and I endorse the comments she made about the proposals. Since then, we have been working very closely with DCMS to try to understand the policy objectives and to assist the Government in identifying legislative options that meet those objectives without sacrificing the strong protection of rights that people in the United Kingdom enjoy.

Q282 Chris Clarkson: Finally, we have heard in various other sessions that most people think that the existing legal framework is there and what is needed is better guidance. How much would you agree with that?



HOUSE OF COMMONS

John Edwards: I do think that the existing framework is robust. I also agree that it requires a bit of explanation to make it practical, and so that people can apply it more easily. At the ICO, we need to put our money down rather than leaving large portions of the economy guessing about how the law is going to be interpreted.

Having said that, the Government have indicated their intention to provide further legislative clarification, so that is the process we are in at the moment. The UK GDPR will be repealed and replaced with a bespoke data protection regime for the United Kingdom. I am expecting that to provide some element of that clarification in statute but that it will not remove from us the obligation to provide more tailored guidance and to be a more active regulator in the sector, providing certainty for researchers and those others who wish to make innovative uses of data.

Q283 **Chris Clarkson:** Sorry: I said “finally” but I am going to ask a final follow-up. Would you say the Government are moving in a positive direction with these proposals?

John Edwards: Everything I have seen of the proposals I would find positive. There were some reservations that my predecessor expressed in her submission on the proposal, and I shared them. We have worked with the Government and I am looking forward to seeing what the final outcome is. There is a recognition among the community and in government that there are great benefits from the European Commission’s adequacy determination and that that should not be set aside lightly. That is a reasonably guiding principle about the extent of the reforms and I am pretty confident, from what I have seen, that nothing that is proposed would threaten adequacy and that there is much that will improve the legislation and provide more flexibility, particularly for my office, to take a more risk-based approach and to target our resources to ensure that we can assist the economy to get the greatest benefit from safe data sharing.

Chair: I think we have Sir Ian Diamond back on the line, but with audio only. Sir Ian, can you hear us? No, apparently not. We will stick with you, Mr Edwards.

Q284 **Graham Stringer:** You are our star performer this morning, Mr Edwards. How do we define science research for the purposes of data sharing? How firm can we be in our view of the boundaries of science research and anything else?

John Edwards: Sorry, Mr Stringer, but could you elaborate on the question a little?

Graham Stringer: As I understand it, science research is going to have more freedom to interrogate datasets than other people. To enable that, you have to have a definition of science research. I am just wondering how firm the boundaries are and how they would be tested.

John Edwards: The concept of the use of data for research is permitted in most data protection regimes but, as you allude to, there is a range of interpretations of the term research. Some jurisdictions take a very strict



interpretation—for example, saying, “Research is only that which is conducted in university institutions.” I prefer the more common usage, which I suspect is something along the lines of, “The ability to derive new knowledge from existing data. For me, that is fairly enabling.

I think it is less about whether we can pin down that definition of research than about whether we can ensure that it is undertaken in ways that are proportionate and safe and that derive a benefit for society. I hope Sir Ian will be able to join us and elaborate, but if you look at the five safes approach, which is the foundation of the trusted research environment, those safes look at things like whether it is a safe research project—“Is there scientific value in this in the first place?” That is a really important inquiry to make at the outset: is this research project going to provide a valid and useful output?

Q285 Graham Stringer: Are you saying it is an emphasis more on the quality of the research than the researcher? Both are important in terms of the integrity of the process.

John Edwards: Exactly. Those are two of the five safes, I think. We want quality in the research and proposal. Is this actually going to derive a significant benefit for the community, and does the researcher have the integrity necessary to be entrusted with this data?

Q286 Graham Stringer: We have been told that if the Government’s starter proposals threaten the United Kingdom’s agreement with the EU on the data adequacy agreement, that has the potential to cost business £1.6 billion. Do you accept both that there is a threat and the quantification of that threat?

John Edwards: To come to your second question first, I have seen a number of different numbers. I am not an economist, so I am not really in a position to challenge them, but there is quite a significant range. Everybody who has looked at that question agrees that there is quite a significant economic benefit to retaining adequacy status and that there is a concomitant risk to losing it.

To the first part of your question, on whether there is a risk, I suppose there is, but I do not believe that, viewed objectively, there is anything in the Government’s proposed data reform that, with some tweaking, can be demonstrably shown to be essentially equivalent to the European standard of GDPR. That is the test that the European Commission will have to apply. Is the output of the law reform process essentially equivalent? In my view, when a European asks the question, “Is my data safe and protected in the United Kingdom?”, the answer, I would hope, would be, “Yes, it is as protected as it is back home.”

You mention risk, and it is a good word to use, because there is a spectrum, and there are a few issues in the initial consultation proposal that I was concerned could risk adequacy. For example, aspects of the proposal could be seen as impinging on the commissioner’s independence. My predecessor, in her submission, pointed out that if those were to carry through into legislation, that could be taken by the European Commission



HOUSE OF COMMONS

as undermining a fundamental aspect of the safe regulatory environment required to represent adequacy.

Q287 Graham Stringer: I don't want to distort your answer, but I am keen to get clarity. You are saying that there is not a lot for us to worry about in this area of our relationship with the EU?

John Edwards: I don't believe so. The Government still have a couple of decisions to make, and I think these are questions that the Minister would be best placed to answer when you get the opportunity later today. As I say, when the European Commission comes to look at the law and examine it, I hope that it will see everything that guarantees a sufficient level of protection of Europeans' data in the United Kingdom.

Q288 Chris Clarkson: I want to briefly touch on the use of artificial intelligence, which is obviously a developing and increasingly changing environment. What changes will you as an organisation be making to ensure better regulation and understanding of how AI works within the data biosphere, as it were? Secondly, I would be interested to know your thoughts on the Government's proposals for AI. Do you think they are adequate? Do you have any concerns or thoughts?

John Edwards: We are very conscious of the potential for AI to release a great deal of value from data. We are conscious of the ethical complexities involved in some AI applications, and we are working closely with the Alan Turing Institute and others to ensure that we are at the cutting edge in understanding those issues and are there to provide a resource for those working in AI.

To the second part of your question, there were a couple of proposals in the consultation document that was issued last year by DCMS that touched on AI, and my predecessor made a submission saying that, for example, she would be concerned if the right to have human review of an automated decision were removed. That was one of the proposals. I would share that reservation. We have to be able to give people the right to challenge the AI.

There is also quite a complex discussion in the paper about the difference between regulating AI for fairness of process versus fairness of outcome. I wonder how practical it is to distinguish between those, particularly when you look at fairness being almost a golden thread that runs right through the GDPR. It seems to me quite a troublesome and complex idea to seek to carve out a fundamental thread in respect of one set of technological applications, but to leave that as the foundation for the regulation of the rest of the economy and things. That seems complex, and I am looking forward to seeing the Government's response to submissions on that point.

Q289 Chris Clarkson: Very briefly on that, obviously legislating has to be done in the moment. I think we all accept that AI is going to look and act very differently as it develops. Do you think the regulations are adequate to keep pace with that, or do you think it is something that is going to need constant review?



John Edwards: I think everything needs constant review in this area of technological change. It will be the role of my office to be patrolling that boundary, if you like, and applying the existing regulation to the new technologies and notifying legislators if we believe that one has outpaced the other. At the moment, our legislation is, in the main, technologically neutral, which I think is really important and gives a bit of agility. Once you start identifying an emerging technology and legislating specifically for it, that is when I think you run the risk of getting outpaced, because legislation does not happen quickly, but technological development does.

Q290 **Chair:** Just on AI, you said that you attach great importance to a human review. How is it possible to have a human review when the essence of AI is that it is, in effect, a black box that is making use of massive computing power and connections that are literally beyond the human brain? Surely, a human review does not make sense.

John Edwards: Thank you, Chair, for that question, which highlights what I think is one of the great problems with the discourse in this area: if you have five people in a room discussing artificial intelligence at this moment, each one probably has a different conception of what you are talking about. At one end of the spectrum, there is automated decision making based on a series of rules or criteria that have been proven to equal a particular outcome—an eligibility for a benefit, or a likelihood of having conducted oneself in a particular way. At the other end, which I think you are describing, it can involve quite deep neural networks and machine learning, and what is sometimes described as designed unpredictability, in which the person who has written the code for the programme is unable to explain an output from a set of data inputs. I agree with you that that is extremely complex.

For that reason, I think we should have rigorous review and some scepticism of that level of automation being deployed in applications that can be making decisions that affect people in a way that might require human review, if you understand what I am saying. I do not think we should be releasing that kind of AI to make a whole lot of administrative categorisations or decisions about what people are entitled to. I do not think we are at that point yet; I do not think we are deploying AI to predict criminality and arrest people before they get the chance to offend, or that sort of thing. We would need to have a strong ethical framework for the commissioning of the applications of AI, and a strong framework of legal oversight.

Q291 **Chair:** You say that we need to be sceptical about it. You are a regulator; you have enforcement powers. How does scepticism translate? You do not have any powers to stop the deployment of AI. Your powers are around data and its collection and sharing, not the analysis of it, surely.

John Edwards: I sketched out two points at each end of a spectrum in those examples. I think there is a midpoint where there are AI offerings where there is a black box, and a vendor is motivated to extol the virtues of the box and apply it in commercial or governmental settings, but they



assert a commercial imperative not to be accountable for the nature of the algorithm or the input.

We have seen this in the United States, for example, in situations where sentencing guidelines for judges use a commercial algorithm, which has in one case resulted in an individual being sent to jail for eight years based on an assessment of risk generated by the algorithm—sentenced to eight years in jail for stealing a car, for example. In that case, the individual said, “On what basis?” and sought access to the algorithm all the way to the Supreme Court. The Supreme Court—I think it was of Wisconsin—upheld the right of the vendor to withhold the proprietary information inside the box over the individual’s right of explainability and the right to see how a decision had been made that directly affected his liberty.

There is a need for transparency. I would favour an approach that allowed visibility of those algorithms in some way by some independent regulator such as myself before they were deployed, in order to provide some assurance that they had been robustly tested. There needs to be the development of a market in auditing algorithms, for example, and AI applications.

Q292 Chair: Is that not literally impossible? It is not a question of algorithms that are written down for inspection. As you say, these are deep neural networks that are making connections that are literally beyond writing down.

Take a perfectly benign purpose such as medical research. You could imagine the detection of a disease based on millions of data points not just from this country, but from around the world; a certain set of conditions leads strongly to a conclusion that a person has or is at risk of having a disease. It might be completely impossible to write that down in a paper. That can be deployed, but you do not have any powers to prevent that from being deployed, do you? Let me clarify what I am saying. Your powers are around the collection of data, not the analysis that is made of it. Is that right?

John Edwards: I think the law is flexible enough to touch on both. You give an interesting example, because we are seeing fantastic deployments of artificial intelligence that learn how to read pathologic slides, for example, and to detect carcinomas far more accurately than human review, right? If I try to think of that in the data protection framework, there is a great benefit to the individuals involved, isn’t there, because their slide is being read more accurately?

If there was a deployment of a technology of that sort before it had been proven by a scientific method and before the clinicians were able to have the confidence in its accuracy, and that resulted in someone’s diagnosis being missed, there would be a data protection component. You would say, “Did the system accurately process the data that was presented by the individual patient?” If that technology was not fit for purpose, I think you would say that the organisation needed to be accountable for that.

Q293 Chair: What are your powers in that particular instance? What are your



HOUSE OF COMMONS

statutory powers to intervene?

John Edwards: We can intervene where there are complaints about the accuracy of personal data.

Chair: But that does not trigger it. The data might be completely accurate; it is about what is being done to it.

John Edwards: Accuracy and fairness: is this a fair processing of data given the steps that were taken to deploy the technology in this particular state? There would be scope for that.

Q294 **Chair:** Of course—if it is done fairly without any bias. What are the triggers for your ability to intervene if it is done fairly and the data is being held privately?

John Edwards: Again, an individual is submitting their personal data for a process of diagnosis, and they are entitled to have that processed accurately. They may make a complaint that they have suffered an adverse outcome because of flaws in the processing, and that could be through a human intervention or an AI intervention. I think that the kinds of questions you are raising are really important. I do hesitate a little about going too far into hypotheticals, and I apologise for my inability to respond on the full range of potential applications of AI and our ability at this stage, under the current regulatory regime, to respond to imagined potential harms.

Q295 **Chair:** They are important questions, and these things are not in the future—this is not tomorrow's world—but are happening now. One purpose of the Committee's inquiry, and one point of public policy, is to anticipate what might be needed. As you know, it can sometimes take time to think about very hard questions such as what the appropriate regulatory regime might be. Generally, the danger of regulators—this is not particularly directed at you—is that you are regulating a world of a few years ago rather than the world of now or the world that is about to dawn imminently.

John Edwards: I agree.

Chair: So what are your preparations as an organisation to get your head around, and think of the right responses to, these imminent questions and problems?

John Edwards: May I first return to the previous question about some of the hypotheticals? You are right about the time to think about it. I wonder if I could beg the Committee's indulgence to take a little more time, workshop some of the kinds of scenarios that we have started to explore in this conversation and return with further written evidence.

Chair: Yes.

John Edwards: In terms of how we are readying ourselves, we are working with colleague regulators internationally—this is the cutting edge of data protection regulations, so we are working on international



standards—and domestically, both with research partners, such as the Alan Turing Institute, and with the Digital Regulation Cooperation Forum. AI is a cross-cutting issue. I know that your focus is on medical and science-based research at this stage, but AI will transform our economies and a range of industries. At the DRCF, I am working with my colleagues at Ofcom, the CMA and the Financial Conduct Authority, and AI is on our work plan for that joint work to identify the cross-cutting issues and the regulatory responses to them that we might bring in an interconnected way.

Chair: Thank you very much indeed. I think we now have Sir Ian Diamond on the line. Can you hear us, Sir Ian?

Professor Diamond: I can hear you, Chair.

Q296 **Chair:** I am sorry we have not had the benefit of your presence throughout, but I will ask you a couple of questions now. On the question of artificial intelligence and its challenges for transparency, which is very dear to the UK Statistics Authority, is a world in which there is more of a black box as to why conclusions are made something that you are thinking about at the UK Statistics Authority?

Professor Diamond: One hundred per cent. I think it is important to make a couple of points. First, as has already been said, the ethical issues here do need to be thought through. One issue that the UK has to address is that a large number of groups are thinking about the ethics of AI, and there needs to be some convening. I hope that DCMS will do that, and it would be honest to say that we have thought about doing it—bringing groups together so that we have a set of clear ethical procedures.

Secondly—and I will come to one of your key points in a moment, Chair—at the end of the day, what any algorithm does, in the main, is search through data. Those data could be qualitative, as opposed to just quantitative. If I take quantitative data as an example, it is absolutely critical that you are aware of the biases and inconsistencies that may exist in your data. That is fundamental. There is a real role, I think, for analysis and approvals powers—of the sort that the ONS has through its trusted research environment—to really address in a transparent way whether the data that you are going to be looking at have the reliability that is needed to address the question. You may think, “Gosh, this is no different to the kind of extremely expert systematic reviews that have been going on for a while,” but it is just that you can look at those data much more quickly and with much greater granularity than you can simply when you are reading through a hundred or a thousand different papers from around the world.

With regard to the question of the algorithms themselves, I think it is incredibly important that we do not allow ourselves to delve into a world that says, “This is a black box.” I think it is and should be possible to explain what the algorithm is doing and how it is working. There has been some very good work by the great Australian statistician Ray Chambers that looks at the distinction between understanding the social processes



and the model you are looking at against actually just taking an algorithm that goes through searches. I think there are enormous benefits to be gained from the ethical use of AI, but it does require good regulation, and it requires that the regulators have a very strong and sophisticated understanding of the models and, critically, the data in order to assure ourselves that the results that are coming out are the best that they can be.

I would also say that it is absolutely critical that we do not allow ourselves to get into a world where, if you like, the computer is saying the answer is 0.8, but actually the answer will never be 0.8. There will be uncertainty around it, and we need to understand the uncertainty and always be looking at it. I am sorry that that was a long answer.

Q297 **Chair:** It is comprehensive, but it points to the fact that there is a lot of work and a lot of thinking to be done on this.

Professor Diamond: There is a lot of work to be done—no question—but there is also a lot of work being done.

Chair: That is good to hear. Perhaps you might, as John Edwards has committed to doing, write to the Committee to outline some of that work that is being done—in your case with the UKSA. I am afraid we need to move on to our next panel. I am sorry we lost you for most of that session, Sir Ian, but you are always generous with your time and I am sure you will be back before the Committee before long. Thank you very much indeed.

Examination of Witnesses

Witnesses: Lord Kamall, Dr Mark Toal and Jennifer Boon.

Q298 **Chair:** I am now going to invite the next panel to join us. We have Lord Kamall, the Minister for Technology, Innovation and Life Sciences at the Department for Health and Social Care. Thank you for joining us. Accompanying the Minister is Jennifer Boon, the deputy director of data policy at the Department for Health and Social Care, and Dr Mark Toal, deputy director of research systems at the National Institute for Health and Care Research. Thank you, all three, for coming to us and helping us with our inquiry. Perhaps I could start with a couple of questions to Lord Kamall. In July 2021, the Government published for consultation the report “Data saves lives” with a set of proposals. When do you expect the response to that to be published?

Lord Kamall: As you rightly say, it was published in draft in June or July 2021. That was before either I or my boss, the Secretary of State, Sajid Javid, were in post. One of the things that we were very clear about—I am very happy that they did this—is that it is important that we engage with stakeholders and take on board questions and issues. There was a really good programme of public engagement. Officials met with stakeholders. They ran online events. They had other forms of engagement and so on. They had to assemble that.



HOUSE OF COMMONS

Three things are going to come out of that: first, increased transparency; secondly, better information to improve how the public understand how their health and care data is going to be used; and thirdly, continuous engagement with the public.

I am afraid that I can't give an exact date at the moment, but I really hope it is as soon as possible. Personally, given that next week is London Tech Week, I think that would be a wonderful time to publish it. I know that it is imminent.

Q299 **Chair:** Okay, so as soon as that. It is imminent. The hope is that it is nearly ready to come out.

Lord Kamall: I can't say, obviously. You have been a Minister yourself, so you know what Government processes are like.

Q300 **Chair:** Absolutely. It is helpful to know that it is imminent.

We have the Data Reform Bill that will go through Parliament. Have you and your Department had an input into that?

Lord Kamall: DCMS is the lead Department—as with any cross-Government initiative, there is a lead Department, and DCMS is definitely the lead on this. I know it has engaged with a number of Departments, including my own. My officials in the NHS transformation directorate and officials in the Department have been working with their DCMS counterparts as part of that cross-Government work. That will form part of the upcoming Data Reform Bill. It is not only the Department of Health and Social Care and the NHS; there are other Departments. My officials definitely are involved. I had a brief phone call with Minister Lopez to make sure that we were completely co-ordinated and understood.

Q301 **Chair:** Your official, Jennifer Boon, was nodding throughout that. Ms Boon, you are deputy director for data policy at DHSC. Would I be right in assuming that you have had a close personal involvement on behalf of the Department with the team in DCMS that is taking the Bill through?

Jennifer Boon: That is absolutely right. As the Minister says, DCMS is the lead Department, but I and my team work really closely with our equivalents. It is something that we have been working on hand in glove.

Q302 **Chair:** Thank you. Minister, you may have seen that we discussed in previous sessions the GP Data for Planning and Research initiative, which was subject to a rapid erosion of public trust, leading to people withdrawing consent. That was some months ago now. Have you had a chance to look back at that and to consider what lessons should be taken from that episode?

Lord Kamall: I don't think you will be surprised to hear that there have been a lot of conversations around what we have learned from GDPR. It is really important that we take on board those lessons. It highlighted that we have to have, first of all, the right safeguards in place. We also have to have the right narrative and engagement. We have to make sure that we adhere to principles of transparency and security, but also trustworthiness.



HOUSE OF COMMONS

The trust is a really important issue. That is probably the main learning, as well as all the others, about health and care data.

I know most Departments would say, “Our data is really important,” but if we think about health and care data, it tells a story about a person, and quite a personal story. Therefore, people will be concerned about how that data is being used. In some ways, if you think about it, in this age of Facebook, we have all gone up a learning curve. When people first started going to Facebook, it was, “Isn’t it wonderful? I can share my data—I can meet old friends” and so on. Then they realised that there was something behind this free service—“What is being done with my data?” We have seen stories about that recently. Clearly, we have to make sure there is trust and transparency and a complete understanding.

On top of that, what I am really concerned about personally is that we make sure that we take advantage of the UK’s diverse population. If people particularly from minority communities do not trust the systems and the data—we saw that with vaccine hesitancy in some communities—we will not have a rich, diverse set of data. When you combine that with so-called AI, you may find bias. It is not intended biases, but unintended biases. Overall, if I could quote one of my favourite philosophers, Jimi Hendrix, we want to build a system that works and in which people have trust. If there is no trust and people opt out, it will be a castle made of sand. As Hendrix said, castles made of sand fall into the sea eventually.

Q303 Chair: Indeed they did, and that programme did. Given the importance—for the reasons you say—of having a very high level of engagement and, as I say, few opt-outs, which in turn requires trust, that must have been known in advance. How was it that, knowing how important trust and the basis of trust was, the Department blundered into something that did not command trust?

Lord Kamall: The straight answer is that I genuinely do not know. The important thing about this is that we can have all these inquiries about where it went wrong, but we have to learn the lesson. I do agree with that slogan, “data saves lives”—that is really important. We have to modernise the health service, but I also worry that if we do not get this right, we will have an imperfect system and people will opt out. I genuinely do not know what went wrong. I have worked in technology in the past, and sometimes people get carried away with a technology solution and think it is wonderful, but do not necessarily take the people with them and do not really understand it.

However, they have learned. In a number of conversations I have had, people have said, “We’ve got to make people aware of how their data is being collected and what it is being used for.” At the end of the day, no matter how much you explain, some people will want to opt out; we have to understand that, and also the responses. One of the responses that has been really interesting is the shift from data sharing to access to data and the concept of trusted research environments, but we have to be clear: we can say it is trusted, but the public also have to trust it.



Q304 **Chair:** Indeed. Obviously, you have the benefit of the Goldacre review. You responded positively to that, and that may be sufficient to address the particular problem, but you are now a Minister with many responsibilities in the Department. That misplaced instinct—that it could be bundled through, and there was not enough thought about public confidence—could apply in other areas of the Department’s work and the NHS’s work. Are you learning the lessons, not just putting right this particular project but making sure this could not happen again in other areas where it is absolutely essential to maintain people’s trust, which are legion in the NHS and healthcare?

Lord Kamall: I think that is really important. What we have to say is that we haven’t just learned from this one incident; we have got to continue learning.

One of the issues with engagement is that there are people who probably warned about this at the time, and I am very keen that we engage with them. For example, I have had meetings with Ben Goldacre. I have spoken to the National Data Guardian, and last week I met with Phil Booth from medConfidential. If these people who spend their life looking at these issues are warning us about something, we have to listen to them, take their advice, and see how we can address their concerns.

Chair: Thank you very much. I will turn to my colleagues: Katherine Fletcher first and then Graham Stringer.

Q305 **Katherine Fletcher:** Thank you, Chair. Thanks for your time, ladies and gents. Dr Toal, can I come to you first? You are a member of the—not easy to read, if you are dyslexic—National Institute for Health and Care Research, which is obviously an incredibly august body. What are you doing to try to understand the trust deficit, if I may put it like that?

Dr Toal: There are several ways in which the NIHR is engaging with the public and service users around data and research questions, working in partnership with others.

For example, alongside other funders, the NIHR contributed funding and support to the Understanding Patient Data initiative, which was hosted by the Wellcome Trust. Through commissioning and undertaking public attitudinal research, UPD helped to bolster our understanding of public perspectives in relation to data sharing and research. Through the work of UPD, we know that the public willingness for data to be linked and shared is high, because people want to see the quality of their health and care improved through evidence-based research. However, that willingness relies on our ability to explain clearly what people’s data is going to be used for, who is going to be using that data and what the safeguards are for the public around patient data, and we need to take a diversity of views with us. That is one method through which we gained insight.

We have also worked with others to support citizen juries in and around data questions. For example, in 2020, one of our patient safety translational research centres and the Information Commissioner’s Office ran a number of juries that considered the use of AI for decision making in

healthcare, and the learning from that is feeding into ICO policies and decision making.

Q306 **Katherine Fletcher:** Could you summarise some of the learning?

Dr Toal: A number of diverse learnings came from that. As well as looking at the AI piece, we have funded work that looked at data sharing during the pandemic. The juries concluded that the Government were right to use emergency powers to share patient data to support research objectives during the pandemic, but that greater transparency is needed. It comes back to the question of trust and transparency. That is the main overriding learning from the juries that we have helped to support in the past.

Q307 **Katherine Fletcher:** When they say transparency, what do they mean? Because one of the questions we are grappling with is, if the data is ultimately transparent, then they are going to know what is wrong with me—they will be able to look it up because the information will be on the internet. I grant you the trust question, and I will return to that, but what is their definition of transparency?

Dr Toal: I think it is not just transparency about particular data; it is knowledge and transparency about the research questions that are the focus of our studies. It is about who is involved in taking forward that research and which specific research groups are involved in accessing and using data. It is about transparency across the board.

Q308 **Katherine Fletcher:** So it is not about the transparency of individual data? I want to make sure that I am following you. People want to know that there is not someone—think of the best Marvel Comics baddy—who has access to the data and is developing some drug that will kill everyone called Katherine from Manchester. They want to have a sense of what is going on.

Dr Toal: Absolutely. There are checks and balances in the system that make sure that, in terms of research that is largely data driven, the Health Research Authority and other parts of the system provide information to the public about ongoing and upcoming research. It could maybe be made more accessible, but it is certainly out there.

Q309 **Katherine Fletcher:** Was there any view on a dynamic opt-in and opt-out? Is that something that you have spoken to your bodies about? I can see a situation where, down the pub, somebody goes, “I am happy to have my data used for cancer research, but I am not happy to have my data used for”—I can’t think of anything right now. It would be a categorisation of what people are happy for their data to be used for. Have you got that far down?

Dr Toal: In terms of considerations on the opt-out, that is sort of outside the space of NIHR. Maybe I could turn to Jen.

Lord Kamall: I am happy to take that question. Once again, this is part of our learning. What we want to be aware of, and we want to make sure that people are aware of this as well, is that if you opt out, you opt out forever. What if you want to opt back in? There has to be some dynamism



there. The second thing is that there have been a number of ideas. All I can say is that we are considering that—if that makes sense. We want to talk to the public about how best it would work in practice, and how it can balance difficult needs. People want the NHS to function effectively, they want data to improve their healthcare, which means they are happy to give access to some of their data—not personal stuff—to make sure we have the appropriate data set, and they want the NHS to be able to plan and function more effectively, rather than people making phone calls and finding the care home is engaged.

Katherine Fletcher: Fax machines.

Lord Kamall: Exactly. At the same time, people also have concerns—as we learned from the GDPR process. It is about getting that balance right. One of the things I can say is that we are looking at a number of different options as we consult on opt-out.

Q310 **Katherine Fletcher:** That is cool—thank you. I will turn to you again, Lord Kamall. We are politicians, and when we go, “Trust me; the data is fine,” the British public, in their complete marvellousness, stand there with their arms folded going, “Yeah, right.” How are we going to bridge the trust gap? What are we going to do to allay the trust gap where we say, “A private enterprise is using your data, but it’s for the public good”?

Lord Kamall: I think, again, we are all on a learning curve. I was talking to a doctor the other day and she said she had asked one of her patients why they had opted out and they said, “Because I don’t want my data shared with Facebook.” And she said, “Well, there was never any chance of your data being shared with Facebook.” It comes down to that level of understanding. I don’t mean to be patronising, but we have to educate the public: “We are using your data. Sometimes it will be your data; sometimes it will be anonymised and there will be large datasets.” We have to go down to that sort of level of detail. We have to say what we are using it for, as Mark says, in terms of research or planning.

I think that is the sort of level we have to go down to, and one way of doing that is this. There are people who have been working in this space for a long time—you have had some of them before the Committee today—and we are engaging with them. I had a conversation last week with Phil Booth from medConfidential and said, “Look, tell us about trust. What’s your definition of trust?” Also, there is lots of public engagement: there are focus groups; there are citizens juries. We have to learn from the public what they mean by trust. As you say, politicians saying “Trust me” is not going to go down very well, frankly—sorry to the other politicians.

Q311 **Katherine Fletcher:** The great British public—I keep referring to them, but this is a national collective effort. That is the whole point of the data. If we can find a way of us all piling in together, that makes it easier to anonymise and more effective in its outcomes. We value and cherish the NHS hugely, and I have heard quite an interesting idea, which is that there are potentially commercial benefits to be had from this dataset.



HOUSE OF COMMONS

Why can't we say to the British public, "Opt in. Here are the safeguards. And this is how we will pay for the NHS," because we will have people paying for access to that data? What assessment have you made of what are almost the incentives to asking people to join in, as well as closing the trust gap?

Lord Kamall: That is an interesting proposal, but generally on the terms of commercialisation of data, or if this is seen to be commercialising people's data, I think we have to be very careful. In some ways, there are, if you like, two extremes. Some people say, "Of course we should be making commercial deals and charging for the sharing of data, because we can invest that money back into the health service." Others say, "Whoa, whatever you do, do not sell that data. This is the NHS for sale; this is my data for sale." We have to get the right balance. There are partnerships. The NHS cannot do everything on its own; it has to work with others—right across the spectrum, in research etc. What is important is that people trust—I know I am coming back to this word and we have to understand what trust is—what is being done with their data and the benefits of sharing the data. For example, if Cancer Research explain why they need to capture large datasets and how that will improve cancer research—

Q312 **Katherine Fletcher:** I do not want to put words in your mouth, but it sounds like your reaction to commercial companies—and they are commercial companies that are looking for solutions to healthcare problems; they are not using the data for daft stuff. It sounds like your reaction to them making a financial contribution to use the data is a bit "no".

Lord Kamall: I think it is a bit more nuanced than that. What I mean by that is that I know, for example, that NHS Digital does try to recover the cost of getting the data in the first place for the partners, but it doesn't seek to do that at a profit. There is a debate in the public sphere. I have heard this. Some people say, "Charge for it. You can invest the money back in the NHS." And others say, "No. Whatever you do, you're commercialising the NHS." We have to tread that very difficult tightrope.

Q313 **Katherine Fletcher:** Oh, so there is a viewpoint that by using individual people's data in a collective way, you are commercialising the British public. I am not a member of the NHS; it wouldn't be my data that we're sharing. Or is that seen as a joint—

Lord Kamall: I think genuinely there are scare stories quite often about "the NHS for sale" and any commercialisation would raise some red flags—on parts of the digital spectrum.

Q314 **Katherine Fletcher:** Okay. I am going to slightly play devil's advocate.

Lord Kamall: I thought that was your job!

Katherine Fletcher: Well, I'm just making sure people don't think I totally think this. We have a gold standard, diamond Crown jewel in the fact that we all came together after the war and said, "We're going to create this thing called the NHS and it's great." We have data there that



could potentially unlock major health challenges not only for our population, but for global populations. But we can only have a public sector solution to analysing that data and we are not going to use any of the potential commercial revenues for access to that data to help to keep funding the NHS as it moves forward.

Lord Kamall: To some people, that will sound like a reasonable proposition, but others are really worried about any commercial aspect to data. I think the balance is this: the NHS cannot do everything itself and so has to work with partners, and with some of those partners it is going to have to give limited access to data—not to share it, but to give limited access—so that, for example, drug companies and charities can use it to develop better medicines and devices. The benefit will be more than commercial; it will be better healthcare in the end.

Katherine Fletcher: Okay. Thank you.

Q315 **Graham Stringer:** I have a couple of questions on trust, to follow up on what Katherine was asking about. Dr Toal, how do you choose your citizens juries?

Dr Toal: We are one organisation that was part of a collective that helped to support citizens juries. We also have patients and members of the public who helped to shape our research and to choose which research questions to follow up on, so that is embedded throughout our research portfolio. In effect, we put out a public call to ask for patients and members of the public to put their hands up and volunteer to work with us, to ensure that patient and public involvement is there to help shape research. So we put out open calls to the public to come and join us in helping shape our research portfolio.

Q316 **Graham Stringer:** Would it be fair to say that there is an element of self-selection?

Dr Toal: There is. One thing that we are very conscious of is the need to work harder to diversify and make more representative the cohort of volunteers who have very kindly spared their time to come and help us shape our research portfolio. That is part of our EDI strategy, which will be coming out later this year.

Q317 **Graham Stringer:** Minister, how does merging NHS Digital with NHS England help to gain the public's trust, given that NHS Digital is there as the guardian of the data?

Lord Kamall: I think that is a reasonable question. It came up in the Health and Care Bill, and at first I didn't quite understand the issue. When I engaged with the peers who were raising it, I managed to arrange meetings so that they could sit down with my officials and the NHS, and they completely understand the safeguard that has now been given. Simon Madden—I know he has appeared before this Committee previously—sat down with those peers and went through how we will shift the arrangement we have now and how it will operate after the merger, to make sure that the NHS is not marking its own homework.



HOUSE OF COMMONS

Q318 **Graham Stringer:** But if they are merged and somebody running NHS England says, "Do this" or, "Do not do that," they are not independent, are they?

Lord Kamall: This is why in the Lords I made the commitment that when the Government transfer NHS Digital's powers to NHS England, we will have to use regulations to provide a statutory safe haven for all data. One of the peers who was, quite rightly, sceptical about this came away from the meeting saying, "Look, I am far more reassured," because they will be part of the process and we will consult to understand their concerns. In fact, Simon Madden actually said to me that he wants to be more ambitious than that; he wants it to be even more trusted than the current arrangement.

Q319 **Graham Stringer:** I feel like a stranger in a strange land, and this world is developing quickly. Can you just explain to me in simple terms how the trusted research environments are procured, and how you will ensure that there is transparent competition and that they are fully resourced and sustainable?

Lord Kamall: One thing that I had to understand when I came to the Department was Ministers' role in procurement. I don't go out and procure stuff. When I talk to the NHS and the Department about public sector procurement, they tell me that it encourages free, fair and open competition, and value for money; that it is in line with the Public Contracts Regulation 2015; and that it is assured through the NHS England and innovation spend controls for digital, and that includes the Department and the Government Digital Service. If you want more details, I do not have them in my head, but I would be happy to write to you.

Q320 **Graham Stringer:** If you could send us that and cover the point about resourcing and sustainability.

Lord Kamall: That does not necessarily fit within my portfolio, but I oversee it, so I would be very happy to write and task my officials to look into that for you.

Q321 **Graham Stringer:** Finally, Dr Toal, how can you ensure that the trusted research environments will be optimised to ensure that they are networks and can share data safely and securely between them?

Dr Toal: Commissioning and supporting our TREs is not a NIHR lead, but it is a lead for the transformation directorate of NHSE. They will use both technical and operational solutions to support networking and linkage of data across TREs. That will ensure that TREs are not silos, but can work with one another to support cross-cutting projects. This will be co-ordinated through a central team in the transformation directorate of NHSE and they will ensure that TREs act as a network, because of the rules in place around the structuring and functioning of TREs. TREs are going to use common practices and common data standards that support interoperability.

The approaches to interoperability might vary, and they will depend on data type, technical infrastructure and relevant governance. They will vary



HOUSE OF COMMONS

between use cases. But this is all under consideration in the transformation directorate in NHSE, working with other partners, and the proposals for how interoperability can be embedded in TREs will be brought out later this year.

Q322 Chris Clarkson: Lord Kamall, turning to the Goldacre report, how will the Government address the recommendations on pseudonymised data in pages 83 to 95 of the report which call for its elimination?

Lord Kamall: That is a really important issue, because we do want datasets, and once again it is the issue of trust. There are some concerns and a debate about pseudo-anonymisation or de-identification. As you all know, there is a spectrum on this. You can remove the direct identifiers—name, date of birth, address and so on, and Ben Goldacre did so—but there is still a risk that users may in theory be able to use additional information to re-identify individuals. We hope that secure data environments will tackle some of those problems and reduce the risk.

Can you eliminate all the risk in anything in life? I am not sure, whether that is here or elsewhere, but we have to try to minimise that risk as much as possible. Hopefully, it will be by having a secure data environment, which is kind of analogous to a reference library for those of us who can remember books and libraries. When you went into a reference library, you couldn't take books out of the reference section or take them home with you: you could consult them and take notes for your research purposes. That is the kind of analogy, but please do not ask me to push it too far as I am sure it would break somewhere. I want to make sure that we reduce the risks.

The problem is that if someone can extract the data and re-identify it with other information—especially for a public figure—it is a real concern. We want to understand the risks, and we are working with the Information Commissioner's Office—I think you have heard from them previously—which is consulting on pseudonymised data. I know that Minister Lopez from DCMS may be able to talk about this—I assure you I am not trying to pass the ball and run away—in terms of the context and wider UK GDPR.

Q323 Chris Clarkson: Do you have a metric in mind for reducing that level of risk, or is that still under discussion?

Lord Kamall: Yes, and I want to work with people who have been sceptical. For example, I know we are engaging with Ben Goldacre, who prefers TREs to sharing data. Phil Booth is an extraordinary civil libertarian and it is important that we engage with him. I have had conversations with the national data guardian and will continue to do so.

Q324 Chris Clarkson: The report also makes a recommendation for a dedicated NHS analytical function. How will that be resourced? Will it be part of the structural reforms that the Health Secretary is talking about? We have heard this week about "a Blockbuster system" in a Netflix age. Is this going to be one of our new streaming options?



Lord Kamall: Generally, any organisation has to adapt to the world around it. My PhD was in organisational change, and one of the things about organisations is that you should start with the environment that you are facing and with what is the ideal organisation to face those challenges, but also how can you improve the service that you give. At the end of the day, how do we improve health and social care? It is obvious in many ways how sharing data and access to data can improve that.

Clearly, as I have seen in my time here, a number of people not only from the public sector, but also from the private sector have come into the NHS with those skills. I am very impressed by the level of technical knowledge we have in the NHS. Clearly, at all levels of the NHS there will need to be more tech-savviness, whether that is clinicians themselves being tech-savvy or bringing in analysts and data scientists overall. That will be at various different levels in our system of health and social care.

Q325 **Chris Clarkson:** But this will be a specific new function, so where is that resource going to come from? Is it existing resource that is being reallocated? Is it new resource? Is it going to be a result of the changes?

Lord Kamall: I cannot answer that question in detail. Resource is a sort of NHS question, but I do know from my conversations—I sometimes sign off submissions to recruit so many people in this area of technology or data. Clearly, the NHS, like any other organisation, has to be more tech-savvy, has to improve its services, and it needs tech-savvy people, data analysts, data scientists, coders, etc.

Q326 **Chris Clarkson:** Turning to NHS Digital, how will the NHS Digital safe haven role for patient data be maintained after its merger with NHS England?

Lord Kamall: I thought Mr Stringer asked that question as well.

Graham Stringer: I did.

Chris Clarkson: Could you flesh the answer out for me?

Lord Kamall: I don't like this answer, to be honest, but this is the answer: it's with the lawyers at the moment. I gave an assurance at the Dispatch Box that we wanted to make sure that, after the merger, the provisions for safe haven were as good as what we have now. Actually, Simon Madden wants to go beyond that. A lot of that is technical detail, but also consultation not only with some of the people you have had before the Committee, but with the peers who raised it. I have always been very clear that I want to take the politics out of it. For example, during the Bill, I made sure that Simon Madden and others sat down with officials to hear their concerns and to address them. One of the peers who was very critical at the beginning came up to me and said, "That's brilliant. If he does what he says, I am going to be very reassured by that."

Q327 **Chris Clarkson:** Would you say you now have a clearer understanding of what the end product is going to look like?



HOUSE OF COMMONS

Lord Kamall: Yes. In conceptual terms, the NHS cannot be marking its own homework when it comes to sharing data, and there must be someone who says, "Whoa, you cannot do that for these reasons." That is what has to happen at the conceptual level.

Q328 **Katherine Fletcher:** You mentioned that your PhD was in organisational change, and I did a bit as well. Sometimes organisations cannot really change themselves. I am hearing you talk about signing off resources to start to create this. Are you convinced that this answer is solvable within the structures of the NHS, or does it need external advice and resources to sort it?

Lord Kamall: I think it needs both. The NHS knows that it cannot do everything itself on a whole range of issues. You will remember the national IT project—

Katherine Fletcher: I remember that a lot of people made a lot of money off it and it never happened.

Lord Kamall: Once again, that comes to trust. The system clearly did not work, and therefore the NHS cannot do it itself. It has to work with partners. Some of that will be about internal skills to ensure that they are working with whichever companies are being worked with.

Q329 **Katherine Fletcher:** Is the key to this making the NHS a good customer?

Lord Kamall: Making the NHS a good customer?

Katherine Fletcher: My background is in delivering IT, so I did have colleagues who were working on the NHS IT problem. The thing that is hardest to do is not to apply technological innovation; it is to make sure that people don't ask for A on Monday and then D by the time you've got to Friday, because that is really expensive, very disruptive and nothing happens. Is the key to this whole environment making sure that the NHS is a good customer, as in it is being clear about what it wants?

Lord Kamall: I would say a couple of things. One is that clearly it has to be a good customer, but you and I, having worked in consulting or technology, know that clients are not always clear and change their mind throughout the process. The NHS is very clear that it wants to offer better healthcare, whatever that means. One way of doing that is via more digitisation and sharing data and access to data, with all the benefits of that. But at the end of the day it has to be about the patient. It has to be patient-centred. One thing that I am very aware of, having come into this role last year and from the House of Lords debates we have, is that we have to champion patients. It all has to be about patients. Yes, doctors and nurses are wonderful—in what they have done throughout covid and will continue to do—but it has to be about patients and how we make sure we deliver a better service for them.

Q330 **Chair:** Just a final point from me to Dr Toal. We obviously cannot afford to repeat the mistakes of the GP data loss of confidence. The trusted research environment that Professor Goldacre is recommending is, I think



HOUSE OF COMMONS

you would agree, the principal safeguard or, if not, one of the principal safeguards that is going to be proposed.

Graham Stringer asked about the citizens juries that had been convened. It is true that Professor Goldacre told us that that series of citizens juries found that "TREs were well understood and also strongly supported by the public." But he went on to say that those juries involved "days and days of hard labour, giving people information so that they could really dive deeply into the issues. I think that is the way we should do it." That gives me concern, because if you have a citizens jury that is, as you said to Mr Stringer, to a certain extent self-selecting people who, I assume, are interested in these things, and it involves days and days of talking them through the complexities and solutions, and they then come out strongly in favour, that is great, but it may not be a very good representation of what is possible to convince the public.

Have you thought about how communicable and how reassuring you need to be about trusted research environments, which I imagine will seem to most people a rather unfamiliar and opaque term? Are you placing too much emphasis on this fashionable device of citizens juries? Are you aware of the risk that you could be tiptoeing into another incidence of a loss of confidence that would be even worse than the first?

Dr Toal: Citizen juries are a really valuable tool, but they are only one tool among several in the toolbox that is public engagement. We do need to get better at communicating what is by definition a really technical area that is at the cross-section between data, technical detail and public trust and people's stories. I do not think citizen juries are the one and only way to crack that; we have to find different ways to communicate with the diversity of our population to really build that trust.

Q331 **Chair:** Let me turn to the Minister. When this is relaunched, I imagine you will be fronting it, as the Minister responsible in the DHSC. You are a very good communicator—

Lord Kamall: That is very kind of you.

Chair: Are you applying your talents and skills in anticipation of what is needed to achieve this very important target to convince people that their data is safe and that they needn't withdraw their consent? It needs some careful anticipation. Are you engaged in that personally?

Lord Kamall: I am not personally engaged in necessarily being front-facing on this, but I have lots of conversations in the Department and with the NHS about the processes. First, we have to respond to suggestions, such as Ben's suggestion. I would also echo the point that there is no one silver bullet on this.

One of the concerns that people have about citizens juries is that it is the people who always put themselves forward and have the time to put themselves forward for these things, so are you going to get a biased view? We have to make sure that we embed the rich diversity of the UK population when we are trying to understand what we are doing.



I clearly get the point about communication. We have conversations and I hear all sorts of ideas, not only about citizens juries, directly online, both push and pull, but also about working with civil society organisations—people who are on the ground and who understand some of the communities that are hard to reach. How do we engage with them? As you say, you cannot make it too onerous, either. If you had a local community group who said, “Let’s work with this particular community in our local area,” you couldn’t do it for three days; those people have lives to get on with and they have to go to work. So you have to get the right balance.

Chair: It has been very helpful to have all three of you to illuminate this subject for us. Thank you for your attendance today.

Examination of Witnesses

Witnesses: Felicity Burch and Joanna Davinson.

Q332 **Chair:** I welcome our next panel of witnesses—a pair of witnesses, in fact. Felicity Burch is Executive Director at the Centre for Data Ethics and Innovation—a Government body that advises the Government and specialises in questions about enabling the trustworthy use of data and also artificial intelligence. Thank you very much, Felicity Burch, for joining us. Joining her at the table is Joanna Davinson, who is Executive Director of the Central Digital and Data Office, which is part of the Cabinet Office. I begin with a question to Joanna Davinson. The Cabinet Office obviously takes a cross-Government role—that is its purpose. What are the key bottlenecks across Government to the effective sharing of data?

Joanna Davinson: As you said, Chairman, the Central Digital and Data Office is accountable for setting strategy and standards for data, digital and technology across Government. I would say there are probably three technical bottlenecks that we come across. The first is the ability to find the data that you want. We don’t have very many externally published catalogues, so often people depend on personal knowledge or have to do quite a lot of work to discover where the data is that they want to access. That also creates a challenge in that we repeat datasets, which is clearly not as efficient as it should be.

The second is the ability to access that data technically. We have a lot of different systems across Government, a long legacy in our estate, and we don’t have a consistent data technical architecture. So, often, when you have found the data that you need, you have to do a lot of work in order to build the technology so that you can extract it and use it in the way you want to use it.

The third issue is that we don’t have consistent standards in terms of how we hold data across different Departments, so, having found the data, having found a way to extract it from its source systems, you then need to do manual work on it to ensure that you can combine it in the ways that you want in order to bring it up to a consistent basis. Those are the technical issues. There are obviously some non-technical issues, which I know you have discussed in previous sessions, around culture, around



attitudes to risk, around people's understanding of legislation versus guidance, that also get in the way.

Q333 **Chair:** Can you describe your role in uncorking those bottlenecks?

Joanna Davinson: In the Central Digital and Data Office, we are accountable for the leadership of the digital data and technology function across Government, so that is about ensuring that we have the right capabilities in place to work across digital data and technology. We set strategy across Government—we set standards. We are actually a relatively new organisation, just over a year old. We are also accountable for assuring and performance managing Departments against the strategy and the standards. We do that through the exercise of a spend control power. So we have powers to assure new programmes and projects that use digital data and technology and we are accountable for assessing those to ensure that they are consistent with the standards that we have set and in terms of value for money. In the context of data, the Central Digital and Data Office is accountable for implementing mission three of the national data strategy, which is the part of the strategy that talks about improving data shares across Government.

Q334 **Chair:** In your assessment, sitting at the centre of Government, obviously you can monitor how well things are being done—you can evaluate it. Is there a muscular central push to implement the necessary steps to allow the safe sharing of data?

Joanna Davinson: I would say there hasn't always been. One of the reasons why we established the Central Digital and Data Office was to create more of a central focus on what we need to do in order to improve data sharing across Government, and we have done a number of things in the last year to facilitate that. We have established a post of chief data officer, so there is now a Government chief data officer. We are actually in the process of recruiting permanently to that post now.

The other thing I would point out is that we have established a chief data officers council, so we bring together all chief data officers across Government to get a consistent perspective and to drive a consistent view of standards and architecture, and to actually hold them to account for delivering that out in their respective Departments. But it's soft power rather than hard levers.

Chair: I see. Let me turn to my colleagues, starting with Graham Stringer.

Q335 **Graham Stringer:** The evidence we have heard is that guidance, rather than legislation, would achieve more in order to address the problems of data sharing. Would you agree with that?

Joanna Davinson: I think we have a legislative environment that does enable data sharing. My organisation is responsible for the Digital Economy Act, which enables data sharing across Government. There are some specific cases where there are legal impediments, just because of other primary legislation, that we need to address over time. But in most cases, I would say, we have the legal powers to do the data sharing that



HOUSE OF COMMONS

we need. The challenges are more around the technical barriers that I just described, but also people's views around risk and just their understanding of what it takes to share data.

Q336 Graham Stringer: Is there a measure of risk aversion that is inhibiting that?

Joanna Davinson: It is really important that people understand the risks associated with data sharing, because there are risks. It is really important that we encourage people to take those very seriously and to think about this. One of the things that we have been doing in the last year is that we have established, together with the Office for National Statistics, something that we call our data sharing playbook, which is looking at difficult examples of where people have not been able to share data and really investigating whether the issue was attitudes towards risk or there was actually an issue that was more technical. Generally, what we have found is that it's people's understanding of the guidance and of how they manage risk that is the issue, rather than it necessarily being just about risk aversion. It's more about getting that understanding. Where there is risk, it is important that we understand that, but there are ways in which it can be managed. Generally, we have found, as we have worked through the cases, that we have been able to find a way to manage it.

Q337 Graham Stringer: Professor Goldacre was worried that there was a major privacy risk even where the data was pseudonymised, if that's the right word. How do you deal with that? What assessment have you made of the risk, and how can it be mitigated? How can you deal with Ben Goldacre's worries?

Joanna Davinson: The first thing to say is that anonymising data does not remove the risk. There is always risk when you are sharing personal data and, as you combine more and more datasets, the opportunity to discover or to de-anonymise gets greater. So the risk is real. I know you have talked to other witnesses about the use of trusted research environments. The UK Statistics Authority is accountable for accrediting those and ensuring that they are operating to standards that manage and mitigate risk.

The other thing I would say is that you have to treat anonymised data in the way you would treat any personal data: you need to make sure that you have done a proper data protection impact assessment on how you are using it, that you have understood the risk involved and that you have mechanisms in place to manage it. There is no silver bullet for how you fix it. You just have to be really conscious of the risk and really conscious of the mechanisms that you have put in place to manage it.

Q338 Graham Stringer: How would members of the public be reassured by that? You are saying there is a risk and we need to understand it, but how can you reassure members of the public whose data is with the NHS or wherever that that data is not going to become public?

Joanna Davinson: Where we have trusted research environments, it is about ensuring that we have the right accreditation and assurance



processes in place so that people can be confident that those environments are being managed appropriately to protect their data. Behind that, I would go back to the data protection impact assessments: when any data is being shared, whether it is anonymised or not, there needs to be a proper and public data protection impact assessment that gives assurances around how the data is being handled.

Graham Stringer: Thank you very much.

Q339 **Chris Clarkson:** Felicity, how can greater transparency be achieved in the use of AI with personal data?

Felicity Burch: Thank you for asking me that question. At CDEI, we believe that transparency is a really key factor to build trust. I know that is something that has been quite a theme in the evidence to date. It is really underlined by a lot of the research we have done with citizens. For example, 63% of citizens polled said that they wanted to know more about how the Government use their data, so there is quite a strong call for that from society, but it has also come from civil society and international groups.

On the CDEI's role in enabling that, as an organisation we essentially exist to enable organisations to innovate with AI and data in a trustworthy manner. We help them to be more transparent in two key ways: first, on a project-by-project basis, and secondly, looking at it on a bit more of a system-wide level. I will just say a little bit about both of those.

First, on a project-by-project basis we work with a number of Government bodies and help them to design their data ethics frameworks. That helps them to understand how they are going to use data better. We have worked with, for example, Bristol City Council and the Greater Manchester Combined Authority on that. We have also worked with Police Scotland on their approach to data governance and their data ethics strategy as well. Those are the individual pieces of work we have done specifically to help those organisations.

At a system level—this is really important—we have been working with our colleagues at CDDO to develop a public sector algorithmic transparency standard. That is a bit of a mouthful—if anyone has a catchier name, let me know—but it is a really important tool to enable transparency. It has been a really nice combination of CDDO expertise and CDEI bringing our public attitudes research.

Over last year, we engaged with the general public to understand with them what meaningful transparency with the use of algorithms might look like. Interestingly, initial polls pointed to the public wanting to have available to them all the information about how algorithms are being used and the information that goes into them—they suggested that they might want really detailed information—but we followed up with deliberative research and presented examples of what that might look like, and quite quickly we realised that actually that is very difficult to understand.



HOUSE OF COMMONS

So what we have developed and are piloting at the moment is almost a two-tier approach. The first tier of information is very-high-level information about how the algorithm is applied and how to find more. The second level gives a lot more detail about how the algorithm works. That plays quite an important role in delivering and enabling scrutiny, but it is not necessarily something that members of the general public would want to go and look at on a daily basis. But it is there and available should organisations want to look at it.

That approach is being piloted at the moment. We launched it as a pilot because we really wanted to see how it worked in the real world, to see what is practical, and we have just published some of the initial findings from the pilot. There is some really good news, I think. First, we found that the organisations who were piloting it found that the tool helped them to have better conversations with the people whose data they were using—it helped them to think that through.

Secondly, it enabled better internal conversations about what the risks with using any particular algorithm might be and how they might put in place the correct mitigations. That really demonstrates the value that transparency can bring in a wider, more accountable use of AI.

We did find, though, that organisations wanted more information about who was responsible for filling this in and more information and specifics about exactly what information was being provided. We will be inspecting the standard over time. We will be working with other pilots and hope to be able to publish more about those soon. The next step will be to consider broader roll-out across the public sector.

Before I finish on that point, this really has been an excellent example of collaboration with the CDDO. This really is a world-leading example. The UK is the first country to publish a national algorithmic transparency standard. We hope that it will lead the way in how our public sector uses algorithms. We have already seen examples of other organisations around the globe looking at this to inspire their own approaches.

Q340 Chris Clarkson: It is very reassuring that that data will be accessible to people in an acceptable format. Because you could obviously just give somebody 600 pages of code and explanations and it might mean nothing to them. Conversely, when people look at how AI has been used to analyse their personal data, what mechanism is there for them to challenge the assumptions that that has produced? Is there any progress on developing a framework where they can do that easily?

Felicity Burch: Yes. Again, this question about challenge, accountability and redress is an important one. It is probably worth starting by saying that AI does need to follow existing regulations. Accountability for decision making and redress does already exist within regulations. I know regulators, such as the ICO and the CMA, have been working on how they interpret regulations in the context of AI. That regulatory framework plays quite an important role there.

At the same time, colleagues at DCMS at the Office for AI are looking at the governance of AI in the broader sense, and will be publishing the AI governance White Paper later this year, which I hope will get into some of that. I know that the aim of that is to enable effective research and innovation, but at the same time ensure citizens' trust.

In terms of what the CDEI has been doing in this space, we provide advice and guidance to help organisations when they are implementing some of these technologies practically. We have really focused on areas where we know from our research that the public sees the highest risk and is most concerned about implementation of AI.

An example is the recruitment sector. There are lots of opportunities for using data and AI in recruitment, to help organisations find more diverse candidates, the best candidates for the job, and potentially improve the candidate experience as well. The public are really concerned about bias creeping in through the process. For individuals who are, for example, neurodiverse the systems can be very difficult to interact with.

What we did was to partner with the Recruitment & Employment Confederation, so that we worked directly with the industry group to produce advice and guidance on how to use AI and data-driven technologies in the recruitment process. We looked at every stage of that process, right from just considering whether you are using it, through to that kind of end and follow-up side of things. We did recommend that organisations consider how they would explain to candidates how decisions were made, and whether they need to put in place right of redress.

I have a final point on this. I apologise, because I know this is a long answer. I think it is really important for organisations when they are innovating to ensure that they consider these factors that build in public trust. The more the public trust the way you are using their data, the more likely they are then to share their data and engage with these products and services. Long term, if we want a more innovative, data-driven economy, building that trust as a foundation is hugely important. We are there to help organisations innovate with trust.

Q341 **Chris Clarkson:** Thank you very much for that incredibly comprehensive answer. It headed me off at the pass on a couple of things. Is there anything further you think could be done to increase that trust? Is it a legislative remedy? Do you think it is more about empowering people to understand how the data is used? Is it a more robust legislative framework? Is it a voluntary framework? What is the optimal scenario for you?

Felicity Burch: It's a bit of a combination of factors.

Legislation and regulation can obviously play an important role in driving public trust, but there are other tools beyond that that can make a big difference. Everything, when it comes down to trust, starts with being trustworthy. Our research with the general public shows that the public trust in governance arrangements around the use of their data is the



biggest driver of their confidence in that technology. Making sure that the right governance arrangements are in place is valuable, as is giving the organisations the tools to do that.

One of the things that we are working on to drive that forward is delivering a mature AI assurance ecosystem in the UK. When we talk about assurance, we mean things like standards and audit. At the back end of last year, we published our AI assurance road map, which was a world first. It set out how we want to grow this industry in the UK. We believe that it is a real opportunity to help organisations, when they are procuring AI, to understand how the AI works—that is, whether it works effectively, but also whether it works the way that it says it does and avoids bias and so on. It will really help people to be better customers of AI from a business perspective, from a Government perspective and from a consumer perspective. Assurance can play a valuable role, in addition to legislation, alongside organisations taking steps themselves to build more trustworthy governance.

Joanna Davinson: May I add to that? CDDO has just published the data-sharing governance framework, which is intended to help organisations. It includes a set of principles on how to set up an effective governance that is open and transparent, and enables trusted data sharing. More consistency in how people apply such standards and frameworks will help.

Chris Clarkson: Almost like a kitemark standard for the application of the algorithm.

Q342 **Katherine Fletcher:** Thank you, ladies. I apologise for having to pop out.

Privacy-enhancing technologies are otherwise known as PETs—a wonderful acronym. If we are going to close the trust gap, it will be partly about framework and partly about being able to explain simply to people how their privacy is enhanced by technology. Felicity, can I come to you first? What role do PETs have in allowing people to have confidence in data sharing?

Felicity Burch: If you think PETs is a good acronym, wait until you hear about the CDEI's PETs adoption guide. It was one of the first projects I heard about when joining the CDEI. I loved the pun and I knew it was the right place for me.

Katherine Fletcher: For a gold star, get "Battersea" in there as well.

Felicity Burch: I will take that under advisement.

PETs can play quite an important role in unlocking responsible innovation. There is a range of privacy enhancing technologies that are variously difficult to pronounce, as I know the Committee has already experienced. They play a really important role in enabling sharing and analysis of data in a way that protects people's privacy.

Q343 **Katherine Fletcher:** Can you describe a simpler version for me, the tape and the general public?



Felicity Burch: A really good headline about PETs is that they enable you to learn from people without learning about people. There are lots of different examples of how you might do that. Let's take homomorphic encryption, for example. At the moment if you are trying to do a bank transfer, your encrypted data will get sent to the bank and they would have to decrypt it in order to process your information. With homomorphic encryption, they do not have to decrypt that data; they can process the transaction on the back of the encrypted data, so at no point would they see your personal or private information. That is a bit of an example of how they work.

Katherine Fletcher: So they are robust until quantum computing really gets going.

Felicity Burch: It is fair to challenge the efficacy of these technologies. We think they play a really important role. We have already heard about some of the risks around pseudonymisation. In fact, any individual PET might have a risk associated with its implementation. Our adoption guide goes into some of the risks and limitations, but we think PETs are quite an important tool to have in your toolkit, and we are really trying to drive forward their adoption. Indeed, we are partnering with the US on an international challenge to deliver more PETs.

Q344 **Katherine Fletcher:** Joanna, please come in at any point. Do you have any problems with driving adoption?

Felicity Burch: Sorry, what do you mean?

Katherine Fletcher: You are pushing for these things to be adopted across the board, so I was just wondering if you had encountered any barriers?

Felicity Burch: There are barriers to their adoption. They tend to sit around low awareness of the existence of the technologies themselves, a lack of expertise and technical limitations. Some of these technologies are still in their fairly early days, so the work we do with the US will, we hope, push forward some of the technical capabilities and demonstrate some use cases for the technology as well.

Q345 **Katherine Fletcher:** It is important that people appreciate the complexity going into the data security. I wonder whether you could give us one "for instance" about the technical challenges. It might require a written evidence piece to follow it up, but it is important that people have at least some sense of the levers that are being pulled in the background to try to keep data private.

Felicity Burch: I am happy to give an example and I am sure that we can follow up with more if you would like. One example during the covid crisis was the OpenSAFELY initiative, which was a collaboration between research universities and the NHS, so CDEI was not involved, but I think it is a good example. That was a project that looked at 24 million records and the purpose was to help identify risk factors for covid. That was done using a combination of tools, including federated analytics and



deanonymisation. That enabled us to move forward in our understanding of risk factors for covid while keeping that huge volume of data safe.

Q346 Katherine Fletcher: Is the technical challenge you are addressing as this moves forward literally that we cannot get laptops with powerful enough chips to process 24 million lines of data or is it something else?

Joanna Davinson: Can I comment on that? I think it is more that each Department is trying to do this on its own, or has historically done that, so the level of visibility we have on the adoption of technologies like PETs or any innovative new technology is not great in the centre. Equally, the knowledge sharing—lots of people try to address the same problem in siloes. I mentioned earlier that my organisation has established a chief data officer council. The purpose of that is to get the visibility of what is happening and to be able to identify the areas where we need to provide additional guidance or support people to share information. One of the sub-groups that we have just established within the chief data officer council is—another acronym—a data and technical architecture design authority. That is the technical people in the data world coming together and defining the key issues they need to work on together to make sure that we promulgate the right approaches across Government to technologies such as PET. At the moment, we in my organisation do not have a point of view on that. We have not done the work on it, but it is an example of something we would put into that architecture group.

Q347 Katherine Fletcher: So the technical challenge, as always with IT implementation, is as much about making sure that you have the right governance and structures and that you are not building something that then needs a lot of retrofit with a ton of interfaces or whatever—I am dating myself now in terms of the technical challenges. So it is not so much that there is a problem writing the code, it is about making sure that we are not writing code in 10 different places that do not talk to each other afterwards.

Joanna Davinson: Absolutely.

Q348 Katherine Fletcher: That is incredibly helpful, thank you. One final one, if I may, Chair. Trust is such a nebulous word, but Felicity, what actions will you take to make sure that the Government data reforms keep to ethical standards and keep that public trust? It is incredibly complicated, so how are you going to explain technical architecture to me, never mind—

Aaron Bell: Chris—*[Laughter.]*

Katherine Fletcher: Chris, who speaks a million languages and is a genius in his own way.

Felicity Burch: It is worth saying that while many among the general public do not necessarily understand how the technologies work, they do have views on how those technologies might impact on them in their day-to-day lives and are quite optimistic about use cases around health, for example. I have already signalled concern about recruitment. It is really



HOUSE OF COMMONS

important that we address those concerns. The data reforms that you mentioned play a really important part in that. I have already said that I think that legislation and regulation are one of the things that helps to build public trust. The CDEI works closely with the data policy team. When the CDEI was previously operating as an independent advisory body our work around algorithmic bias and decision making fed into the consultation around the national data strategy. Our team has also worked to support engagement with a range of experts in the run-up to the Bill.

As a wider point, our work actively demonstrates trustworthy uses of data, and it helps to give organisations technical solutions—for example, the PETs challenge, but also the new approaches and the right governance mechanisms that can help them to innovate with trust. We view our role there as complementary to the regulatory work that our colleagues are doing.

Q349 Katherine Fletcher: Cool. Just to simplify it, you are basically understanding what people are worried about; you are monitoring what we are up to; and then you are trying to demonstrate positive examples of what is going on. Is that fair?

Felicity Burch: Yes.

Joanna Davinson: I just wanted to add a comment. You make the point about communicating. It is a complex technical area, and data specialists in my world struggle to communicate to other technologists what we are trying to deliver. One of the things that we are doing is creating something that we call the digital data essentials—upskilling all civil servants in the basics around digital and data. The thinking behind that is that we are all going to have to get better at understanding how to use these capabilities going forward. It does not matter if you are a policy professional, data professional or a finance person, you need to understand it.

Part of the thinking behind that is that the more we are able to upskill all of us the better we will be able to communicate and design our policy so that it takes into account the need to explain and be transparent about what is happening. We will also get better at communicating because we will have more people able to speak in plain language.

Katherine Fletcher: You are identifying opportunities and spotting potential issues. You are increasing your capability: brilliant.

Chair: Thank you very much indeed, Katherine. Thank you to our two witnesses, Joanna Davinson and Felicity Burch.

Examination of witnesses

Witnesses: Julia Lopez and Jenny Hall.

Chair: We now move to our final panel of witnesses: Minister Julia Lopez, who is the Minister for Media, Data and Digital Infrastructure at the Department for Digital, Culture, Media and Sport. I am sure she has her job title off pat.



Julia Lopez: It has taken me a while.

Q350 **Chair:** The Minister is joined virtually by one of her officials, Jenny Hall, who is the director for data policy at DCMS.

Thank you, Minister, for attending previous sessions. You will have detected that the Committee is supportive of the use of data, especially for scientific research that can have a big impact on human lives in this country and around the world, but there are important questions of privacy that arise from that. An underpinning level of public trust is crucial to that. Are you tracking levels of public trust on data sharing?

Julia Lopez: Yes, and hopefully you garnered from Felicity that that is something that CDEI is doing. It has a tracker survey of about 4,000 citizens, making sure that we can understand as a Government whether trust is increasing or decreasing.

To pull back, I would like to underline how important trust is to whether we can get economic benefits from data. We are democracy, and I want to make sure that democracies have distinct approaches to data as a means to empowering citizens, rather than trying to control them. The trust that citizens have in how their data is used is fundamental to whether they feel confident about sharing the data, with not only all the scientific and economic benefits that derive from that but all the benefits that will come to citizens themselves in more personalised services that work better for them.

In my previous role in the Cabinet Office—my first ministerial role—I had data and digital in my portfolio. It was within that portfolio that we set up the centre for data and digital—sorry, I cannot remember what it was. The CDDO, basically. I was never that keen on the acronym.

In one of your earlier questions to Joanna, you asked whether there is a central push within Government to improve data use within Government, and there definitely is. That is being driven by Steve Barclay, particularly in his former guise as Chief Secretary to the Treasury, but also by Alex Chisholm and, when I was at the Cabinet Office, through me as a Minister, through Lord Agnew and through Michael Gove when was the CDL.

A large number of things are under way within the Cabinet Office to improve that and to make sure that we have higher quality datasets within Government and better data sharing between Departments, not because the Government are trying to track people but because we believe that better data sharing will lead to better outcomes for citizens.

Q351 **Chair:** Absolutely. Given your point of view from your current portfolio and from your previous portfolio at the Cabinet Office, how would you characterise the current level of trust in the use of people's data?

Julia Lopez: It is one of those funny things that people are quite happy to share their data with companies they don't know all that much about—for instance, Facebook—but we have to be aware as a Government that people will always, quite rightly, have a scepticism. As citizens we should always have a sceptical eye on Government, but with things that would



directly benefit the services they would receive, people are concerned about that. I always felt that my role in the Cabinet Office was to make sure that that public trust was maintained.

We want to continue that within the private sector, and there are various different layers to this. I don't think you can do one thing that suddenly creates trust in the system. This is about having data infrastructure that is secure and has cyber-resilience; skilled people, both within and outside Government, who understand how to use data in a responsible way; strong governance; a strong regulator in the ICO; and rules that are clear and that people understand. All of these things build trust, layer upon layer, and we as a Government have to make sure that all those different aspects of data, privacy and control are in place so we can build that trust from our citizens.

Chair: Thank you very much, Julia. Let me turn to colleagues, starting with Katherine Fletcher and then Aaron Bell.

Q352 **Katherine Fletcher:** Thank you, Chair, and thanks for your time, Minister. On the issue of getting the public to consent to what is being done with their data, I am interested in exploring what we can do with private companies. Before we get into that, is there a single decision to be made that says the data of individuals that is held within the public sector is definitely going to be shared with private companies? It almost comes back to the NHS question. We just heard from Lord Kamall and he is saying, "Yes, we will do bits with private companies but we won't do loads because it is the NHS's data." Where do you stand on that spectrum of everybody should have secure access to data because of the opportunities it creates versus we should be incredibly careful of what we share and hold it within the battlements of various baronetcies of Whitehall?

Julia Lopez: I do not want to be annoying, but there are different types of data and different layers of protection that should be adopted for each of those types of data and data use. You can have an individualised patient record—I am sort of moving into somebody else's territory here—that you might not want lots of other people to see, but would I mind if somebody, made me a generic person according to age, gender and some of the diseases I have had, helped to improve scientific innovation using large national datasets? We should be more relaxed about that because it is harder to identify who I am, and I think the public broadly consents to the idea that scientific innovation and better health treatments because you have good quality datasets is a public good.

One of the challenges that I have as a Minister is trying to bring this to life, and that is something I have pushed my officials on a lot. For instance, Government sharing public data with private companies during the pandemic might sound frightening to people, but what does that mean in reality? They have to do it on public interest grounds, and they have to be explicit about what that public interest ground is.



HOUSE OF COMMONS

One of the examples from the pandemic is the list of vulnerable people who need better slots in supermarkets to get their food. First, the Government had to say whether a person is vulnerable or not. Would that person mind if Sainsbury's knows that they tick that box so they can get a better slot in the shop? I don't think they would. Sainsbury's does not have to know everything about them; they just have to go with an API to a database that says, "Is this person vulnerable—yes or no?" I don't think the public would mind that as a means of data sharing, but they should also be reassured that there is a specific process that has to be adhered to in doing that. While the person has not provided explicit consent, there is a public interest test that Sainsbury's has to meet to get the data, and if a company is not setting out the grounds for that public interest, we have a strong regulator in the ICO, which can hold them to account for that. So there are various different layers to this.

Q353 Katherine Fletcher: That is really interesting, and very clear and helpful. Do you have a body that is checking what the code of the API is? For the uninitiated, that is the piece of code you write to join dots in databases. Do you have somebody within Government checking that the API is only checking what it says it is?

Julia Lopez: This is one of the reasons we set up the CDDO—to be able to bring some standards, transparency and similar ways of working. What you have to understand about Government is that it is not one entity. It is several different businesses or Departments with their own data and digital teams, and they may be doing things to different standards. One of the things we were trying to do in the Cabinet Office was to bring some level of standard to that process in order to bring high-quality people into the system who understand how to do these things well, and to bring better collaboration between Departments so that you can have greater standardisation and a professional approach to things. That is not the easiest task.

Q354 Katherine Fletcher: I can believe that. So the transparency and consent within the public data and how it is shared with the private companies—does the public need more control over its data? If it does, how is it going to do it?

Julia Lopez: It does in some circumstances and not others, and it is for us to try to work that through. I think you will start to see new technologies come online that give people better control over the information that they have. At the moment, a lot of people just give a lot of information away without deriving any benefit from that—whether that is economic or through a better personalised service. It is for us to encourage the market to start to deliver things that give people a greater sense of control over what information is held about them—for instance, web browsers, where you might want to start to see things that say, "Do not track me" when you go browsing, rather than having to always fill out an individual form for each website that you go to.

Q355 Katherine Fletcher: Okay. How will we make sure that the public are more aware of those risks? You have alluded to them a couple of times,



HOUSE OF COMMONS

and I was struck by a particular example earlier whereby somebody had opted out of GP patient records because they did not want Facebook to have their data. That was given as an item of evidence that suggests that the public need to be better educated, and that the Department needs to make the public aware of the risks associated with data sharing and put them in the different brackets, as you have so helpfully articulated. Health data is one thing, but what your likes and preferences are to support a DCMS tourism drive is something very different. How will you get the public to better understand both the risks and opportunities?

Julia Lopez: Without wanting to blow smoke, these kinds of inquiries are a helpful part of that. It is about having public debate and discussion. Parliament will have a role as things such as the data Bill go through the House, and that will stimulate greater public discussion within the media as well. Obviously, we have our own consultation processes as well—we had about 3,000 responses to that. We are trying to stimulate discussion and debate, but these are really quite difficult concepts to articulate. All of us are in the foothills of what this means. We have greater awareness as some of the debates over Facebook, Google and the power of data develop, but I also want to try to make sure that it is not some sort of black and white debate, and that people understand that nuance. I guess my sitting here and being tested by you guys is one of the means by which I can do that, but I also think the pandemic was a great moment in terms of building public awareness of what better sharing of health data means for outcomes for you as a citizen. One of the systems that I was developing within the Cabinet Office was something called One Login for Government.

Q356 **Katherine Fletcher:** That has been a holy grail for about 20 years, hasn't it?

Julia Lopez: Yes, but I feel like we may be getting there—I may come to regret those words. When people start to understand what it means to have better data sharing within government, you will get a more personalised service so that you do not always have to fill in a new form every time your circumstances change when receiving universal credit. These sorts of things will help to bring these debates to life. People will have a better understanding of them in this more gradual process rather than this being some big bang situation where we try and have a public education programme or whatever.

Q357 **Katherine Fletcher:** Okay, I take that point. If we are going to head towards, as I heard my colleague Aaron just say, the fusion of the Government data world, which is a single point of login—which I think is a reasonable analogy—we have Professor Goldacre's recommendations about data sharing. He talks about whether sharing pseudonymised data across government would be stopped. I know this gets very technical, but where are you on Professor Goldacre's recommendations, and how does that sit within that idea of simplifying it to make it easier to both control and make safe?



HOUSE OF COMMONS

Julia Lopez: This is a fairly recent recommendation, I have to confess. Because it was in the health space, I am not all over this, so perhaps Jenny might come in and help, but I think pseudonymisation should only ever be seen as one part of a toolbox of privacy enhancing measures. Jenny, do you want to say more on that?

Jenny Hall: Yes, I am happy to. I apologise to the Committee for not being there in person, but unfortunately I have covid, so I am stuck at home.

The Minister is absolutely correct: pseudonymisation should be seen as one tool in the privacy toolkit. When you think about the requirements of GDPR, a key one is data minimisation, so you only use the data that is strictly necessary. You use it in as slimmed down a way as possible. Already under the legislation, data controllers are obliged to have due regard to the risk of re-identification, which is the key risk with using pseudonymised data before sharing it onward. So there are limitations, and we agree with the recommendation in that respect.

We asked questions about this in our data reform consultation recently, and the ICO has a call for views out at the moment on pseudonymised data, anonymised data, and the role of privacy enhancing technologies. It is one tool among many. It is definitely not a silver bullet and should not be seen as such. I think it should be considered alongside a range of other options, some of which you have heard about through other evidence today. I hope that helps.

Katherine Fletcher: That is very helpful. Jenny, is there anything you want to say on the questions I have just been asking? I'm sorry to hear you have covid—hope you're not feeling too crap.

Jenny Hall: No, that's absolutely fine. Thank you.

Q358 **Katherine Fletcher:** Just one final one from me, Chair, if that's all right. We have heard lots of evidence during this inquiry from scientists chomping at the bit to get at this golden nugget of data that the NHS holds, for what I judge to be entirely noble reasons. They want to get on with doing research and they can see a fantastic toolkit to do it. Where do you stand on allowing that to happen if those scientists are in a private sector organisation? Have you got any views on that? I know it is slightly out of your ministerial portfolio.

Julia Lopez: I would go back to an earlier answer that I gave: it depends on what type of data we are talking about. Are we talking about anonymised data where you can tell something about a population but not an individual?

Q359 **Katherine Fletcher:** Let's take it down to the ultimate kind of postcode, which is your DNA. In theory, if you shared an entire DNA sequence, it would be relatively easy to get back to the individual quite quickly, as some of the other things show, but sharing large sections of DNA in a data way is an enormous opportunity to identify disease causes and risk factors and all of those other wider health benefits. I am just trying to



HOUSE OF COMMONS

test the limits of where you are in the process, because ATCGGC is just data at the end of the day. Biology has just invented a pretty spiral to put it in.

Julia Lopez: To be honest, some of this stuff really is in the health space. We are here to set the general policy. When it comes to the thorny debates about who holds the DNA database, would I feel comfortable with the police sharing with scientists details of the DNA database? I think that would be a very difficult debate, probably one we would have to have very publicly so that there was a sense of the pros and cons to that. We are not at the stage of that happening. Jenny, I don't think there's anything you want to add. You may have had conversations with—

Chair: We need to move on; we have a lot to get through. Aaron Bell.

Q360 **Aaron Bell:** Thank you, Chair. And thank you, Minister. It has been a great inquiry and it is very good to have you as our final witness. I want to talk a bit more about the consultation, "Data: A new direction". First, is this something that the Government are seeing as a benefit of Brexit.

Julia Lopez: Yes.

Q361 **Aaron Bell:** Excellent. And the consultation closed in November. Are we expecting a formal response to the consultation soon?

Julia Lopez: Yes.

Q362 **Aaron Bell:** If so, could you give us a guideline on that?

Julia Lopez: The problem is that you have to go through these rather laborious processes. Obviously, you need to analyse and compile all the responses you have received, but then you have to—it will be very much linked to the legislation when it comes out. You have seen that it was introduced in the Queen's Speech, so we will aim to publish the consultation and then the legislation, because the two are linked. And we have to go through the process of write-round to make sure that Departments are fully comfortable with the proposals being put forward. So I am in that slightly frustrating period, certainly for you as a Committee, where I can't give the full details of precisely what has been decided.

Q363 **Aaron Bell:** Sure. Will it be before the summer recess?

Julia Lopez: Goodness, don't try to do this to me! This is one of those things where, once you put it out into the ether as a Minister, you lose control as to the timing of it. All I can say is that I've done all my bit.

Q364 **Aaron Bell:** Fair enough. There are obviously lots of reasons why you want to do this. You want to increase the competitiveness and efficiency of the UK economy. You talk about boosting trade, reducing barriers to responsible innovation, better public services, less burden on business, and better outcomes for ordinary people. What assurances can you give this Committee that that is not going to dilute privacy?



HOUSE OF COMMONS

Julia Lopez: Well, I don't see those two things as mutually exclusive. You can do things to enhance privacy and to increase economic growth precisely because there are high levels of trust in the system. We will be looking at things like smart data and how we encourage privacy-enhancing technology, so that you can get the economic growth and scientific innovation without diluting privacy. A lot of what we want to try to do within our legislation is to provide greater clarity as to what you can and cannot do. I know that this Committee is interested in guidance versus legislation. One of the things we hoped to achieve with our consultation on data was to get scientists, businesses and others to tell us what the areas are where there is insufficient clarity from guidance and where you believe it requires legislation for you to be told what you can and can't do. Otherwise, you have to go through a lot of legal processes and you have to have a larger data team than you might otherwise have, because it's not entirely clear what you can and can't do.

Q365 **Aaron Bell:** One of the proposals was to insert into legislation a clear test for determining when data would be regarded as anonymous. What would that look like?

Julia Lopez: Jenny, maybe you would be best to come in on this.

Jenny Hall: I'm happy to. That is exactly one of the things that I referred to previously. We are looking at and working very closely with the ICO on how we would come at working out that definition, because as you know, as soon as data is deemed credibly to be fully anonymous, the data protection laws will no longer apply to it. As the Minister says, without going into the full detail, because that will be coming down the track, that is one of the areas where we are trying to get absolutely clear and leave minimal room for interpretation and uncertainty as to exactly when data is anonymous, when it might be pseudonymous and when therefore special safeguards and categories might apply. It's a live discussion and one that you will see more on in due course.

Q366 **Aaron Bell:** Thanks, Ms Hall. Minister, in addition, there are proposals to remove data protection impact assessments and the need for data protection officers, and there is the introduction, potentially, of a fee regime for subject access requests. How confident are you that these proposals will enhance and not diminish the independence of the ICO?

Julia Lopez: What we are trying to do with our reforms is to move to an outcomes-based approach, rather than a tick-box approach. You won't necessarily be required to have a data protection officer, but you will still have to have a privacy management programme, within your business or organisation, where you need to have proper accountability and proper reporting. You need to be able to assure the ICO about the measures in place within your own organisation to make sure that people's data is fully protected and compliant with legal requirements. So I don't think that there should be any concern that we are diluting standards. We are simply giving companies greater flexibility. For instance, if you are a small business, it is quite a big deal to have to employ a separate person to be the data protection officer. I think having somebody within your senior



HOUSE OF COMMONS

management team whose role is to be the data protection person is a more reasonable approach. We are trying to put a stronger governance structure around the ICO so that you have a proper board and chairman, and that the codes and guidance are developed with greater consultation with those affected.

Q367 **Aaron Bell:** Sticking to the fee point, do you not think that introducing a fee, whatever level it is set at—even if it is just nominal—sends the wrong message regarding transparency?

Julia Lopez: I don't think there has been any announcement on fees. On the subject access request—

Q368 **Aaron Bell:** I think you were considering introducing a fee regime.

Julia Lopez: You will find out the decision on that in due course.

Q369 **Aaron Bell:** In due course. I was going to ask why the data reform Bill is required, but you have clearly answered that already.

Julia Lopez: Just on the subject access request, it has been a policy debate within Government, and I very much want to test it as a Minister. We want to ease the burden on businesses without diluting people's rights. It is not always an easy balance to strike but, for instance, GPs in my constituency have complained to me about the cost and time of doing subject access requests. We want to get the right balance between not having a large burden on businesses and not diluting people's rights. We are not going full pelt at putting lots of fees and costs on to the individual to get things that they have a right to. I hope that gives a level of reassurance.

Q370 **Aaron Bell:** Understood. This is presumably going to be a fairly chunky Bill, because it needs to cover a lot of circumstances. How confident are you that it will actually end up reducing legal complexity, which is obviously the goal? How much legislation are we going to repeal at the same time as bringing in this new legislation?

Julia Lopez: I don't think it is going to be the chunkiest DCMS Bill out there, given that there is the Online Safety Bill. On the size of the Bill, Jenny might be better placed to answer, because she is involved in all the drafting.

Jenny Hall: I am happy to jump in there. The first point to make is that the Bill is designed to build on existing legislation. We are trying to improve the legislation that we have largely inherited from the EU, and that is done intentionally in a way to make sure we are not adding costs and a burden on people from having to look at a whole new set of requirements. That is one of the core intentions behind the Bill.

The point was made earlier that, sometimes, having clarity in legislation leaves less room for interpretation and potential differences of opinions among those actually operating, as you heard in the previous evidence sessions, in a very complex and fast-moving space. Having that legislative

clarity, even if it takes a bit of time to take a Bill through, pays off in terms of how it is implemented and how the ICO monitors it.

Q371 Aaron Bell: Thank you. You have obviously demurred already on when you are going to publish the consultation, so you are not going to tell me when the Bill is coming. You will do the two consecutively—the consultation, swiftly followed by the Bill.

Julia Lopez: The consultation will be published before or at the same time as the Bill is published. It is one of those difficult things. A lot of people are frustrated by the EU data system and might be inclined to have a completely blank sheet, but you have to bear in mind that creating something entirely new would create burdens for businesses, so we are trying to take what is there but make it better, and in so doing reduce the burdens on businesses, rather than increase them.

Q372 Chris Clarkson: We have just heard that the chance to review this is a benefit of Brexit, so I want to ask you how the Government's proposals for data sharing will maintain the UK's data adequacy agreement with the EU.

Julia Lopez: I know this will be of real interest to a lot of parliamentarians on both sides of the Brexit divide. We are confident that it will maintain adequacy. I have a fantastic team in DCMS, and they are in regular touch with their EU counterparts. There is a debate within the EU about the GDPR requirements too.

As I alluded to in my response to Aaron, we are taking the existing system and improving it, and therefore we are not starting with a blank sheet of paper in a way that might risk adequacy with the EU. At the same time, we are not going to be dictated to, in terms of what we can and cannot do. It should be borne in mind that there are 13 non-EU countries that have adequacy agreements with the EU. In so far as people are concerned about adequacy lost, they should bear that in mind. You should also bear in mind that our officials are in regular contact with the EU, on a "no surprises" basis.

It should also be borne in mind that there are two different adequacy deals: one on law enforcement and one on GDPR. A lot of the time, these two things are conflated with the idea that if we lose adequacy, criminals will be pinging around, here, there and everywhere. To put it on the record, there are two different strands to this as well.

Q373 Chris Clarkson: What protections will be in place to make sure that individuals are not subject to decisions based solely on automated processing? I am thinking about profiling, for example.

Julia Lopez: You will still have a right to human review. Again, this is something that will probably be fully clarified when we can talk about our consultation response and legislation. However, you will still have the right to human review, and you can then make a complaint if you feel that that has not been given to you—to the organisation itself and subsequently to the ICO.



Q374 **Chris Clarkson:** Has the framework for that review already been drawn up or is it under discussion as to the basis on which you will be able to challenge a decision?

Julia Lopez: Jenny, I don't know if you could come in on that one.

Jenny Hall: Yes, I am happy to. The thing to bear in mind here is that when we are looking at this through the data protection legislation, we are looking purely at article 22 and the specific right and safeguard that is incorporated within that. The Minister is absolutely right: we have been looking at that extremely closely and, as you will know, we consulted on it and have been looking on how we maintain that going forward.

I think this needs situating, though, in some of the wider work happening across Government at the moment, some of which you heard about earlier this morning, on AI governance more broadly and making sure that some of these rights to review and understanding how your data is being used are seen holistically and not just through a data protection lens. You will see a bit more detail about this in our consultation response, but you will see far more on it in the Government's White Paper a little bit later this year. We are trying to develop our approach in a data protection field or sphere fully in line with that, so we end up with the most coherent approach possible.

Julia Lopez: Just to zoom out a bit, privacy is incredibly important; we see it as absolutely fundamental. Trust and privacy are fundamental to where we want to go with this, but we also need to be mindful of where the world is moving on things like AI and of the need to maintain competitiveness with people who are doing things differently in this space. You think about China, which is refining its AI on vast datasets. I am not suggesting that we move towards a Chinese regime by any stretch of the imagination; it is simply to say that we have to be mindful of the need to be economically competitive and to allow our scientists and innovators to have access to high-quality datasets. While maintaining privacy and trust, we should not create fear about those things in a way that undermines our businesses and our scientists' ability to innovate.

Q375 **Chris Clarkson:** Would you contemplate something like the algorithmic standard, which Felicity Burch mentioned in the last panel—almost a kitemark for the application of some of these algorithms?

Julia Lopez: Felicity is a part of Government. The CDEI is something that we created and we have tasked with doing things like creating algorithmic standards.

Q376 **Chris Clarkson:** So that would be underpinning how people would be able to see—

Julia Lopez: Potentially. We will see how this piece of work develops. This is a world-leading piece of work, where we are trying to drive forward a distinctly British approach to maintaining high standards while innovating.

Q377 **Chris Clarkson:** Would you include a requirement to publish code or algorithms where they have been applied to people's personal data?



HOUSE OF COMMONS

Julia Lopez: To be honest with you, this is getting a bit techy. Felicity would be best to answer where she intends to take that piece of work. Jenny, do you want to say anything about whether that is something that has been discussed at official level?

Jenny Hall: As Felicity said earlier, this is quite a recent thing, and I think they have some really good feedback from the first couple of iterations. I am sure that is something that they will be considering going forward. The overarching point is being as transparent as possible, looking back to where we started, to enable us to build and maintain public trust in everything that Government are doing in relation to data sharing.

Q378 **Chair:** Finally, on this point of international practice, you rightly said that we need to have an eye on our competitiveness and our ability to make scientific advances. Are there regimes in the world that you have looked at in order to learn from? Are there any particular ones that commend themselves, on the AI?

Julia Lopez: We have close working partnerships with countries like Singapore. We are in very close touch with the US and the Australians. Jenny is part of the wider data team at official level that has ongoing conversations. Jenny, do you want to add anything on this front?

Q379 **Chair:** On AI specifically, new regulation is required. Every country is grappling with it. Are there any that are doing it particularly well or doing it in an interesting way, in your view?

Jenny Hall: The Minister is right. We engage really actively and regularly on this. The people we have most recently spoken to, on AI but more specifically data, which is more our area, are our G7 colleagues last year. You will have seen under the UK's G7 presidency last year that we did a huge amount of work with partners within that grouping, looking at a range of issues, including data free flow with trust and also wider issues on AI and digital technologies. I and my team were in the room for some of those negotiations, and there was a real acknowledgment that we are all facing similar challenges here. There are some good examples in different areas. There are different strengths in different areas. Some of the work that we did over the past year was to really share learnings in some of those areas.

Q380 **Chair:** Can you give us an example of a country that you think shows strength in future AI regulation?

Jenny Hall: Regulation specifically is really quite novel. There is not a huge amount of it out there. It's more the sort of approaches that are being taken in terms of trialling new technologies and investing in different infrastructure, working with organisations. The UK's version is the CDEI; there are other organisations in different countries doing similar things, albeit not as forward-leaning as some of the work we are doing.

The communication we get from counterparts in foreign Governments is that they are trialling a range of different things and trying to draw together some of the aspects of digital regulation that touch on AI,



HOUSE OF COMMONS

whether that is data protection or digital competition, because it is not one single thing. There are a range of aspects feeding into that and a range of approaches as a result, some of which are rooted in different legal systems and different cultures and different approaches. It is for us to pick and spot where there are real strengths that we can learn from and build from for a UK system. That is certainly the approach that we are trying to take.

Chair: Okay. On that note, Ms Hall, thank you very much for persevering through covid and not taking to your sickbed but giving evidence to us today. We very much appreciate that. We very much appreciate the Minister's attendance too, to wrap up what has been a fascinating inquiry. We will now have the difficult job of reflecting on the richness of the evidence that we have had and making some recommendations to the Government that I hope you will be able to respond to positively. To the Minister, her official and all the witnesses we have heard from this morning, thank you very much indeed. That concludes this meeting of the Committee.