THE ROYAL SOCIETY

Royal Society Submission to the Data: A New Direction consultation

Key Points

- The current General Data Protection Regime leaves substantial room for interpretation of its articles, which creates considerable uncertainty around a number of issues laid out below. This results in many actors erring on the side of caution which in some cases be legitimate, but can also restrict data sharing and present barriers to innovation.
- Clearer guidance on interpretation of the regulation will be beneficial to enable use of data for research, innovation, societal benefits and economic growth.
- Research and innovation is broad and varied, and guidance relevant to different disciplines and sectors is needed, while ensuring interoperability across areas.
- Public trust and confidence in the regulatory system are key to enabling trusted use of data for research and innovation. The Royal Society and British Academy's Data Management and Use report made recommendations for a governance framework for data management and data use which should:
    - Anticipate, monitor and evaluate,
    - Build practices and set standards,
    - Clarify, enforce and remedy.

  Any review of data regulation should ensure that these different functions are satisfied adequately in order to build and maintain trust.
- The experience of the Covid-19 response, both positive and negative, highlights the need for the private and public sectors to adopt a duty to safely share data in an emergency, and indeed for ongoing public benefit. It is however essential that this sharing and linking of data is done in a safe and well-governed way to protect the rights of data subjects and commercial interests of companies.
- The EU remains a hugely significant R&D partner for the UK, and it is important to ensure the UK retains data adequacy with the EU, in order to enable collaboration.
- Fairness in AI is an important challenge to address. Regulation relating to AI and algorithmic decision making should be sensitive to context of use.

Content

1. Background

1.1. The Royal Society is the National Academy of science for the UK. Its Fellows include many of the world's most distinguished scientists working across a broad range of disciplines in academia, industry, charities and the public sector. The Society draws on the expertise of the Fellowship to provide independent and authoritative advice to UK, European and international decision makers.

1.2 The Society's fundamental purpose, reflected in its founding Charters of the 1660s, is to recognise, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity.

1.3 In partnership with the British Academy, the Academy of Medical Sciences, and the Royal Academy of Engineering, the Society hosted a workshop for the consultation on the implications of data reform to researcher access to data. Other positions are drawn together from previous Society publications—a bibliography of these is included at the end

2. Research purposes

4.1 GDPR has the potential to enable use of data for research and innovation. However, much depends on how it is interpreted and implemented. Depending on the culture of implementation, it has the potential to act as an impediment to researchers accessing data, repurposing it, and linking of data sets.

4.2 This is primarily due to the culture surrounding the implementation of GPDR, and the interpretation of it, rather than the detail of the legislation itself. The regulations already make provisions for greater researcher access, but use of these is limited by a lack of precision in terminology, a culture of caution surrounding permitting access to data and a lack of clear guidance on interpretation.

4.3 Reasons for this risk-averse culture include:

- A lack of clear guidance on researcher access
- Guidance that is specific to certain disciplines being applied in other areas
- The personal liability of data processors in the case of data breaches
- GDPR potentially being used as a reason for denying researchers access to data, when there are other motivations in play

4.4 For example, in some cases, researchers (e.g. in humanities/social sciences) have to grapple with guidance designed for different disciplines to their own (e.g. medical sciences).

4.5 Public trust and confidence in the regulatory system are key to enabling trusted use of data for research and innovation. Language that implies any regulatory reform is done primarily with the aim of loosening protections is liable to undermine this public trust.

4.6 Lack of enforcement of GDPR also limits public trust. In the last ICO annual report there were only three fines and one Enforcement Notice, despite over 30,000 complaints. As set out in the Royal Society and British Academy report Data management and use: Governance in the 21st century, one of the fundamental functions of data governance should be to *enforce* regulation and provide remedy for harms. This is an essential aspect of a trustworthy governance framework.

4.7 There are currently provisions within the GDPR that can support use of data for research and innovation, if the regulation is interpreted and implemented correctly. For example, GDPR enables individuals to allow the transfer of privately-held data about them to researchers in the public sector. This could facilitate access to high quality data for researchers. However, there are concerns that organisations that hold personal data, such as social media companies, are not allowing this access to data despite it being lawful.

4.8 The present data access regime benefits large companies with the ability to collect large amounts of personal data, while disadvantaging publicly funded research carried out for public benefit, and SMEs that could help drive innovation. Government has a role in brokering data sharing between such private and public actors and addressing data monopolies.

4.9 Trusted Research Environments and pseudoanonymisation of data are effective means of protecting privacy in shared data and limiting the risk of data breaches. The development of Privacy

Enhancing Technologies (PETs) will improve this even further. However, this risk can never be totally eliminated, and focus on it should be balanced against potential benefits.

4.10 Sharing data with researchers in Europe and elsewhere is essential to enabling international research collaboration and for the UK to continue as a leading research nation. Therefore maintaining data adequacy with the EU is crucial. While existing EU data adequacy agreements with third parties display some degree of divergence, most of these agreements pre-dated the introduction of GDPR and were 'grandfathered' into the system. They may therefore not be a sound guide to future adequacy decisions.

Q1.2.1. To what extent do you agree that consolidating and bringing together research-specific provisions will allow researchers to navigate the relevant law more easily?

4.11 Greater clarity and guidance in this area is a priority. Given the substantial room for interpretation of the GDPR, many researchers and data controllers take a risk averse approach to data access and sharing, erring on the side of caution. This risk aversion can be observed across sectors (companies, government, universities). One driver may be the personal responsibility for data breaches held by individuals who handle data within GDPR. As a result, data which is to be used for scientific research and the public interest may not be shared by a data controller. This constitutes significant barriers to innovation and international competitiveness for the UK.

4.12 The tolerance for risk is closely tied to commercial benefit that can be gained from holding and processing data and the legal advice an organisation may be able to obtain. As a result, actors in the public sector and smaller organisations are likely more risk averse, while larger companies with in-house legal expertise and business models based on data value may be able to act more freely.

4.13 The culture around data sharing in research is also highly dependent on discipline and established practices within a specific area of research. A challenge for researchers may be around practical (in addition to legal) aspects of GDPR adherence. Data governance regimes and data platforms may not be set up in some research disciplines, or may be less developed in some than others. For example, practices in the humanities (eg history) may be very different to those in biomedical sciences.

4.14 The interpretative uncertainty of the GDPR and risk aversion may be used by some actors as a negotiating or gatekeeping tactic to prevent data sharing or as a barrier to broader collaboration. Clarification of GDPR articles would address attitudes to risk. However it may not be capable of mitigating gatekeeping on the basis of GDPR, particularly in the international context where UK regulations would not apply.

Q1.2.2. To what extent do you agree that creating a statutory definition of 'scientific research' would result in greater certainty for researchers?

Q1.2.3. Is the definition of scientific research currently provided by Recital 159 of the UK GDPR ('technological development and demonstration, fundamental research, applied research and privately funded research') a suitable basis for a statutory definition?

4.15 The widely used Frascati definitions are a standard reference for R&D surveys and data collection. These principally define R&D as consisting of basic research, applied research and experimental development.

4.16 The definitions used should encompass areas GDPR has particularly impacted access. These include the sharing of soft (qualitative) data which particularly affects the social sciences. Non-textual data, such as image data, may also be subject to particularly strict regulation if pertaining to a natural person (eg, web-based images of people) and if copyright is unclear and data subjects have not provided consent to data sharing for research. This can cause challenges for machine learning research which may use large data sets of web-scraped images, and here the concern is more likely to be the uncertainty relating to such publicly available but personal data rather than the nature of the research.

4.17 In some instances, researchers have had to grapple with guidance which was initially developed for another discipline from their own, which may result in infringements on academic freedom of expression. One example pertains to social research (such as on corruption and other forms of illegal activity) being confronted with guidance established for biomedical research.

4.18 Further, novel forms of data sets will emerge in the coming decade which will pose new questions for data protection regimes. As a result, any reform should not be bound by specific use cases but should rather be allowed flexibility in the face of novel data types.

Q1.2.4. To what extent do you agree that identifying a lawful ground for personal data processing for research processes creates barriers for researchers

Q1.2.5. To what extent do you agree that clarifying that university research projects can rely on tasks in the public interest (Article 6(1)(e) of the UK GDPR) as a lawful ground would support researchers to select the best lawful ground for processing personal data?

Q1.2.6. To what extent do you agree that creating a new, separate lawful ground for research (subject to suitable safeguards) would support researchers to select the best lawful ground for processing personal data?

4.19 As previously noted, the GDPR already makes provision for greater researcher access, but use of these is limited by a lack of precision in terminology, a culture of deny-access-by-default and a lack of clear guidance. The changes to article 6 of GDPR proposed in the consultation (adding 'research' as a legal basis for processing personal data) may not significantly change what is already legally permissible, as most research has a legal basis as a 'public benefit', but would have the useful effect of making the existing situation less ambiguous. Such an unambiguous legal basis that could be referenced for ethics reviews would greatly simplify research management.

5 AI and machine learning

5.1 As noted in the Royal Society's work on Explainable AI and AI, society and social good, fairness refers to the equitable treatment of different social groups and of individuals. It also means meeting individuals' reasonable expectations of how their data is used to make decisions about them (see AI technologies and their implications for society). Combating algorithmic bias and discrimination are therefore vital to ensure fairness. Machine Learning systems may inherit biases embedded in society which are reflected in the data used to train these models. This may occur in the form of either conscious or unconscious biases in human decision making or it can occur at the point of data collection if sampling error, inaccuracies or missing entries render the data unrepresentative for certain groups and individuals.

5.2 Technical fixes are insufficient to address questions surrounding fairness in the context of AI. Rather, a broader understanding of the AI system and the social influences that shape AI systems is needed (see AI technologies and their implications for society).

Q1.5.1. To what extent do you agree that the current legal obligations with regards to fairness are clear when developing or deploying an AI system?

5.3 Definitions of fairness and bias may be context dependent and specifying desired outcomes and subjective functions can therefore be challenging (see AI, society and social good and Public views of Machine Learning). Clearer interpretative guidance with respect to the definition and operationalisation of fairness as well as the legal obligations regarding fairness specifically within the context of AI is needed. Given the differences between development and deployment phases of an AI system, different considerations with regards to fairness may apply. Moreover, the legal obligations of actors who may be differently involved throughout the process of the training and use of an AI system, respectively, should be clarified, for instance in cases where an organisation may deploy a system developed by another actor.

Q1.5.3. What legislative regimes and associated regulators should play a role in substantive assessments of fairness, especially of outcomes, in the AI context?

5.4 Independent and active monitoring of the AI landscape is key to enable a proactive approach in responding to new challenges surrounding issues of fairness, bias and discrimination with regards to the development and deployment of AI systems. Next to technical approaches to test whether AI systems adhere to fairness, other organisational approaches are needed, such as safeguards, audit functions and other means of accountability and assurance which have previously been developed in the use for human decision-making processes (see the Royal Society's work on Explainable AI and AI and society).

5.5 A system is not fair if it uses data for decision making in a manner that the data subject could not reasonably expect, or when the use of individual's protected characteristics or sensitive data by an AI system results in consistent disadvantages in terms of profiling or decision making. A requirement for developers is therefore to provide a priori information on how personal data is used and to assess the presence of bias in their algorithms throughout training (see also the Royal Society's report on Machine Learning). Risk assessment and mitigation steps should therefore include a use case specific consideration of the implications of the deployment of an AI system and should include the requirement for continuous monitoring of the performance of a model to mitigate drift or other processes that might have implications for the fairness of the system. Clearer guidance and tools to assess whether an organisation or actor adheres to fairness principles is needed.

Q1.5.5. To what extent do you agree that the government should permit organisations to use personal data more freely, subject to appropriate safeguards, for the purpose of training and testing AI responsibly?

5.6 Data access is key to the training and testing of machine learning models. However, wider access to data also bears greater risk to individuals in terms of data protection and privacy. Trade-offs between privacy and access have to be made transparently, accountably and inclusively (see the Royal Society and British Academy Data management and use report). It is essential to provide clear guidance on appropriate safeguards to mitigate the risks associated with the use of personal and sensitive data.

5.7 Computer vision belongs to the field of machine learning and relies on vast amounts of image data to train algorithms. This type of data can be obtained through automatic web scraping or from databases of third-party providers. Interpretation of GDPR can pose a particular challenge for this type of research which uses publicly available but personal information.

5.8 Certain risks can potentially be mitigated and managed with a set of emerging technologies and approaches often collectively referred to as 'Privacy Enhancing Technologies' (PETs). Whilst cybersecurity is focussed on protecting data so that other people cannot access it, PETs, in data analysis, are focussing on enabling the derivation of useful results from data without giving other people access to all of the data. This nascent but potentially disruptive set of technologies, combined with changes in wider policy and business frameworks, could enable significantly greater sharing and use of data in a privacy-preserving, trustworthy manner. It could create new opportunities to use datasets, including in training, without creating unacceptable risks. It also offers great potential to reshape the data economy, and to change, in particular, the trust relationships between citizens, governments and companies. Further recommendations on how government can support the innovation and use of these technologies can be found in the Society's ongoing work in this area.

Q1.5.10. To what extent do you agree with the proposal to make it explicit that the processing of personal data for the purpose of bias monitoring, detection and correction in relation to AI systems should be part of a limited, exhaustive list of legitimate interests that organisations can use personal data for without applying the balancing test?

5.9 Detection and correction of algorithmic bias is essential to prevent discrimination. An algorithm may find that particular attributes of certain persons in the data result in better predictions and in some

cases these attributes may be deemed as inappropriate to use because they result in discrimination. Even if explicit data relating to age, race, gender and membership of other groups is not being used in a given AI system, other factors may be tightly linked to these attributes and may then be used by the system as proxies. It can be challenging to identify these proxies and ensure they are not resulting in unfair models (see Royal Society work on Explainable AI, Data management and use, Machine Learning and AI, society and social good).

5.10 In the development and testing phase of AI systems, one way of mitigating bias may be to determine whether outputs and classifications of algorithms are correlated with sensitive characteristics such as race. If the same types of errors are consistently made, this may constitute consistent discrimination against certain groups of people. Favouring processes that produce inconsistent errors or diversity of outcome could achieve a more equitable distribution of errors. The availability of sensitive data may allow developers to identify proxy attributes which may contribute to biased algorithms. The use of sensitive data in the development and deployment of AI systems bears risk to data protection whether by means of accidental disclosure or re-identification or as a result of targeted attacks on data and algorithms (see Royal Society work on Privacy Enhancing Technologies). It is therefore essential to only permit the use of sensitive data in cases where the risk is outweighed by the benefits for society and individuals.

Q1.5.11. To what extent do you agree that further legal clarity is needed on how sensitive personal data can be lawfully processed for the purpose of ensuring bias monitoring, detection and correction in relation to AI systems?

5.11 Clarity is necessary to ensure that organisations and actors who develop and deploy AI systems are aware of their legal requirements, follow best practices, and only use sensitive data where the benefits outweigh the risk to the individual. In the case of AI systems, identification of lawful purposes for the processing of sensitive data may be particularly challenging given that AI can exacerbate existing and create new risks to data protection (see Royal Society work on Explainable AI, Machine Learning).

Q1.5.19. Please share your views on what, if any, further legislative changes the government can consider to enhance public scrutiny of automated decision-making and to encourage the types of transparency that demonstrate accountability (e.g. revealing the purposes and training data behind algorithms, as well as looking at their impacts).

5.12 Explainability and transparency are essential in creating trustworthy AI systems. To achieve this, both technical means and other measures such as assurances are required. Trustworthy governance also requires conversation between governing bodies and the public to ensure that citizens feel heard (see the Royal Society's work on the role of public engagement in policy making).

5.13 The Royal Society's report on explainable AI recommends any actions taken by the government should consider transparency and explainability as key factors when developing means to enhance public scrutiny of automated decision making. This involves clear language as to the collection and use of personal data, how decisions are being made about individuals and what options members of the public have to request human intervention and justification. Active dialogue between the public and regulators is needed to give citizens an opportunity to voice their concerns and for regulators to respond to these concerns adequately. The Royal Society's public engagement work has shown automated decision making and profiling to be particular concerns for members of the public (see Public views of Machine Learning and Public views of machine learning: Digital Natives).

5.14 As set out in the Royal Society's work on Machine Learning and on Explainable AI, it is important to consider the context in which algorithmic decision making takes place, when considering whether existing legislation is sufficient. There are significant and important unresolved issues for some applications of machine learning, such as whether algorithms need to be interpretable in particular use cases, when humans should be involved in decision processes, and when algorithms should be

held to a higher standard of accuracy or interpretability than human decision-makers. The answers to these questions will vary with the application area.

6 Data minimisation and anonymisation

Q1.6.4. Please share your views on whether the government should be promoting privacy-enhancing technology, and if so, whether there is more it could do to promote its responsible use.

6.1 Privacy Enhancing Technologies (PETs) have an important role to play in enabling well-governed use of data. However, there are many other approaches to risk minimization that are already in use, and no technological intervention will eliminate risk entirely.

6.2 The following actions are set out in the Society's report on privacy enhancing technologies, Protecting privacy in practice. Government should take the following steps to promote responsible use of privacy enhancing technologies (PETs):

- Accelerate the research and development of PETs.

  Funders, government, industry and the third sector can work together to articulate and support the development of cross-sector research challenges, alongside providing continued support for fundamental research on PETs.

- Promote the development of an innovation ecosystem.

  UK Research and Innovation (UKRI) have a role in encouraging data-handling companies to engage with the start-ups and scale-ups developing PETs, to support research and early trials. This will help UK investors and businesses realise the extent of the market opportunity for PETs.

- Drive the development and adoption of PETs.

  Government can be an important early adopter, using PETs and being open about their use so that others can learn from their experience. Government departments should consider what existing processing might be performed more safely with PETs and how PETs could unlock new opportunities for data analysis, including opening up the analysis of sensitive datasets to a wider pool of experts whilst fully addressing privacy and confidentiality concerns.

- Support organisations to become intelligent users of PETs.

  There is a need for Government, public bodies and regulators to raise awareness further and provide guidelines about how PETs can mitigate privacy risks and demonstrate compliance with GDPR and other data regulation. For example, the Information Commissioner's Office (ICO) should provide guidance about the use of suitably mature PETs to help UK organisations minimise risks to data protection, and this should be part of the ICO's Data Protection Impact Assessment guidelines. Such guidance would need to cover how PETs fit within an organisation's overall data governance infrastructure, since the use of PETs in isolation is unlikely to be sufficient.

- Give public sector organisations the level of expertise and assurance they need to implement new technological applications, enable a centralised approach to due diligence, and assure quality across the board.

  The National Cyber Security Centre should act as a source of advice and guidance on the use of suitably mature PETs, as part of a network of expert organisations. Such a network of expertise would support the development and evolution of best practices and also provide access to advice on specific cases of data use or sharing. Ultimately, this could also serve as a point of engagement for academics and industry bodies working

in the space and provide a portal from which private sector organisations interested in learning about PETs could access information on existing case studies.

- Create the skilled workforce needed to develop and implement PETs.

  Funding should be made available so that the capacity to train UK PhD and Master students in cryptography, statistics, systems engineering and software development increases with the level of demand for well-trained, high-calibre candidates. This could be an outcome of the National Cyber Security Programme and the cybersecurity centres of excellence scheme by the Engineering and Physical Sciences Research Council. Universities should consider adding privacy engineering to the curriculum of software engineering and data science courses, treating the need to protect data as core knowledge in data analysis.

- Promote human flourishing by exploring innovative ways of governing data and its use that are enabled by PETs.

  The Department for Digital, Culture, Media and Sport (DCMS), the Centre for Data Ethics and Innovation (CDEI), Office for AI, regulators and civil society should consider how PETs could become part of the data stewardship infrastructure, underpinning governance tools such as 'data trusts' and other initiatives for the governance of data use.

7 Further Questions

Q1.8.2. In addition to any of the reforms already proposed in 'Reducing barriers to responsible innovation' (or elsewhere in the consultation), what reforms do you think would be helpful to reduce barriers to responsible innovation?

7.1 The GDPR regime should foster rather than inhibit innovation. As well as regulatory change, this requires the right skills and infrastructure are in place to support innovation. The Society's response to the National Data Strategyconsultation highlighted how government can lead by example and share best practice, encouraging private sector organisations to collect and maintain high quality data. Where government carries out 'pathfinder' or 'lighthouse' projects to connect and use data, lessons can be learned about what it is possible to do in a well-governed way.

7.2 As argued in the Royal Society's Protecting privacy in practice.such government-led projects can also make use of the technologies that enable secure and privacy-preserving use of data. Government carrying out these lighthouse projects can also stimulate research and development in privacy enhancing technologies. Ensuring the UK has the highly skilled workforce to deliver on the promise of data-driven approaches should be a priority for the government, as it is not guaranteed that this training will happen to the level required without this support. In addition to training, viable career routes in research need to be maintained given the demand from the private sector.

8 Adequacy

8.1 In a world in which research is carried out on a truly global basis, international interaction is important to scientific success. The UK is a world leader in science, and researchers move and collaborate to pursue scientific excellence; collaboration and mobility are a key part of the business of science, and they are distinct and complementary. Mobility ensures a circulation of skills and ideas around the world, and 'brain circulation' in the global research system sees scientists follow the best science and the best resources.

Q3.2.2. To what extent do you agree that the government should consider making adequacy regulations for groups of countries, regions and multilateral frameworks?

8.2 The EU remains a hugely significant R&D partner for the UK, and it is important to ensure the UK retains data adequacy with the EU. Data sharing is an essential part of modern research. Pooled data on individuals are often needed to ensure sufficiently large study numbers, and to replicate findings and identify complex pathways, or to allow for international comparison. The consequences of not solving this problem are already impeding and will continue to impede health research in many critically important fields.

8.3 Scientists have a long history of working together, but the level of international collaboration is increasing. When UK based researchers publish internationally collaborative papers, they are more highly cited, a measure of scientific impact, than papers published by only UK-based authors. This gap has widened over time.

8.4 Pooling UK expertise with other European countries has led to advances in medicine and public heath, cleaner energy, environment and transport innovations, and the creation of jobs in UK regions. For example:

- Horizon 2020 funding is supporting UK-EU collaborative research into COVID-19. This includes obtaining data regarding the nature of the virus and how it spreads, and providing infrastructure and co-ordination across clinical research networks to ensure all nations involved have access to the best available evidence on COVID-191 .
- Cure rates for British children with leukaemia are being improved as a result of the IntReAll project involving researchers from Germany and the University of Manchester.

8.5 Further evidence on the scientific and societal benefits of close work with other European countries can be found in the Society's position statement on The UK and Horizon Europe.

8.6 The Society is a member of the All European Academies (ALLEA), a forum for science and humanities academies in the Council of Europe region, who have recently published a report that encompasses the challenges to health data sharing GDPR presents: International Sharing of Personal Health Data for Research. This highlights legal challenges that have resulted in impediments to data sharing with researchers outside the EU/EEA, affecting both the direct transfer of data to non-EU/EEA countries and remote access to data at its original location.

9 Use of personal data in the COVID-19 pandemic

Q4.3.3. To what extent do you agree with the proposal to clarify that public and private bodies may lawfully process health data when necessary for reasons of substantial public interest in relation to public health or other emergencies?

9.1 As argued in our response to the National Data Strategy, access to both health data and data relating to everyday interactions was challenging during the pandemic. The Royal Society's RAMP and DELVE initiatives found that timely access to data has been the greatest barrier to providing the best possible scientific advice and has frustrated exploiting the UK's extraordinary data science capability to fully support the response to the pandemic. This includes access to data such as mobility and transport data, held largely in the private sector.

9.2 The experience of the Covid-19 response, both positive and negative, highlights the need for the private and public sectors to adopt a duty to safely share data in an emergency, and indeed for ongoing public benefit. It is however essential that this sharing and linking of data is done in a safe and well-governed way to protect the rights of data subjects and commercial interests of companies. Extending the powers of the Office for National Statistics as a trusted data processor for appropriate types of data could be a key step in delivering that public duty safely.

9.3 The Royal Society supported the DELVE initiative, which brought a data-focused and cross-disciplinary response to the evidence needs in the early stages of the pandemic. Work by the DELVE group highlighted key steps to enable access to data in an emergency:

- Government should update the statutory objective of the Office for National Statistics (ONS) to accommodate trustworthy access to happenstance data to generate national and local statistics. Such statistics are required on very short time frames to facilitate fast decision-making for the nation in the rapidly evolving circumstances of a national emergency.
- The ONS should collaborate closely with the Information Commissioner's Office (ICO) to formulate a standardized qualification for data access, equivalent to a 'data driving license' that would demonstrate trustworthiness and ensure that qualified experts can get rapid access to different data types with the appropriate standardized ethical and legal training in place.
- Government should fund interdisciplinary pathfinder data projects. These projects should require collaborations between industries, run across government departments and integrate different academic expertise. Each project should target a specific policy question. Beyond the pathfinder role, the projects will leave a legacy in the form of expertise and guidance in understanding the stages of the data-sharing pipeline.

References

Reports

Explainable AI

Protecting privacy in practice.

Machine Learning

Data management and use

Other outputs

2021 consultation response to the National Data Strategy

2020 position statement on The UK and Horizon Europe

2019 You and AI: Conversations about AI technologies and their implications for society

2018 Note of discussion on AI, society and social good

Public views of Machine Learning

Public views of machine learning: Digital Natives

Other referenced material

ALLEA 2021 report on International Sharing of Personal Health Data for Research

Frascati manual guidelines for collecting and reporting data on research and experimental development