

# Smart Meter Energy Data: Public Interest Advisory Group

A policy dialogue and work programme  
led by  
Centre for Sustainable Energy & Sustainability First

PIAG Workshop 1 – 19 March 2018

Stimulus paper 3  
Data Ethics – A review of the landscape  
Author : Maxine Frerk

## Status of this Document

This paper was prepared as an input to the work programme of the Public Interest Advisory Group on access to smart meter energy data.

Last Revise – 10 August 2018

## Overview

While the focus of the smart metering debate to date has been based around the issues of “privacy”, more widely there is a new focus developing on the broader topic of “data ethics”. With the growth of big data and the potential for new insights to be developed from combining data from different sources and for automated decision making using AI, new issues and challenges are arising.

For example, with big data even where data is anonymised there are risks that other data sources can be used to re-identify data or that policy decisions can be taken on the basis of aggregated data that are discriminatory or disadvantage particular groups.

While the debate on smart metering is still essentially around getting access to data, the work on data ethics which relies more on a principles-based approach may provide some useful insights. And as we look to develop principles for the use of smart metering data there may be some direct read across.

The ICO argued at the Tech UK Data Ethics Summit last year<sup>1</sup> that the principles of fairness, transparency and accountability, which underpin what they do on the Data Protection Act, also provide the framework for data ethics. They argue that the GDPR, in looking to modernise data protection, covers issues such as automated decision making. However, the broad consensus is that something wider is needed and that the current data governance concepts of privacy and consent are under increasing strain.

This work was highlighted by attendees at the introductory PIAG workshop with a plea that we do not reinvent the wheel. This note therefore summarises the main work that is taking place in other sectors.

While the most direct read across is in relation to the big data issues that have developed a momentum, the thinking here is at an early stage. This note therefore also covers the analogous debates around ethics in relation to health data where there are existing governance arrangements, and the existing frameworks for ethical social research.

Potential implications for PIAG would seem to be:

- issues of data ethics need to be thought of alongside the conventional privacy questions that were addressed in the smart meter Data Access and Privacy Framework (DAPF);
- while half-hourly meter data in and of itself may not be that revealing the issues around big data point to the ways that it could be used in combination with other data to reveal additional insights;
- from a research perspective with the SMRP one of the challenges is that it is not possible to state clearly exactly what research will be done. This problem is common with big data and is precisely why the focus is on the use of ethical principles (which is a common approach already for research involving individuals);
- SMRP may wish to work with PIAG to develop a data governance framework akin to that used by the Biobank covering recruitment as well as use of the data. While the Biobank is dealing with highly

---

<sup>1</sup> <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/12/techuk-data-ethics-summit/> (accessed 30/1/18)

sensitive data having such a framework that has been subject to some external scrutiny could add to public trust in the project;

- as noted at the last PIAG the data that sits outside the “walled garden” is actually more sensitive and yet is not subject to any controls beyond standard data protection legislation. As well as looking at how access to the half-hourly data for public interest purposes might be facilitated, PIAG may want to consider arguing for some principles to be agreed around use of the more granular data. Maintaining public trust and confidence in the use of all smart meter data will be important in facilitating access to data inside the “walled garden” for public interest purposes.

### **Part 1: Data Ethics in the context of AI and Big Data**

A number of organisations have been looking at the issues in this area with a clear consensus developing around the need for a new stewardship body. The Nuffield Foundation working with others has established a Data Ethics Convention to explore the issues. Government has also now announced the creation of a Centre for Data Ethics and Innovation. The precise roles remain to be clarified and there are some concerns being voiced about potential overlaps.

#### ***DDCMS’s Centre for Data Ethics and Innovation***

The Department for Digital, Culture, Media and Sport (DDCMS) has announced that it is establishing a new Centre for Data Ethics and Innovation:

*“Government is investing £9 million in the centre as part of plans to make the UK the best place in the world for businesses developing artificial intelligence (AI) to start, grow and thrive. It will advise on the measures needed to enable and ensure safe, ethical and innovative uses of data-driven technologies. The UK already benefits from a world-class regulatory regime and the centre will build on this by making sure we understand and respond to the rapidly evolving ways in which data is impacting our lives.”*

The original announcement was made in the Budget last year but the PM at Davos (Jan 2018) launched the search for an interim chair to establish the body.

#### ***Royal Society and British Academy - Data management and use: Governance in the 21st century***

The Royal Society and the British Academy have conducted a major project on the needs of a 21st century data governance system<sup>2</sup>.

The report identifies that the amount of data generated from the world around us has reached levels that were previously unimaginable. Meanwhile, uses of data-enabled technologies promise benefits,

---

<sup>2</sup> <https://royalsociety.org/~media/policy/projects/data-governance/data-management-governance.pdf>

from improving healthcare and treatment discovery, to better managing critical infrastructure such as transport and energy. These new applications can make a great contribution to human flourishing but to realise these benefits, societies must navigate significant choices and dilemmas.

In the report they use 'data governance' to mean everything designed to inform the extent of confidence in data management, data uses and the technologies derived from it, together with the purpose to which it is put. These stages, which used to be more separate, are now often tangled with each other across applications and across the world.

They make the point that the traditional data lifecycle was clear, relatively sequential, predictable, often managed by a single organisation and made it comparatively easy to erase data sets that were no longer needed. These characteristics also meant that data governance could focus on a specific point in the 'cycle', such as on collection, and use this to control the broader process. However, this traditional approach is now under considerable strain because of the increased complexity of data lifecycles. Instead of individual organisations operating independent data lifecycles, with linear approaches to processing, data is now generated and exchanged across many organisations, often without the subject's awareness.

In addition, consent is proving an increasingly difficult concept. Consent is one of the legal grounds for processing personally identifiable data in the current data protection regime. However, genuine consent is difficult to achieve. The application of consent suffers from what is often referred to as the 'transparency paradox'. Consent requires transparency of what is being consented to. Such transparency has to be meaningful. Anything too long or complex is unlikely to be broadly understood or read yet making a summary widely comprehensible often discards the details that people care about. As data collection become less about the active 'giving' of information and more about information captured as a by-product of interactions with products, services, the physical environment and each other, it is increasingly difficult for individuals to provide meaningful consent to share data.

The report presents a Framework for social and ethical tensions which includes:

- Using data relating to individuals and communities to provide more effective public and commercial services, while not limiting the information and choices available.
- Making use of the data gathered through daily interaction to provide more efficient services and security, while respecting the presence of spheres of privacy.
- Providing ways to exercise reasonable control over data relating to individuals while encouraging data sharing for private and public benefit.
- Making the most of the ability of algorithms to provide accurate outcomes beyond the human ability while ensuring appropriate levels of interpretability and transparency, and allowing for systems of accountability to be put in place.

The report argues that in this fast-moving landscape, governance challenges need to be addressed in a timely manner if the overall system of governance for data management and data use is to maintain public trust. It argues that two responses are required:

- Firstly, a set of high-level principles to help visibly shape all forms of data governance and ensure trustworthiness and trust in the management and use of data as a whole.
- Second, the creation of a body to steward the evolution of the governance landscape as a whole.

The principles that they came up with are very broad. The “promotion of human flourishing” is the overarching principle that they argue should guide the development of systems of data governance (notwithstanding some criticism of this as a difficult concept). There are then 4 principles beneath this:

- protecting individuals but also thinking about collective rights and interests;
- having any trade-offs discussed transparently, accountably and inclusively;
- having systems in place to learn from the good practice of others and also to learn from bad practice and to correct errors or mistakes in the system;
- enhancing existing democratic governance to be sufficiently flexible and robust to evolve as these technologies continue to develop.

#### *Supporting materials*

The project involved a significant amount of research and dialogue and the supporting materials that they produced are a useful resource:

- Case Studies: They present a set of detailed case studies<sup>3</sup> including one on smart metering which draws primarily on academic research. It talks about the privacy risks of real time metering data, including the claim that one can identify which appliances are being used and even which TV programme is being watched from such data (which actually requires more granular data than is even available over the HAN and certainly is not possible from the half-hourly data). They discuss the potential for privacy enhancing technologies such as Cryptographic computation which allows mathematical operations to be carried out while data is still encrypted, and then allows the decryption of the final result but not any of the individual parts. There is some criticism of DECC for not having explored these more fully. They also compare the central DCC model in terms of institutional arrangements with those in Holland and Germany but they do not really seem to understand the DCC model (ie that it does not store data) or the data access and privacy framework that is in place.
- A landscape review<sup>4</sup> setting out the privacy legislative framework.

---

<sup>3</sup> <https://royalsociety.org/~media/policy/projects/data-governance/data-governance-case-studies.pdf?la=en-GB>

<sup>4</sup> <https://royalsociety.org/~media/policy/projects/data-governance/data-governance-landscape-review.pdf>

- A literature review of past consumer engagement exercises on data and its use<sup>5</sup> which highlights low levels of awareness (of the technology, of how much data is collected and of the regulatory protections in place), the importance of transparency, different views on use of data by commercial bodies versus public sector. It notes a general desire for people to have greater control over their data and to know who was holding data on them. In addition, studies have shown that people are generally much more comfortable when data is anonymised and aggregated, and cannot be traced back to them or used to target them. Studies also showed that greater acceptance and participation in research studies can be gained if there is an output which leads to perceived benefits at the individual, local, regional or national level. Concerns were raised about organisations holding data for longer than they should, inaccurate data, and the use of linked administrative data to justify controversial policies such as the Bedroom Tax. Mixed views were recorded on the idea of surveillance. The ability of organisations to keep data securely was vital for trust and there was a lack of trust in particular for commercial organisations.
- The output<sup>6</sup> from a seminar held in October 2017 including papers from a wide range of different perspectives. Participants raised a concern that the discussions are taking place in silos when the opportunities and challenges come from combining data from very different sources.

### ***Nuffield Foundation's Convention on Data Ethics***

The Nuffield Foundation is developing plans in partnership with the Royal Society, British Academy, Royal Statistical Society and Alan Turing Institute, to establish an independent Convention on Data Ethics. This deliberative body will bring together academics, policy makers and public and private sector interests to improve public understanding and to explore solutions for fairer and safer data-use arising from technological innovation, regulation, or changes in public behaviour.

With a working title of Convention on Data Ethics and Artificial Intelligence, this independent body will seek to ensure that the power of data – and the automated technologies that serve to augment it – is harnessed to promote human flourishing, both for society as a whole and for different groups within it. The convention will have three main aims:

- **Represent** the interests of society in debates on ethical data use at a national and international level.
- **Promote** and support a common set of data practices that are deserving of trust, and are understandable, challengeable and accountable.

---

<sup>5</sup> <https://royalsociety.org/~media/policy/projects/data-governance/data-governance-public-engagement-review.pdf?la=en-GB>

<sup>6</sup>

[https://royalsociety.org/~media/policy/Publications/2017/Data\\_management\\_and\\_use\\_governance\\_in\\_the\\_21st\\_century\\_2017\\_seminar\\_report.pdf](https://royalsociety.org/~media/policy/Publications/2017/Data_management_and_use_governance_in_the_21st_century_2017_seminar_report.pdf)

- **Convene** different interests to develop shared terminology for data ethics and promote human flourishing.

They are currently engaging with stakeholders and will then appoint a Chair, with the aim of establishing the Convention by the end of 2018.

This body would consider questions, problems and opportunities arising from uses of data and AI, which are not unlawful but have potential to:

- cause widespread or profound economic or social harm – or good- to society, or different groups within society;
- challenge or destabilise social and democratic norms or principles (such as ownership; consent; privacy; professional expertise; regulation);
- facilitate future developments with unknown consequences;
- introduce inconsistencies between treatment or rights in the offline and digital spheres, or between different domains;
- effect change in the UK but inform thinking internationally.

### ***Wendy Hall's Report to Government on the AI industry***

DDCMS commissioned Dame Wendy Hall to produce an independent report on how to support AI<sup>7</sup>. The report talks about building on the Turing legacy and making the Turing Institute responsible for promoting AI which they now are. While most of the recommendations are about how best to promote the sector (building expertise, promoting research and increasing demand) it does also include a recommendation to establish “data trusts” (proven and trusted frameworks and agreements) to improve trust and ease around sharing data.

The report's recommendations for improving access to data are:

- Open data (government making its data open where possible) without implicating personal privacy. The Open Data Institute operates in this space driving government (and private sector) to make data available.
- Managing sensitive data. Managing access to sensitive data has two key aspects: - providing trust and confidence and reducing transaction costs. The main example is healthcare data where even anonymised datasets could be of value to innovators but there are difficulties in knowing how to build the trust for data sharing. The security challenges can be overcome by agreements, but at the moment agreements are made on an ad hoc basis, incurring large transaction costs and making it difficult for smaller companies to compete. Solutions that provide trust and confidence also need to be affordable and sustainable. If legal and procedural costs remain high for each transaction, a particular solution for data-sharing will not be taken up widely.

---

<sup>7</sup> <https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk>

- To enable this to be done more easily and frequently, it is proposed to develop terms and mechanisms for these parties to form, between them, individual “data trusts” to enable AI to be developed to meet the needs of the parties involved and allow data transactions to proceed with confidence and trust. These trusts would not be a legal entity or institution, but rather a set of relationships underpinned by a repeatable framework, compliant with parties’ obligations, to share data in a fair, safe and equitable way.
- A support organisation, the Data Trusts Support Organisation (DTSO) could be developed, which would lead on the development of tools, templates and guidance for those who want to share and use data. A data stewardship body as proposed by the Royal Society and the British Academy could provide advice and oversight.
- Government agencies already have experience in secure data-sharing for researchers which may be relevant to the development of data trusts, for instance the Office for National Statistics Virtual Microdata Lab and ‘Five Safes’ framework.

### ***The Alan Turing Institute***<sup>8</sup>

The Alan Turing Institute focuses on the potential for data science to change the world for the public good. Understanding the ethical and societal implications of data is one of The Alan Turing Institute’s key research priorities and they have created a Data Ethics Group to lead their research in this area. The Group works in collaboration with the broader data science community, supports public dialogue on relevant topics, and sets open calls for participation in workshops, as well as public events.”

The Chair of the Data Ethics Group, Luciano Floridi, recently co-edited a special theme issue of the Philosophical Transactions of the Royal Society, ‘The ethical impact of data science’<sup>9</sup> The theme issue included contributions from several members of the Data Ethics Group. Luciano’s introductory paper on “What is data ethics?” says that the ambition is to establish *data ethics* as a new branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g. right conducts or right values

It argues that *social acceptability* or, even better, *social preferability* must be the guiding principle for any data science project with even a remote impact on human life, to ensure that opportunities will not be missed. On the other hand, overemphasizing the protection of individual rights in the wrong contexts may lead to regulations that are too rigid, and this in turn can cripple the chances to harness the social value of data science.

---

<sup>8</sup> <https://www.turing.ac.uk/data-ethics/>

<sup>9</sup> <http://rsta.royalsocietypublishing.org/content/374/2083>



In this context, key issues concern possible re-identification of individuals through data-mining, -linking, -merging and re-using of large datasets, as well as risks for so-called 'group privacy', when the identification of types of individuals, independently of the de-identification of each of them, may lead to serious ethical problems, from group discrimination (e.g. ageism, ethnicism, sexism) to group-targeted forms of violence. Trust and transparency are also crucial topics in the ethics of data, in connection with an acknowledged lack of public awareness of the benefits, opportunities, risks and challenges associated with data science. For example, transparency is often advocated as one of the measures that may foster trust. However, it is unclear what information should be made transparent and to whom information should be disclosed.

### ***Big Data Dilemma Inquiry by Science and Technology Select Committee (2016)*<sup>10</sup>**

This inquiry primarily focussed on the opportunity for the UK and issues around skills and open data. It did include a recommendation for a Council on Data Ethics to look at balancing privacy, anonymisation, security and public benefit.

### ***Royal Statistical Society***

This is another group that were calling for a Council for Data Ethics. They have highlighted in particular a trust deficit where people are less trusting in how companies will use their data than they are in that company generally<sup>11</sup>. They argue this trust deficit will reduce the potential for data to be used for public good.

The consumer research<sup>12</sup> that they published is also helpful context. It includes, for example, the finding that consumers are very supportive of the use of their data by university researchers for government funded research.

---

<sup>10</sup> <https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2015/big-data/>

<sup>11</sup> <https://www.statslife.org.uk/features/3417-trust-deficit-could-hinder-efforts-to-use-data-for-good>

<sup>12</sup> <http://www.statslife.org.uk/images/pdf/rss-data-trust-data-sharing-attitudes-research-note.pdf>

**EU Commission**

While the UK may have an ambition to lead the thinking on data ethics it is clearly also on the EU agenda. The European Data Protection Supervisor has produced a report on digital data ethics<sup>13</sup>. And at a recent event on GDPR where the EDPS was speaking (via video link) they talked a lot about the rise of data ethics. Their aim is to identify the questions that need wider debate across society – not to try to develop a parallel legal regime.

In considering how to deal with these data ethics issues the report returns to the fundamental human values of dignity, freedom, autonomy, solidarity, equality, democracy, justice and trust.

**US**

In the US a Council for Big Data, Ethics and Society<sup>14</sup> was established in 2014 to bring together researchers from a range of disciplines thinking about these issues. In some sense it is the model for the sorts of governance / advisory bodies that are being established now in GB. That said nothing has been posted on their website since 2016 which suggests attention is focussed elsewhere.

---

<sup>13</sup> [https://edps.europa.eu/sites/edp/files/publication/18-01-25\\_eag\\_report\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf)

<sup>14</sup> <https://bdes.datasociety.net/>

## **Part 2: Data Ethics in a health context**

### ***Care.data lessons***

In mid-2013 NHS England initiated the development of care.data, an ambitious programme to pool all GP patient information into one central database to be held by the newly created Health & Social Care Information Centre (HSCIC), now NHS Digital. The aim of the programme was to produce a single resource accessible by statisticians and researchers who could use it, among other things, to help healthcare providers find new ways to improve and enhance practices across the NHS. But delays and controversy, including privacy concerns, led to a decision to close the project down.

Ahead of that decision a formal review was carried out by the National Data Guardian (Dame Fiona Caldicott). This proposed new standards for opt-out and consent in relation to the use of health data<sup>15</sup>. The report proposes a clear opt out from the sharing of data and sets out possible wording for allowing individuals to make their choices (eg whether to share data to help in running the NHS or for research) while keeping the options simple. It also concludes that people should not be able to opt out from providing anonymised data (given the importance in a health context of having fully representative data). It recognises that if NHS digital could provide a linked but anonymised data set this could reduce the need for individual commissioning bodies to hold large databases to enable them then to link the data themselves. The report also stresses the need for public engagement on the issues and a focus on security to maintain trust.

The care.data project fell short in a number of ways. A leaflet was sent to every household in England but it was not made clear how the information would be used or what to do if you wanted to opt out. Particular concerns were around whether and on what basis the information could be provided to insurance companies. Although the National Data Guardian review did not look explicitly at the care.data arrangements the reviews recommendations clearly went beyond what was being done on care.data and contributed to the project being pulled.

Following the care.data debacle there are still seen to be some tensions between the idea of having a large centralised database (even if anonymised) versus more local solutions where local engagement can be more effective. For example, a Connected Health Cities project has been launched in the North of England which has been undertaking consumer research to understand how people feel about sharing their health data with the ambition of linking health and social care data to improve local services.

More generally, the National Data Guardian (NDG) advises and challenges the health and care system to help ensure that citizens' confidential information is safeguarded securely and used properly. The NDG's role is to help make sure the public can trust their confidential information is securely safeguarded and make sure that it is used to support citizens' care and to achieve better outcomes from health and care services.

---

<sup>15</sup> [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/535024/data-security-review.PDF](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF)

The three principles that have informed the NDG's work are:

- To encourage sharing of information in the interests of providing care to individuals.
- There should be no surprises to citizens and they should have choice about the use of their data.
- There must be dialogue with the public, helping to increase their knowledge and choices about how data is used to improve health and care.

An Understanding Patient Data<sup>16</sup> taskforce was established after the care.data fiasco and has emphasised the importance of public engagement. Understanding Patient Data supports better conversations about the uses of health information. Its aim is to explain how and why data can be used for care and research, what's allowed and what's not, and how personal information is kept safe. It works with patients, charities and healthcare professionals to champion responsible uses of data.

### ***Health and Google Mind***

When Google acquired Deep Mind it committed to setting up an ethics board. Despite the initial commitments it appears this is just an internal board and there is no visibility of what it is doing.

However in the specific context of health and the partnership between Deep Mind and the NHS (Deep Mind Health) an external independent review panel was set up involving academics in the data sphere and health practitioners<sup>17</sup>. The group produce an annual report. The latest report argues that the bar needs to be set particularly high because of the links with Google and the huge amount of data they have access to. Although health data is not being shared with Google the public perception remains important for continuing trust and confidence in the project. The difficulties are highlighted of getting messages out through the media (who aren't interested in reassuring stories).

### ***UK Biobank***

UK Biobank is a large long-term study in the UK which is investigating the respective contributions of genetic predisposition and environmental exposure (including nutrition, lifestyle, medications etc.) to the development of disease. The study is following about 500,000 volunteers in the UK, enrolled at ages from 40 to 69. Initial enrolment took place over four years from 2006, and the volunteers will be followed for at least 30 years thereafter. The data collected involves an initial assessment visit (capturing weight, height, blood pressure, lung capacity etc) plus follow-ups, linking with health records and also some diary type surveys.

---

<sup>16</sup> <https://understandingpatientdata.org.uk/about-us>

<sup>17</sup> <https://deepmind.com/applied/deepmind-health/transparency-independent-reviewers/independent-reviewers/>

From 2012, researchers were able to apply to use the database (though they are not given access to the volunteers, who will remain strictly anonymous). A typical study using the database might compare a sample of participants who developed a particular disease with a sample of those that did not, in an attempt to measure the benefits, risk contribution and interaction of specific genes, lifestyles, and medications.

The UK Biobank project operates within the terms of an Ethics and Governance Framework<sup>18</sup> The Framework describes a series of standards to which UK Biobank will operate during the creation, maintenance and use of the resource and it elaborates on the commitments that are involved to those participating in the project, researchers and the public more broadly. The independent UK Biobank Ethics and Governance Council provides advice to the project and monitors its conformity with the Framework.

The Framework<sup>19</sup> covers issues such as the approach to recruitment, understanding and consent, confidentiality, principles of access by research users and governance.

The Ethics and Governance Council acts as independent guardian of the Framework. This means that it has responsibility for advising on revisions and monitoring and reporting publicly on the conformity of the UK Biobank project with the Framework. They review any changes to the framework, explore any significant or novel issues (which Biobank flags to them) and produce an annual report. They do not approve individual research proposals but would expect to discuss with Biobank any novel or contentious applications and to look more generally at how the range of research undertaken sits against the Framework.

### **Part 3: Social research ethics**

While the data ethics issues explored above are probably most relevant to smart metering, it is worth also looking at the existing guidelines around social research ethics, some of which may be of direct relevance to the UCL Smart Meter Research Portal (SMRP).

#### ***EU<sup>20</sup>***

At an EU level, guidance on socio-economic research ethics sets out general principles about voluntary participation; the decision to participate being based on informed consent; all data being treated with appropriate confidentiality and anonymity; and an endeavour to ensure participants are protected from undue intrusion, distress or other harm.

---

<sup>18</sup> <https://egcukbiobank.org.uk/Ethics-and-governance-framework.html>

<sup>19</sup> <http://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf?phpMyAdmin=trmKQIYdjjnQIgJ%2CfAzikMhEnx6>

<sup>20</sup> <http://www.respectproject.org/ethics/412ethics.pdf>

The guidance does then include a discussion of the tensions that arise in some of these areas such as the problems with informed consent given the difficulties in saying in advance all the potential research uses the data could be put to. The paper suggests making general statements about usage; not viewing this as a one-off process and in particular taking care if the data could be used to reach policy conclusions that might disadvantage some participants (so they would be indirectly disadvantaged).

### **UK – ESRC**

The ESRC expects all research that is carried out using ESRC funding to comply with ethical principles. It provides guidance on ethics<sup>21</sup> which includes six key principles for ethical research:

- research should aim to maximise benefit for individuals and society and minimise risk of harm
- the rights and dignity of individuals and groups should be respected
- wherever possible, participation should be voluntary and appropriately informed
- research should be conducted with integrity and transparency
- lines of responsibility and accountability should be clearly defined
- independence of research should be maintained and where conflicts of interest cannot be avoided they should be made explicit.

Universities will often have ethics committees and ESRC publishes case studies where potential issues are explored.

### **UK - ADRN**

The Administrative Data Research Network (ADRN) allows accredited researchers to access de-personalised administrative data for social and economic research.

The Network provides a single point of access through which researchers can be screened to ensure they are properly trained to handle potentially sensitive information. It also looks to identify ways of linking different data sets together without compromising privacy (which might be usefully looked at as part of the technical options work in PIAG). A public consultation was undertaken (“Dialogue on Data”) – which informed the founding principles.

The ADRN has its own ethics framework<sup>22</sup> recognising that consent will not have been used to collect the data. The ADRN relies instead on legal powers to collect and then anonymise the data. Although the anonymised data is not counted as personal data it is still treated as sensitive data in order to protect the integrity of the overall process. The principles that are followed focus on maintaining the

---

<sup>21</sup> <http://www.esrc.ac.uk/funding/guidance-for-applicants/research-ethics/our-core-principles/>

<sup>22</sup> [https://adrn.ac.uk/media/174021/ethics-and-administrative-data-guidance\\_00\\_09\\_pub.pdf](https://adrn.ac.uk/media/174021/ethics-and-administrative-data-guidance_00_09_pub.pdf)

privacy and security of the data rather than the ethical content of the research itself which is covered by ESCRC and university ethics frameworks.

The ADRN framework involves the concepts of:

- **safe projects** – projects have to be approved by an Approval panel
- **safe people** – researchers have to be from academia, charities, third sector, government or an independent research organisation recognised by the ADRN governing board. They will not currently consider proposals from researchers from commercial organisations. Researchers have to be trained in the safe use of data.
- **safe data** – data is de-identified (ie all names and other identifiers are removed and are not accessible by the researchers). Report findings are reviewed by ADRN to check for privacy.
- **secure environments** – research can only be carried out in an ADRN secure centre with no mobiles or other data devices allowed in the room.

The ethical considerations that the ADRN expects to be considered for each project are:

- the additional moral responsibility to safeguard the data given that individuals will not have given explicit consent for use of their data for research;
- the risk of re-identification
- the sensitivity of the data (even though it has been de-identified).

This is very similar to the approach of the UK Data Service at the University of Essex which utilises a “5 Safes” framework (but covering the same basic ground).<sup>23</sup>

## **Part 4: Other**

### ***Geospatial commission***

A new Geospatial Commission that will start by establishing how to open up Ordnance Survey mapping data to UK-based small businesses, committing £80 million over the next two years. While this is a useful example of the government’s commitment to opening up access to data it does not seem at this stage to raise the sorts of data ethics issues that are flagged elsewhere.

---

<sup>23</sup> <http://blog.ukdataservice.ac.uk/access-to-sensitive-data-for-research-the-5-safes/>