Administrative Data Research Network

An ESRC Data Investment

Legal Issues for ADRN Users

ADRN Publication

Authors: Jessica Bell and Heather Gowans

Series editors: Elaine Mackey & Mark Elliot

Better Knowledge Better Society

Legal Issues for ADRN Users

Written by Jessica Bell and Heather Gowans

Edited by Elaine Mackey and Mark Elliot

ADRN Publication - December 2016

This guide, 'Legal Issues for ADRN Users' was created by Jessica Bell and Heather Gowans, edited by Elaine Mackey and Mark Elliot and published by the Administrative Data Research Network.

Contents

1. Introduction 1

1.1. Description of ADRN 1
1.2. Purpose of this guide and who it is aimed at 2

2. Data protection legislation and its impact on data sharing 3

2.1. The Data Protection Act 1998 (DPA) 3

2.1.1. The eight data protection principles 4
2.1.2. Fairly and lawfully 5

2.2. Research, history and statistics (Section 33) 6

2.3. Other law and legislation pertaining to data sharing 7

3. Lawful data access and data sharing 9

3.1 Data Linkage 10
3.2 Data Security 11

4. Issues beyond the law 12

5. The future: EU General Data Protection Regulation 2016 13

6. Further reading 15

7. Legislation 17

8. Useful websites 18

9. Glossary of non-standard terms 19

1. Introduction

1.1. Description of ADRN

The Administrative Data Research Network (ADRN) is a UK-wide partnership between academia, government departments and agencies, national statistical authorities, funders and the wider research community to facilitate new economic and social research based on routinely collected government administrative data.

The Network is establishing a new legal, secure and efficient pathway for the research community to access de-identified linked administrative datasets.

This will potentially benefit our society by providing a greater evidence base to inform policy.

The Network consists of:

- four Administrative Data Research Centres (ADRCs):
  - ADRC England: led by the University of Southampton
  - ADRC Northern Ireland: led by Queen's University Belfast
  - ADRC Scotland: led by the University of Edinburgh
  - ADRC Wales: led by Swansea University
- an overarching Administrative Data Service, which is the co-ordinating body of the Network
- administrative data owners
- the Economic and Social Research Council (the funding body)
- the UK Statistics Authority (chairing the ADRN Board)

The Network has commissioned this guide to support the development of knowledge and skills in the subject topic area.

1.2. Purpose of this guide and who it is aimed at

This guide requires no previous knowledge or experience of the topic. It is aimed at academic researchers who have an association with the Administrative Data Research Network as well as at a general audience interested in the subject matter. For Network researchers, the document will serve as useful background for the legal aspects of the certification training they receive before they can access the service. Wider audiences might be interested in how the Network's secure environment ensures the data access we allow is fair and lawful.

The guide sets out the legal background to data protection laws in the UK, and offers a broad explanation of the current law relating to data sharing and linkage, as well as a consideration of the implications of the impending EU General Data Protection Regulation 2016 (GDPR). There is also consideration of some non-legal issues surrounding the topic. Readers can refer to the Network's website for further information, and should consult professional legal advice on any specific legal points.

EU law will still apply to the UK as an EU Member State until the UK formally withdraws from the European Community and enters into a 'withdrawal agreement' in accordance with Article 50 of the Treaty on the European Union. The EU Treaties will cease to apply to the UK once the withdrawal agreement is implemented. The withdrawal agreement is likely to repeal the European Communities Act 1972 (which forms the basis of the UK membership of the EU) and the UK may have to take steps to make sure EU laws which are currently directly effective in the UK continue to apply.

It has already been confirmed by the Information Commissioner's Office (ICO) that the GDPR will apply in the UK from 25 May 2018 and the ICO has also confirmed that the Data Protection Act 1998 remains in force in the UK irrespective of the referendum result. The government has confirmed that the UK's decision to leave the EU will not affect the commencement of the GDPR.

2. Data protection legislation and its impact on data sharing

When considering the legal issues surrounding data protection and the use of data for research in the UK, it is necessary to understand the relationship between domestic (national) law and European Union law. For the purposes of this guide, it is useful to subdivide the legal structure in the UK into national law, European law (which has been implemented into UK law), and UK statutes (written laws that have been passed by an Act of Parliament). Wherever there is conflict between the provisions of EU law and domestic (national) law of an EU Member State, then EU law will (for the time being) prevail according to the principle of supremacy.

There is no single source of law that regulates the use and sharing of personal data. The collection, use and dissemination of personal data are governed by several different laws. The principal ones in relation to data sharing are most prominently the Data Protection Act 1998, the common law tort of breach of confidence; the Human Rights Act 1998 (and the European Convention on Human Rights); and the Statistics and Registration Services Act 2007.

2.1. The Data Protection Act 1998

The Data Protection Act (DPA) became law in the UK on 1 March 2000, and implements EU Directive 95/46. It is the keystone in UK information law which regulates the uses of personal data held manually or on computer. The aim of the Act is to protect the rights of the individual about whom data is obtained ('the data subject'), stored, processed or supplied. The Information Commissioner is responsible for overseeing and promoting compliance with the DPA; and the ICO has a wide range of criminal and non-criminal tools available for regulatory action.

The DPA applies to 'personal data' (Section 1), which is defined as data relating to a living individual who can be identified either from that data alone, or from that data in conjunction with other data in the data controller's possession. The Act also refers to 'sensitive personal data' (DPA Section 2), such as race, politics or religion to which more stringent rules apply. Personal data must be processed in accordance with the eight data protection principles set out in the DPA. 'Processing' includes the holding, obtaining, recording, use, disclosure and sharing of information.

The ICO has provided further guidance on the standard to apply to decide what is personal data and identifiability:

"The starting point might be to look at what means are available to identify an individual and the extent to which such means are readily available. For example, if searching a public register or reverse directory would enable the individual to be identified from an address or telephone number, and this resource is likely to be used for this purpose, the address or telephone number data should be considered to be capable of identifying an individual. When considering identifiability it should be assumed that you are not looking just at the means reasonably likely to be used by the ordinary man in the street, but also the means that are likely to be used by a determined person with a particular reason to want to identify individuals. Examples would include investigative journalists, estranged partners, stalkers, or industrial spies. Means of identifying individuals that are feasible and cost-effective, and are therefore likely to be used, will change over time. If you decide that the data you hold does not allow the identification of individuals, you should review that decision regularly in light of new technology or security developments or changes to the public availability of certain records. Taking this into account, a person who puts in place appropriate technical, organisational and legal measures to prevent individuals being identifiable from the data held may prevent such data falling within the scope of the Directive" (2012(b):9).

2.1.1. The eight data protection principles

The 'data protection principles' (set out in DPA Part 1 of Schedule 1) require that personal data should be:

- processed fairly and lawfully
- obtained only for one or more specified lawful purpose
- adequate, relevant and not excessive for its purpose(s)
- accurate and where necessary, kept up to date
- not kept for longer than is necessary
- processed in accordance with the data subject's rights under the DPA
- secure
- not transferred outside the EEA unless an adequate level of data protection is ensured

2.1.2. Fairly and lawfully

In order to 'fairly' process personal data (DPA Part II of Schedule 1), regard must be had:

- to the method by which the personal data are obtained
- that the person obtaining the information is authorised, or required by, or under any enactment, to obtain it
- that the following information is made 'readily available' to the data subject:
    1. the identity of the data controller or any nominated representative
    2. the purpose(s) for which the data are intended to be processed
    3. any further information necessary to enable the processing to be fair

The personal data will only be deemed to be 'lawfully' processed where it is in accordance with at least one of the conditions set out in Schedule 2 of the DPA:

- the data subject has given consent to the processing
- the performance of a contract
- in compliance with a legal obligation, other than a contract
- in the vital interests of the data subject
- where it is necessary for the administration of justice or for the exercise of certain specified functions; or where the processing is necessary for the legitimate interests pursued by the data controller or third party to whom the data are disclosed, without infringing the rights of the data subject

With regard to schedule 2, the Network lawfully processes on the basis of legitimate interest or consent.

It is worth particularly noting that – in terms of compliance with the DPA (and the Human Rights Act 1998) – one does not necessarily always need consent of the data subject to share their personal data. The DPA sets out seven possible criteria under Schedule 2 for the legitimate processing of personal data (and sharing, like using, is for the most part just another form of processing), and if any one of the criteria is met, the DPA test is satisfied.

Furthermore, consent in relation to personal data does not need to be explicit – it can be implied. More stringent rules apply to sensitive personal data, when consent does need to be explicit if that criterion is used, but criteria other than consent can still be used for sensitive personal data. Even without explicit consent for the sharing of sensitive personal data, it is still possible to share the data legitimately if this is necessary in order to exercise any statutory function, or to protect the vital interests of the data subject where, for example, consent cannot be given.

If the information to be processed is sensitive personal data, the proposed processing will need to satisfy a further condition in Schedule 3 of the DPA, in addition to the Schedule 2 condition. The most common Schedule 3 conditions likely to be satisfied will be where the data subject(s) has given "explicit consent" for the processing (Schedule 3 condition 1); or where the information "has been made public ... by the data subject" (Schedule 3 condition 5); or specifically with regard to the processing of sensitive personal data for research purposes, the data must relate to research activities that are "in the substantial public interest".

This last condition is the one that is most relevant to the Network, and is why – as part of any application for access to data through the Network – a researcher has to demonstrate the potential public benefit of his or her research when the information being processed contains sensitive personal data.

2.2. Research, history and statistics (Section 33)

Section 33 of the DPA provides specifically for certain exemptions in respect of the processing (or further processing) of personal data for research purposes. The term 'research purposes' is not defined in the DPA, other than clarifying that it includes statistical or historical purposes. This exemption is important to the Network because data-holding government departments often rely on it to be able to share data, even in the highly secure setting of the Administrative Data Research Centres.

For this reason the wording of this exemption – what it allows and what it does not – is one factor in determining the design of services like the Network. It is therefore valuable for researchers to understand section 33 so they understand why the Network does things the way that it does.

Section 33 only applies as an exemption where the (further) processing of personal data for research is exclusively for such purposes and, also, where:

- the personal data are not processed to support actions affecting particular individuals
- the processing does not take place 'in such a way that substantial damage or distress is, or is likely to be, caused to any data subject'

This so-called 'research exemption' is quite narrow and it only affects three of the data protection principles:

- the second (relating to the purpose for which the data were obtained)
- the fifth (the duration for which the data can be kept)
- the sixth (relating to a data subject's right of access under s.7 DPA)

Where the exemption applies, therefore:

- the further processing of personal data for research is permitted, irrespective of the original reason for collecting the data
- the personal data may be kept indefinitely, despite the fifth data protection principle
- it will be exempt from the data subject's rights of access (Section 7) if they are processed in compliance with the relevant conditions
- the results of the research or any resulting statistics are not made available in a form which identifies or causes harm to the data subject(s)

The section 33 'research exemption' does not give a blanket exemption from all of the data protection principles which apply to personal data that are provided and/or used for research. Researchers wishing to use personal data must be aware that most of the data protection principles will still apply and must be adhered to.

2.3. Other law and legislation pertaining to data sharing

If personal data is shared without the consent of the data subject, but in a manner that is in accordance with the DPA, this may nevertheless contravene other laws such as the Human Rights Act 1998 (HRA, which embeds the European Convention on Human Rights (ECHR) in UK law) and the common law duty of confidentiality.

The duty of confidentiality provides that where there is a confidential relationship, such as doctor/patient, the person receiving the confidential information is under a duty not to pass on the information to a third party. To be protected by this duty, the information has to have a 'quality of confidence' and it has to have been given in circumstances giving rise to an expectation of confidentiality. However, confidentiality is not an absolute right and information can be shared without breaching the common law duty if:

- the information is not confidential in nature
- the person to whom the duty is owed has given explicit consent
- there is an overriding public interest in disclosure; or
- sharing is required by a court order or other legal obligation

There is no single definition of what constitutes public interest. In the context of information disclosure, it must be decided case by case whether the public interest is achieved by disclosing the information or keeping it confidential. Public interest justifications may include protecting individuals from the risk of serious harm, or to enable secondary uses of information, such as for research purposes, which will benefit society over time (for further information see the Scottish Health Informatics Programme website).

If, after considering these factors, it is decided that the information should be disclosed, the disclosing body must only disclose the minimum of information necessary to achieve the objective of the research project, in accordance with the principle of proportionality.

In addition, sharing information may be a breach of an individual's Article 8 'right to respect for private and family life' (found in Article 8(1) of the HRA). The scope of Article 8 is broad and covers the collection, use and exchange of personal data. When considering Article 8, it is necessary to take into account not only the content of the data, but also the anticipated use of the data.

The right to private life is not absolute, though, as Article 8(2) HRA justifies interference with the right if it is:

1. in accordance with the law
2. in pursuit of a legitimate aim
3. necessary in a democratic society

Therefore, when disclosing information, article 8(2) requires a balancing exercise between the rights of the individual to whom the information relates and the pursuit of legitimate interests – and, where the balance is in favour of the latter, whether sharing the information is a proportionate response to this interest.

Thus, neither the common law duty of confidence nor the Human Rights Act 1998 prevent the sharing of personal information, as long as the common law duty of confidence and the rights enshrined within the Human Rights Act 1998 are balanced against the effect on the individual or others of not sharing the information. Any information disclosure, even if justified, must also comply with the DPA 1998 (described above).

The Statistics and Registration Service Act 2007 (SRSA) applies specifically to data designated as 'official statistics' and includes statistics produced by the Office for National Statistics (ONS), central government, and devolved administrations. It provides that 'personal information' which relates to and identifies an individual or business, and which is held or disclosed by the Statistics Authority, is confidential (SRSA section 39(1)).

The SRSA does permit certain specified disclosures of personal information held by it (SRSA section 39(1)), which includes to an 'approved researcher', i.e. an individual to whom the Statistics Authority has granted access for the purposes of statistical research (SRSA section 39(5)). Other specific statutory gateways may permit data sharing.

3. Lawful data access and data sharing

If sharing personal data for research is to be lawful, the organisations holding the data need to address the data protection compliance issues that are likely to arise. Depending on the nature of the access, the researchers seeking access to the data may also need to address these issues. This requires a consideration of the legal constraints on data sharing in the UK mentioned above.

The main questions are:

- Would the data sharing breach any common law duty of confidence?
- Would the data sharing amount to a breach of any human rights?
- Would the researcher applying for access to personal data need to acquire an 'approved researcher' status?
- Would the data sharing be in accordance with the Data Protection Act 1998, and in particular with the data protection principles contained therein?

A decision has to be made as to whether it is appropriate to share personal data for a particular purpose. In this context, a secure setting such as the Network can provide researchers access to data, which would otherwise be unlikely to be shared.

These proven methods of safely accessing and sharing datasets in a safe-setting environment improves the level of access to data for certified researchers; and reduces the risk of a breach of security of the personal data that is being shared.

In the Network, researchers are only given access to de-identified administrative data (i.e. data that have had information that directly identifies individuals removed) in a secure setting. Such data have already been functionally anonymised for use by researchers, so that the user is not – at the point of use – processing personal data.

Statistical outputs from the researcher may be released for publication, but these are checked by an output checker who works with the researcher to make sure such outputs are non-disclosive.[1] This process in turn makes sure no personal data is released.

1 See Lowthian and Richie's forthcoming guide to disclosure control in this series for further details about this process.

3.1. Data linkage

Data linkage is an increasingly important feature of research in the health and social sciences. When conducting research, we gather an immense amount of information, yet there may be administrative data (routinely collected government data such as health, education, or economic records) that can enhance, compliment or improve this collected information. Administrative data can provide researchers with a rich source of information, but it is often not designed for research or statistical use. Linking different data together can provide more comprehensive and useful material for data analysis. Data linkage may be between administrative data from different sources, or it may be between administrative data and other non-administrative data (such as longitudinal survey data).[2]

The Network seeks to establish relationships with data holders to facilitate data access, which may result in linked data being securely analysed by researchers. In the ADRN, data are linked together by a trusted third party (TTP) so that neither Network staff nor researchers see any data that identifies data subjects. The mechanism also prevents personal data being shared between the various organisations (e.g. government departments).

Despite these safeguards, data linkage still represents a form of processing personal data (from the view point of the data holder who must carry out data processes to facilitate the linkage) – so the whole process still falls within the gamut of the legislation described above. Whether the linkage is deemed fair and lawful may also depend on specific legislation pertaining to particular organisations (for example HMRC's capacity to share data is also delimited by the Commissioners for Customs and Revenue Act 2005).

It is advisable to investigate the potential for linkage while designing your research by speaking to Network staff and/or the data owners to consider the feasibility of linkage, potential identifiers, consent permission wording (if applicable) and ethical issues.

2 See Harron's (2016) guide to data linkage in this series for more details on data linkage processes.

3.2. Data security

Sharing information can bring many benefits, and in the context of academic research, it is a time-efficient way of facilitating the dissemination of information. But sharing information also presents risks.

The seventh data protection principle (set out in DPA Part 1 of Schedule 1) requires that personal data must be kept secure. In order to comply with the DPA, it is crucial that data is properly protected and the identity of data subjects safeguarded. As information systems and the worldwide web become more complex and widespread, so the potential for information about individuals' private lives to become known increases. As a practical consequence of this, when deciding to share personal data, both the party seeking access and the data provider need to identify the objective of the share, and then consider the potential benefits and risks.

The privacy risk involved in any data sharing will depend on the specific circumstances of the data and the situation. A risk assessment should be carried out whenever a decision is to be made about whether or not to disclose personal data. The Network advocates having carefully considered procedures in place that meet the principles of the five safes (safe settings, safe data, safe projects, safe researchers and safe outputs – see Desai et al (2014) for a full description of this framework). Each Administrative Data Research Centre in the Network has a set of controls in place to assess and minimise the risk of disclosure occurring when a researcher is using administrative data. In addition, Centre staff must clear a researcher's final data analysis as safe, before the results are released back to the researcher. It is only at that point that the research results are permitted to leave the secure environment. (The data collection itself never leaves the secure environment.)

4. Issues beyond the law

Compliance with the law only provides a partial explanation as to why data holding organisations do not disclose information to third parties when they have the legal power to do so.

The importance of public perception and trust have been identified as reasons why organisations and those working in them do not disclose information, irrespective of whether they have the power to do so in law. Reported levels of trust vary depending on whether an individual is asked whether public bodies should "collect and use" personal information or whether they should "share" personal information. Use of the term "data sharing" appears itself to reduce public trust and confidence. Trust is fragile and easily undermined by breaches of data security. Trust and confidence are also undermined where there is a perception that information is supplied for profit, even if the benefit goes to public services. Individuals may have different attitudes regarding their trust in: the overall system, individual officials they deal with, different institutions and particular sectors, for example local government or the NHS. It is also difficult to measure in advance, or predict the impact of particular proposals on, public trust and confidence.

Other non-legal issues include:

- Risk aversion and the common belief that it is more risky to disclose than not to. This may be increasingly misleading given the potential obligations on public bodies to share data that is in the public interest; and the potential legal avenues to hold public bodies to account in promises to share data (e.g. via judicial review of decisions that are not made appropriately transparent/not adequately consulted upon etc.).
- Exaggerated fears of enforcement action, in particular a fear of individual liability when in fact responsibility will rarely fall with an individual – it will usually fall higher (i.e. with an employer etc.).
- IT and organisational systems differing between institutions.

All these issues mean that facilitating access to data can be a challenge, even through the secure mechanism of the Network. These issues also emphasise the need for identifying and communicating the public benefit deriving from Network projects.

For the UK research community, the Network represents an important opportunity to inspire confidence in trustworthy data sharing through the rigorous training and certification of safe researchers and the transparent and secure operating procedures. If public trust is not inspired, this can have major implications for the success, reputation and longevity of a public-facing project such as the Network. The same applies to building relationships with data-sharing Government departments.

5. The future: EU General Data Protection Regulation 2016

The final text of the General Data Protection Regulation (GDPR) was published in April 2016.[3] The GDPR replaces EU Directive 95/46/EC, and will take effect on 25 May 2018, after a two-year implementation period. It will regulate the use of personal data and the free movement of such data across all sectors.

The 1995 Directive was a legislative act that set out an operating framework that all EU countries had to deliver, but left it to individual countries to implement their own laws and regulate their own interpretation of those laws. In contrast, the GDPR is a directly binding legislative act that must be applied in its entirety across the EU. Under the GDPR, individual member states may maintain or introduce additional conditions or limitations relating to certain issues, such as the processing of genetic or health data, so long as these are not incompatible with the Regulation itself.

How the Regulation will be implemented in practice is still being worked out. However, the provisions relevant to the Network include:

- Enhanced direct obligations and liabilities on data processors, including to maintain certain records of all processing activities and take appropriate security measures to protect personal data: "to ensure a level of security appropriate to the risk".
- Data protection safeguards to be built into organisations' products and services from the earliest stage of development (this is also known as 'privacy by design'). Privacy-friendly techniques such as pseudonymisation will be encouraged.
- Legal requirements for data processors to notify data controllers without undue delay after becoming aware of a data breach.
- A 'Data Protection Officer' must be appointed by all public authorities and those companies that perform certain risky data processing operations.
- Citizens will gain more rights to control their data, such as the right for their data to be destroyed, the right to be forgotten, and being asked to give unambiguous consent for their data to continue to be used or shared.
- Broad consent for scientific research is specifically allowed although more specific consent options should be offered if compatible with the research.
- Further processing for 'scientific research' is to be considered a compatible lawful basis.

3 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC', referred to as the General Data Protection Regulation (GDPR).

The Network already demonstrates best practice under the GDPR in several ways. For example, only certified researchers are given access to de-identified data, in a secure setting, and using TTPs to match data (see Data Security section above for further details). However, enhanced obligations on data processors and accompanying accountability and liability provisions mean it will be important for the Network to take steps to implement the requirements of the GDPR over the coming years. Further material on issues such as this will be published on the Network's website in due course.

6. Further reading

ADLS (2012) A review and assessment of best practice to link studies and surveys with administrative records, available at: http://tinyurl.com/huecv9z [Accessed: 15/6/2016].

DESAI, T. RITCHIE, F. & WELPTON, R. (2014) Five Safes: designing data access for research. University of West of England Economics Working Paper 1601, available at: http://tinyurl.com/jqz7wef [Accessed: 15/6/2016].

GRAY, M. (2010) A review of data linkage procedures at NatCen. National Centre for Social Research, available at: http://tinyurl.com/ gmpv3ao [Accessed 15/6/2016].

HARRON, K. (2016) An introduction to data linkage. ADRN Guide series.

JENKINS, S.P. LYNN, P. JACKLE. & SALA, E. (2008) The feasibility of linking household survey and administrative record data: New evidence from Britain. International Journal of Social Research Methodology 11: 29-43.

KILLEN, F.2016) EU Referendum: The Results, Insight. https://andersonstrathern.co.uk/news-insight/eu-referendum-the-legal-implications-of-brexit/ [Accessed: 6/10/2016].

LAURIE, G. & STEVENS, LA. (2014) The Administrative Data Research Centre Scotland: A Scoping Report on the Legal & Ethical Issues Arising from Access & Linkage of Administrative Data. Edinburgh School of Law Research Paper No. 2014/35, available at SSRN: http://ssrn.com/abstract=2487971 or http://tinyurl.com/zeflhgu [Accessed: 15/6/2016].

LAW COMMISSION (2014) Law Commission Report on Data Sharing Between Public Bodies: A Scoping Report. Her Majesty's Stationery Office, available at: http://tinyurl.com/jzsjy95 [Accessed: 15/6/2016].

LIGHTFOOT, D. & DIBBEN, C. (2013) Approaches to linking administrative records to studies and surveys – a review. Fife: Administrative Data Liaison Service, University of St Andrews, available at: http://tinyurl.com/gq8qwf8 [Accessed: 15/6/2016].

LOWTHIAN, P. & RITCHIE, F. (Forthcoming) Ensuring the confidentiality of statistical outputs from the ADRN. ADRN Guide series.

SAKSHAUG, J. W. & KREUTER, F. (2012) Assessing the magnitude of non-consent biases in linked survey and administrative data. Survey Research Methods. 41: 535-569.

SALA, E. KNIES, G. & BURTON, J. (2014) Propensity to consent to data linkage: experimental evidence on the role of three survey design features in a UK longitudinal panel. International Journal of Social Research Methodology. 17(5):455-473.

SCOTTISH HEALTH INFORMATICS PROGRAMME (SHIP) WEBSITE http://www.scot-ship-toolkit.org.uk/legal-concepts/public-interest

STEVENS, L. (2015) The Proposed Data Protection Regulation and Its Potential Impact on Social Sciences Research in the UK European Data Protection Law Review 1(2): 455-473

UK: INFORMATION COMMISSIONER'S OFFICE (2012a) Anonymisation: managing data protection risk code of practice, available at: http://tinyurl.com/ICO-ANON [accessed 25/5/2016].

UK: INFORMATION COMMISSIONER'S OFFICE (2012b) Determining what is personal data. Version 1.1, available at: http://tinyurl.com/ICO-WHATISPD [accessed 30/5/2016].

7. Legislation

EU: EUROPEAN CONVENTION ON HUMAN RIGHTS (1950) Strasbourg: Council of Europe [Online], available at: http://tinyurl.com/hmafux7 [Accessed: 15/6/2016].

EU: REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL (General Data Protection Regulation) (2016) Strasbourg: Council of Europe [Online], available at: https://tinyurl.com/ju74dm4 [Accessed 15/6/2016].

UK: STATISTICS AND REGISTRATION SERVICES ACT (2007) London: The Stationery Office [Online], available at: http://tinyurl.com/UK-SRSA [Accessed: 30/5/2016].

UK: DATA PROTECTION ACT (1998) London: The Stationery Office [Online], available at: http://tinyurl.com/UK-DPA98 [Accessed: 20/5/2016].

UK: HUMAN RIGHTS ACT (1998) The Stationery Office [Online], available at: http://tinyurl.com/zuc322x [Accessed: 15/6/2016].

8. Useful websites

EU: EUROPA.LAW AND PUBLICATIONS, available at: http://tinyurl.com/ju74dm4 [accessed 15/6/2016]. EUR-LEX section of the EU official site provides free access to EU law including Directives, Regulations and other related publications.

UK: ADMINISTRATIVE DATA RESEARCH NETWORK, available at: https://adrn.ac.uk/ [accessed 15/6/2016]. This is official site of the ADMINISTRATIVE DATA RESEARCH NETWORK (ADRN). It provides information on administrative data through its data catalogue, as well as help and advice on how to apply to use the ADRN for research using linked administrative data. The legal section of the ADRN website, available at: https://adrn.ac.uk/adrn.ac.uk/getting-data/resources/legal/ [accessed 21/12/2016] provides information on the legal framework that governs how the ADRN uses administrative data for research.

UK: INFORMATION COMMISSIONER'S OFFICE, available at: https://ico.org.uk/ [accessed 15/6/2016]. Is the UK'S independent authority set up to uphold information rights in the public interests. It provides information and support on data protection law and freedom of information law.

9. Glossary of non-standard terms

Certified researcher: A researcher from academia, the public sector or a research organisation on the Research Councils UK (RCUK) list of eligible Independent Research Organisations (IROs) that:

(i) is a trusted 'fit and proper' person i.e. they must be capable of carrying out the research either independently or under the direction of an appropriate supervisor or lead investigator
(ii) has been granted access to de-identified Data by the respective Data Controller(s) on an ADRN research project
(iii) has successfully completed ADRN certification training,
(iv) is backed by an Institutional Guarantor, and
(v) has signed up to the ADRN Terms of Use (ADRN021).

Data breach: An incident in which data is unlawfully removed from a secure environment. A personal data breach means "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed in connection with the provisions of a public electronic communications service

Data processor: In relation to personal data, means any person (other than an employee of the data controller) who processes the data on behalf of the data controller.

Data controller: A 'person' recognised in law, i.e. individuals, organisations, other corporate and unincorporated bodies of persons, who (either alone or jointly or in common with other persons) determines the purposes for which, and the manner in which, any personal data are, or are to be, processed.

Data linkage: The process by which records about a single data subject across multiple Datasets are associated with each other to create a new 'linked' Dataset.

Data matching: The process by which direct identifiers from different data sources are used to identify common data subjects. Alternatively, multiple sets of direct identifiers can be linked to a population spine (e.g. a count by age, sex and small area) to identify common Data Subjects. Within the ADRN data linkage process, the output of the data matching process is typically a set of matched IDs that can be used by the linker to link the research data.

De-identification: The process of removing or masking the direct identifiers with a dataset. Common strategies for de-identifying datasets are deleting direct identifiers, such as name, postcode and date of birth. Such identifiers might also be concatenated and replaced with a unique reference- a process called pseudonymisation.

De-identified data: Extracts from data which have undergone the process of de-identification.

Direct identifier: Any data item that, on its own, could uniquely identify an individual case. It is sometimes referred to as a formal identifier, examples of which include a data subject's name, address and unique reference numbers e.g. their social security number or National Health Service number.

Directive: A legislative act that sets out a goal that all EU countries must achieve. However, it is up to the individual countries to devise their own laws on how to reach these goals.

EEA: European Economic Area

Functional anonymisation: is a process which controls disclosure risk by considering the totality of a data situation.

Indirect identifiers: These can in principle include any piece of information (or combination of pieces of information). For example, consider the following combination of information a 'sixteen year old widow', whilst age and marital status are not immediately obvious identifiers, our implicit demographic knowledge tells us that this is a rare combination. This means that such an individual could potentially be re-identified by, for example, someone spontaneously recognising that this record corresponded to someone they knew.

Institutional guarantor: An individual, at a Network researcher's institution, with the legal status to act on behalf of that institution.

Statutory gateway: Legislation which enables access to personal data.

Personal data: Data which relate to a living individual who can be identified: (a) from those data or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

Regulation: A binding EU legislative act, which must be applied in its entirety across the EU.

Secondary data: Data that has already been collected for a purpose other than the current research project which may also have value for research.

Sensitive data: In the DPA this is data consisting of information as to a data subject's: (a) racial or ethnic origin, (b) political opinions, (c) religious beliefs or other beliefs of a similar nature, (d) whether they are a member of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992, (e) physical or mental health or condition, (f) sexual life, (g) commission or alleged commission by him of an offence, or (h) any proceedings for any offence committed, or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.

Survey data: Data collected through surveys, typically for the purpose of producing statistics. By participating in a survey, data subjects may give general consent to using the data collected for research purposes.

Trusted third party (TTP): A data linkage environment that adheres to the principle of separation of identifying data and payload data, though the means of this separation may be operationalised differently at the different places. In the Network's data linkage process, a TTP typically produces a set of matched IDs that can be used by the linker to link the research data. TTP activities can be carried out at a data controlling organisation, at an Administrative Data Research Centre or a separate organisation. In each case, there needs to be a clear separation of roles between matchers, linkers and certified researchers.

ADRN PUBLICATION
Produced by the Administrative Data Service
University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ
T. +44 (0) 1206 87 2976 E. help@adrn.ac.uk W. adrn.ac.uk

Funded by

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL