

Balancing information governance obligations when accessing social care data for collaborative research

Malkiat Thiarai, Sarunkorn Chotvijit and Stephen Jarvis
Department of Computer Science, University of Warwick, Coventry, UK

Abstract

Purpose – There is significant national interest in tackling issues surrounding the needs of vulnerable children and adults. This paper aims to argue that much value can be gained from the application of new data-analytic approaches to assist with the care provided to vulnerable children. This paper highlights the ethical and information governance issues raised in the development of a research project that sought to access and analyse children's social care data.

Design/methodology/approach – The paper documents the process involved in identifying, accessing and using data held in Birmingham City Council's social care system for collaborative research with a partner organisation. This includes identifying the data, its structure and format; understanding the Data Protection Act 1998 and 2018 (DPA) exemptions that are relevant to ensure that legal obligations are met; data security and access management; the ethical and governance approval process.

Findings – The findings will include approaches to understanding the data, its structure and accessibility tasks involved in addressing ethical and legal obligations and requirements of the ethical and governance processes.

Originality/value – The aim of this research is to highlight the potential use of use new data-analytic techniques to examine the flow of children's social care data from referral, through the assessment process, to the resulting service provision. Data held by Birmingham City Council are used throughout, and this paper highlights key ethical and information governance issues which were addressed in preparing and conducting the research. The findings provide insight for other data-led studies of a similar nature.

Keywords Public sector, Ethics, Governance, Information management, Legislation

Paper type Case study

Introduction

Researchers wanting to work with or use personal data within their research will be familiar with the requirement to ensure that the use of the data meets both legal and ethical standards. In support of these requirements, there are established legal and ethical frameworks within which research using personal data must be conducted.



In the UK, the Data Protection Act 2018 (which repealed the Data Protection Act 1998 in May 2018) is the principal piece of legislation governing the use of personal data. Several additional sources of guidance and good practice are available for researchers in the UK, including the *UKRIO Code of Practice for Research: Promoting good practice and preventing misconduct* (UKRIO, 2009) and the *RCUK Policy and Guidelines on Governance of Good Research Conduct* (RCUK, 2013) that set out key principles to effective research governance.

Background

The data associated with this research project is held by Birmingham City Council and are recorded as part of the Council's statutory duty to provide social services for a "child in need": This term is statutory and is defined in the Children Act 1989 (the "1989 Act"). Any child can be a "child in need", even if they are living with their family. A child is someone who is under 18 years of age.

There is no need for a court order to be made for a child to be deemed in need. It is the role of a local authority's children's services department to assess and provide services in this regard. Being a child in need is therefore broader than just those children in local authority care further to a care order, or those provided with accommodation by the local authority under section 20 of the 1989 Act.

Section 17 of the 1989 Act defines a child in need as follows:

- He is unlikely to achieve or maintain, or to have the opportunity of achieving or maintaining, a reasonable standard of health or development without the provision for him of services by a local authority under this Part.
- His health or development is likely to be significantly impaired, or further impaired, without the provision for him of such services.
- He is disabled.

Under the 1989 Act, "every local authority shall take reasonable steps to identify the extent to which there are children in need within their area", as well as to publish information about the services it provides to children in need (and other groups) and to "take such steps as are reasonably practicable to ensure that those who might benefit from the services receive the information relevant to them".

In addition to the services that local authorities provide for all children, the 1989 Act specifies the range of services that can be made available for a child in need:

- advice, guidance and counselling;
- occupational, social, cultural, or recreational activities;
- home help (which may include laundry facilities);
- facilities for, or assistance with, travelling to and from home for the purpose of taking advantage of any other service provided under the 1989 Act or of any similar service;
- assistance to enable the child concerned and his family to have a holiday;
- maintenance of the family home if the child is in need (but is not a looked after child) and is living apart from their family to either to enable the child to live with their family, or to promote contact between the child and their family;
- day care if the child is under 5 years of age but is not yet attending school;
- care or supervised activities (either outside school hours or during school holidays) for a child attending any school; and
- providing accommodation to a child and their family.

Consequently, the depth, breadth and sensitivity of data that may be recorded as a result of a local authority carrying out its statutory functions are significant. The volume of data is likely to be much greater than that collected either from an adult receiving social care services or an individual receiving health or healthcare services. A further distinction in the recording of children's data is likely to be the absence of consent for the collection, use and processing of the data given the statutory obligations imposed upon the local authority.

An international review by [Deloitte \(2016\)](#) into the "Secondary use of health and social care data and applicable legislation" found that across the six countries reviewed (England, Netherlands, New Zealand, Israel, Canada and Australia), all had started to recognise the value of collected health records, but that there is no separate legislation for secondary use of health and social care data in any of these countries, rather privacy laws in each country defined how the personal health and social care records can be used.

It is important for organisations to have clarity as to the legal basis for processing personal data. Where legal obligations are being fulfilled then the basis for processing is provided. Changes to EU law following the introduction of the General Data Protection Regulation (GDPR) in 2016 ([EUR-Lex, 2016](#)) have provided for a higher standard for consent compared to previous legislation. It requires that, where there is a requirement for consent, an indication of consent must be unambiguous and involve clear affirmative action and furthermore, be able to be withdrawn at any time.

The GDPR requires that there should be distinct ("granular") consent options for distinct processing operations. It could be argued, therefore, that processing data for research purposes is a distinct processing operation, separate from the original purpose for which the data was collected and, as such, the relevant obligations would apply.

The Information Commissioner's Office's (ICO) guidance on "consent" concludes:

Consent is one lawful basis for processing, but there are alternatives. Consent is not inherently better or more important than these alternatives. If consent is difficult, you should consider using an alternative.

As such, an early determination in this research was that the basis on which the data could be processed for research purposes as set out in Section 33 of the Data Protection Act (DPA) 1998. The DPA 1998 does not define "research". The ordinary meaning of "research" is therefore used when determining whether personal data is being processed for research purposes – research is a systematic investigation intended to establish facts, acquire new knowledge and reach new conclusions.

Section 33 of the DPA makes it clear that "research purposes" includes statistical and/or historical research:

- (1) "research purposes" includes statistical or historical purposes; "the relevant conditions", in relation to any processing of personal data, means the conditions:
 - that the data are not processed to support measures or decisions with respect to particular individuals; and
 - that the data are not processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject.
- (2) For the purposes of the second data protection principle, the further processing of personal data only for research purposes in compliance with the relevant conditions is not to be regarded as incompatible with the purposes for which they were obtained.

There are, however, several challenges that researchers face in understanding the social care data they wish to utilise as well as adhering to a range of different guidance, whether on governance or ethics, issued across the sector.

First, there is a plethora of sector and professional specific guidance that provides a framework that those working with social care data need to follow. Examples include:

- (1) The [Department of Health \(2010\)](#) resource pack, which is designed to support the implementation in social care of the *DH Research Governance Framework*. The Framework contains information, guidance and a range of resources for supporting research governance and is primarily aimed at those involved in setting up and running governance systems in local authorities or for people who take part in the review of relevant research.
- (2) The Information Governance Review ([Caldicott, 2013](#)) concluded, as part as its review that accredited safe havens should be required to meet requirements for data stewardship that included:
 - Robust governance arrangements that include, but are not limited to, policies on ethics, technical competence, publication, limited disclosure/access, regular review process and a business continuity plan including disaster recovery.
 - Clear conditions for hosting researchers and other investigators who wish to use the safe haven.
- (3) The Department for Children and Youth Affairs in Ireland published specific guidance on research involving children, in part driven by the principles of the United Nations Convention of the Rights of the Child, in particular Articles 2, 3, 4 and 6. The guidance, recognising that there was no single regulatory system and no body responsible for research ethics in the country, set out a number of core ethical principles and concepts that needed to be followed by all those who carry our research with, and for, children in Ireland.
- (4) The International Medical Informatics Association developed and published a distinct code of ethics for health information professionals, individuals who, in their professional capacity, provide health informatics services, arguing that they play a unique role and occupy a unique position that is distinct from that of informatics professionals who do not specialise in health care data and who do not work in the health care setting.

The Archives and Records Association published a Code of Ethics setting out the standards of professional behaviour expected of archivists, archive conservators, records managers and those occupied in related activities who are individual members of the association ([Archives and Records Association, 2018](#)).

These examples of ethical and governance practices within different sectors and organisations reflect the views of the UK ICO, who noted in the paper on “Big data, artificial intelligence, machine learning and data protection” when discussing the ethical approaches emerging in the field of data protection that it was:

Notable that these ethical frameworks have been developed not by regulators but by companies and other organisations themselves.

A further example of an organisation developing its own approach is that of the Nuffield Foundation where the Trustees have decided to establish the Nuffield Family Justice Observatory ([Broadhurst et al., 2018](#)). The aim of the Observatory is to support the best possible decisions for children by improving the use of data and research evidence in the family justice system in England and Wales. This work is also linked to the Children in Family Justice Data Share ([Ministry of Justice, 2019](#)), a collaborative project that has resulted in a database of child data linked from across the Ministry of Justice (MoJ), the

Department for Education (DfE) and Cafcass' management information systems. The MoJ's report into this project includes a reflection on the "significant legal issues to overcome in relation to compliance with the DPA" and whilst a formal agreement to share the data was established, challenges remain, not least the unforeseen technical issues associated with IT infrastructures.

Nevertheless, many aspects of these frameworks echo key data protection principles and demonstrate the strong link between these ethical approaches and data protection law.

Ethical considerations

The EU Data Protection Directive of 1995 made no mention of a human right to data protection. In contrast, the General Data Protection Regulation (GDPR) is framed in terms of rights, with the protection of "fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data" set out in Article 1 of the Regulation.

Furthermore, with the enactment of the Lisbon Treaty, the Charter of Fundamental Rights of the European Union in 2009 saw, for the first time, a stand-alone fundamental right to data protection. The significance of this right is discussed by [McDermott \(2017\)](#), examining the parameters of this right and its links to key values of privacy, transparency, autonomy and non-discrimination in other legal European statutes.

The opinion of the [European Data Protection Supervisor \(2015\)](#)[1], called for a four-tier "big data protection ecosystem" to respond to the digital challenge requiring a collective effort, underpinned by ethical considerations. The proposed ecosystem encompassed:

- *Future-oriented regulation*: urging simpler rules for handling of personal data which stay relevant for a generation.
- *Accountable controllers*: putting in place internal policies and control systems that ensure compliance and provide relevant evidence.
- *Privacy conscious engineering*: empowering individuals who wish to preserve their privacy and freedom through anonymity.
- *Empowered individuals*: focussed on a "prosumer" environment, consent and control and data ownership.

The opinion called for dignity to be at the heart of new digital ethics arguing that "better respect for, and the safeguarding of, human dignity could be the counterweight to the pervasive surveillance and asymmetry of power which now confronts the individual".

In its submission to the Select Committee on Artificial Intelligence (AI), the ICO commented that:

Despite robust data protection compliance, the law only takes us so far. We believe that it can be highly challenging to apply certain data protection concepts such as fairness and relevance to advanced AI applications. For example, empathic computing involves the use of AI to examine an individuals' on-line behaviour. It considers the vocabulary individuals use, the way they input type and the pictures they look at longest to assess that individual's mood and deliver content accordingly. This certainly involves the processing of personal data and therefore engages data protection law. However, whilst the pure data protection compliance aspects of using AI in empathic computing and other contexts can be addressed using the compliance steps outlined in the annex, the use of AI raises wider ethical issues of significant public interest.

The range of problems to be found in the practical use of social care data are explored in Gillingham and Graham's paper ([Gillingham and Graham, 2017](#)) that highlights issues such as data integrity, subjectivity in decision making at the recording stage and hidden biases. Furthermore, the paper considers the impact of the extracted data if it is rendered

meaningless because of removing it from the system as well as the fact that narrative accounts or in-depth social explanations of complex problems being lost or replaced by simpler descriptions.

The paper also highlights several ethical concerns with consent to use personal data being one and confidentiality being another. In the context of children's social care data, as explained above, the basis for collecting the data is to meet statutory obligations, so consent at data collection stage is not the issue. One ethical issue is whether the statutory obligations extend to other uses, such as research, which from a legal perspective would be considered a compatible purpose.

Accessing data for research purposes also raises issues of the quality of the data and the methods for records management in the curation of the data for research, particularly, where it is for a secondary purpose.

A review of "Digital records management in Australian government" suggests that the basic tenets for records management – create, capture, manage, access, secure, describe and dispose – remain the same, the methods seem to have stagnated and that the method for how we keep and manage records needs to be adjusted for digital times (Stuart, 2017; Maroye *et al.*, 2017).

The challenges in utilising data obtained in the course of child social work or welfare provision is discussed in Naccarato's paper on Child Welfare Informatics (Naccarato, 2010). It proposes the possible evolution of this topic as a sub-specialty in social work, highlighting that "concerns exist as the discipline has ties to face-to-face interactions and there is a minimal amount of time available for practitioners and policy makers to focus on data-related needs" whilst at the same time needing to protect and manage sensitive information and data linking capabilities.

The ethical challenges posed by "Big Data" are also discussed in the Metcalf and Crawford paper (Metcalf and Crawford, 2016) on the emerging ethics divide. The paper argues that big data is stretching the concept of ethical research and that existing ethical regulations promote a particular approach towards "research subjectivity" that is being eroded by data science. Further, it suggests that the traditional concept of a "human research subject", that is, what constitutes an intervention, when and how consent should occur and what types of harm are relevant, are out of step with large-scale data practices. The paper questions, "Who is the data subject in a large scale data experiment and what are they owed?"

The Metcalf and Crawford paper offers a preliminary examination of how critical data studies may generate a theory of data subjectivity, to enable responsible scientific practice with Big Data methods and thereby address some of the ethical issues that exist and avoid the human subject becoming invisible or irrelevant to data science.

Organisational approaches/controls

The Royal Society and British Academy's report *Data management and use: Governance in the 21st Century* (The Royal Society, 2017) addressed the changing data landscape, recommending a principled approach to data governance and calling for stewardship of the entire data governance landscape. The Academies hosted a seminar to explore the priorities across sectors for such a stewardship body. Discussions at which, set out governance needs, practical challenges and conceptual concerns that any such a body could take on.

The UK Government announced in its budget statement on 22 November 2017 the creation of a new Centre for Data Ethics and Innovation (The Royal Society, 2017) to enable and ensure safe, ethical and ground-breaking innovation in AI and data-driven technologies.

Training needs have been expressed, for example, the introduction of ethical thinking at the core of and during the continuing education of all professionals, so that they can develop

the necessary analytical tools to respond to ethical situations as and when they arise (Lacovino, 2002). It is also argued that codes should be used to focus on professional duties and virtues and as a collective consensus of professional values.

Organisations are also investing in and implementing internal control systems which include a set of elements such as integrity and ethical values (Rubino *et al.*, 2017). This reflects changing attitudes in management in the field of data control and this can, and perhaps should, shape the way data is viewed as part of the overall control objective.

Design/methodology/approach

This paper discusses the approaches taken to identify and obtain approval for the use of social care data for a research project titled “Data continuity analysis of the assessment process in children’s social care” within Birmingham City Council and the steps taken to demonstrate compliance with legal and ethical frameworks.

The city of Birmingham is the UK’s largest and most populous city outside of London. Birmingham has a population of over 1.1 million people, and the population is growing faster than the UK average. Birmingham is a young and diverse city; half of the population are aged 30 or under. It is the sixth most deprived local authority in the UK; 40 per cent of the city is ranked in the most deprived 10 per cent of areas in England. There are significant levels of child poverty; 30 per cent of the city’s children live in a deprived household.

Birmingham City Council (BCC) is the largest local authority in Europe. Income and expenditure in 2016/17 was £3.094bn, of which £782m was spent on schools, £550m spent on benefits, £805m spent on services for people and £287m spent on housing (Birmingham City Council, 2016). Managing BCC’s priorities has been difficult in the context of recent fiscal challenges. Birmingham City Council is expected to make total savings of £815m over the 9 year period 2011/12 to 2019/20.

The most recent Ofsted inspection report (Ofsted, 2016) for Birmingham City Council showed that at 31 August 2016, 1,816 children were being cared for by the local authority (a rate of 64 per 10,000 children); this is similar to that seen in March 2015 (also 64 per 10,000 children).

Methodology

The research seeks to use data analytic methods to examine the flow of children’s social care data from referral through the assessment process to the resulting service provision. In this example, the research is collaboration between a UK local authority and an UK University. The methodology used to address the information governance and ethical issues aims to balance the respective obligations of the partner organisations. For example, accessing the proposed dataset for research purposes for the local authority represents a secondary use of data that is deemed to be compatible with the original purpose for which the data was obtained. For the University, however, the data is being processed for the primary purpose of research.

The determination that s33 of the DPA is the appropriate gateway for the legitimate use of the data for research purposes means that adhering to the “relevant conditions” must be met throughout the whole life of the data. It also informs partner organisations at the onset of the research of these conditions and the subsequent limits imposed on the use of the data.

As such, each party needs to document and demonstrate to the other how they meet these respective obligations. Using this research as an example, the Table I shows a series of activities undertaken and by whom to demonstrate compliance with the information governance and ethical requirements.

Table I.
Data custodian and collaboration partner activity

Steps	Data custodian activity	Joint activity	Collaboration partner activity
Step 1	Research governance application		
Step 2		Collaboration agreement	
Step 3		Training in data ethics and governance	
Step 4	Ethical approval application		
Step 5			Ethical review and questions
Step 6	Additional information provided		
Step 7			Ethical approval granted with actions
Step 8		Data sharing agreement	
Step 9	Data obfuscation and transfer		
Step 10			Data receipt and security measures applied
Step 11			Data retention for research purposes
Step 12		Ongoing monitoring	

The key components within the methodology are the initial research governance application within the Local Authority (Data Custodian) and the ethical review process within the University (Collaboration partner) as they establish the boundaries for the research partners in meeting their respective obligations. Both organisations required appropriate training to be undertaken.

For the research governance application, this incorporates outlining the legal framework and setting out the scope and structure of the data. The ethical review process allows for the collaboration partner to assess the research proposal in the context of their own organisational frameworks. How these respective obligations are managed are discussed in further detail below.

Legal framework

As highlighted above, the Data Protection Act is the principal piece of UK legislation governing the use of personal data. The application for ethical approval was in progress as this legislation was being amended because of the General Data Protection Regulation ([Regulation GDP, 2016](#)).

However, both the 1998 and 2018 Data Protection Acts ([The National Archives, 2018](#)) make provisions for research.

This exemption can apply if you process personal data for:

- scientific or historical research purposes; or
- statistical purposes.

It does not apply to the processing of personal data for commercial research purposes such as market research or customer satisfaction surveys. It exempts you from the GDPR's provisions on:

- the right of access;
- the right to rectification;
- the right to restrict processing; and
- the right to object.

The GDPR also provides exceptions from its provisions on the right to be informed (for indirectly collected data) and the right to erasure. However, the exemption and the exceptions only apply:

- To the extent that complying with the provisions above would *prevent* or *seriously impair* the achievement of the purposes for processing.
- If the processing is subject to appropriate safeguards for individuals' rights and freedoms (see Article 89(1) of the GDPR – among other things, you must implement data minimisation measures).
- If the processing is not likely to cause substantial damage or substantial distress to an individual.
- If the processing is not used for measures or decisions about particular individuals, except for approved medical research.
- As regards the right of access, the research results are not made available in a way that identifies individuals.

There are some notable changes in the new laws that it is important to consider when using data for research purposes. First, the GDPR requires that “technical and organisational measures” are put in place to ensure that data controllers process only the personal data necessary for the research purposes, in accordance with the principle of data minimisation outlined in Article 5(c). In addition, Recital 33 states that controllers should act “in keeping with recognised ethical standards for scientific research”. Whilst not elaborating any further on what those recognised ethical standards are, the Regulation makes the linkage between research and ethical re-use of personal data. Furthermore, GDPR Article 89(1) specifically references the use of “pseudonymisation” as a method by which a data controller could comply with the mandate for technical and organisational measures obligations.

Understanding the structure of the data

One of the first tasks in preparing the research governance application was to understand the nature of the data held, its format and structure. The data required in this case are primarily collected and stored to support the Local Authority's social care case management. As [Best \(1990\)](#) sets out, there are challenges in utilising information that has been collected for a specific purpose in a related level of the organisation and its subsequent re-use for a secondary purpose within another level of the organisation.

To access this data, several meetings took place with Local Authority employees who had detailed technical knowledge of the system, the structure of the data and the changes that had occurred over time (whether technical, legislative or organisational) that impacted the way the data was recorded and categorised. Data was extracted using an Open Database Connectivity (ODBC) connection through Crystal Reports, which is the application in use in the Local Authority. A series of extracts were produced, and the data was then joined together using IDEA (Data Analysis and Extraction) software. The information generated was validated throughout and checked against the live application to prove its integrity.

As [Gillingham and Graham \(2017\)](#) highlight, data extraction presents one of the key ethical challenges in converting data into a suitable form for re-use. Notably, big data requires that the “raw data” is available in a format that is better suited to statistical analysis and computation. The problem with this process, as their paper points out, is that the decisions taken at this stage may not only affect the research but may also be invisible in the findings. This is true to some degree here, in that to understand the structure of the data,

we are reliant on the knowledge and expertise of those familiar with the technology to be able to describe what data is held and the relationships between different data types.

The data

The data extract is of administrative data related to the social care assessment and agreement process (Table II).

Assessment data details the flow of information through the system of a referral to Children’s Social Care and the outputs from this process. *Agreement data* details those referrals that result in an “agreement” which relates to a service that is subsequently provided to the child.

The data are pseudonymised prior to its release to the researchers. Individual names are removed, however, the “Person Identifier”, which is key to making the data identifiable, is retained. This is to allow for analysis of the data and the number of unique individuals receiving the services. The importance of this level of detail in the research governance application cannot be overstated. This establishes both the intent of the research but also the risks posed to those individuals whose data is the subject of research. For example, retention of the full postcode will allow for much more granular analysis of the data at a geographical level but loss or unauthorised access to the data could allow for the identification of the individual possibly causing substantial damage or substantial distress to any data subject, thereby contravening one of the “relevant conditions” in s33 DPA that established to legal basis for the research.

Ethical approval process

The ethical approval process at our collaborating university requires all those involved in delivering, supervising, or supporting research (whether staff, students and/or their supervisors)

Record	Description	Assessment	Agreement
ASM_ID	Assessment Identifier	X	
PERID	Person Identifier	X	X
DOB	Date of birth	X	X
AGREEMENT_START	Start date of the agreement	–	X
AGREEMENT_END	End date of the agreement	–	X
ASSESSMENT_START	Start date of the assessment	X	–
ASSESSMENT_END	End date of the assessment	X	–
SERVICE	Alphanumeric coding of the service	–	X
SERVICE_DESCRIPTION	Description of the service	–	X
ELEMENT	Alphanumeric coding of the element	–	X
ELEMENT_DESCRIPTION	Description of the element	–	X
QSA_DESCRIPTION	Quality standard assessment process description	X	–
QSA_GROUP	Quality standard assessment process group	X	–
ASSESSMENT_REASON	Reason of taking assessment	X	–
FORM_OUTCOME	Outcome of assessment form	X	–
SERVICE_TEAM	Corresponding assessment team	X	–
SERVICE/ELEMENT_COST	Cost of service/element where available	–	X
POSTCODE	Postcode (unit level)	X	X
GENDER		X	X
ETHNICITY	Ethnic classification	X	X
CLIENT_GROUP	Disability status	X	X
WARD		X	–
CONSTITUENCY		X	–

Table II. Administrative data fields for social care assessment and agreement extraction

and research support staff (technical and administrative), to complete data governance and research ethics training.

Scrutiny with regard to research ethics and governance is undertaken for each individual research project. If ethical approval is required, this involves completion of an application form and a supporting protocol. The protocol includes two mandatory fields for completion, one of which is the Ethical Considerations, which as a minimum, should contain subsections examining Informed Consent, and Participant Confidentiality and Data Security.

It should be noted that in the application for ethical approval, confirmation was required that the research was “limited to the use of previously collected identifiable data” which was reflective of the whole data held in the social care system.

The data selected were solely to support the aims of the research, which is to understand the flow of data through the assessment process and identify any temporal or spatial analysis, errors, data quality issues or other patterns in the data that may inform current practice and processes.

Data analysis was to focus on the aggregate number of users. Using Excel, R, Python and spatial mapping tools including QGIS, the research sought to generate results that showed:

- service assessments/agreements over time;
- service assessments/agreements by time and by postcode district;
- number and type of assessments/agreements at different stages in the assessment/agreement workflow;
- analysis of assessments/agreements by user type, including age group, gender, ethnicity and disability, and by commissioning team; and
- an assessment of how this analysis might support Council priorities.

This level of detail is an important component in a collaborative research environment as it provides the collaboration partner with background information as to the origin of the data, its structure, how it will be accessed and the how the data will be manipulated.

Information governance issues

The review of the submitted ethical application resulted in several information governance issues requiring further clarification ranging from queries on:

- data minimisation;
- data security;
- damage and distress to data subjects;
- data loss;
- data sharing; and
- data retention.

In reflecting on the information governance issues raised by the reviewers in respect to this ethical application, it is interesting to note that the nature of the issues raised equate to those that the ICO included in their consultation document on what should be considered, assessed and addressed as part of a Data Protection Impact Assessment (DPIA).

This process involves assessing the necessity and proportionality of research plans to achieve its stated purpose, including:

- your lawful basis for the processing;
- how you will prevent function creep;

- how you intend to ensure data quality;
- how you intend to ensure data minimisation;
- how you intend to provide privacy information to individuals;
- how you implement and support individuals rights;
- measures to ensure your processors comply; and
- safeguards for international transfers.

The next stage of the process required the identification and assessment of risks where consideration of the potential impact on individuals and any harm or damage that might be caused by the processing – whether physical, emotional or material. The assessment considered whether the processing would possibly contribute to:

- inability to exercise rights (including but not limited to privacy rights);
- inability to access services or opportunities;
- loss of control over the use of personal data;
- discrimination;
- identity theft or fraud;
- financial loss;
- reputational damage;
- physical harm;
- loss of confidentiality;
- re-identification of pseudonymised data; and
- any other significant economic or social disadvantage.

The third step required how to identify mitigating measures to the risks identified. For example:

- deciding not to collect certain types of data;
- reducing the scope of the processing;
- reducing retention periods;
- taking additional technological security measures;
- training staff to ensure risks are anticipated and managed;
- anonymising or pseudonymising data where possible;
- writing internal guidance or processes to avoid risks;
- adding a human element to review automated decisions;
- using a different technology;
- putting clear data sharing agreements into place;
- making changes to privacy notices;
- offering individuals, the chance to opt out where appropriate; and
- implementing new systems to help individuals to exercise their rights.

Many of these issues would now be captured by a DPIA, in particular in relation to detailed assessment of the mitigation of risks and harm to individuals and addressed this at an earlier stage.

Conclusions and recommendations

In setting out the experience of preparing, submitting and gaining approval for research using social care data, this paper seeks to highlight some of the key issues identified in this process and their relationship with current legal processes and requirements. These include, understanding and demonstrating the key legal basis for accessing the data for the purposes of processing it for research, understanding the structure of the data, how it can be extracted and the way it needs to be secured, navigating through the ethical approval process and responding to the issues raised by the reviewers of the application.

The information governance issues raised by the reviewers of the ethical application demonstrate consistency between the Council and the university collaborators and the level of scrutiny provided by each was extremely supportive to the research team involved.

The frameworks and structure for the ethical use of social care data continues to be developed and the challenges highlighted may in part be addressed not just through sector specific frameworks but through the approaches organisations take to the capture, use and re-use of data as part of their wider control measures. For example, the impact of an IT governance framework on the internal control environment (Rubino *et al.*, 2017), which states that the:

Essence of a firm is effectively controlled and represented by the attitude of its management. If top management believes that control is important, the other members of the organisation will feel so and will respond with a conscientious respect for the controls established.

This could facilitate a more clearly defined approach to use of data from the onset.

This statement could apply equally to an information governance framework and the approach taken within an organisation, particularly as one of the seven categories of factors in the control environment are the ethical values and integrity standards within the organisation. In a local authority context, approaches to managing ethical and information governance standards could be incorporated into the constitution of a local authority to raise the profile of this topic.

With the introduction of new data protection legislation, there is the possibility of considering how the new law may be applied. For example, while research is not explicitly designated as its own lawful basis for processing, in some cases, it may qualify under Article 6(1)(f) of the GDPR as a legitimate interest of the controller.

Gillingham and Graham (2017) propose that to counter the inaccuracy and incompleteness of datasets, it may be worth developing policies that promote the collection of ever more detailed and full data sets about service users and service activity. From a data protection perspective, this would be considered as “data protection by design and by default”. The GDPR requires organisations that process personal data to put in place appropriate technical and organisational measures to implement the data protection principles and safeguard individual rights. This means integrating data protection into processing activities and business practices, from the design stage right through the lifecycle and that would include subsequent re-use. This concept is not new; however, as the key change with GDPR is that it is now a legal requirement. Adopting this approach should also assist organisations in their obligations to be transparent as to the use of the personal data they process and, in turn, demonstrate how they fulfil their ethical obligations to this data.

The GDPR also contains specific provisions that adapt the application of the purpose limitation and storage limitation principles when personal data is processed for scientific, historical or statistical purposes and as discussed there are specific obligations within the GDPR that relate to the use of personal data for research purposes. The absence of the consideration of these obligations in research proposals could therefore result in breaching data protection law.

In addition to the GDPR, the UK Parliament has enacted legislation, through Chapter 5 of Part 5 of the Digital Economy Act 2017, that facilitates the linking and sharing of datasets held by public authorities for research purposes.

The power set out in Chapter 5 (“the Research power”) broadly enables information held by one public authority to be disclosed to another person for the purposes of research. In making use of this power, public authorities must be able to demonstrate and meet a number of conditions. These include, establishing processes for “de-identifying” personal information to be shared under the power; adhering to a published Code of Practice, containing seven principles of data sharing for research purposes, concerning the disclosure, processing, holding or use of personal information intended to collectively ensure that the provision of personal information is ethical and legal under this provision; requiring parties involved in the disclosure of this information to be accredited.

With the introduction of these new legislative obligations and codes of practice in different settings and the range of issues being raised as to the ethical use of personal data, this is an appropriate time for the UK regulator to exercise powers within the Data Protection Act 2018 to publish a Code of Practice specifically related to the ethical re-use of data, rather than allow these standards to emerge as part of either organisational or sector activity.

Note

1. European Data Protection Supervisor (EDPS) is an independent institution of the EU, responsible under Article 41.2 of Regulation 45/2001 ‘With respect to the processing of personal data [...] for ensuring that the fundamental rights and freedoms of natural persons, and in particular their right to privacy, are respected by the Community institutions and bodies’, and ‘[...] for advising Community institutions and bodies and data subjects on all matters concerning the processing of personal data’.

References

- Archives and Records Association (2018), “Code of ethics”, available at: www.archives.org.uk/images/ARA_Documents/ARA_Code_Of_Ethics.pdf (accessed 20 September 2018).
- Best, D.P. (1990), “The future of information management”, *Records Management Journal*, Vol. 2 No. 2, pp. 50-60.
- Birmingham City Council (2016), “Budget for Birmingham 2016/2017”, BCC, Birmingham, available at: www.birmingham.gov.uk/download/downloads/id/1695/budget_for_birmingham_2016_to_2017.pdf (accessed 9 December 2018).
- Broadhurst, K. Budd, T. and Williams, T. (2018), “The Nuffield family justice observatory for England and Wales: making it happen”, available at: www.nuffieldfoundation.org/sites/default/files/files/Nuffield_Family_Justice_Observatory_making_it_happen_v_FINAL_13_02_18.pdf (accessed 7 December 2018).
- Caldicott, F. (2013), “Information: to share or not to share: information governance review”, Information: to share or not to share.
- Deloitte (2016), “International review secondary use of health and social care data and applicable legislation”, available at: www.sitra.fi/julkaisut/Muut/International_review_secondary_use_health_data.pdf (accessed 20 September 2018).
- Department of Health (2010), “Research governance framework for health and social care: second edition”, available at: www.hra.nhs.uk/documents/379/research-governance-frameworkresource.pdf (accessed 20 September 2018).

- European Data Protection Supervisor (2015), "Towards a new digital ethics: data, dignity and technology", available at: https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf (accessed 11 December 2018).
- EUR-Lex (2016), "General data protection regulation", available at: <https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04> (accessed 11 December 2018).
- Gillingham, P. and Graham, T. (2017), "Big data in social welfare: the development of a critical perspective on social work's latest 'electronic turn'", *Australian Social Work*, Vol. 70 No. 2, pp. 135-147.
- Lacovino, L. (2002), "Ethical principles and information professionals: theory, practice and education", *Australian Academic and Research Libraries*, Vol. 33 No. 2, pp. 57-74.
- McDermott, Y. (2017), "Conceptualising the right to data protection in an era of big data", *Big Data and Society*, Vol. 4 No. 1, pp. 1-7.
- Maroye, L., van Hooland, S., Aranguren Celorrio, F., Soyeze, S., Losdyck, B., Vanreck, O. and de Terwangne, C. (2017), "Managing electronic records across organizational boundaries: the experience of the Belgian federal government in automating investigation processes", *Records Management Journal*, Vol. 27 No. 1, pp. 69-83.
- Metcalf, J. and Crawford, K. (2016), "Where are human subjects in big data research? The emerging ethics divide", *Big Data and Society*, Vol. 3 No. 1, pp. 1-14.
- Ministry of Justice (2019), "Children in family justice data share: the public law applications to orders (PLATO) tool", available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696108/children-in-family-justice-data-share.pdf (accessed 7 December 2018).
- Naccarato, T. (2010), "Child welfare informatics: a proposed subspecialty for social work", *Children and Youth Services Review*, Vol. 32 No. 12, pp. 1729-1734.
- Ofsted (2016), "Birmingham city council inspection report", available at: <https://reports.ofsted.gov.uk/provider/44/80429>
- RCUK (2013), "RCUK policy and guidelines on governance of good research conduct", available at: www.ukri.org/files/legacy/reviews/grc/rcuk-grp-policy-and-guidelines-updated-apr-17-2-pdf/ (accessed 20 September 2018).
- Regulation GDP (2016), "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", *Official Journal of the European Union (OJ)*, Vol. 59, pp. 1-88, p. 294.
- Rubino, M., Vitolla, F. and Garzoni, A. (2017), "The impact of an IT governance framework on the internal control environment", *Records Management Journal*, Vol. 27 No. 1, pp. 19-41.
- Stuart, K. (2017), "Methods, methodology and madness: digital records management in the Australian government", *Records Management Journal*, Vol. 27 No. 2, pp. 223-232.
- The National Archives (2018), "Data protection act 2018", available at: www.legislation.gov.uk/ukpga/2018/12/notes/division/6/index.htm (accessed 20 September 2018).
- The Royal Society (2017), "Data management and use: governance in the 21st century – a British academy and royal society project", available at: <https://royalsociety.org/topics-policy/projects/data-governance/> (accessed 10 December 2018).
- UKRIO (2009), "UKRIO code of practice for research: promoting good practice and preventing misconduct", available at: <http://ukrio.org/wp-content/uploads/UKRIO-Code-of-Practice-forResearch.pdf> (accessed 20 September 2018).

Further reading

- Chotvijit, S., Thiarai, M. and Jarvis, S.A. (2018), "A study of data continuity in adult social care services", *The British Journal of Social Work*, Vol. 1, pp. 1-25.

Information Commissioners Office. (2017a), “Big data, artificial intelligence, machine learning and data protection”, available at: <https://ico.or.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf>

Information Commissioners Office. (2017b), “Select committee on artificial intelligence – submission from the information commissioner”, available at: <https://ico.or.uk/media/about-the-ico/consultation-responses/2017/2172538/select-committee-ai-ico-submission-20170906.pdf>

The Department for Children and Youth Affairs. (2012), “National guidance for developing ethical research projects involving children”, available at: www.dcy.gov.ie/docs/EN/Publications-copy-dcy-gov-ie-2019/200/1865.htm (accessed 20 September 2018).

The International Medical Informatics Association. (2016), “IMIA code of ethics for health information professionals”, available at: <https://imia-medinfo.org/wp/imia-code-of-ethics/> (accessed 20 September 2018).

Corresponding author

Malkiat Thiarai can be contacted at: m.s.thiarai@warwick.ac.uk

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com