

The Changing Research Data Landscape and the Experiences of Ethics Review Board Chairs:
Implications for Library Practice and Partnerships

Brian Jackson

Mount Royal University

4825 Mount Royal Gate SW

Calgary, Alberta, Canada

T3E 6K6

bjackson@mtroyal.ca

Abstract:

Academic libraries have to a large extent taken the lead in facilitating new approaches to research data management, but changes to the research data landscape have had an impact on numerous areas of academic work, including ethics review. Using interpretive phenomenological analysis of interviews with chairs of Canadian research ethics boards, this study explores how ethics review boards have experienced changes to data policy and related technologies in order to describe the ethical implications of new approaches to data management and to explore ways in which the library, ethics review boards, and other campus partners might harmonize efforts to support emerging data practices. While ethics review boards in Canada are keenly aware of open data policies, data publishing in practice is still nascent. There is uncertainty about the adoption of changing technologies for research and their impacts on privacy protection. Where responsibility lies for addressing these uncertainties is often unclear. Academic libraries and research ethics boards are well-suited to engage in mutual knowledge transfer and to integrate data management planning and ethics review processes. Institutional-level oversight that includes all campus departments impacted by changes to the research data landscape may facilitate improved communication and reduce role ambiguity.

Keywords: research data management; research ethics

Introduction

The research data landscape is changing. Data management planning is becoming a more prominent step in the research process, while a growing list of funders and publishers have tied strict data archiving and open release requirements to their agreements with researchers. Driving the evolution toward better data management practices, in part, is a dynamic technological infrastructure. Researchers must see beyond research methodologies to consider how the devices and software they use will impact the collection, storage, security, publication, and disposition of their data. Without adequate thought to the technologies involved in a project, data may be suboptimal or inaccurate, accidentally lost or retained, inappropriately shared, published in inaccessible formats, or otherwise unfit for future use. Events such as these not only impact the outcomes of research, but may have serious ethical consequences, particularly when the data describe human subjects.

On-campus advocacy and support for research data management (RDM) and data publishing initiatives come from numerous departments, including libraries, research administration, information technology services, legal offices, and, of course, researchers, with librarians and research administrators most often taking the lead in promoting and supporting RDM (Cox, Kennan, Lyon, & Pinfield, 2017). While there is potential for jurisdictional conflict where these services overlap (Verbann & Cox, 2014), there is also potential for new partnerships. Adequate support for RDM requires expertise and resources from disparate campus departments, and the pooling of these assets will only benefit the larger community (Jones, Pryor, & Whyte, 2013).

This collaborative environment requires an understanding of the strengths and struggles of other campus units that support RDM. A sympathetic view of other campus departments can lead to the reconciliation of differing perspectives and more efficient achievement of shared goals (Latham, 2017). Much of the discussion to date around roles and perspectives in the RDM realm has focused on the policy and technology concerns of the library, IT, and research administration (Pinfield, Cox, & Smith, 2014). One area of concern that has been largely overlooked is research ethics.

Although ethical compliance often falls under the general purview of research administration, ethics review boards as a sub-group have received little mention, outside of ethics specific journals, with respect to their role in the changing research data landscape. This role is an important one, given the complexity of new data handling environments. The evolution of technologies for collecting, analyzing, and storing data and the push for open data raise questions about data security, participant privacy, and informed consent. These questions need to be considered by all campus partners involved in data management planning and infrastructure development. Ethics review boards, which are usually composed of experienced researchers and reviewers, are well-positioned to provide expertise in this area.

This study seeks to tap into that expertise by examining the experiences of ethics review boards at Canada's most research intensive universities. Using a phenomenological approach and semi-structured interviews, this study describes how eight chairs of ethics review boards have encountered changes to the research data landscape and how they perceive the ethical issues surrounding those changes. It is hoped that this research will contribute to a deeper

understanding of the ways in which ethical compliance processes should be situated with respect to the work of librarians on data planning, management, archiving, and publishing, and help illustrate the role of ethics review in campus-wide research data management activities.

Background

Ethics Review and Research Data in Canada

The human research ethics review process at Canadian universities is structured through a joint policy of the three federal research funding agencies, the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Social Sciences and Humanities Research Council of Canada (SSHRC), often referred to as the Tri-Council or Tri-Agency. Institutions receiving Tri-Council funding are required to sign a memorandum of understanding to abide by the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (Canadian Institutes of Health Research, 2014), or the *TCPS 2*. The *TCPS 2* establishes the policy framework for institutional ethics review in Canada, including the structure of research ethics boards (REBs), as well as the scope, principles, and even some processes for ethics review. The *TCPS 2* applies to all human research occurring under the auspices of institutions that have signed the MOU, not just Agency funded research.

Questions of ethics that apply to the research data lifecycle are primarily covered in the chapter of the *TCPS 2* on privacy and confidentiality. In it, the document outlines expectations that researchers describe procedures for maintaining participant confidentiality and for safeguarding private research data against “unauthorized access, use, disclosure, modification, loss or theft.” It also places a very broadly-stated responsibility on institutions to develop safeguards that should include “adequate physical, administrative and technical measures, and should address the full life cycle of information.” According to the policy statement, then, responsibility for data security throughout its cycle is shared by the researcher and the institution.

The Tri-Council’s Interagency Advisory Panel on Research Ethics (n.d.) has, since the original publication of *TCPS 2*, clarified its vision of stakeholder responsibilities through an interpretation of the chapter on privacy and confidentiality. In its interpretation, the panel places overall responsibility for safeguarding participant data and for anticipating potential data breaches on the researcher, with responsibility for review of the researcher’s data security measures falling on the REB. The same interpretation expanded institutions’ responsibilities to include “creating and maintaining a supportive research environment, establishing appropriate institutional security safeguards, training researchers and REBs regarding best privacy practices and implementing processes and policies that guide and support researchers and REBs in protecting participant confidentiality.” This structure places the REB in a role of gatekeeper (Cook, Snyder, & Calvert, 2015), in a position to observe trends in approaches to research data management and the degree to which institutions are meeting their policy obligations *vis-à-vis* infrastructure and training for data security.

REBs are also responsible for ensuring that privacy measures are communicated to participants through the informed consent process. Broadly, informed consent procedures are meant to establish the terms of participation - that it is fully voluntary with full knowledge of the purpose,

risks, and benefits of the research. With respect to data, the *TCPS 2* requires that participants be informed about the nature of the data that will be collected and the purposes of collection, who will have access to information about participant's identity, how confidentiality will be maintained, and the anticipated uses of the data.

The *TCPS 2* provides a big umbrella approach to ethics review but, as some critics have noted, it can be challenging for boards to consistently apply the document's principles across unique cultures and projects. As a result, REBs develop local rules and customs beyond what is recommended in an effort to ensure full compliance and consistent application of ethical principles within the institution (Bell, 2016; Schrag, 2010). Board members themselves may be unsure about what is required by policy and what is local practice (Cook et al., 2015). In some cases, identical research protocols sent to multiple REBs have received very different reviews (Warrell & Jacobsen, 2014), suggesting that the approach to ethics review in Canada is much more siloed than what is intended by the Tri-Council policy.

The significant responsibility for scrutinizing data management practices is not unique to Canadian ethics committees, of course. Speaking about American ethics boards, Hardy, Hughes, Hulen, & Schwartz (2016) noted that "one of the most difficult charges for IRB committees is the ability to anticipate a full range of potential conflicts or dangers that could result from improper data collection, storage, or maintenance, and to ensure that researchers develop plans to offset any risks from the outset of a project." This is exacerbated by the continuous evolution of the technologies and methods of research, an evolution that has occurred while the protocols used to protect participants largely have not changed (Buchanan & Ess, 2009). Reviewers are faced with the challenge of applying broad ethical principles to projects that include complex and fluid tools for data collection, storage, and security, while limited training resources for boards tend to be devoted to general research ethics rather than specific technological challenges (Buchanan & Ess, 2009).

Ethics and New Data Environments

How research data are collected, stored, shared and destroyed is heavily impacted by evolving technological, methodological, and philosophical approaches to research. These changes naturally raise questions about the suitability of the existing ethical framework within a complex landscape. This section will outline some of those questions.

Internet-based data collection has complicated ethics review by blurring lines between public and private information. In 2008, a working group of the Interagency Advisory Panel on Research Ethics, noting gaps in the first iteration of the *TCPS*, made a number of recommendations in order to address concerns around internet research, including differentiating between non-intrusive data collecting and participatory human research, as well as outlining requirements that researchers announce themselves and obtain consent when gathering data from online spaces where there is some expectation of privacy, such as chat rooms (Social Sciences and Humanities Research Ethics Special Working Committee). While the *TCPS 2* did fill some gaps around the scope of ethics review with regard to Internet research, it has been criticized for failing to

address other complexities of Internet research, including recruitment, estimating risk, and informed consent (Warrell & Jacobsen, 2014).

More recently, issues around consent and privacy have been raised in the context of big data analytics involving human data. Metcalf and Crawford (2016) note that, because the field of data science stems from disciplines such as mathematics and computer science that have traditionally been removed from human research, the use of human data collected from Internet sources by data scientists presents new challenges for both researchers and ethics review boards more accustomed to behavioural research. As there are few existing frameworks for the ethical handling of large datasets about people, researchers in the discipline are grappling with issues of risk and re-identification, and differentiating between public and private data. In health sciences research, as well, the case has been made for the reuse of medical and biobank data without consent (Bialobrzeski, Ried, & Dabrock, 2012). Ioannidis (2016) noted that proponents of the use of medical data without consent use the contradictory arguments that the data are routinely collected and, therefore, their additional use does not constitute research, and that they are valuable tools for discovery that can “mitigate, affect, modify, even force perceptions and interventions for participants and/or for their environments.” The author calls this contradiction the “oxymoron of research that is not research.”

These issues are pushed to the fore where there are increasing expectations that research data will be published in open repositories for secondary analysis. One of the core principles of informed consent is that participants will be made aware of the purposes of data collection. This principle is codified not only in ethics policy like the *TCPS 2*, but also in legislation such as the European Union’s new General Data Protection Regulation (n.d.).

While these issues are relatively new to some disciplines, other fields such as genetics have used open data practices for decades. GenBank was an early repository of DNA sequencing data, and one of the first to grapple with the challenges of informed consent and the protection of personal information in an open data environment (Mauthner & Parry, 2013). As a solution, an open consent process was developed in which participants would be made aware that their personal data could be used for indeterminate future applications. Open consent has been criticized on the grounds that many participants are not in a position to understand potential future uses of data and, therefore, the risks of participation (Kaye, de Vries, Heeney, Hawkins, & Boddington, 2009), and that participants may object to future uses on moral grounds if they had adequate information and opportunities to do so (Carusi & Jirotko, 2009; Parry & Mauthner, 2004).

Compromises to the problem of open consent have been proposed, including the option for participants to participate in the current research but opt out of inclusion in open datasets (Kozlakidis, Cason, Mant, & Cason, 2012; Yardley, Watts, Pearson, & Richardson, 2014), a dynamic process in which consent is obtained prior to each use of the data (Zarate et al., 2016), and creative commons licensing for open data that reflects the wishes of participants (Childs, McLeod, Lomas, & Cook, 2014). Carusi and Jirotko (2009) put forward the idea that participants be shown examples of archived data and provided the opportunity to edit their data before deposit. Unfortunately, these procedures are not practical for many research projects. Selective inclusion and retention of participant contact information may add risk to privacy (Lunshof,

Chadwick, Vorhaus, & Church, 2008) and participants may become difficult or impossible to contact over time (Childs et al, 2014).

While the protection of participant data is in most cases an ethical and legal imperative, informed consent as an absolute requirement has been questioned. The standard practice, in which researchers alone set the terms for data collection and storage, has been criticized as paternalistic, removing agency from research participants (Carusi & Jirotko, 2009). The terms of confidentiality may instead need to be a negotiation between researcher and participant (Bishop, 2009). In some cases, participants may wish to waive their rights to confidentiality or, going further, may even desire to be personally identified with their data (Carusi & Jirotko, 2009). The Personal Genome Projects (PGP), for example, openly publish identifiable medical records and genome sequencing data of participants in order to maintain the integrity of and add value to the data. Despite the fact that very intimate details about PGP volunteers are publicly available, participants tend to view the risks as minimal in light of the contributions they feel they are making to science (Zarate et al., 2016).

Novel approaches to confidentiality notwithstanding, the obligation researchers have to safeguard participant identities is a cornerstone of ethical research data management. How that stewardship looks in practice, though, has been impacted by both policy and technological change. On the policy front, changes to data management practice has been driven by the widespread adoption of open data mandates by funding agencies and publishers. While open data policies have many well-documented benefits, including transparency, accountability, and a reduction of the burden on participants (Borgman, 2012), they also increase the stakes for participants and researchers by imposing greater risk to the privacy of both parties (Mauthner, 2016). Researchers must be capable of fully de-identifying data prior to publication, and ethics boards must have adequate expertise to review procedures for safeguarding data in a variety of formats.

Qualitative data is particularly challenging in this regard. In some cases, the nature of a qualitative dataset may be such that the removal of identifying information to an extent sufficient to protect participant identities will render the data meaningless (Antes, Walsh, Strait, Hudson-Vitale, & DuBois, 2018). The interpretation of qualitative data often relies on the context of its collection and the relationship between researcher and participant; removing that context changes the nature of the data (Mauthner & Parry, 2013). Qualitative data are defined by these contexts to an extent that some have wondered whether novel questions can reasonably be answered by secondary qualitative data (Yardley et al., 2014), and, consequently, if there is value in archiving and publishing these types of data. Others see this challenge as one that can be overcome with support, in part, from the library through educational initiatives and selective publishing (Mannheimer, Pienta, Kirilova, Elman, & Wutich, 2018). Still, research by Tenopir, Sandusky, Allard, and Birch (2014) suggests that relatively few libraries are offering this level of data preparation and technical support.

Open publishing policies also add a layer to already complex and contradictory legal frameworks for data management. In addition to open data requirements, researchers may have to sort out questions of data ownership, particularly if researchers are working closely with community groups or sponsors (Hardy et al., 2016). There may also be existing institutional requirements for

data storage and retention, as well as provincial, national, or intergovernmental legislation that regulates the conditions under which personal information, including that contained in research data, may be held and distributed. And there may be differing practices within or between disciplines that influence a researcher's data management practices (Carusi & Jirotko, 2009). As libraries participate and lead both institutional policy and infrastructure development for data publishing (Cox, Pinfield, & Smith, 2014), they will be impacted increasingly by the complex issues of data ownership and the legal framework surrounding data collection and distribution.

Like the evolution of policy frameworks, changes to the technological landscape have added additional layers to procedures intended to protect participant data. Researchers must be aware of and account for the surreptitious collection and transfer of data, as when IP addresses are collected along with ostensibly anonymous data or when tools used to collect or house data automatically save those data to remote servers. Any participant data stored on web-connected servers, in fact, may be vulnerable. According to *The Times* (Yeung & Bennet, 2017), universities in the UK received 1152 unauthorized server access attempts in 2016/2017 alone. While the vast majority of these attempts are unsuccessful, some, like a ransomware attack on University College London in 2017, do cause significant disruptions and pose a risk to confidential data stored on internal servers (Hern, 2017).

Technologies also play a role in international research and, specifically, the cross-jurisdictional transfer of data. The complexity of electronic data flows makes it difficult to trace the movement of data sent by email or transferred by other means. Even when data are transferred within a narrow locality, those data may cross international borders (Clement & Obar, 2015). Laws and ethical standards may differ between countries or even provinces, so that data handling practices that have been cleared in one location might not mesh with the ethical or legal frameworks in another (Mauthner & Parry, 2013). Data that cross borders, electronically or physically, may also be subject to search and seizure laws that differ from those of the originating country.

Anonymization and de-identification of research data mitigates some of the risks to privacy imposed by technological change, but both of those procedures come with competing interests that must be balanced (Parry & Mauthner, 2004). When data are fully anonymized, participants cannot withdraw from a study or be contacted for follow up or reuse of the data. As we have seen, data that are anonymized may be stripped of important contextual information making further analysis impossible. When data are de-identified, on the other hand, withdrawal and follow up are possible, but there is a greater risk that confidentiality will be breached. Technological security measures, such as encryption, may provide added protection, but encryption imposes an additional layer of complexity to the management of data and, as Sebastian van Baalen (2018) recommends, "researchers should remain cautious about employing technical solutions to what are essentially ethical problems."

In addition to traditional domain and methodological awareness, researchers need to be cognizant of changing technological, ethical, and regulatory structures that impact the research data lifecycle. Navigating this complex environment requires the support of diverse nodes of knowledge throughout the campus. Libraries have to a large extent taken the lead in supporting research data management and publishing practices, but knowledge transfer among all of the campus units supporting RDM is crucial to providing comprehensive and concordant support.

The study seeks to facilitate one aspect of that knowledge transfer by examining the experiences and perceptions of ethics review board chairs in light of the above changes to the research data landscape.

Methodology

Interpretive Phenomenological Analysis

This study uses interpretive phenomenological analysis (IPA) as an approach to examining the experiences of human ethics board chairs in the face of significant changes to the research data landscape, with particular attention to recent technological, philosophical, and regulatory transformations. IPA was developed in 1996 by Jonathan Smith and has since been used widely in the psychological and health sciences as a means of examining the experiences of individuals with distinct phenomena. More recently, the method has been applied in a broader range of disciplines.

John Budd (2005) sees a role for phenomenological thinking in library and information studies, noting that “phenomena such as relevance judgments are understood in various ways, but the phenomenological approach offers understanding in a way that others cannot.” He goes on to recommend phenomenology as a means of understanding the experience of using the tools and technologies of LIS. IPA as a methodology was introduced to LIS through the discipline’s linkages with health sciences in a small number of papers examining the information seeking behaviours of patients experiencing various medical conditions (VanScoy & Evenstad, 2015). It has since been employed further as a framework for studying experiences related to technology and reference services (VanScoy & Evenstad, 2013). In studies of higher education, IPA has been used more widely to examine experiences ranging from online learning by mature undergraduate students (Meyers & Bagnall, 2017), the transitions of practitioners to classroom instructors (Wood, Farmer, & Goodall, 2016), classroom observations (Tenenberg, 2016), student evaluations (Leary, 2017), and teaching in a post-1992 UK university (Holland, 2014).

IPA as a qualitative method is shaped primarily by three elements from the philosophy of knowledge: phenomenology, hermeneutics, and ideography (Smith, Flowers, & Larkin, 2009). Phenomenology stems from Edmund Husserl’s ideas about experience, specifically the ways in which we might better understand experiences by setting aside thought of our activities in the world in order to examine the essence or characteristics of the experience itself. It was Husserl’s student Martin Heidegger who first applied hermeneutics, or the theory of interpretation, to phenomenological inquiry. Heidegger (1962) posited that human beings cannot escape the context of their own worlds, and that this context inevitably influences their descriptions of their world. In this framing, descriptions are better conceptualized as interpretations framed by prior experiences. In phenomenological research, researchers interpret the sense-making activities of participants, which are themselves interpretations of experiences. This is often referred to as double hermeneutic.

IPA is also ideographical in that it focuses on the particular. Rather than focusing solely on the commonalities in the experiences of a group, IPA retains the individual experience, highlighting divergences between participants as much as similarities. Smith et al, (2009) are careful to note

that ideography “does not eschew generalizations, but rather prescribes a different way of establishing those generalizations.” Like other qualitative methods, IPA includes the development of themes, but it presents those themes as they relate to each individual as well as the group.

For this study, IPA provides a suitable framework for analysing the experiences of ethics board chairs in the face of significant shifts in the research data environment. Through a deep analysis of these experiences, we can better understand the impacts of both organic and intentional changes on the process of ethics review. It is hoped that this will contribute to a broader discussion about the way longstanding ethical principles are applied in new contexts and how library’s and other campus departments can support the management of research data throughout its lifecycle. Moreover, the hermeneutical approach intrinsic to IPA allows the researcher to acknowledge prior experiences with ethics review and research data and incorporate these into the interpretation of the data.

Participants and Recruitment

Research ethics board chairs were selected both for their capacity as leaders of discussions and decision making processes around ethics and for the likelihood that they have the greatest amount of experience among REB members in the realm of ethics review. After approval for the study was obtained from the Human Research Ethics Board at [researcher’s affiliated university], the contact information of REB chairs at U15 Group of Canadian Research Universities member institutions was gathered from publicly facing websites. U15 universities were chosen as the most research intensive in the country, receiving 79% of the competitive research funds allocated to Canadian universities (U15 Group of Canadian Research Universities, n.d.), and therefore likely to generate the widest array of circumstances related to the ethics of managing human research data. REB chairs at primarily French speaking universities were excluded due to language limitations of the researcher. In total, 35 chairs and co-chairs of 33 research ethics boards were invited to participate. Of the total, 14 boards (16 chairs) were classified by the researcher as primarily medical/clinical in scope, while the remaining 19 boards (19 chairs) were classified as social/behavioural in scope.

The chairs were invited to participate in semi-structured telephone interviews in two waves during the spring and summer of 2017. A total of 8 chairs consented to be interviewed for the study, 2 from medical/clinical boards and 6 from social/behavioural boards. As the number of participants in phenomenological studies normally ranges from 5 to 25 (Creswell & Poth, 2018), 8 chairs were deemed sufficient to explore the changing data landscape in this context. The 8 participants were located in 5 Canadian provinces. The number of years in which participants sat on ethics boards in any capacity ranged from 6 to 17 years.

IPA studies typically attempt to gather the experiences of a homogeneous group (Smith et al., 2009). Although in this case there was some variation in discipline, geography, and years serving on boards, the participants shared enough characteristics, including their roles, institution size and type, background knowledge, and understanding of the regulatory environment of ethics,

that they had all experienced the changing research data landscape within a narrow context and, the researcher believes, a meaningful comparison of experiences was possible.

Data Collection and Analysis

The interviews lasted approximately 40 - 45 minutes each. Participants were asked about the context of their work and how they have perceived changing policy and cultures around open data, the increasing complexity and variety of data storage, transfer, and security technologies, and evolving approaches to data collection and retention. Participants were asked to consider specific events or circumstances and to use those to illustrate their responses to discussion topics, where appropriate.

The primary investigator recorded and transcribed the interviews verbatim. Participants were assigned pseudonyms (Don, Robert, Emily, Lois, Laura, Susan, Matthew, and Heather) and the transcripts de-identified. This process was followed by multiple close readings of transcripts with initial note taking using qualitative data analysis software. Notes and comments were initially descriptive, focusing on the explicit meanings of the text. Additional readings and comments drifted away from the explicit to the conceptual, attempting to situate the text within its context and draw an understanding of how this context shaped the meanings of participant responses.

Note taking was followed by the development of emergent themes. This process involved close analysis of notes and transcripts both together and separately and the bottom-up development of themes from the data. All themes were imported to a concept mapping tool and, through a process of abstraction, patterns among themes were identified. Themes were arranged along the lines of these patterns and superordinate themes created that describe the groups of subordinate themes. Table 1 outlines superordinate and subordinate themes developed through this process.

Once the themes had been established and clustered, profiles of each participant were developed, outlining their experiences and approaches to the changing data landscape. Profiles and transcripts were then closely compared to the thematic structure in order to establish relationships between the themes and the experiences of each chair. This iterative examination of the part and the whole is referred to as the hermeneutic circle, and is a key aspect of interpretive analysis. The nature of the relationships between the themes and the chairs are described in the next section.

Table 1: Superordinate and Subordinate Themes

Superordinate Themes	Subordinate Themes
The regulatory environment	Legal and policy framework
	Data retention
	Open data polcies
	Qualitative data

Reacting to uncertainty	Unformed data landscape
	Open data and consent
	Data security approaches
Threats	Borders and travel
	Motivated hacking
	Accidental loss
	Re-identification
Oversight	Compliance review
	Definitions
	Rerouting
	Muted skepticism
Situating ethics review	Institutional resources
	Institutional systems
	National-level policies and procedures

Results

Five superordinate themes emerged from the interviews: the regulatory environment, reacting to uncertainty, threats, oversight, and situating ethics review. Together, these themes shed significant light on the experiences of ethics review board chairs with respect to changes to the research data landscape.

The Regulatory Environment

Most of the participants commented heavily on the complexity of the regulatory environment surrounding research data. As a group, they are well aware of the legal and policy frameworks that guide data collection, retention, and storage, even those that do not fall directly under the ethics umbrella. Both Lois and Susan discussed limitations that provincial legislation imposes on research data. For Lois, those limitations were related to the transfer of personal data, “In [province] we have this privacy law where private information is not allowed to leave the country. So if it’s a cloud storage or any kind of storage that’s outside of the country, then we’re not allowed to use that for storage of personal information.” Susan was more concerned with the impact of provincial legislation on retention periods, “Some jurisdictions like [province] don’t allow you to store anything indefinitely. You have to be able to give real reasons why. They put it into the law, so we’re a little stuck with sort of legal parameters about protection of data.”

Susan, Don, and Emily also discussed the US Patriot Act and its ramifications for the privacy of data stored on American servers. Neither Emily nor Don factor it into ethics review considerations, though, citing approval by institutional privacy officers for data to be stored in the US and the widespread use of American cloud-based software for institutional email and data storage. Susan is more cautious with respect to data storage in the US. She encourages researchers to collect data using tools located in Canada and prefers that participants are informed when data will be housed in the US. For her, “the concern is that [US government agencies] have the *right* to grab data,” and this constitutes a degree of risk to participants.

In addition to legal frameworks, participants also cited complex and at times contradictory policy frameworks that impact data handling. Robert saw these contradictions within the institution, “The senate of [the university] has a policy with respect to research data storage, as does our collective agreement. The university also has a data storage piece. Those pieces are not congruent one hundred percent.” While Robert emphasized the challenges of working amidst contradictory regulatory frameworks, others were more comfortable doing so. Both Emily and Susan discussed the additional layer of policies imposed by funding and oversight agencies. When asked about reconciling conflicting requirements, Susan summarized her approach plainly,

There can be differences, but when there’s differences, you just need to follow the highest minimum. So the *Tri-Council Policy Statement*, you have to follow that. Whatever applicable laws also have to be followed. So if the law sets a higher standard, you have to follow the higher standard. But if the law is a lower standard or silent, you just follow the *Tri-Council Policy Statement*.

As Emily noted, the *TCPS 2* does not prescribe retention period for human research data. When asked about data retention, the chairs expressed differing expectations, some influenced by institutional policies and some by longstanding practices. None were rigid with respect to specific time periods for data retention, but both Matthew and Emily cited a seven year retention period as a starting point that can be adjusted based on the requirements of a project. Heather, on the other hand, felt that data storage for an extended period of time is unnecessary in most cases, “I tried to communicate to researchers that if the data are going to be destroyed, then you should destroy them in a timely fashion once the study is complete.” Many of the chairs were unaware of the origins of local data retention period practices, but Don speculated that they stemmed from institutional records retention policies.

The chairs were also split on their perceptions of indefinite or perpetual retention of research data. Susan cited legislative requirements that personal data be destroyed after a specific period of time, ruling out perpetual retention as an option. Emily, Robert, and Matthew all felt that indefinite retention periods are too ambiguous, with Matthew asking “what does indefinitely mean? It could mean I’m going to get rid of it in six weeks, I’m going to hold onto it until the day that I die.” Others saw value in perpetual retention. Don, for example, was of the opinion that “as long as the researcher has plausible long-term plans for securing the data I think we ought to be encouraging people to keep their data.” Laura, likewise, noted that “ten years ago, the norms were that you had to specify the date for the destruction of the data. And if you said that you were going to keep it, then there were a lot of questions about why would you do that.

And now we've gone a complete circle, and if you gave us a date for destruction, we'd be like, 'Are you really sure you want to do that?'"

Much of the discussion of the retention of data was done in the context of open data mandates by research funders and publishers. Nearly all of chairs were very well aware of open data policies and their potential impact on ethics review. Don, for example, referred multiple times indirectly to the *Tri-Agency Statement of Principles on Digital Data Management* in which Canada's major federal granting agencies outline their support for open data and data management practices. For him, though, open access policies have not changed the fundamental practices of the board, "Our role hasn't changed at all. I mean our role is when it comes to this kind of stuff, is privacy protection." Heather also recognized the potential repercussions of open data mandates on ethics review, but had not yet seen many situations in which researchers were compelled by policy to publish data.

Still others felt that they had already been impacted by open data policies, primarily through intensive discussion and modification of board guidelines in preparation for the changes these policies would bring. Lois, for example, explained,

We have to think about how will those be stored, where will they be stored, who might be potentially using those data and why, and do the participants, have they been made aware that these data are going to be stored? Is it going to be potentially re-identifiable? So it raises all these questions, whereas before, pretty much typically everyone said, 'I'm going to collect these data, do these interviews, transcribe them, and as soon as the project's done they'll be destroyed.'

Lois, like Heather, had seen few situations in which open data policies were present in the ethics review process.

Heather noted that open data policies seemed to apply more to quantitative research and lamented what she viewed as the exclusion of qualitative data from serious consideration for open retention, "I have to say they've been profoundly uninterested in qualitative data, in terms of would we keep it forever...The quantitative people said you can't really use it for much once you've completed your study...and it was like, we certainly used all the data across researchers and so on, but there does seem to be not much interest in perpetual data storage and availability of qualitative research." Heather was confident that qualitative data can in many cases be de-identified and made available for secondary use, and was primarily concerned that these valuable data were being discounted.

Reacting to Uncertainty

Despite their in-depth awareness of changes to data related policies, every participant expressed a sense of uncertainty about the current research data landscape. The complexity of ethical decision making, the absence of guidelines, the array of people involved, and changing technologies were all cited, often in tandem, as reasons for uncertainty. Laura, for example, described the environment as "a case of something's happening too quickly for all of the pieces to get caught up together." Both Robert and Susan, who rely heavily on established policy for

decision making, referred to the challenges of keeping their guiding documents current. Robert, for instance, explained that “we have talked about them in our meetings, we have tried to figure out ways of dealing with these issues, but we recognize every time we start looking at getting better policy around these kinds of issues, that as soon as we get a policy, other things have changed.” While Robert expressed frustration at trying to keep ahead of changes, Susan was more resigned, noting that standards and laws in this realm rarely keep up with current practices.

For Matthew, changes to board policies and approaches happened so regularly that he struggled to describe the current policy. He noted, “I feel like I don’t always have my finger on the pulse on all the guidelines and all the requirements, because they’re actually changing a fair bit each time...and so a lot of that stuff I don’t always have on me, on the tip of my tongue.” Don experienced similar challenges but was confident that the future would bring more certainty, “So it’s kind of a daily issue for me but it’ll disappear over time as the practices change and the understandings change and also the resources become available for researchers to be able to do what’s expected of them now with the help of the institution.”

Amidst the generalized sense of uncertainty among the chairs, some referred to specific issues facing their boards. One of these issues is that of the transition to an open data environment and its implications for informed consent. Laura, in particular, articulated the impact of the change,

There’s generally a time lag between collecting the data, writing the paper, getting it published, and now having the open data issue arising. So some of the problems that we’ve encountered is consent forms that have the old language about data will not be shared with anybody and/or it will be destroyed after a certain date. And then some of that data being in the paper that ends up getting published where there’s a request if possible to make that data freely accessible. So then you can’t post publicly data unless you told the participants that that was a possibility, even if it’s anonymous.

Many of the chairs, like Laura, alluded to the necessity of modifying guiding documents to incorporate consent for open publishing of data. Some of these participants had already engaged in discussions and made the decision as a board to permit open consent for secondary use of data. Although both referred to complicating factors, Heather and Lois were comfortable with open consent for low-risk research with the proviso that participants are notified. According to Lois, “it’s much cleaner if they can include at least some kind of a statement on the consent forms that lets the person know that this data is going to be stored indefinitely at [University] and it may be used in the future for other research. And then if the person signs off on the consent, then good, you have that statement in there that kind of covers you for some future research.”

Don described his board’s approach to open consent similarly, “We have to change the template for our consent form to better anticipate this situation where the participants are entitled to be told that anonymous data from the research they’re participating in will have this kind of open ended, in-perpetuity use. In other words, essentially the data are going to exist on the public domain and participants are entitled to know that.” On a personal level, though, Don expressed some uncertainty about open consent, “On the other hand, some people would say ‘well, because it’s a distinct right people have [to know how their data will be used] even for their anonymous information and so they should be given the opportunity to [opt out of secondary use], and to not

give that to them is disrespectful of them.’ And I could be easily persuaded that that’s where we should come down.”

In addition to the philosophical challenges of informed consent, participants conveyed their reactions and approaches to the uncertainties of technology, primarily those technologies used to collect and store identifiable data. At Don’s institution, the Board adopted a blanket policy requiring all identifiable data to be encrypted, “If there is any question at all about whether or not the data are really that clearly anonymous or not, there’s just one other category for digital information - that is it has to be encrypted and stored in places where the encryption key would not be known to the people.” For him, this approach has reduced some of the uncertainty around data security.

For others, though, the range of tools used to collect, transport, and store data are more confounding. Matthew expressed the need for continuous education for the board itself, “Especially this day and age when we have a lot of things like, people are using cloud storage and things like that. We are still actually learning quite a bit because it’s always changing.” While Susan also acknowledged the need for education among board members, she also put some of the onus for understanding and conveying details about the security of different technologies on researchers, “If it’s electronic data, it’s been a struggle for a lot of years to figure out how to do that, store that well with safety issues, but we expect people to be able to tell us how they can manage the electronic data.”

Threats

As a group, the chairs viewed threats to research data and, ultimately, participant privacy from several sources. A theme that occurred repeatedly, both prompted and unprompted, was the elevated threats associated with borders and travel. Many used data collection in far-flung locations, including Africa, Japan, Korea, and Saudi Arabia, as illustrative of higher-risk situations where additional ethics scrutiny is required. Don’s example reflects the theme clearly, “If you’re a graduate student wandering around Uganda with a memory stick in your pocket on a regular basis...if you can’t encrypt it on your memory stick, you better get it some place quickly from the memory stick where you *can* encrypt it.” Emily pointed out the need for greater data security when travelling, but noted that the data were still vulnerable to loss, “We have had people of course going to other countries with data and there we expect encryption, not just password protected, even though that’s not even a guarantee, right? Someone could take your laptop\ . Like, the border control can take your laptop.”

Most of the chairs viewed any transportation of identifiable data on portable storage devices as highly risky. Matthew, for example, found that “a lot of people wouldn’t think twice about, ‘Here’s my data file. I’m going to throw it on my flash drive here and bring it to my home computer’ or something like that, and don’t realize at that point that it’s no longer in that safe zone it was before.” But others were less stringent. Heather, for example, does not require encryption of data on portable devices, “We decided after talking about this to require password protected. Many of the faculty use encrypted data storage devices. Students don’t always do that. And of course encrypted devices are more expensive. And so it’s at least password protected. I don’t specify anything further than that.”

In addition to accidental loss or theft during transportation, some participants raised the threat of intentional intrusion by third parties. Matthew and Robert both invoked the image of the unknown, technologically savvy hacker trying to access private data. For Matthew, “the bad guys are always one step ahead,” while Robert noted that “new devices are available, new kinds of encryption are available, new ways to hack through the data are available.” Although both chairs did see this type of intrusion as a viable threat, all discussions of the topic were speculative, and Matthew conceded that “it would have to be a motivated person” who would attempt to steal the data by electronic means.

The theme of the motivated intruder was also raised by Don, but in the context of re-identification of data. For him, there is a risk that multiple variables in a de-identified, public dataset may be used to triangulate the identity of research participants, although he acknowledges that it is an unlikely threat. Matthew, in contrast, sees re-identification of participants as a more significant risk in an open data environment, “That’s where the big risk comes in...there might be freely available shared data on a certain minority group of individuals and there are four pieces of demographic data in that file. And with those four pieces of demographic data...that might narrow it down to three people. Three people, which could easily happen. And so that’s where I think the potential issue is.”

Matthew also sees accidental exposure of private data by researchers as inevitable, “It’s not that they might happen, mistakes will happen at some point where non-anonymized data will be shared.” Heather expressed a similar sentiment about the inevitability of inadvertent disclosure of personal information, “I do think most researchers are pretty careful with their data. I mean, for all the obvious reasons...And one day someone somewhere will screw up, and then we’ll all have to accommodate.”

One of the key principles of the ethics review process emphasized by nearly all of the chairs is that of proportionate review - if a greater risk is present due to the use of a particular methodology or the vulnerability of participants, there is a higher level of scrutiny given to the proposal. For many of the chairs, the duty of concern for certain populations extends to their attention to threats of re-identification, particularly when the data will be published. Participants cited children, sex trade workers, victims of violence, students, and small indigenous communities as those about whom data may be more difficult to fully de-identify or who may be at higher risk if they are personally identified. Emily noted, for example, “we certainly have people who are working with very small groups of indigenous people in a very specific part of Canada. And there’s like a hundred of them. So you’re going to be able to identify those people. So maybe it shouldn’t be on a public repository.”

Oversight

All of the chairs referred in a variety of ways to the oversight their boards provide in advising, directing, teaching, and protecting those involved in the research process. In part, these roles are a natural development of the ethics review regime in Canada - REBs are structured as gatekeepers to human research and given a significant degree of oversight over the application of

the principles of ethics review. But, as the participants made clear, some non-traditional elements of this role have fallen on REBs due to uncertainties in a changing data landscape.

Ethics review boards, of course, are not the only groups who have been impacted by changes to data regulations. Researchers also face changing obligations imposed by sponsors or publishers and, in some cases, may be unaware of those changes. Some chairs described the occasional need to remind researchers of their own data requirements, even when those obligations fall outside of the scope of ethics review. Susan, for example, was keenly attuned to data retention requirements for Health Canada sponsored studies, “So if the researcher has come to us and said we’re going to store it for 7 years because University Whatever has a policy that you store everything for 7 years or 5 years or 10 years...but it’s a drug study, we’ll say ‘No, this is a product that going to go forward for regulation, to be regulated for drugs. So you have to store it for 25 years.’” Emily, as well, recounted a scenario in which she reminded a graduate student of the need to retain data until their work is finalized, “A PhD takes 4 or 5 years to finish. You don’t want to keep it for only three years when the PhD is not even done.”

Participants also discussed the need to help some researchers understand their own data and the terminology that defines them. Language related to the identifiability of data was described as particularly confounding. Lois outlined the challenges she faces in this regard, “Because people use those terms incorrectly all the time. So they’ll say it’ll be totally anonymous, but they’re interviewing them in person, they’ve got written transcripts, they’ve got codes, they say they’re going to send them a copy of their transcripts, they’re going to, you know, so it’s not anonymous, right? It’s de-identified, hopefully.” While most of the chairs were cautious about definitions used by researchers, Heather and Matthew both acknowledged the need to check their own language at times.

Many of the chairs spoke about rerouting researchers to other campus departments before signing off on ethics applications. In the context of research data, redirections were most frequently toward information technology support for assistance with data security. For some, like Robert, these referrals are a product of the range of technologies used by researchers, “For somebody who says ‘Well, I’ve got my data on such and such a platform,’ we send them to the encryption people, not us, to have them explain what kind of encryption is available for your particular device.” Laura, Lois, and Susan also discussed the role of the library to support researchers in planning data stewardship and archiving activities of the project. Interestingly, many of the participants used language of redirection - ‘point them,’ ‘send them,’ ‘put them in contact’ - when discussing technological issues, while using a more collaborative tone - ‘working with,’ ‘give advice,’ ‘helping’ - when referring to the types of support provided by librarians.

Amidst discussion of the educational and referral functions of the board, there was a muted skepticism expressed by some chairs that researchers adhere to the protocols for data handling that they describe in the ethics review process. Both Susan and Emily made perfunctory remarks suggesting that some researchers use boilerplate language just to satisfy ethics review. Both quickly emphasized that they do trust that researchers intend to handle their data according to stated protocols, but there were intimations of doubt that researchers pay close attention to ethical data handling practices. Robert was more direct in his assessment, “We can put all kinds of safeguards in place and the researchers will do whatever they want.”

Situating Ethics Review

Many of the participants, when speaking about their experiences in the changing data landscape, situated the ethics review processes within larger networks or systems. Most were finely attuned to the mutual influences between local processes and resources, and national-level concerns about the conduct of research.

Several participants referred to the necessity of sufficient institutional resources in supporting the ethical conduct of research. For Susan, this meant providing leadership for adequate management of data, rather than strictly providing funding,

It doesn't mean that the institution has to provide all the means. They're not the funder of the research, although they support it. To some extent a research institution ought to help researchers do the right thing. And so I think it's quite proper, for instance, when there's a lot of sensitive research going on that...the institution kick in and help educate their research teams, be supportive of their research teams, provide flexibility in some respect, but also a little bit of oversight.

Others were more specific about the places from which support should come. IT services mentioned most frequently, but the need for support from the library for data publishing was also discussed by some. Laura included support for understanding privacy issues, as well, "There must be institutional resources available. So perhaps there's a privacy officer at their university who could also guide them on whether some of the information that they're including that they think is anonymous might actually not be."

Most of the chairs were broadly aware of the technological infrastructure in place at their universities to manage the storage, transfer, and archiving of data, but some struggled to name the systems or provide details. Matthew, for example, was able to generally describe an internal file transfer system, but had to look up the name of the system during the interview, "We have recommended, and I'm not going to remember the name for this right now, but we actually have secure ways of transferring information on campus that doesn't use email." Still, as a group, the chairs were attuned to the existence of these institutional supports, who has responsibility for them, and how the systems can assist researchers engaged in the ethics review process.

The chairs situated their boards within the broader context of ethics review in Canada in different ways. Don pointed out the tendency for boards to develop their own approaches, "Although we have the *TCPS* and we have these frameworks and all that kind of stuff, institutions have and tend to exercise a lot of their own discretion about how they're going to deal with these things." Others, faced with some uncertainties, expressed the desire for national guidance and infrastructure to support review of data handling practices. Matthew, in particular, felt that more cross-institutional clarity would help the work of his board.

Some participants also spoke of shifting responsibilities between national-level structures and individual institutions, and the impact these have on the board. Speaking of impending Tri-Council data management and archiving requirements, Lois explained,

The Tri-Councils have said we're going this way and so it's up to us as universities who have funding from those agencies to suddenly implement it. And so different people in different parts of the university are kind of working towards it. And so we're trying to figure out who are those people and what are they working towards, what would they recommend, how do we provide that information to our board members so they can use it in assessing the applications and how can we provide this to the researchers so that they have some clue of what to say or what to consider at least.

Discussion

Of the changes to the research data landscape on university campuses, it is the adoption of new technologies to collect, transport, and store data that have had the greatest immediate impact on ethics review, according to the chairs in this study. Primary collection of data, in particular, was singled out as one of the more problematic phases of the data lifecycle. The participants highlighted the challenges of understanding the implications for privacy of the myriad tools available to researchers to collect and store data and, in turn, to provide adequate guidance to researchers seeking to understand the implications themselves.

Many of the chairs conveyed a general sense of security about data that are de-identified and housed on campus computers and servers, and conversely, a sense of unease about the risks around identifiable data stored in any format, as well as data that are being transported on portable devices. They expressed confidence in institutional infrastructure for managing data but identified numerous threats to participant privacy outside of those structures. Among those threats, unauthorized intrusion (i.e. hacking) was cited by several participants although few details were provided about the conditions under which this was most likely to occur. Accidental loss or theft of portable devices was also cited as a significant threat, but many were also confident that password protection and encryption of these devices were sufficient to mitigate the risk. Government intrusion was not a significant concern, except in cases where data are being physically transported across borders. Although the privacy risks of recruiting participants publicly through social media were mentioned, none of the participants expressed concern about access to data by the proprietors of software used to collect data. Given what we know about the commodification of user data by a wide range of internet companies, this may be an area requiring further education and guidance.

Some of the chairs are indeed seeking broad guidelines that will help facilitate the assessment of risk, but it is not ethical guidance that most of the chairs are seeking. Primarily, they would like to understand the technology better - the paths that data travel, who has access, the lifespan of data on particular platforms - in order to more acutely apply their existing knowledge of the ethical conduct of research. Although information about the provenance of data processed by proprietary technologies can be elusive, many libraries are well-situated to provide some degree of support in this regard. Using appropriate technologies for research data storage, analysis, and sharing is already a part of the suite of information provided through library data management services. By working with ethics review boards to integrate technology recommendations into ethics training or ethical guidelines, librarians working in this area would not only ease some of the uncertainty faced by the boards, they might also reach new audiences for RDM advocacy

initiatives and sharpen their own awareness of privacy concerns related to research technologies. Going further, coordinating data management planning with ethics review has the potential to improve adherence to research protocols by fleshing out procedures for secure stewardship of data.

Navigating data archiving and publishing requirements is another area in which libraries are well-placed to collaborate with ethics review boards. At the time of the interviews, the impact of open data initiatives on ethics review was largely one of anticipation. Although most of the chairs had already done significant work to understand the implications of data publishing on ethics review and to modify forms and guidance for researchers, they reported few instances of open data publishing in practice. There is potential, though, for a significant burden to be placed on the ethics review board to respond to questions around data archiving and publishing. Some of the chairs in this study saw the library as a natural leader in this area, opening the door to partnerships and added support where open data policies impact ethics review.

In order to integrate supports to this extent, campuses should consider developing a framework outlining the types of skills required and where they reside on campus.. The chairs conveyed a strong sense of role ambiguity in the changing landscape. Verbann and Cox (2014) noted the potential for jurisdictional conflict when it comes to research data management, and it may be that the ambiguity felt by the chairs is a result of this jurisdictional overlap. While it was fairly clear that issues related to encryption, data transfer, and some aspects of archiving fall in the realm of information technology services, the chairs saw other pieces related to data policy, de-identification, training, and data publishing as being scattered across other realms including the library, privacy officers, the research office, funders, publishers, and government. Policy was characterized as a moving target, and it was not always clear who is responsible for its fulfillment and for related educational initiatives.

There are competing interests at play between the ideal of coordinating approaches to research data management across institutions, and the influence that local cultures and customs have on the research data environment. Researchers are encouraged to use internal infrastructure for storing and sharing data, and many of the policies and norms guiding data retention and disposition come from the inner workings of the university. At the same time, there may be a clash of expectations when external policies from funders or publishers, or legal requirements apply. This clash is often first encountered by the ethics board, which is in a position to provide initial assessment of a research plan and is frequently aware of competing requirements for data management. The library community is heavily involved in developing shared standards and shared data infrastructure for cross-institutional benefit. While it makes economic sense to pool resources in this way, it is important that librarians account for the unique characteristics of their institutions and be aware of conflicting policies so they help ensure researchers on their campus are more fully informed of research data requirements. Working with local ethics review boards to identify areas of potential conflict would be a positive step toward reconciliation of regulations and practices between the local and broader communities.

The REBs in this study encounter many questions about the responsible management of research data, some of which are arguably better suited to other departments. Along with the establishment of frameworks of expertise and knowledge transfer between campus services,

centralized oversight for institutional research data practices and initiatives would help reduce uncertainty and deflect questions to those with the best answers. Many campuses have already delineated responsibility for RDM initiatives, but it often lives in one department, usually the library or the research office. The degree of change to the research data landscape may now be such that oversight needs to be in the form of a permanent body with representatives from the library, research administration, ethics review board, information technology, privacy office, and other relevant departments.

Conclusions

As the research data landscape shifts, in terms of technological innovation, novel uses of existing tools, changing cultures of practice, and policy adjustments, it is important to consider the impact that the new environment has on those most heavily impacted. Much of the conversation to date has focused on library initiatives to support data management planning and data publishing, IT's role in developing and supporting the infrastructure for data security and archiving, and the work of research administration in navigating funder policies around open data. But the ground is also shifting underneath the feet of university ethics review boards. There is a sense of uncertainty among many of the chairs who participated in the study, primarily around the technologies used in the research process and the ways in which those technologies handle confidential data. At the same time, changes to the regulatory environment complicate efforts of REB chairs to maintain currency, provide accurate guidance, and to fully understand the implications of policy on core issues like informed consent and the responsible stewardship of research data. While some have adopted approaches that address these fundamental concerns to their satisfaction, others are less certain. Institutional support in the form of training and coordination of departmental initiatives, with centralized oversight, may ease some of the uncertainty currently felt by those boards.

References

- Antes, A. L., Walsh, H. A., Strait, M., Hudson-Vitale, C. R., & DuBois, J. M. (2018). Examining data repository guidelines for qualitative data sharing. *Journal of Empirical Research on Human Research Ethics*, 13(1), 61-73. <http://doi.org/10.1177/1556264617744121>.
- Bell, K. (2016). The more things change, the more they stay the same: The TCPS 2 and institutional ethical oversight of social science research in Canada. In W.C. van den Hoonaard, & A. Hamilton (Eds.) *The ethics rupture: Exploring alternatives to formal research ethics review* (pp. 189-205). Toronto: University of Toronto Press.
- Bialobrzeski, A., Ried, J., & Dabrock, P. (2012). Differentiating and evaluating common good and public good: Making implicit assumptions explicit in the contexts of consent and duty to participate. *Public Health Genomics*, 15(5), 285-292. <http://doi.org/10.1159/000336861>.

- Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of Social Issues*, 44(3), 255-272. <http://doi.org/10.1002/j.1839-4655.2009.tb00145.x>.
- Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059-1078. <http://doi.org/10.1002/asi.22634>.
- Buchanan, E., & Ess, C. (2009). Internet research ethics and the institutional review board: Current practices and issues. *ACM SIGCAS Computers and Society*, 39(3), 43-49. <http://doi.org/10.1145/1713066.1713069>.
- Budd, J. M. (2005). Phenomenology and information studies. *Journal of Documentation*, 61(1), 44-59. <http://doi.org/10.1108/00220410510578005>.
- Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences & Humanities Research Council of Canada. (2014). *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*. Retrieved from http://www.pre.ethics.gc.ca/pdf/eng/tcps2-2014/TCPS_2_FINAL_Web.pdf
- Carusi, A., & Jirotko, M. (2009). From data archive to ethical labyrinth. *Qualitative Research*, 9(3), 285-298. <http://doi.org/10.1177/1468794109105032>.
- Childs, S., McLeod, J., Lomas, E., & Cook, G. (2014). Opening research data: Issues and opportunities. *Records Management Journal*, 24(2), 142-162. <http://doi.org/10.1108/RMJ-01-2014-0005>.
- Clement, A., & Obar, J. A. (2015). Canadian internet “boomerang” traffic and mass NSA surveillance: Responding to privacy and network sovereignty challenges. In M. Geist (Ed.), *Law, privacy, and surveillance in Canada in the post-Snowden era* (pp. 13-44). Ottawa: University of Ottawa Press.
- Cook, K., Snyder, J., & Calvert, J. (2015). Canadian research ethics board members' attitudes toward benefits from clinical trials. *BMC Medical Ethics*, 16(1), 84. <http://doi.org/10.1186/s12910-015-0079-8>
- Cox, A.M., Kennan, M.A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, 68(9), 2182-2200. <http://doi.org/10.1002/asi.23781>.
- Creswell, J.W., & Poth, C.N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Thousand Oaks, CA: Sage
- European Union. (n.d.) GDPR information portal. Retrieved from <https://www.eugdpr.org>.

- Hardy, L. J., Hughes, A., Hulen, E., & Schwartz, A. L. (2016). Implementing qualitative data management plans to ensure ethical standards in multi-partner centers. *Journal of Empirical Research on Human Research Ethics*, 11(2), 191-198. <http://doi.org/10.1177/1556264616636233>.
- Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, trans.). New York: Harper & Row.
- Hern, A. (2017, June 15). University College London hit by ransomware attack. *The Guardian*. Retrieved from <http://www.theguardian.com>.
- Holland, F. (2014). Teaching in higher education: An interpretive phenomenological analysis. *Sage Research Methods Cases*. <http://doi.org/10.4135/978144627305014528646>.
- Interagency Advisory Panel on Research Ethics. (n.d.). *TCPS 2 interpretations*. Retrieved from <http://www.pre.ethics.gc.ca/eng/policy-politique/interpretations/Default/>
- Ioannidis, J. P. A. (2013). Informed consent, big data, and the oxymoron of research that is not research. *American Journal of Bioethics*, 13(4), 40-42. <http://doi.org/10.1080/15265161.2013.768864>.
- Jones, S., Pryor, G. & Whyte, A. (2013). *How to Develop Research Data Management Services - A Guide for HEIs*. Edinburgh: Digital Curation Centre. Retrieved from <http://www.dcc.ac.uk/resources/how-guides>.
- Kaye, J., de Vries, J., Heeney, C., Hawkins, N., & Boddington, P. (2009). Data sharing in genomics - re-shaping scientific practice. *Nature Reviews Genetics*, 10(5), 331-335. <http://doi.org/10.1038/nrg2573>.
- Kozlakidis, Z., Cason, R. J., Mant, C., & Cason, J. (2012). Human tissue biobanks: The balance between consent and the common good. *Research Ethics*, 8(2), 113-123. <http://doi.org/10.1177/1747016112442031>.
- Latham, B. (2017). Research data management: Defining roles, prioritizing services, and enumerating challenges. *Journal of Academic Librarianship*, 43(3), 263-265. <http://doi.org/10.1016/j.acalib.2017.04.004>.
- Leary, T. D., IV. (2017). *The lived faculty experience with formalized assessment initiatives: An interpretive phenomenological analysis* (doctoral thesis). Northeastern University, Boston, Massachusetts.
- Lunshof, J., Chadwick, R., Vorhaus, D. B., & Church, G. M. (2008). From genetic privacy to open consent. *Nature Reviews Genetics*, 9(5), 406-411. <http://doi.org/10.1038/nrg2360>.
- Mannheimer, S., Pienta, A., Kirilova, D., Elman, C., & Wutich, A. (2018). Qualitative data sharing: Data repositories and academic libraries as key partners in addressing

challenges. *American Behavioral Scientist*. Advance online publication.
<http://doi.org/10.17605/OSF.IO/7Z2BT>

Mauthner, N. S., & Parry, O. (2013). Open access digital data sharing: Principles, policies and practices. *Social Epistemology*, 27(1), 47-67. <http://doi.org/10.1080/02691728.2012.760663>.

Mauthner, N.S. (2016). Should data be regulated? In W.C. van den Hoonaard, & A. Hamilton (Eds.) *The ethics rupture: Exploring alternatives to formal research ethics review* (pp. 206-229). Toronto: University of Toronto Press.

Metcalfe, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1), 1-14.
<http://doi.org/10.1177/2053951716650211>.

Meyers, C. A., & Bagnall, R. G. (2017). The challenges of undergraduate online learning experienced by older workers in career transition. *International Journal Of Lifelong Education*, 36(4), 442-457. <http://doi.org/10.1080/02601370.2016.1276107>.

Parry, O., & Mauthner, N. S. (2004). Whose data are they anyway? Practical, legal and ethical issues in archiving qualitative research data. *Sociology*, 38(1), 139-152.
<http://doi.org/10.1177/0038038504039366>.

Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLOS One*, 9(12), e114734.
<http://doi.org/10.1371/journal.pone.0114734>.

Schrag, Z.M. (2010). *Ethical imperialism: Institutional review boards and the social sciences, 1965-2009*. Baltimore: Johns Hopkins University Press.

Smith, J.A. (1996). Beyond the divide between cognition and discourse: Using interpretative phenomenological analysis in health psychology. *Psychology & Health*, 11(2), 261-271.

Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative phenomenological analysis: Theory, method and research*. Los Angeles: SAGE.

Social Sciences and Humanities Research Ethics Special Working Committee. (2008). *Extending the spectrum: The TCPS and ethical issues in internet-based research*. Retrieved from http://www.pre.ethics.gc.ca/policy-politique/initiatives/docs/Internet_Research_-_February_2008_-_EN.pdf

Tenenberg, J. (2016). Learning through observing peers in practice. *Studies In Higher Education*, 41(4), 756-773. <http://doi.org/10.1080/03075079.2014.950954>.

Tenopir, C., Sandusky, R., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library and Information Science Research*, 36(2), 84-90. <http://doi.org/10.1016/j.lisr.2014.11.003>

U15 Group of Canadian Research Universities. (n.d.). Our Impact. Retrieved from <http://u15.ca/our-impact>.

van Baalen, S. (2018). 'Google wants to know your location': The ethical challenges of fieldwork in the digital age. *Research Ethics*, Advance online publication, 1-17. <http://doi.org/10.1177/1747016117750312>

VanScoy, A., & Evenstad, S. B. (2015). Interpretative phenomenological analysis for LIS research. *Journal of Documentation*, 71(2), 338-357. <http://doi.org/10.1108/JD-09-2013-0118>.

Verbaan, E., & Cox, A. M. (2014). Occupational sub-cultures, jurisdictional struggle and third space: Theorising professional service responses to research data management. *Journal of Academic Librarianship*, 40(3-4), 211-219. <http://doi.org/10.1016/j.acalib.2014.02.008>.

Warrell, J. G., & Jacobsen, M. (2014). Internet research ethics and the policy gap for ethical practice in online research settings. *The Canadian Journal of Higher Education*, 44(1), 22-37.

Wood, C., Farmer, M. D., & Goodall, D. (2016). Changing professional identity in the transition from practitioner to lecturer in higher education: An interpretive phenomenological analysis. *Research In Post-Compulsory Education*, 21(3), 229-245. <http://doi.org/10.1080/13596748.2016.1195173>.

Yardley, S. J., Watts, K. M., Pearson, J., & Richardson, J. C. (2014). Ethical issues in the reuse of qualitative data: Perspectives from literature, practice, and participants. *Qualitative Health Research*, 24(1), 102-113. <http://doi.org/10.1177/1049732313518373>.

Yeung, P., & Bennett, R. (2017, September 5). University secrets are stolen by cybergangs: Scientific research targeted by hackers. *The Times*, pp.1,2.

Zarate, O. A., Brody, J. G., Brown, P., Ramirez-Andreotta, M. D., Perovich, L., & Matz, J. (2016). Balancing benefits and risks of immortal data: Participants' views of open consent in the Personal Genome Project. *Hastings Center Report*, 46(1), 36-45. <http://doi.org/10.1002/hast.523>.