Data Ethics, Privacy, and Governance Workshop

A summary report

Organised by Te Rourou Tātaritanga and Social Wellbeing Agency

April 4, 2022

1 Background

The online workshop was organised and hosted by Associate Professor Barry Milne (University of Auckland), Professor Colin Simpson (Victoria University of Wellington), and Associate Professor Andrew Sporle (University of Auckland and iNZight Analytics Ltd), in partnership with the Social Wellbeing Agency. Participants were from government agencies, industry and academia.

This report summarising the discussion is the primary output of the workshop, although the workshop will also help inform the themes of a symposium scheduled for early 2023.

Initial scene-setting by Milne, Sporle and Simpson encompassed several issues, designed to provide context for discussions:

- the existing uneven governance and ethical review of population data research;

- the absence of fit-for-purpose ethical review processes;

- Māori data sovereignty; and

- a governance and process model from Scotland (eDRIS).

Workshop participants, in small groups, addressed two main topics, followed by a whole group discussion:

1. Ethics and governance; what are the key issues?

2. What does an ideal system look like and what are the steps required to make it happen?

The Chatham House rule applied to all discussions in the workshop, with no attribution of comments to individuals or organisations. The purpose of this was to increase the openness of discussion.

2 Workshop discussion points

This report does not attempt to reflect every perspective raised in the workshop discussions, rather, it captures common themes. Similarly, not all participants unanimously agreed with the themes that follow.

Key themes in the order discussed during the workshop:

1. Key issues

i. Currently, there is wide variability in how, or even if, governance, ethics, privacy and data sovereignty (referred to in this report from here on as "data ethics") are handled in relation to the collection, use (including secondary analysis) and sharing of population data. This creates uncertainty around when data ethics should be considered, and what, if any, of the existing processes (like institutional or Health and Disability Ethics Committees) are appropriate. Data users may have relatively little experience with existing processes. There is generally an absence of useful recorded precedents on which users can base decisions about how to proceed.

ii. With existing data, there is variability in people's understanding of consent for current, alternate or future use, and any consent can be difficult to confirm.

iii. There is a broad spectrum of activity associated with population data, from the highly operational (e.g. providing summary data about situation x affecting subpopulation y), to that which is clearly recognised as research (e.g. individual-level regression analysis).

iv. There is a range of users of population data, from academic researchers to Ministry and agency officials, to community groups and private industry. Across the range of users, there is variability in experience and understanding of the governance, ethics, privacy and data sovereignty issues in relation to the use of that data. This variation in understanding results in variation in practice/implementation.

v. Challenges in facilitating broader access to and use of population data, for example by communities, include the suboptimal IT infrastructure, uncertainties around permissions to collect, use and share data, and the lack of analytic capability.

vi. Under the new health reforms, it will be important to monitor ethical standards regarding the collection, storage and use of data in genomic medicine. Ethical standards will involve ensuring the accurate data are collected, and this is particularly important for major diseases that are prevalent in the Pacific population, both here and in the Pacific region.

vii. Currently, there is limited statutory oversight around data collection and use, except for some specific circumstances. An ethical framework may be more flexible and adaptable than a legislative approach.

2. What does an ideal system look like?

vii. The workshop generally acknowledged the value of developing, instituting and sustaining a coherent and comprehensive framework covering data ethics.

viii. Some suggested that there should be a statutory foundation for the structure and operation of any data ethics framework (this is in contradiction to point vi).

ix. A good framework will support sound decision-making, and even allow "challenging" uses of data to be considered, where at present that is often avoided because of uncertainty about risk.

x. Data has a life cycle, and so any framework needs to deal with not just access but also secure storage and destruction of sensitive data.

xi. Currently, processes around access to and use of IDI data provide an example of how part of the framework might operate, but this is not fit for the broad range of work and users to which the framework will be applicable.

xii. The framework needs to identify:

a. what data ethics issues are relevant,

b. in what circumstances processes to manage these data ethics issues are required,

c. processes that facilitate decision-making on the appropriateness of management of the data ethics issues, and

d. a sustainable structure that mandates the work of groups involved in decision-making, oversight and advisory activities.

xiii. A framework to manage processes around data ethics must recognise that there will initially be significant variation in expertise and familiarity with the issues amongst users, even of the need for processes. Education will be an important component of any system, both for practitioners but also for "customers". The workshop did not further develop the forms of education that might be required, but this would be important future work.

xiv. There is a mixed understanding of the implications of, and distinctions between, data sovereignty, data ownership, and custodianship/stewardship.

xv. The implications of Māori data sovereignty, its oversight, what it means practically, and how it should be handled, prompted several perspectives.

a. It is variably, but generally poorly, managed currently.

b. It was suggested that consideration of issues around data sovereignty for Māori in any proposed population data collection and use should be integrated into one framework alongside the issues of governance, ethics and privacy.

c. It was noted that it was important to ask Māori about how data sovereignty should be considered in any potential data collection and use.

xvi. There may be other groups, as well as Māori, that have specific expectations around data sovereignty, and this would have to be determined as any framework was developed.

xvii. Any framework needs to be deeply embedded, such that its use is normalised, and not an "extra step" or a compliance hurdle. It was noted that a requirement for ethical review by institutions, funders and publishers (for a clearly defined subset of research activity) has normalised this activity, but this same requirement may be more difficult to institute for some settings of work involving population data.

xviii. While the framework should provide coherence and harmonisation of processes taking account of data ethics, it is important to recognise that a "one size fits all" approach is unlikely to be appropriate. Different uses of population data, and different users, will have different risk profiles; consideration of some uses will need to be quite detailed, whereas consideration of other uses could be quite a light touch.

xix. Transparency of decision-making will facilitate acceptance and use of the framework.

xx. There were differing perspectives on how "government" should be involved; while some saw government as a key partner, others suggested some independence was important. For example, basing management of any framework within a Ministry runs a risk that the Ministry's other work priorities may supersede work around running the framework. This incomplete consideration in the workshop of the role of government applies to all three phases of the work identified in xxi-xxiv below—developing, instituting and sustaining—and deserves fuller discussion.

xxi. Looking ahead, there are four phases of work, scoping of which might form themes for the planned symposium later in 2022:

(a) scoping existing frameworks[1]

(b) adopt and expand the framework (or develop a new one)

(c) instituting the framework

(d) sustaining the framework

Simply expressed there are three groups of people who should be involved in the three phases of work; organisations and people managing the data, people using the data, and the people the data is about.

It will be worth looking at data ethics systems from other jurisdictions, and also at local systems, such as the Health and Disability Ethics Committees, to understand what works best.

xxii. *Developing the framework*: The workshop supported wide representation during this work, by including those whose data is held, or from whom it may be collected, with an emphasis on the diversity of participants (e.g. ethnicity, culture, gender, age, religion, disability and sexual orientation). Users and stewards of data should also be included. Who should lead the work was not considered in depth. An important component will be setting the boundaries on what data, and the types of use, sharing or analysis are covered by a framework. Multiple attendees agreed that any system might need to allow

---

[1]This assumes the creation of a new framework. There is a need for an earlier step to review and assess the existing frameworks. It will likely be more effective to expand the use of a framework that already has traction than to introduce another (competing) framework.

for a "scientific" review (for some types of work) to confirm the validity and value of the proposed work, this potentially being separate from consideration of issues of ethics, privacy and data sovereignty.

xxiii. *Instituting the framework*: The workshop emphasised the importance of an educational role while any framework is instituted, given the differing levels of understanding and expertise amongst the public, Ministry and agency officials, researchers, and groups requesting data for analysis. As well as clear guidance and easy to understand processes, the framework might be accompanied by a toolkit so users can decide when it is relevant, what must be considered, and what information is needed to support decision making on the appropriateness of provisions made around data ethics. The framework must recognise different realities of working with population data, and processes and decision-making timelines should be designed and implemented with this in mind. For example, processes should accommodate expedited decision-making on the appropriateness of arrangements around data use that fits within clearly defined low-risk parameters. Equally, geography should not be a barrier to access (compared to limited sites to access IDI currently).

xxiv. *Sustaining the framework*: The workshop recognised the need for resourcing if the framework was to be sustained. This includes the management and operation of the system, but should also facilitate capability development amongst users, those administering processes and those involved in decision-making. Some considered that being at some distance from central government was important so that the framework could be seen to be operating with independence. A sound data ethics framework, well implemented and managed, will generate evidence of the value of using population data, supporting the case for sustaining the resourcing of the system.

xxv. Part of the work, especially, but not only, during the development phase will be to build social and cultural licenses for data use. There are clearly quite disparate views currently in the public domain and it will require a quite deliberate workplan to achieve general acceptability. Part of this might include the question of whether commercial groups could access and use data, or what types of use they might be allowed.

xxvi. There is currently work within government (for example, the Government Data Strategy and Roadmap (www.data.govt.nz/leadership/strategy-and-roadmap)) that is likely relevant to any future data ethics framework. If a data ethics framework is to be developed it cannot be in a vacuum and must recognise existing and planned government work.

Acknowledgements