THE GENERATION OF KNOWLEDGE · TRUST, QUALITY AND THE FUTURE

Ethical Governance of Knowledge Generation in Survey, Research, and Intelligence Contexts

How the ethics of generating knowledge about people is governed in practice across survey and market research, academic and public research, and intelligence work, covering ethics review, data protection, professional codes, oversight of intelligence, the limits of anonymisation, Indigenous data governance and the new challenges of big data and artificial intelligence.

ModQu research article · 2,749 words · 27 references (APA 7th edition)

Ethical principles tell us what researchers owe the people they study: respect, protection from harm and fair treatment. But principles do not apply themselves. They must be built into laws, institutions, codes, procedures and habits, so that they shape what actually happens when data are collected, analysed and used. This is the work of **ethical governance**. It differs across contexts. A market research company, a university research team and an intelligence agency all generate knowledge about people, but they operate under different laws, answer to different overseers and face different temptations. This article examines how ethical governance works in each of these contexts, what has been learned about its strengths and weaknesses, and the challenges that new forms of data and analysis pose for all of them.

From principles to governance

Modern research ethics grew out of responses to abuse. The Belmont Report in the United States set out three principles, respect for persons, beneficence and justice, and translated them into requirements for informed consent, the assessment of risks and benefits, and the fair selection of participants (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). In Canada, the Tri-Council Policy Statement adopts the closely related principles of respect for persons, concern for welfare and justice, and sets out how they apply to research funded by the federal research agencies (Canadian Institutes of Health Research et al., 2022). Emanuel et al. (2000) argued that ethical research must also have social value and scientific validity, be reviewed independently and respect participants throughout the study, not only at the point of consent.

Governance turns such principles into enforceable practice through several layers: laws that set minimum standards and penalties; external bodies that oversee compliance; professional codes that define expected conduct; internal controls within organisations; and openness that allows the public to see what is done. The balance between these layers differs across contexts.

How knowledge generation is governed in three contexts

Layers of governance, from law to everyday practice, as they apply to surveys, research and intelligence

	Surveys and market research	Academic and public research	Intelligence
Law	Privacy and data protection law	Privacy law; research and funding rules	Statutes setting powers and limits
External oversight	Privacy commissioners and regulators	Funding agencies; research ethics boards	Independent review bodies; courts
Professional codes	Industry codes of conduct	Ethics policies; disciplinary codes	Codes of practice; professional ethics
Internal controls	Quality systems; data security; audits	Ethics review; data management plans	Legal advice; compliance checks; audit trails
Openness	Methods disclosure with published results	Publication; open data where possible	Public reports by review bodies

Based on CIHR, NSERC and SSHRC (2022), ICC/ESOMAR (2016) and Born and Leigh (2005).

Figure 1. Layers of ethical governance in three contexts of knowledge generation. Based on CIHR, NSERC and SSHRC (2022), ICC/ESOMAR (2016) and Born and Leigh (2005); diagram by ModQu.

Governance of academic and public research

Ethics review

In academic and publicly funded research, the central mechanism of governance is **prospective ethics review**. Before research involving people begins, a committee, called a research ethics board in Canada and an institutional review board in the United States, examines the proposal and decides whether its risks are justified, whether consent procedures are adequate and whether participants' privacy is protected.

Ethics review has clear benefits. It forces researchers to think through the ethical aspects of their work before it starts, and it provides an independent check on judgements that researchers might make in their own favour. But it has also attracted criticism. Abbott and Grady (2011) systematically reviewed empirical studies of how review boards work and found wide variation in their decisions about the same proposals, considerable delays, and little evidence about whether review actually improves the protection of participants. Haggerty (2004), a Canadian sociologist, described a process of **ethics creep**: the steady expansion of review into areas of research that pose little risk, accompanied by increasingly bureaucratic requirements, which he argued could hinder valuable social research without making participants safer. Schrag (2010) traced the history of ethics review in the United States and argued that rules designed for biomedical experiments had been applied, often awkwardly, to interviews, surveys and observation in the social sciences.

These criticisms do not suggest abandoning ethics review, but they show that governance must be proportionate. The Tri-Council Policy Statement adopts a principle of **proportionate review**: research that poses minimal risk can receive a lighter review than research that poses greater risk (Canadian Institutes of Health Research et al., 2022).

Indigenous data governance

Governance also concerns who controls knowledge about communities. In Canada, First Nations have articulated the principles of **ownership, control, access and possession** (OCAP), which assert that First Nations communities own information about themselves collectively and have the right to control how it is collected, used and shared (First Nations Information Governance Centre, 2014). Internationally, the CARE principles for Indigenous data governance call for data practices that serve collective benefit, recognise Indigenous authority to control data, embody responsibility, and centre Indigenous ethics (Carroll et al., 2020). These principles complement the FAIR principles for data sharing, by asking not only whether data can be found and reused but who benefits and who decides.

Governance of survey and market research

Commercial survey and market research is generally not subject to formal ethics review in the way academic research is. Its governance relies more heavily on **law**, particularly privacy and data protection law, and on **self-regulation** through professional codes.

Data protection law

Privacy and data protection laws set rules for the collection, use, disclosure and retention of personal information. They generally require that organisations have a lawful basis or consent for collecting personal data, collect no more than they need, use it only for the purposes for which it was collected, keep it secure and allow individuals to see and correct information about them. The European Union's General Data Protection Regulation, in force since 2018, is the most influential example, and it applies to organisations anywhere that process the personal data of people in the European Union (Voigt & von dem Bussche, 2017). In Canada, federal privacy law applies to the handling of personal information in commercial activities, and is enforced by the Office of the Privacy Commissioner of Canada.

The Office of the Privacy Commissioner of Canada and the Information and Privacy Commissioner for British Columbia jointly investigated Facebook after it emerged that data from millions of users had been obtained through a personality quiz application and used for political targeting. They found that Facebook had failed to obtain meaningful consent from users and their friends and had not adequately safeguarded their information (Office of the Privacy Commissioner of Canada, 2019). The case showed how data collected under the appearance of research could be diverted to other uses when governance was weak.

Professional codes

The research industry has long maintained its own codes of conduct. The ICC/ESOMAR International Code, adopted by national associations around the world, requires researchers to be honest and transparent with participants about the purpose of research, to respect their right to refuse or withdraw, to protect their personal data, to take particular care with children and vulnerable people, and, crucially, to keep research clearly distinct from commercial activities such as

advertising and direct marketing (ICC/ESOMAR, 2016). The code of the American Association for Public Opinion Research adds requirements for the disclosure of methods whenever results are published, so that the public can judge the quality of the findings (American Association for Public Opinion Research, 2021).

Self-regulation has strengths. Codes are written by people who understand the work and can be updated quickly as practices change. But it also has limits: codes bind only members of the associations that adopt them, and enforcement depends on complaints and on the willingness of associations to discipline their members.

Governance of intelligence

Intelligence work presents the hardest governance problem. It often requires secrecy, may involve collecting information without the knowledge or consent of those concerned, and can affect liberty and security directly. The ordinary tools of research governance, such as informed consent and publication, cannot apply in the usual way.

Bellaby (2012) argued that the ethics of intelligence collection should be judged by the harm it causes to people's vital interests, especially privacy and autonomy. More intrusive methods cause greater harm and therefore require stronger justification, a just cause, proper authority, proportionality and the exhaustion of less harmful alternatives. Omand and Phythian (2018) set out similar principles, drawing on the tradition of just war thinking, and argued that intelligence agencies can be both effective and ethical only if their powers are clearly defined in law and subject to independent oversight.

Born and Leigh (2005) reviewed legal standards and best practice for intelligence oversight in democracies. They identified several layers of control: a clear legal mandate defining what agencies may do; internal management and legal advice; control by ministers who are accountable to parliament; review by parliamentary committees; independent bodies such as inspectors general or review agencies; and judicial authorisation for intrusive methods. Gill and Phythian (2018) emphasised that oversight must have access to secret information, independence from the agencies it reviews and the ability to report publicly, at least in general terms, if it is to provide real accountability. Canada strengthened its own arrangements in this period, creating a committee of parliamentarians in 2017 and an independent review agency in 2019 with access to classified information across government.

Pressure from those who commission knowledge

A form of ethical risk shared by all three contexts comes from the people who commission and use knowledge. Clients of market research may hope for results that support a product or a campaign. Governments may prefer research that supports their policies. Political leaders may press intelligence agencies for assessments that fit decisions already made. In each case, the pressure is rarely explicit. It works through the choice of questions, the framing of reports, the selection of which findings to emphasise and which to leave out.

Governance addresses this pressure in several ways. Professional codes require researchers to report findings honestly and not to misrepresent them, and to make clear when results are released who commissioned the work (ICC/ESOMAR, 2016; American Association for Public Opinion Research, 2021). In intelligence, the separation of analysis from policy, and the independence of analysts to report unwelcome judgements, are treated as basic safeguards (Gill & Phythian,

2018). Within organisations, clear written agreements about who owns the data, who may publish results and how findings will be reported, settled before a project begins, protect researchers from later pressure to change what they found.

The limits of anonymisation

All three contexts rely heavily on **anonymisation**, the removal of information that identifies individuals, to protect privacy while allowing data to be used and shared. Research has shown that anonymisation is much harder than it seems. Sweeney (2000) found that a large majority of the United States population could be uniquely identified by the combination of their postal code, gender and full date of birth, details often left in "anonymised" datasets. Narayanan and Shmatikov (2008) showed that individuals in a supposedly anonymous dataset of film ratings released by Netflix could be identified by comparing a few of their ratings with public reviews posted elsewhere. Rocher et al. (2019) estimated that almost all Americans could be correctly re-identified in any dataset using fifteen demographic attributes.

Ohm (2010) concluded that the law's reliance on anonymisation rested on a false promise, and that rules should focus instead on limiting the risks of harm from data use. One technical response is **differential privacy**, a mathematical framework that adds carefully calibrated random noise to statistics so that the presence or absence of any individual in the data cannot be inferred with confidence (Dwork & Roth, 2014). The United States Census Bureau adopted differential privacy for the 2020 census to protect respondents' confidentiality (Abowd, 2018). Its adoption also showed that privacy protection involves trade-offs: more protection means less precise statistics, particularly for small areas and groups, and decisions about that balance are themselves matters of governance.

New challenges: big data and artificial intelligence

Traditional governance assumed that researchers collect data directly from people who know they are taking part. Much modern data is different. It comes from digital traces, such as social media posts, mobile phone records and online transactions, collected for other purposes and repurposed for analysis. Metcalf and Crawford (2016) described an emerging **ethics divide**: research using such data often falls outside the scope of ethics review, because it is not considered to involve human subjects, even though it can affect the people whose data are used. Zook et al. (2017) proposed ten rules for responsible big data research, including acknowledging that data are people and can do harm, recognising that privacy is more than a binary value, guarding against re-identification and debating the tough choices openly.

Artificial intelligence raises further questions. Jobin et al. (2019) reviewed 84 sets of ethical guidelines for artificial intelligence published by governments, companies and other organisations, and found broad agreement on a small number of principles, including transparency, justice and fairness, avoiding harm, responsibility and privacy, but wide differences in how these principles were interpreted and applied. Mittelstadt (2019) argued that principles alone cannot guarantee ethical practice. Unlike medicine, he noted, the field of artificial intelligence lacks common aims, professional history, established methods for translating principles into practice and robust mechanisms of accountability. Floridi and Taddeo (2016) proposed that **data ethics** be developed as a field in its own right, concerned with the ethics of data, of algorithms and of the practices that surround them.

Toward effective governance

Several lessons emerge across the three contexts.

Governance must be layered. No single mechanism is sufficient. Law sets minimum standards; oversight bodies check compliance; codes guide professionals; internal controls embed ethics in daily work; openness allows public scrutiny.

Governance must be proportionate. Rules designed for high-risk research can impose unnecessary burdens on low-risk work, while rules designed for low-risk work may be inadequate for intrusive methods. The intensity of governance should match the potential for harm.

Governance must follow the data. Ethical obligations do not end when data are collected. They apply to storage, sharing, reuse, linking and deletion, and to the uses that others make of data and findings.

Governance must include those affected. Communities whose data are collected, especially those that have historically been studied without benefit or consent, should have a voice in how knowledge about them is governed.

Governance must be accountable. Oversight bodies need independence, access to information and the ability to report publicly. Organisations should be able to show, not merely assert, that they follow their own rules.

Implications for research organisations

For an organisation that conducts surveys and research, ethical governance is a practical matter. It should comply fully with privacy and data protection law in every jurisdiction where it collects data. It should adopt and follow recognised professional codes. It should have internal procedures that consider ethical risks before each project begins, proportionate to the sensitivity of the topic and the vulnerability of participants. It should protect data through security, limited access and careful anonymisation, recognising the limits of anonymisation. It should keep research separate from selling and never use research as a pretext for marketing. It should respect the rights of communities to govern data about themselves. And it should be open about its practices, so that clients, participants and the public can see how knowledge is generated and judge whether it deserves their trust.

A short set of questions, asked at the start of every project, helps to put this governance into practice. What personal information will be collected, and is all of it necessary? How will participants be told about the purpose of the study, and how will their consent be recorded? Are any participants children or otherwise vulnerable, and what extra protections do they need? Could any questions cause distress, and how will that be handled? Where will data be stored, who will have access, and when will identifying information be deleted? Could the published results, alone or combined with other data, identify individuals or small groups? Does any community have a claim to govern data about itself? And who commissioned the work, and has it been agreed in writing how the findings will be reported? Recording the answers creates a simple ethics file for each project, which can be reviewed by a colleague and shown to a client or regulator if needed.

Conclusion

Ethical governance turns principles of respect, welfare and justice into practice. In academic research, it relies on ethics review, which provides valuable independent scrutiny but must be proportionate to risk. In survey and market research, it relies on data protection law and professional codes that require honesty with participants and the separation of research from selling. In intelligence, where secrecy limits the ordinary tools of governance, it relies on clear legal mandates and independent oversight with access to secret information. Across all three, the limits of anonymisation, the growth of repurposed data and the spread of artificial intelligence pose new challenges that principles alone cannot meet. Effective governance is layered, proportionate, attentive to the whole life of data, inclusive of those affected and accountable. Organisations that generate knowledge about people earn the right to do so by showing that such governance is real.

References

Abbott, L., & Grady, C. (2011). A systematic review of the empirical literature evaluating IRBs: What we know and what we still need to learn. *Journal of Empirical Research on Human Research Ethics, 6*(1), 3–19. https://doi.org/10.1525/jer.2011.6.1.3

Abowd, J. M. (2018). The U.S. Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (p. 2867). Association for Computing Machinery. https://doi.org/10.1145/3219819.3226070

American Association for Public Opinion Research. (2021). *AAPOR code of professional ethics and practices*. AAPOR.

Bellaby, R. (2012). What's the harm? The ethics of intelligence collection. *Intelligence and National Security, 27*(1), 93–117. https://doi.org/10.1080/02684527.2012.621600

Born, H., & Leigh, I. (2005). *Making intelligence accountable: Legal standards and best practice for oversight of intelligence agencies*. Publishing House of the Parliament of Norway.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2022). *Tri-Council policy statement: Ethical conduct for research involving humans – TCPS 2 (2022)*. Government of Canada.

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for Indigenous data governance. *Data Science Journal, 19*, Article 43. https://doi.org/10.5334/dsj-2020-043

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science, 9*(3–4), 211–407. https://doi.org/10.1561/0400000042

Emanuel, E. J., Wendler, D., & Grady, C. (2000). What makes clinical research ethical? *JAMA, 283*(20), 2701–2711. https://doi.org/10.1001/jama.283.20.2701

First Nations Information Governance Centre. (2014). *Ownership, control, access and possession (OCAP™): The path to First Nations information governance*. First Nations Information Governance Centre.

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A, 374*(2083), 20160360. https://doi.org/10.1098/rsta.2016.0360

Gill, P., & Phythian, M. (2018). *Intelligence in an insecure world* (3rd ed.). Polity Press.

Haggerty, K. D. (2004). Ethics creep: Governing social science research in the name of ethics. *Qualitative Sociology, 27*(4), 391–414. https://doi.org/10.1023/B:QUAS.0000049239.15922.a3

ICC/ESOMAR. (2016). *ICC/ESOMAR international code on market, opinion and social research and data analytics*. ESOMAR.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

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

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy* (pp. 111–125). IEEE. https://doi.org/10.1109/SP.2008.33

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health, Education, and Welfare.

Office of the Privacy Commissioner of Canada. (2019). *Joint investigation of Facebook, Inc. by the Privacy Commissioner of Canada and the Information and Privacy Commissioner for British Columbia* (PIPEDA Report of Findings No. 2019-002). Office of the Privacy Commissioner of Canada.

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review, 57*, 1701–1777.

Omand, D., & Phythian, M. (2018). *Principled spying: The ethics of secret intelligence*. Oxford University Press.

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications, 10*, Article 3069. https://doi.org/10.1038/s41467-019-10933-3

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

Sweeney, L. (2000). *Simple demographics often identify people uniquely* (Data Privacy Working Paper 3). Carnegie Mellon University.

Voigt, P., & von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR): A practical guide*. Springer. https://doi.org/10.1007/978-3-319-57959-7

Zook, M., Barocas, S., boyd, d., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B. A., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLoS Computational Biology*, *13*(3), e1005399. https://doi.org/10.1371/journal.pcbi.1005399