

Catalogue no. 11-633-X — No. 051
ISSN 2371-3429
ISBN 978-0-660-72776-9

Analytical Studies: Methods and References

The Foundations of Ethical Reviews at Statistics Canada

by Guillaume R.-Maranda

Release date: December 20, 2024



How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "[Standards of service to the public](#)."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2024

Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An [HTML version](#) is also available.

Cette publication est aussi disponible en français.

Table of contents

Introduction	4
Ethical Reviews	4
Six Guiding Principles	7
Benefits for Canadians	7
Fairness & Do no Harm	9
Quality	10
Transparency & Accountability	10
Trust & Sustainability	11
Privacy & Security	11
Conclusion	12
Bibliography	13

Introduction

Data ethics is a branch of ethics that raises questions about the appropriate use of data across its life cycle¹ and identifies permissible practices and actions. This discipline is operationalized by Statistics Canada's Data Ethics Secretariat (DES) through ethical reviews. The ethical review process is a direct consequence of the adoption of the Necessity and Proportionality Framework.² The aim of this paper is to describe the foundations and the purpose of such reviews. This can help Canadians understand the work of the DES and how Statistics Canada justifies its data acquisitions.

For many years now, there has been a steady expansion both in terms of the nature of the data that is acquired, and the methods used to acquire and analyze the data at Statistics Canada (Arora 2018). Such an expansion can be attributed to four main causes: 1- a need for more granular, complete, and timely data, 2- declining response rates of traditional surveys methods, 3- the prohibitive costs of surveys, and 4- technological developments and expertise. All causes combined accelerated the acquisition of non-survey data and, in some cases, the introduction of non-probabilistic survey methods, which is a significant shift in paradigm³ for a National Statistical Organization (NSO). This change of paradigm not only raises methodological questions but is also raises several ethical challenges that need to be addressed formally through ethical reviews. This is why ethical reviews have become a necessity.

Ethical reviews are meant to be complementary to Statistics Canada's internal processes, policies and directives. Given that Statistics Canada is a public organization, it is important that its actions not only fall within its legal mandate,⁴ but also that it can provide a coherent rationale to explain why it is undertaking any project involving data at various stages of its life cycle. This is where ethical reviews come into play. Ethical reviews aim to evaluate the ethical justification for the acquisition and proper use of the data. This is a necessary condition to ensure that Statistics Canada can communicate why it needs the data it is acquiring and how it will ensure its proper use. This can help establish a wider social acceptance and help maintain public trust for the use of these data that goes beyond the traditional use of focus groups.

This paper contains two main sections. In the first, ethical reviews are defined with three steps that ultimately rely on six guiding principles. In the second, these guiding principles are expanded upon. It is shown that the principles are not independent from one another, which means that a review is more than checking a box next to each of them.

Ethical Reviews

An ethical review can be defined with three main operational steps:

1. Obtaining the relevant information about a given project.
2. Identifying the strengths and weaknesses of the project.
3. Determining if a weakness outweighs the strengths of the project and making recommendations accordingly.

1. This includes the acquisition, retention and analysis of the data, all the way to the dissemination of statistics: gather, guard, grow, and give (Rancourt 2019).

2. [Principles of Necessity and Proportionality \(statcan.gc.ca\)](#).

3. The expressions "shift in paradigm" is used to mean that the standard set of problems and solutions used by a group of scientists are being replaced by a new one. It is very close (if not identical) to some interpretations of the Kuhnian notions of "shift in paradigm" (Bird 2018).

4. [S-19.pdf \(justice.gc.ca\)](#).

In this section, each step and the conditions that must be met to complete them are clarified.

The first step is completed by gathering information according to six guiding principles that are described in greater detail in the next section:

- Benefits for Canadians,⁵
- Fairness & Do no harm,
- Quality,
- Transparency & Accountability,
- Privacy & Security,
- Trust & Sustainability.

The information collected with respect to each of these principles will vary depending on whether the data has been acquired through a survey or not because the ethical challenges will be different. For example, when comparing the two, the acquisition of non-survey data is often less transparent, and the data can be more privacy intrusive because it can contain more personal information than we need. Moreover, the data itself has usually not been acquired for statistical purposes in the first place and this can bring specific quality challenges that can impact the appropriate use of such data.

This is one of the reasons why ethical reviews at Statistics Canada are usually performed at the acquisition stage⁶ through various processes. The fact that they are done at the acquisition stage also allows for ethical reviews to bring every aspect of the data life cycle within its scope. It should be noted however that some ethical aspects of a given project (especially those that are more salient during the later stages of a project) are also covered by other types of reviews conducted by different groups within the organization (see the guidelines on responsible machine learning⁷ for example).

The successful completion of this first step must meet at least three conditions:

- i. Necessary and sufficient information must be collected. Forms and guiding principles can always be revised to make sure that the necessary and sufficient information is collected. This approach could lead towards a form of reflexive equilibrium (Daniels 2020) where even the foundations of ethical reviews not only guide the reviewers, but where the growing experience of the reviewers can help improve those very foundations (i.e., a two-way/reflexive process) until a lasting consensus is reached (i.e., an equilibrium).⁸
- ii. The information collected must be true. It can be important to consult with different members of the team working on a given project to obtain a complete picture of the project.
- iii. The information must not merely represent the opinion of a single individual but of the team involved with the project. This is especially important when discussing the purpose and the expected benefits of a project. Consulting with different members of the team working on a given project can provide a global view which sometimes reveal inconsistencies that should be resolved.

5. Canadians is understood very broadly throughout this document so as to include everyone living in Canada.

6. Throughout this paper, the 'acquisition stage' refers to the entire process of acquiring data. Ethical reviews are done before the actual acquisition of the data and their scope includes every part of the data life cycle.

7. [Framework for Responsible Machine Learning Processes at Statistics Canada July 2020 \(statcan.gc.ca\)](https://www150.statcan.gc.ca/n1/pub/95-662-x/2020001/article/00001-eng.htm).

8. For example, experience has taught us, at the Data Ethics Secretariat, that the description of the expected benefits of an acquisition can be too broad and generic. This is one of the reasons why our current documents now emphasize the importance of adding concrete examples. This, in turn, changed the way justifications have been written in many cases. Keep in mind that a reflexive equilibrium does not imply that we can systematically change our principles in order to justify any course of action. Doing so would not change the fact that we could recognize this incoherence and therefore the lack of equilibrium.

The second step of an ethical review consists in identifying the strengths and weaknesses of a given project against the six guiding principles. For example, the data quality assessment could show that the data is representative of the target population and of a good quality overall. However, the collection method could be intrusive because the data is personal and that the consent of the individuals has not been obtained. The adequate completion of this step must meet at least two conditions:

- i. Reviewers must have the appropriate expertise to identify strengths and weaknesses in function of the six guiding principles. Therefore, it is important to consult relevant experts and committees with members of different fields within the organization. This is why some reviews are complemented by consulting the Data Ethics Committee (internal) and/or the Advisory Council on Ethics and Modernization of Microdata Access (ACEMMA), which is external.⁹
- ii. The reviewers must evaluate the quality of the documented justification and not assume that the project is justified by default. This means that the documentation should be self-sufficient to assess the ethical merits of a project. This also means that if the documentation is lacking, a reviewer should not make unverified assumptions about the strengths of a given project. If the documentation is incoherent, one should not simply interpret the documentation in order to make sense of it. Failing to do so can lead to a biased review.

The third and final step of an ethical review consists of making sure that any weaknesses do not outweigh the strengths of the project and that recommendations are made accordingly. For example, if the quality of the data is subpar because it is not representative of the target population, thus impacting the expected benefits, then the project should perhaps be exploratory and minimally intrusive¹⁰ into the private lives of individuals.

The level of coherence sought in the documentation will likely manifest itself in a unique way depending on each case. This is what makes an ethical review a non-trivial task. For instance, each case is likely to carry its own unique expected benefits, level of privacy intrusion and data quality that will have to be jointly acceptable, given the current social context, from an ethical point of view. This implies that the ethical review must be sensitive to the particularities of a given case. This is exactly what it means to say that data ethics is contextual. It does not imply ethical relativism in any way. Our understanding of the six guiding principles does not change in function of a given case, but their coherent implementation will.

The adequate completion of this step must meet the following condition:

- i. The weights that the reviewers put on the strengths and weaknesses should not be arbitrary and dependent on the individuals reviewing a project. The consensus of experts (i.e., the reflexive equilibrium) should alleviate this concern.

There are four possible recommendation outcomes at the completion of this step:

- a. No modifications are necessary.
- b. Changes to the documentation are necessary to better justify the project.
- c. Changes to the implementation of the project are in order.
- d. The project should not proceed any further.

Recommendations should be actionable items with tangible actions. Comments such as “Please keep in mind that...” or “Please consider...” should be avoided. The purpose of ethical reviews is greater than simply raise awareness. They ensure that our data acquisitions and usage throughout the data life cycle are ethical, justified, and well-documented.

9. [Advisory Council on Ethics and Modernization of Microdata Access \(statcan.gc.ca\)](https://www150.statcan.gc.ca/n1/pub/95-631-x/2016001/article/00001-eng.htm).

10. See the section on Privacy and Security.

Six Guiding Principles

The approach proposed in the previous section does not imply the application of a high-level moral theory.¹¹ Such theories are complex and not universally accepted. These attributes make them less suitable for a context where recommendations on a given case need to be reached based on the discussion of a group that does not necessarily include academically trained philosophers. What has been presented so far is much more akin to a mid-level theory (Flynn 2020) where six ethical principles are being introduced as part of a framework to evaluate a given case.

The choice of those principles has not only been informed by the literature on the topic but with the growing experience in Data Ethics within Statistics Canada. The guidance they provide should be informative enough for reviewers to focus on the appropriate questions and flexible enough for them to be able to reach a consensus on a given case. This feature is coherent with the practical experience of philosophers, like Stephen Toulmin, that seems to indicate that “there can be agreement on how to resolve a practical issue amongst those who strongly disagree on which moral theory ought to prevail” (Flynn 2020).

This section presents each principle under which both survey or non-survey data acquisition¹² are reviewed. Details are provided about the kind of information that is being collected during the first step of an ethical review. The reason why each principle has been added to the list is also explained. It should become apparent that the increasing use of non-survey data and non-probabilistic survey methods makes each of these principles more relevant than ever.

Benefits for Canadians

The increasing need for more granular and more timely information, combined with increasing response burden, declining response rates, and the cost of conducting surveys has caused a shift towards an increasing reliance on other modes of data acquisition (e.g., through external data providers or web scraping). In fact, the importance of non-survey data sources has been underscored by the former Chief Statistician Anil Arora:

“One of the most significant changes has been incorporating more alternative data into the production of official statistics. The volume and variety of new data sources are increasing exponentially across the world. National Statistical Offices are well aware of the value these sources offer. As NSOs, we can no longer rely solely on data we collect ourselves. Many organizations have already taken steps to integrate data from administrative and alternative sources into their programs and products.” (Arora 2022)

This shift has definite advantages but it comes with new challenges, one of them being that we can no longer communicate and document the expected benefits of our activities to Canadians in the same way. The goals and expected benefits of a survey are usually conveyed through a communication campaign, Statistics Canada’s Trust Centre,¹³ and they are clearly written on the questionnaires or explained by an interviewer. For example, the following is an extract of what was written on the paper questionnaire of the 2021 Census:

“Thank you for taking a few minutes to participate in the 2021 Census. The information you provide is converted into statistics used by communities, businesses, and governments to plan services and make informed decisions about employment, education, health care, market development and more.”

This type of communication strategy is not available when data is not acquired through a survey. Therefore, it is important to find alternative ways to explain why the data Statistics Canada wishes to acquire is needed. One method is to document the expected benefits of an acquisition and to take proactive steps to communicate them efficiently and in a timely manner. More details will be provided in the “Transparency and Accountability” sub-section below.

11. Example of high-level moral theory include deontology, virtue theory and consequentialism. To learn more about moral theory, please consult the Stanford Encyclopedia of Philosophy: [Stanford Encyclopedia of Philosophy](#).

12. The method of acquisition is determined from Statistics Canada’s perspective. For instance, survey data means data collected by Statistics Canada. Data linkage is considered a ‘non-survey data acquisition’ because the linked data set is a new entity.

13. [Statistics Canada’s Trust Centre \(statcan.gc.ca\)](#).

Benefits for Canadians is thus an important guiding principle that is taken into account when conducting an ethical review. It can be important to document the expected benefits for the environment, businesses or institutions, but ultimately there should be a benefit for Canadians even when the data is about the environment, companies or institutions. The adequate description of the benefits includes:

- i. a non-technical, self-sufficient, explanation of the need for the information that is about to be acquired and
- ii. at least one concrete example of a possible benefit from the point of view of individual Canadians. This condition will ensure that the justification is not generic and therefore more meaningful. When relevant, the examples should reflect the fact that the individuals could be members of a specific group. For instance, if the data collected refers to Indigenous populations, the example should speak to individuals within that community. A consultation with the Center for Indigenous Statistics and Partnerships can be necessary.

It should be noted that the focus here is on the Canadian population. This includes any Canadian and any sub-segment of the Canadian population that can be affected positively by a statistical activity, especially when the data refers to them specifically. This also includes non-citizen members of society like recent immigrants and refugees.

Here is a good example of a description that speaks to a specific group in the population:

The Federal Government has committed to reduce childcare fees by 50% by the end of 2023 and to 10 dollars per day by 2026. Results from this acquisition will enable the Federal Government to monitor its progress on this commitment to ensure affordable childcare for all Canadian Families. Results from this acquisition will also be used to inform policies designed to improve access to affordable childcare and to reduce barriers to childcare for all Canadian families.

This description is good because it is self-sufficient (i.e., it is possible to assess its merits without further research); non-technical; it speaks to Canadian parents; and it gives an idea of the expected usefulness of the data that is to be acquired.

Here is an example that is unsatisfactory:

This acquisition will fill data and knowledge gaps, support more representative data collection (i.e., at various levels of geography), enhance statistics on diverse populations across age groups, and support governmental policies and societal efforts to address inequities and promote fairness and inclusion in decision making.

This example is unsatisfactory because there is no concrete example of how this data can be used to benefit individual Canadians. The description is also vague (How will it promote fairness how? What kind of decision making? Which policies? What societal efforts?) and therefore it is unclear how the data will be used.

Moreover, it uses the expressions ‘knowledge gaps’ and ‘data gaps’ without explaining why we need to fill those gaps. It is not because there is a lack of information on a given topic that it necessarily needs to be addressed. There is probably a knowledge and data gap on the total number of birthday parties that took place last weekend. It is also very probable that we do not need to fill that knowledge and data gap. The inclusion of this guiding principle is to ensure and document that our acquisitions answer a need for information that will ultimately be used for the benefit of individual Canadians and that data is not acquired without a clear purpose.

It can be more challenging to explain the expected benefits for Canadians when the acquisition of the data is meant to improve or maintain the operations necessary for a statistical activity like a survey. For example, when acquiring data in order to build a survey frame, the benefits to specific groups of the population is somewhat remote. In that case, it is recommended to describe the expected benefits of the specific survey that the acquisition will support.

It is not sufficient to claim that more efficient operations will save resources and tax-payer dollars. Eliminating a survey will also save money and resource. The underlying assumption here is that the survey is important enough that it must be done with efficient operations. The key is to explain why it is important in the first place.

In summary, it is especially important to document the need for the information and the expected benefits of an acquisition when the data is not acquired through a survey. But this does not mean that an ethical review of a survey does not assess the need for such a survey and its expected benefits. It might be possible to improve on how such benefits are communicated, especially when a survey is new or when new questions are added to a questionnaire.

Finally, documenting the expected benefits of a project is coherent with political neutrality and scientific integrity. Statistics Canada can explain, without interference, how its analytical products can be used for good and how current policies can use them in that respect. Statistics Canada will also publish results even if they do not necessarily support initial expectations.

Fairness & Do no Harm

In many ways, this principle represents the flipside of ‘Benefits for Canadians’. Statistical activities have the potential to benefit society, but they can also cause harm in some situations, depending on contextual elements such as the topic and the target population. It is important to know the risks and how to mitigate them. For example, biased data can harm Canadians and that is why ‘quality’ is one of the guiding principles. In this section, five examples of harms or inequalities that can be generated or perpetuated, even with good quality data and good intentions, are discussed.

Firstly, when acquiring and linking large amounts of data, often detailed descriptions of fine-grained subgroups within society can be created. In fact, there can be a greater demand for the creation and analysis of these subgroups.¹⁴ Police and ambulance data, for example, can allow an analyst to create a detailed cluster at a specific geographical level. But it is important to keep in mind that statistics are not created in a vacuum. Each estimate is potentially released within a given social context that needs to be considered to make sure that Statistics Canada does not propagate stereotypes, and that the language that is used is appropriate. This is especially important when information is being collected on subpopulations that may be marginalized or that do not have control over their personal information (e.g., under 18 years old, sex workers, illegal immigrants, non-citizens, etc...). This is why it is important to know if a given acquisition allows for the identification of or targets specific sub-populations.

This means that an ethical review must recommend, when applicable, that appropriate context be given when presenting such data. In some cases, a proactive outreach to the concerned communities to communicate and discuss our intentions can be important before Statistics Canada releases results. It can be relevant to consult with various centers of expertise within the organization. When the data refers to Indigenous populations more specifically, it is mandatory to consult with the Center for Indigenous Statistics and Partnerships because they are in regular contact with the Indigenous communities and can use their network to make this determination.

The issues surrounding the propagation of stereotypes can also manifest themselves in the assumptions behind a statistical inference. For instance, using an algorithm in order to infer ethnicity in function of criminal history would be a very sensitive or unacceptable activity. That is why it is not only important to question the end goals of a given project but also the means used to achieve them.

Secondly, data can represent a reality where inequalities are prevalent. Therefore, it is important to make sure that the data is not used to reinforce such inequalities. For example, geographical information can be highly correlated with information such as ethnicity and socio-economic status that can raise such ethical issues. This phenomenon has been discussed in a report by Bloomberg concerning Amazon’s same-day delivery services.¹⁵ By offering their services in areas with existing paying members, Amazon was inadvertently avoiding predominantly Black or African American (expressions used in the article) neighborhoods across various American cities. By making accessibility to the service more difficult for Black or African Americans, the service became fundamentally unfair. An ethical review must harness enough expertise on a given topic to avoid similar challenges.

Thirdly, the way in which the alternative data is acquired can also cause harm. Web scraping, for example, can perturb a website if done improperly and cause economic harm to the website owner. The quality of the web scraped data must also be taken into consideration. It is thus important to underscore the importance of the relevant directives within the organization when conducting an ethical review.

14. [Disaggregated Data Action Plan \(statcan.gc.ca\)](https://www150.statcan.gc.ca/n1/pub/26-267-x/2016001/article/00001-eng.htm).

15. <https://www.bloomberg.com/graphics/2016-amazon-same-day/>.

Fourthly, harm could occur in the event of a data breach. This would result from a failure to maintain privacy. Data breaches can affect individuals and companies. Good practices are covered more specifically under the ‘Privacy and Security’ principle. Avoiding such breaches has always been a priority at Statistics Canada and it can be relevant to raise questions related to security during an ethical review especially when a new kind of statistical activity is being proposed.

Finally, surveys can cause harm due to the human interaction involved in the acquisition of the data and to the sensitive topics of some surveys. Survey respondents might share information or signs of distress that indicate they may require support. Interviewers may have to handle sensitive situations. A good ethical review must address these potential challenges and recommend that appropriate resources and training be in place. Again, preventing this kind of harm has always been a priority within the organization but it can be relevant to raise questions related to potential respondent distress, during an ethical review, especially when a new survey is under review. When trying to prevent such a source of harm, it is good to remember that there can be a dilemma between the need to protect privacy and the need to report.

Quality

One cannot adequately document the expected benefits of a project and how the ‘Fairness and Do no harm’ principle is taken into consideration without having an idea of the quality of the data that is going to be acquired. Benefits must be credible and sources of harm and unfairness can stem from poor quality data.

The expertise and the level of control over the quality of the data coming from probabilistic surveys is extensive. Challenges can be greater when considering non-survey data or data that is generated through a non-probabilistic survey. Regardless, it is important to understand and document the limitations of the data, in terms of quality, and explain how it is possible to reach the objectives of a project and its expected benefits despite such limitations.

This is the reason why an ethical review must ask for a general overview of the limitations in terms of quality and make sure that the objectives and expected benefits are credible. Quality elements, such as coverage, variable definitions, missing data, timeliness, relevance and response rates are of the utmost importance and it is possible to have some information on these even before the data is acquired. The reader can consult the Data Quality Toolkit for more guidance.¹⁶

Transparency & Accountability

The importance of transparency and accountability has long been recognised at Statistics Canada. Former Chief Statistician, Ivan Fellegi, for example, claimed that Statistics Canada needs: “non-political objectivity, quality, vigorous confidentiality protection, and full disclosure of the methodologies used and the limitations of the resulting data. And of course [...] a media strategy which combines visible openness to them with a vigorous defence of the agency in cases of unfair criticism or erroneous reporting” (Fellegi, 2003). Ethical reviews can help make sure that these components are in place by providing recommendations.

The acquisition of non-survey data sometimes does not involve the consent (explicit or otherwise) of the individuals concerned. While the use of such data can reduce the response burden associated with surveys, individuals may be totally unaware of the expanding use of their data. This means that individuals have less control over their data. Thus, taking proactive steps towards providing timely transparency is important in order to mitigate issues related to this lack of control. Transparent communication on the use of the data, its quality, and the expected benefits for Canadians of a project is essential. Examples of transparent measures could be a post on the Trust Center (see sub-section on Trust and Sustainability), making sure that data providers notify their clients, or a proactive outreach to some specific communities (see sub-section on Fairness and Do no harm).

The acquisition of data through a survey is more transparent than the acquisition of administrative data. But some aspects of surveys can go unnoticed, such as the potential linkage that can be done with the results of the survey or the metadata that can be collected during a survey (e.g., IP addresses). Ensuring transparency in this context is important.

The increasing variety of services that Statistics Canada provides to external partners also open the door to ethical questions concerning the neutrality of the Organization and its commitment to collect and use data for

16. [Data quality toolkit \(statcan.gc.ca\)](https://www150.statcan.gc.ca/n1/pub/11-633-x/2016001/article/00001-eng.htm).

statistical purposes only. Transparency is key to ensure that we maintain the trust of Canadians. This is the topic of the next subsection.

Trust & Sustainability

One of the main differences between conducting a survey and acquiring administrative data is the existence of an external data provider, i.e., the entity from which we obtain the data. The relationship with this intermediary can raise ethical concerns. It is important that the agreements that the organization has with data providers does not undermine the trust of Canadians and that they are sustainable, especially if the data is needed for the foreseeable future. Moreover, if there is collaboration with other agencies or external partners, it is important to discuss how the project fits within the mandate of Statistics Canada (see introduction). If a project goes beyond using data for statistical purposes only, this could raise trust issues because Statistics Canada could go beyond its mandate.

More generally, ensuring that the data acquisition/collection is socially acceptable is key, whether the organization is conducting a survey or acquiring data from elsewhere. Good quality data is often a function of the participation of the target population. Sensitive questions could increase partial non-response and have long term impacts on the respondent's participation. Keeping the level of privacy intrusion proportional to the need of information is also important in that respect.

Privacy & Security

Once the information is integrated within Statistics Canada, it is our duty to keep it secure and to preserve confidentiality. Ethical reviews should take this duty into consideration. However, privacy protection and confidentiality are the tasks of different expert groups within the organization. The main focus of ethical reviews is on the notion of privacy intrusion.

The notion of privacy is complex and notoriously difficult to define (See (Nissenbaum 2019), (Martin 2016), (Francis & Francis 2017), (Mulligan, Koopman & Doty, 2016)). Some definitions give the impression that privacy is simply nonexistent and incompatible with Statistics Canada's mandate. Here is one of those definitions: "Privacy is the right to be left alone, to be free from interference, from surveillance and from intrusions". There are not many situations in modern life where these conditions are met.

In this document, privacy is considered as a limiting force that can contextually change with world events on i) the quantity of information; ii) the way in which we can ethically acquire information; and iii) the nature of the entity that is collecting the information (e.g., Statistics Canada). Traditionally, the concept of privacy has been applied to information about identifiable individuals or groups of individuals, but these restrictions can also be relevant when the information is about businesses or institutions. In this section, the issues surrounding each of the three points are presented in order.

Good documentation must give a justification for the quantity of information (e.g., the sample size and the number of variables) that is being acquired. More specifically, the level aggregation (resolution in case of images) and the frequency of the updates needed to attain the objectives of a project should be justified in such a way that it is clear that not more information than needed is being acquired. In cases where more information than needed is acquired, measures should be taken to ensure that the information is being accessed internally on a need-to-know basis. By abiding to this principle of proportionality, privacy as a moderating force on data acquisition is being respected.

However, more must be done to take privacy into account. This idea of proportionality also applies to the way in which we acquire the data: our collection operations should not be more intrusive than they need to. A collection operation can be intrusive because of the time it takes to fill a questionnaire, because an interviewer needs to visit or phone a household (sometimes more than once), or because physical measures are being taken (e.g., blood samples). These dimensions need to be taken into consideration during an ethical review to assess their proportionality. For example, the presence of an interviewer might not be necessary and could even be detrimental to the acquisition of personal information in some specific cases.

A collection method can also be intrusive when an individual loses control over the use and ownership of their personal information. In the case of surveys, this can happen when one individual is being asked to provide personal information on someone else in the household. In case of non-survey data, it should be noted that they

are often used for a different purpose than the one it was originally obtained for. In fact, this is sometimes included in the very definition of “administrative data”. This form of privacy intrusion, recognised by many Chief Statisticians (See (Fellegi 2003)) needs to be addressed in an ethical review. Credible alternative sources must be considered to justify the proportionality of a given collection method.

Finally, the nature of the entity that is collecting the information can restrain the information that it can acquire. For example, the fact that banks have detailed information on the finances of an individual might be considered less privacy intrusive than if Statistics Canada were to obtain the same information simply because Statistics Canada is a governmental entity. This is an important dimension to take into consideration when conducting an ethical review.

In conclusion, one should notice that in order to moderate adequately the quantity of data, the way we acquire the data and if Statistics Canada should collect, every other principle needs to be taken into consideration. That is why ‘Privacy and Security’ has been presented last. It is the most encompassing and difficult component to evaluate in a review

Conclusion

In summary, this document provides an operational definition of an ethical review at Statistics Canada. It has three steps:

1. Obtaining the relevant information about a given project.
2. Identifying the strengths and weaknesses of the project.
3. Determining if a weakness outweighs the strengths of the project and make recommendations accordingly.

Six guiding principles are used in order to complete the first step: Benefits to Canadians, Fairness & Do no harm, Quality, Transparency & Accountability, Privacy & Security, Trust & Sustainability. Each of these dimensions are assessed as a function of the type of data that is being acquired: survey or non-survey data. With this knowledge, one can better understand the requirements of a review and how to better prepare for one (steps 2 and 3). The outcome should yield a better justification for a given project that uses data at any stage of its life-cycle. It should help maximise both privacy protection and the production of information and thus be in line with Statistics Canada’s necessity and proportionality framework.¹⁷

The landscape has changed at Statistics Canada in terms of the implementation of data ethics and its principles. It will continue to do so in order to meet the evolving need for data within the organisation and in society at large. For example, there is a growing need for more precise guidelines on the ethical use of disaggregated data. This document lays the foundations for future work.

17. [Principles of Necessity and Proportionality \(statcan.gc.ca\)](https://www150.statcan.gc.ca/n1/pub/11-633-x/2016001/article/00001-eng.htm)

Bibliography

Arora, A. (2018), “Modernizing the National Statistical System – Stakeholder Consultations”, *Catalogue 89200003*. Statistics Canada. ISBN 978-0-660-31580-5.

Arora, A. (2022), “Speaking notes for 37th Meeting of the Voorburg Group on Service Statistics”, [Speaking notes for 37th Meeting of the Voorburg Group on Service Statistics \(statcan.gc.ca\)](https://www.statcan.gc.ca/speaking-notes).

Bird, A. (2018), “[Thomas Kuhn](#)”, *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2018/entries/thomas-kuhn>.

Daniels, N. (2020), “[Reflective Equilibrium](#)”, *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2020/entries/reflective-equilibrium>.

Fellegi, I. P. (2003), “Official Statistics: Pressures and Challenges ISI President’s Invited Lectures”, *International Statistical Review / Revue Internationales de Statistique*, 72 (1) : 139-55.

Floridi, L. & Taddeo, M. (2016), “What is data ethics?”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374: 20160360.

Flynn, J. (2020), “[Theory and Bioethics](#)”, *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2021/entries/theory-bioethics>.

Francis, L. & Francis J. (2017), *Privacy: What Everyone Needs to Know*, Oxford University Press.

Martin, R. (2016), “Understanding Privacy Online: Development of a Social Contract Approach to Privacy”, *Journal of Business Ethics*, 137: 551-69.

Government Digital Service (2020), Data Ethics Framework, [Data Ethics Framework - GOV.UK \(www.gov.uk\)](https://www.gov.uk/data-ethics-framework).

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L., (2016), “The ethics of algorithms: Mapping the debate”, *Big Data & Society*, 3 (2): 2053951716679679.

Mulligan, D. K., Koopman, C., & Doty, N. (2016), “Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374: 20160118.

Nissenbaum, H., (2019), “Contextual integrity up and down the data food chain”, *Theoretical Inquiries in Law*, 20 (1): 221-56.

Rancourt, E., (2019), “The scientific approach as a transparency enabler throughout the data life-cycle”, *Statistical Journal of the IAOS*, 35: 549-58.