

Proceedings of Statistics Canada Symposium 2021 Adopting Data Science in Official Statistics to Meet Society's Emerging Needs

Empowering analysts to consider the ethics of their work: A case study of the UK Statistics Authority's data ethics framework

by Natalie Shlomo

Release date: October 22, 2021



Empowering analysts to consider the ethics of their work: A case study of the UK Statistics Authority's data ethics framework

Simon Whitworth¹

Abstract

Our increasingly digital society provides multiple opportunities to maximise our use of data for the public good – using a range of sources, data types and technologies to enable us to better inform the public about social and economic matters and contribute to the effective development and evaluation of public policy. Ensuring use of data in ethically appropriate ways is an important enabler for realising the potential to use data for public good research and statistics. Earlier this year the UK Statistics Authority launched the Centre for Applied Data Ethics to provide applied data ethics services, advice, training and guidance to the analytical community across the United Kingdom. The Centre has developed a framework and portfolio of services to empower analysts to consider the ethics of their research quickly and easily, at the research design phase thus promoting a culture of ethics by design. This paper will provide an overview of this framework, the accompanying user support services and the impact of this work.

Key words: Data ethics, data, research and statistics

1. Introduction

The UK Statistics Authority (UKSA) is an independent body at arm's length from government. The UKSA has a statutory objective of promoting and safeguarding the production and publication of official statistics that 'serve the public good' which includes informing the public about social and economic matter, assisting in the development and evaluation of public policy and regulating quality and publicly challenging the misuse of statistics. As set out in the UKSA's latest five year strategy, [Statistics for the Public Good](#), the Authority's mission is "high quality data and analysis to inform the UK, improve lives and build the future" (UK Statistics Authority, 2020). Concerns about data ethics are often cited as presenting barriers to unlocking the power of data for analytical purposes. As a responsible producer and user of data, the UKSA is meeting this challenge by addressing any ethical concerns head on, launching the [UK Statistics Authority's Centre for Applied Data Ethics](#) (the Centre) earlier this year to help empower the analytical community to identify and address ethical issues in data collection and use for analytical purposes. Empowering analysts to efficiently and robustly identify and address ethical issues in their work is crucial to enabling analysts to collect and use data in ethically appropriate ways and therefore address concerns around data ethics.

All of the work of the Centre is guided by an independent advisory committee consisting of experts from government and academia, including representatives from the Ada Lovelace Institute, the Central Digital and Data Office, the Centre for Data Ethics and Innovation, the Office for National Statistics, Administrative Data Research UK and the Royal Statistical Society's Data Ethics and Governance Section. The Committee provides much valued independent advice on the strategic direction, outputs and impact of the Centre.

¹ Simon Whitworth, UK Statistics Authority, 1 Drummond Gate, London, SW1V 2QQ, United Kingdom.

2. Data Ethics Principles

This work builds on the endeavours of the [National Statistician's Data Ethics Advisory Committee \(NSDEC\)](#), which was established in 2016 to provide independent ethics advice and assurance to the National Statistician. NSDEC, consisting of experts in research, data ethics, data law and a lay member to provide a valuable non-expert perspective, transparently considers research projects from across the analytical community. The committee has reviewed a range of innovative uses of data to inform important government policies.

In its early days NSDEC helped to develop the UKSA's six ethical principles, which are listed below. All of our ethical services assess research projects against these principles. The principles are reviewed regularly to ensure that they are consistent with the latest standards and developments in the ever-evolving data landscape and include:

- The use of data has clear benefits for users and serves the public good.
- The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.
- The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.
- Data used and methods employed are consistent with legal requirements such as Data Protection Legislation, the Human Rights Act 1998, the Statistics and Registration Service Act 2007 and the common law duty of confidence.
- The views of the public are considered in light of the data used and the perceived benefits of the research.
- The access, use and sharing of data is transparent, and is communicated clearly and accessibly to the public.

2. Moving beyond ethical principles

Developing high level ethical principles is all well and good, but how do we ensure that they are used across the analytical community and have impact on the ground? After all, there is no point in having ethical principles if they are not used. We have addressed this challenge and moved from words to action through the development of the [UK Statistics Authority's ethics self-assessment tool](#). The tool empowers analysts to identify and mitigate ethical risk in their projects. The tool's generated ethical risk score also identifies projects that are considered to be more ethically contentious so that they can then be considered by the independent experts on NSDEC.

Analysts are not left to navigate using the tool alone. Guidance, training, user support and oversight are all provided by a team of experts. Feedback from users suggests that this makes the tool easy to use which has helped to make it popular across the analytical community. At the time of writing, it has been used over 470 times by analysts from Government, academia, the commercial sector and charities. To enable timely research, it is important that this an efficient process and currently the average turnaround time for a member of the team to provide advice and support on a completed self-assessment tool is down to two days.

Working with analysts across the analytical community on over 470 ethics self-assessments means we are very well placed to identify specific areas where cross-cutting ethical guidance would be useful to the analytical community. We have therefore kicked off a research programme to publish practical user guidance on a range of applied data ethics topics. So far, this has included publishing guidance on:

- [the ethical considerations in the use of geospatial data for research and statistics](#),
- [considering public good in research and statistics: ethics guidance](#),

- [considering public views and engagement regarding the use of data for research and statistics](#)
- [ethical considerations in the use of machine learning in research and statistics](#)
- [a high-level ethics checklist for using data from third parties.](#)

For those short on time, these guidance pieces include an ethics checklist which summarises the main points covered in the guidance which analysts can quickly work through. While these checklists can be used as standalone summaries, they are also designed to provide supplementary support to researchers completing an ethics self-assessment tool for a research project which has a focus on these particular topics. These guidance pieces are published as open early drafts for comment and feedback from the user community, with feedback received informing future iterations and continued development of the guidance.

4. International work

An important aim of the Centre is to be “recognised world-leaders in the practical application of data ethics for statistics and research” (Centre for Applied Data Ethics Strategy, 2021). Although we have been operational for only eight months, we already have a growing international profile. We have presented our work at a range of international conferences including at the United Nations Economic and Social Commission for Asia and the Pacific Stats Café on Big Data Governance and Data for Policy 2021. We are also represented on the UNECE Ethical leadership task team, which is currently exploring issues related to both professional ethics and data ethics at the organisational level and are developing resources to help support National Statistical Institutes (NSIs) in these topic areas, and have led an international collaboration on producing ethics guidance for the use of Machine Learning in official statistics. This is an important strand of broader international work on Machine Learning for official statistics that is being coordinated by the United Nations Economic Commission for Europe.

The Centre is also leading a programme of work to share learning and experiences amongst NSIs to develop collaborative solutions to data ethics challenges. Earlier this year, we conducted an International Survey of NSIs to identify what common ethical challenges they are facing and the ways in which they are responding to these challenges. The following three potential deep dive topic areas have been identified as common data ethics issues that NSIs are grappling with: data ethics policies, ethically maximising the utility of data for research and statistics and distinguishing between statistical and operational uses of data. We plan to lead international collaborations on all three of these important topic areas.

4. Summary

Thinking on data ethics is evolving at pace. One of the Centre for Applied Data Ethics early publications was [a review of the UK’s data ethics landscape](#) (UK Statistics Authority, 2021), which presents the large volume of innovative work that is happening in the UK data ethics space among different organisations. As prolific users of data, often in increasingly innovative and sophisticated ways, data analysts from across the statistical system have the potential to make a significant contribution to shaping wider thinking on data ethics. The work of the Centre will play an important role in enabling the analytical community to make this important contribution.

Our experience has shown that if analysts are empowered to consider data ethics and address any ethical issues early in the research design phase, most concerns about data ethics can be successfully mitigated and do not need to be a significant blocker to enabling the use of data for radical and ambitious analysis. The provision of effective data ethics services, guidance, and user support to enable analysts from across the statistical system to consider the ethics of their work at the research design phase can therefore play a significant role in helping the statistical system to produce high quality research and statistics for the public good.

References

UK Statistics Authority, 2020, Statistics for the public good, <https://uksa.statisticsauthority.gov.uk/statistics-for-the-public-good/>

UK Statistics Authority, 2021 Centre for Applied Data Ethics Strategy, <https://uksa.statisticsauthority.gov.uk/publication/centre-for-applied-data-ethics-strategy-enabling-ethically-appropriate-research-and-statistics-for-the-public-good/>

UK Statistics Authority, 2021, Ethical considerations in the use of geospatial data for research and statistics. <https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-geospatial-data-for-research-and-statistics/pages/1/>

UK Statistics Authority, 2021, Considering public good in research and statistics: Ethics guidance, <https://uksa.statisticsauthority.gov.uk/publication/considering-public-good-in-research-and-statistics-ethics-guidance/>

UK Statistics Authority, 2021, Considering public views and engagement regarding the use of data for research and statistics, <https://uksa.statisticsauthority.gov.uk/publication/considering-public-views-and-engagement-regarding-the-use-of-data-for-research-and-statistics/pages/1/>

UK Statistics Authority, 2021, Ethical considerations in the use of machine learning for research and statistics, <https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-machine-learning-for-research-and-statistics/>

UK Statistics Authority, 2021, Using data from third parties for research and statistics: High level ethics checklist, <https://uksa.statisticsauthority.gov.uk/publication/using-data-from-third-parties-for-research-and-statistics-high-level-ethics-checklist/>

UK Statistics Authority, 2021, Identifying gaps, opportunities and priorities in the applied data ethics guidance landscape, <https://uksa.statisticsauthority.gov.uk/publication/identifying-gaps-opportunities-and-priorities-in-the-applied-data-ethics-guidance-landscape/>