United Nations ECE/CES/2024/14/Add.15

Economic and Social Council

Distr.: General
26 June 2024

English only

Economic Commission for Europe

Conference of European Statisticians

Seventy-second plenary session
Geneva, 20 and 21 June 2024
Item 11 (a) of the provisional agenda
Programme of work of the Statistics subprogramme of the United Nations Economic Commission for Europe: Reports on the work of the Conference of European Statisticians, its Bureau and Teams of Specialists

Report of the Workshop on Ethics in Modern Statistical Organizations

Prepared by the Secretariat*

Summary

This document provides the report of the Workshop on Ethics in Modern Statistical Organizations which was held in Geneva, Switzerland on 26–28 March 2024. The Conference is invited to take note of its contents, and to provide any guidance as appropriate.

* This document was submitted late for processing due to resource constraints.

1. The workshop was organized as part of the Conference of European Statisticians' work programme for 2023, within the context of the High-Level Group for the Modernisation of Official Statistics (HLG-MOS) activity. It was held in Geneva, Switzerland on 26–28 March 2024.

2. There were 42 participants, including representatives of national statistical offices (NSOs) and government agencies of the following 19 countries: Albania, Australia, Belarus, Canada, Finland, France, Hungary, Ireland, Italy, Latvia, Netherlands (Kingdom of the), New Zealand, Nigeria, Poland, Portugal, Republic of Korea, Sweden, Switzerland and the United Kingdom of Great Britain and Northern Ireland.

3. In addition, there were experts from Amherst College (United States of America), the University of Malta (Malta) and the University of Manitoba（Canada). The workshop was also attended by representatives from Eurostat and the United Nations Economic Commission for Europe (UNECE).

4. The workshop was chaired by E. Dhuli (Instat, Albania). The Organizing Committee of the workshop includes: F. Rotundi, A. Leonetti and K. Ambrosino (Istat, Italy), M. Beaulieu and M. Karaganis (Statistics Canada), O. O'Gorman (CSO Ireland), E. MacDonald (Statistics New Zealand) and L. Augustyniak (Eurostat).

5. The agenda included the following substantive topics, each comprising its own session within the meeting:

- Session 1: Ethics in institutional contexts
- Session 2: Ethics in daily work life
- Session 3: Ethics for new data sources and technology
- Session 4: Ethics and proactive communication and
- Session 5: Open discussion for the Reference Book on Ethics.

6. Twenty-four substantive presentations were made within these sessions. The timetable, papers, and presentations from the workshop are available on the UNECE website https://unece.org/statistics/events/Ethics2024.

7. The key messages and takeaways of the sessions and Q&A included (more details available in Annex):

- Working on ethics is not optional but mandatory.
- Sharing examples and pooling resources is an effective way to collectively addressing the issue that many countries lack resources for doing work on data ethics. It was proposed to compile available materials on data ethics in official statistics.
- Sometimes it is not easy to distinguish between ethical and quality principles. It is important to define clear goals for ethical behaviour. We should be specific when discussing ethical dilemmas, especially in statistical data, and clearly define the roles of actors in the scenarios.
- The Generic Statistical Business Process Model (GSBPM) is useful to ensure ethical considerations are taken into at every stage of the statistical production.
- Ethics frameworks should apply to all data sources.
- New data sources present new ethical issues of concern. The pandemic also has raised issues of work-life balance given the consequential increase in teleworking.
- The importance of delineating roles and responsibilities of individuals, as well as the need for more organizational controls was highlighted. The process of ethical decision-making by the data ethics staff and escalation for vulnerable cases to the internal ethics committee are necessary functions of organizations. The establishment and functioning of the ethics committee, including member selection and responsibilities, are important for organizations.
- It is important to clearly communicate ethical values.

8. A number of themes and ideas were identified as future work priorities during and after the workshop (through an evaluation survey):

- Developing a cohesive definition of ethics in official statistics along with its constituent components is imperative. Equally crucial is the creation of a flexible framework that accommodates diverse needs. Moreover, providing guidance and online ethics training will empower NSOs to effectively utilize this framework.
- Expanding the scope to address future challenges, such as the evolving role of NSOs in data governance and stewardship.
- Exploring methods to enhance staff engagement in changing organizational culture through an ethical lens.
- Examining the various roles of NSOs in ethics, including data collection, dissemination, and awareness promotion.
- Considering methods to build social acceptability of official statistics, whether through internal efforts or external assistance.
- Conducting a study on data ethics cases related to statistical production using new data sources.
- Expanding discussions on data ethics and responsible artificial intelligence (AI).
- Conducting further discussions on microdata ethics and AI, focusing on the roles of NSOs.
- Providing concrete examples will clarify ethical considerations in practical contexts and guide ethical decision-making processes.

Annex

Summary of proceeding and discussions

Session 1: Ethics in institutional contexts

1. This session was chaired by F. Rotundi (Istat, Italy) and it included the following presentations:

- Keynote Presentation: Do statistical ethics apply equally to NSOs and to other official statistics producers, whether regional/international or other national statistical authorities? – A. Georgiou (Amherst College)
- Democracy dies in darkness without Official Data – L. Di Gennaro Splendore (University of Malta)
- Structure of ethical issues in new data ecosystems – M. Johnson, T. Koskimäki, M. Sovala (Statistics Finland)
- Revision of the Swiss Official Statistics Charter: opportunities and risks – M. Baumann (Federal Statistical Office, Switzerland) and P. Laube (Swiss Council for Official Statistics)
- UK Statistic Authority's Centre for Applied Data Ethics (CADE) – the first three years – N. Shearman (Office of National Statistics, United Kingdom)
- Investigating Ethical Practices in NSOs – Surveys Results – K. Ambrosino (Istat, Italy)

2. The points raised during the presentations and discussions include:

- The definition of ethics is relatively simple: "how one ought to act". However, it becomes more complicated as we delve into specific areas that we must deal with.
- If there is universal ethics in statistics regardless of the viewpoint of the person, how can statisticians formulate ethics for different users of statistics, such as the media or the public? Do we have the background to develop a universal ethical approach for statistics? Should we focus on the production side, the ethics of producing statistics? – We are not looking to produce ethics for all users in their own lives, but we do have ideas on how they should interact with official statistics. Very important aspects of statistical ethics concern the actions and behaviours of actors beyond the statistical producers (e.g., "recruitment of head of NSO" is a task of law makers). We should, as statistical producers, use what can help us stay independent and produce objective and impartial statistics of high quality.
- Traditional ethical guidelines no longer suffice; they require expansion. Should we address this in the code of practice (COP) and if so where, or should we explore new avenues? Do we need to develop a more comprehensive code, or should we enhance the existing one with specific provisions? This is an important question, and it is necessary to start with a common language before any progress can be made.
- Is the use of the UK Ethics Self-Assessment Tool enforced by staff regulation or is it discretionary? Is it merely a checkbox exercise, lacking genuine reflection? – In ONS UK, use of the tool is required, data collection cannot proceed without conducting the assessment. The users are willing to review and adjust accordingly. Pre-screening is conducted for sensitive data, with specific triggers identified. There is no room for relaxation in this process; priority is given to the more critical cases.
- Is the UK ethics tool viewed as an enabler rather than a barrier? – The 99 per cent of its users perceive it as a helpful tool that shapes research. Users have adapted the tool, shifting the focus of questions as needed. Additionally, consideration is given to alternative data sources during the ethical approach application process.

- Is the public involvement principle taken into account by the UK Statistics Authority? – The external committee, comprising a diverse range of experts, oversees this process. Public representation is carefully considered within this committee. Moreover, consultations with the relevant groups are conducted when collecting data from them, ensuring full engagement and participation.
- Question about complexities of stakeholder relationships in Switzerland, is this a continuous collaborative process or sporadic? The offices in proximity collaborate through various organized groups. There is also a working task force, comprising of representatives from federal level, FSO and Ethics Council, all mandated by their respective committees to which they report afterwards. It represents decentralized model of stakeholder relationship.

Ethics boot camp

3. This part of session 1 was chaired by A. Leonetti (Istat, Italy). The structure of the session was as follows: each group was provided with a short description of scenarios related to ethical dilemma and instructions on how to analyse it with the Cressey's Triangle.

4. The participants noted that Cressey's Triangle was useful for analysing both individual thought processes and organizational processes, offering an effective framework for identifying and addressing issues within the organization. The conclusions from the group discussion during the boot camp underscored the importance of preventive measures such as privacy policies and declarations of conflict of interest. Additionally, the discussions emphasized the significant role individuals play in shaping outcomes, as their actions are influenced by various needs and pressures as well as that of organizations in implementing more effective controls or preventive measures to avoid legal issues in specific situations.

Session 2: Ethics in daily work life

5. This session was chaired by N. Shearman (Office for National Statistics, United Kingdom). It included the following presentations:

- Rules of Professional Ethics in the State Statistics Bodies of the Republic of Belarus. V. Pazharytskaya (National Statistical Committee of the Republic of Belarus)
- Proposals to Promote Change from Compliance to Ethical Commitment in Istat – A. Leonetti (Istat, Italy)
- Incorporating ethics in statistical organizations through GSBPM and GAMSO – I. Choi (UNECE)
- French official statistician and ethics: from law to practice – M. Chaleix and O. Lefebvre (Insee, France)
- Ethics in staff and user satisfaction survey (Case of Albania) – V. Lasku (Instat, Albania).

6. The points raised during the presentations and discussions include:

- GSBPM is used across various statistical domains for different types of data. As such, it is difficult to establish consensus on how to integrate AI, machine learning, and data stewardship into the model in a way that is applicable for varying situations and contexts. Inputs from ethics experts would assist with the aim to refine the wording for broader acceptance.
- The increasing use of microdata and privately held data prompted the need for specific ethical considerations is important to analyse proportionality and incorporate this aspect into ethical perspectives.
- Statisticians must incorporate ethics centrally into their thinking/way of being.
- There is a need for transparency and ethical considerations in data usage, particularly as NSOs increasingly work with granular data.

- There is debate on whether ethics should be part of quality management or addressed separately (is ethics part of quality or quality is part of ethics?) with suggestions for dedicated ethics champions and compliance teams. There should be clear reporting structure for ethical concerns and a team that can escalate issues directly to top management.
- During GSBPM revision, it was suggested either to create a separate sub-process dedicated to ethics, or integrate ethical considerations into all phases, as ethical considerations can be identified in all phases.
- The motivation for pursuing the International Organization for Standardization (ISO) certification in Istat hinged on legislative requirements where the legislation mandated the implementation of measures to combat corruption, hence the necessity for certification. Organizations are mandated to exhibit efforts towards corruption prevention, and ISO certification serves to fulfil this obligation by providing a tangible demonstration of these efforts.
- Insee identified a need for a unified data processing process for greater efficiency and productivity gains. It touched upon issues related to data access, reduction, and processing methods. Overall, the presentation covered diverse aspects of data management and processing.
- Engagement strategies regarding information dissemination and handling public commentary: What are the approaches to addressing incorrect information, emphasizing the importance of adapting responses to different situations? Additionally, how can we gain insights into decision-making processes and challenges faced in providing accurate information? The approach to checking and addressing issues varies by situation: for journalists, it involves direct calls and suggesting alternative perspectives, while for online content, it includes specific rectifications or re-requests; responses are adapted to the situation's needs, and previous strategies are leveraged to explain decisions and challenges effectively.
- Questions were raised about broader communication issues, emphasizing that ethics should not be reduced to PR. There should be ethical considerations in data access and dissemination, It was suggested that communication strategies should be integrated into standard processes.
- Is integrity a part of ethics or is it different? It was suggested the integrity is part of ethics, but further clarification of this topics would be helpful.

Session 3: Ethics for new data sources and technology

7. This session was chaired by M. Beaulieu (Statistics Canada). It included the following presentations:

- Reimagining how we deliver quality data and statistics: Stats NZ Journey – E. MacDonald (Statistics New Zealand)
- The Role of Data Ethics to Maintain and Improve Public Trust: The Statistics Canada Experience – M. Beaulieu (Statistics Canada)
- Towards a data ethics programme for the Australian Bureau of Statistics: Considering privacy, ethics and trust for our innovative data uses – J. Hillermann (Australian Bureau of Statistics)
- Statistics Netherlands ethics committee – purpose, composition and methods – E. de Heij (Statistics Netherlands)
- Ethics of Technology – M. Karaganis (Statistics Canada)
- The role of geo-information in ethics within modern statistical institutions – M. Deva (Instat, Albania).

8. The points raised during the presentations and discussions include:

- Example of voluntary Algorithm Charter for Aotearoa New Zealand it was highlighted that it is not mandatory for government agencies to sign up. Agencies voluntarily commit to following good practices outlined in the charter, focusing on algorithmic transparency and accountability. An algorithm impact assessment was released alongside the charter. The algorithm charter is great, but a lot of people request additional toolkits to understand the principles outlined in the charter better. The charter also serves as one of four critical guidelines for data driven technology.
- Availability of a special budget and whether there is adequate staff capacity to explore new topics within the ten-year data programme. The response provided insights into ongoing projects within the wider transformation work programme, including initiatives focused on registers (business, spatial and people registers), Census 2028, and overall transformation efforts. Additionally, it indicated that innovation is underway, with teams working on projects related to business registers, spatial data, population registers, and technology transformation.
- Interaction between two entities: the Māori data authority and data ethics authority, and whether these are groups separate and how they are managed. It was clarified that the metadata authority and the Statistics New Zealand are distinct entities. Additionally, there is a multidisciplinary group known as the Māori Data Sovereignty Group, which operates independently to provide oversight and ensure accountability in governmental data practices. Collaboration between these entities was highlighted, emphasizing the importance of working together while recognizing their distinct roles.
- A question about integrated data infrastructure in New Zealand and its ethical considerations. The whole ethics team was involved in its development. The "five-safe" framework that was used comprises components that ensure data safety and confidentiality for researchers both government-affiliated and external. Due to the sensitive nature of the data, particularly concerning vulnerable individuals, ethical safeguards within the system are exceptionally strict. The focus remains on addressing concerns about centralized government data and ensuring people's comfort with the consolidation of information from various agencies. Concerns were raised about the increasing use of AI and the need to address ethical issues associated with it. There was also a focus on the importance of staying ahead of technological advancements and ensuring that statistical research remains relevant and inclusive. Additionally, the role of governance in overseeing data practices globally was emphasized, highlighting the need for proactive measures in addressing emerging challenges.
- Various topics were addressed regarding practical approaches and organizational structures for effective monitoring and assessment of ethics. There was an emphasis on clear objectives and strategies to achieve them, particularly in the context of assessment. The importance of well-defined organizational models and the composition of committees were highlighted, with a focus on external expertise. Concerns were raised about the clarity and effectiveness of monitoring processes, especially in evaluating compliance with established principles. Participants emphasized the need for thorough assessment and evaluation of ethics during the monitoring phase to ensure ongoing compliance. Additionally, the importance of integrating privacy and ethics considerations at every project stage, including obtaining external ethics approval for certain endeavours such as health research, was raised. The application of the five safes framework was identified as a key method to ensure data safety, although its implementation may vary depending on project elements.
- Once a project is launched in Canada and has passed through all stages, including privacy considerations, there is typically no further review unless significant changes occur. However, there are intermediate steps in place for projects, and any substantial changes must be presented to relevant committees aware of ethical considerations. Continuous monitoring throughout the project lifecycle is not feasible due to resource limitations. Therefore, monitoring responsibilities are often delegated to project management committees.

- There was a focus on privacy and trust, particularly in relation to maturity models for evaluation. Participants were interested in learning about the key points targeted in defining privacy models and whether examples could be shared for reference. It was noted that having clear models, guidelines, and processes contributes to higher ratings in privacy and trust. However, there was recognition of a gap in the ethics space where documented frameworks and processes are needed for maturity. Some governments, like in Australia, have additional governance layers for projects involving data of Indigenous people to ensure cultural safety and avoid deficit of narratives in research analysis. While not all processes are fully in place, there is a clear understanding of where improvements are needed for better documentation of processes.

Session 4: Ethics and proactive communication

9. This session was chaired by L. Augustyniak (Eurostat). It included the following presentations:

- An ethical approach to the development of social acceptance and its application – J. Byrne (Central Statistics Office, Ireland)
- An assessment of ethics and proactive communication practices in the Nigerian Statistical System – K. Dzaan (Central Bank of Nigeria)
- Ethics and proactive communication: the Istat case – G. Peci and M. Troia (Istat, Italy)
- Building trust culture in the office – examples of ethics-driven proactive internal communication at Statistics Poland – A. Borowska and O. Świerkot-Strużewska (Statistics Poland).

10. The points raised during the presentations and discussions include:

- The discussion touched on the high employee satisfaction in the NSOs, as they provide stimulating and ethical work environments, despite lower-level salaries compared to other government agencies. It is important to recognize the value of employees and to foster a supportive workplace culture.
- The importance of understanding and maintaining trust was highlighted, with an emphasis on the need to consider the perspective of users in assessing the value of official statistics.
- Ethics play a core role in data integrity and organizational practices.
- The discussion touched upon the prioritization of communicating ethics within a complex system and limited resources. The questions were raised on how to identify priority groups for communicating ethics more efficiently. The discussion highlighted the importance of leveraging existing platforms, such as biannual meetings attended by state representatives, to initiate discussions on ethics and ensure accountability for implementing ethical practices. The emphasis was on utilizing influential voices within the system to drive the agenda forward.
- The discussion raised concerns about handling user requests for images in a fair and unbiased manner. Reference was made to recent news about Google Gemini generating images trained for over-correction, leading to misrepresented gender and race.
- Even with a thorough planning, NSOs may encounter unexpected backfire. Active engagement, expressing apologies for any error made and appreciating for reaching out to NSO for the correction could eventually lead to positive responses from users for the transparency and commitment for improvement of the statistics office.
- Why has ethics become more prominent now? It is likely due to the changing landscape of data management and the presence of new actors in the field. We have many new challenges posed by new roles, and we need to emphasize the need for accountability and responsibility in managing the entire data cycle.

- We should stress the importance of internal communication channels and tools in building trust, since trust is integral to ethical practices.
- There is an absence of a distinct ethical code for statisticians. All civil servants in NSOs abide by unified ethical principles, so it would be beneficial to employ a generic model to tackle ethical considerations in statistics, and in particular the concerns around data production and other various organizational domains.
- It is important to consider user perspectives in progressing the understanding of the value of official statistics.

Session 5: Open discussion for the Reference Book on Ethics

11. This session was chaired by F. Rotundi (Istat, Italy) and it had the following structure:

- Introducing the Reference Book
- Adjusting the proposed content based on the workshop findings:
    - Conceptual groundwork
    - Ethics in official statistics
    - Ethical dilemmas
    - Results of the ethics surveys.

12. The following points were raised during the discussion of the Reference Book on Ethics:

- There is a need for a common language on ethics and its different areas.
- It was suggested to extend the scope to include future challenges, such as the role of NSOs in data governance and data stewardship and to address data accessibility as one of the future challenges.
- The need to act ethically amidst rapid innovation was highlighted, suggesting case studies from the countries as examples.
- It is important to cover statistical literacy in the context of distinguishing real and fake statistics.
- Is it necessary to highlight the distinction between ethics in official statistics and ethics in national official statistics?
- For the ethical dilemmas, we should specify the risks, how they were dealt with, possible outcomes and lessons learned.
- It was suggested to highlight in the Reference Book, why there is an urgent need to have discussion on ethics now.
- The analysis on legal versus ethical responsibility of NSOs should be included in the book.
- We need to address ethical culture of the organizations, provide a definition, and note how it is applied in official statistics, versus stewardship and research.
- The Reference Book should contain complex ethical dilemmas to provide guidance on decision-making. However, it is noted that there is no single solution that fits every scenario, and some situations need a nuanced approach.

13. The following examples of ethical dilemmas were mentioned during the discussion:

- Ethical dilemma around the use of data concerning specific populations, such as people with disabilities. It advocates for taking the perspective of the individuals whose data is being utilized and emphasizes the importance of consulting them to determine what constitutes acceptable and unacceptable use of their data. While data is acknowledged as a public good, the ethical dilemma arises from whether it is

ethically acceptable to compare data with that of the population with disabilities without their input or consent.

- Dilemma revolving around the allocation of school funding based on a new algorithm that considers granular family-level data instead of solely relying on geographical area- While this approach may improve funding allocation accuracy, it raises concerns about intruding into families' privacy and potentially disadvantaging those who might receive less funding under the new system. The argument for using higher quality data rests on the need for evidence to justify the change and ensure better outcomes. However, it also requires careful consideration of different perspectives and an examination of the ethical implications, including potential harm to certain individuals or groups. Overall, the dilemma underscores the importance of balancing the benefits of using more accurate data with the potential consequences for those affected by the decision. The is a contradiction between intended and perceived public good.
- An ethical concern is so called Daily Mail question: when something is legal but not ethical. It raises concerns about unequal access to data and its impact on societal benefits, underscoring the need for ethical considerations in data practices. Ultimately, we should reflect on navigating these complexities with sensitivity and integrity.
- Ethical dilemma surrounding access to microdata. The dilemma centres on unequal access to data, with certain groups or individuals, particularly those with funding or institutional affiliations, having greater access to microdata compared to others. This inequality raises concerns regarding fairness, transparency, and the democratization of information.
- More granular data leading to misinterpretation of data. The dilemma revolves around the balance between providing more detailed, granular data to enhance transparency and the risk of misinterpretation of such data. There is a discussion about the need to maintain methodological integrity to prevent public misinterpretation. It is important to find a balance between transparency and preserving privacy and confidentiality, especially concerning personal and sensitive data. This dilemma underscores the complexities in communication and data transparency and highlights the need for careful consideration in determining the appropriate level of data granularity for public disclosure.
- An ethical dilemma surrounding the definition of the population in a particular city. This dilemma revolves around ensuring the accuracy of data regarding the recorded number of citizens, which directly affects resource allocation. The discrepancy in population figures primarily arises from commuters and migrants, prompting discussions on the importance of maintaining data quality and transparency. It emphasizes the significance of clear communication and fair treatment in the collection and reporting of data. This example illustrates the challenges statisticians face in accurately representing population data and highlights the crucial role of ethical communication and transparency in statistical reporting.
- Stigmatization of marginalized groups. There is concern that certain figures from a study could exacerbate political tensions and lead to stigma against these groups. The dilemma revolves around balancing objectivity and providing necessary information without causing harm. Participants should consider consulting affected communities during questionnaire design to ensure acceptance of the results. They should discuss contextualizing data dissemination to mitigate negative impacts. Ultimately, the dilemma highlights the importance of ensuring data accuracy while minimizing harm to vulnerable populations.
- Role of NSOs with respect to informing public on risk of emerging technologies. Do NSOs have a social responsibility to share their expertise to help citizens understand and use of technology, as well as inform them about associated risks?