

Structure of Ethical Issues in New Data Ecosystems

Markus Sovala

Director General, Statistics Finland

UNECE Workshop on Ethics in Modern Statistical Organizations

March 26-28, 2024

Geneva, Switzerland



Outline: Structure of Ethical Issues

- Methodology
- Data and results
- Conclusions
- Discussion

Methodology – study of professions

- Study of professions is an established field of research in social sciences.
- Professional ethics is an established subject in studies of professions.
- Study design:
 - Two recent incidents of distrust analyzed.
 - Public distrust emerges, when perceived behavior (of the professionals) does not comply with the expectations of the audience (expected ethics).
 - Research question: What were the main differences between perceived and expected behavior.
- ISI Declaration on professional ethics used as framework to analyze the incidents
 - which aspects of the ISI code emerged in incidents, and which not.
 - which core values were perceived being violated.





Picture: Helsingin Sanomat

Case 1: Attempted healthcare reform in Finland

(... one of many attempts, this was in 2015-2019 ...)

- New regional administrative layer (welfare areas)
 - To manage public health services and social care (read: old people care).
- Allocation of central government funding
 - To private and public service providers in new welfare areas.
- Capitation reimbursement model
 - An individual health risk coefficient for all citizens
 - Micro-data on health records and socio-economic status.
- Public debate on the approach
 - Plan not accepted by the parliament and Government resigned, but not only for data related reasons.





Picture: E24

Case 2: New data sources

- New Statistics Act of Norway:
 - Powers to NSI to access privately held data
- The data request:
 - Detailed information on consumption, based on cash registers, to be statistically linked with payments data (credit cards).
 - To replace Household Budget Survey and to provide new statistics on health aspects of the diets.
- Public debate of the approach:
 - (Many) Data holders rejected the request, being worried of their own reputation.
 - Data Protection Supervisor gave negative opinion.
 - Plan halted, negotiations with DPA and data holders continue.



ISI declaration of Professional Ethics – perceived violations

- Core values:
 - Respect, professionalism, truthfulness and integrity
- Ethical principles
 - 1. Pursuing Objectivity
 - 2. Clarifying Obligations and Roles
 - 3. Assessing Alternatives Impartially
 - 4. Conflicting Interests
 - 5. Avoiding Preempted Outcomes
 - 6. Guarding Privileged Information
 - 7. Exhibiting Professional
 - 8. Maintaining Confidence in Statistics
 - 9. Exposing and Reviewing Methods and Findings
 - 10. Communicating Ethical Principles
 - 11. Bearing Responsibility for the Integrity of the Discipline
 - 12. Protecting the Interests of Subjects
- Respect:
 - 12. Protecting the interest of the subject
 - Perceived as a risk by the audience
- Professionalism
 - 3. Assessing alternatives impartially
 - 10. Communicating ethical principles
 - No alternative to intrusive methodologies were considered
 - Tertiary users (in healthcare-case) failed to understand the sensitivity of their approach
- Truthfulness
 - 2. Clarifying Obligations and Roles
 - 8. Maintaining Confidence in Statistics
 - 9. Exposing and Reviewing Methods and Findings



Conclusions

- The researchers and institutions acted by the book, no (traditional) professional misconduct.
- The audience perceived the actions quite differently: attempts to create a new, potentially better funding system became difficult or impossible (Case #1). Data providers didn't trust NSI's good intentions (Case #2).
- The stakeholders and supervisors perceived the actions quite differently
 - The legislators (Case #1) and Data protection supervisor (Case #2) disagreed.
- **Conclusion: The current professional code does not support statisticians in emerging new data ecosystem roles.**



Discussion

- **Possible reason for missing guidance:** The basic assumption behind current codes is that all data is collected by NSIs, mainly by surveys. This not true any longer.
 - In Finland about 99 per cent of data points used for statistics are coming from admin registers.
 - Most countries will face the same situation as the censuses and (other statistics as well) are moving rapidly towards use of administrative data.
 - It's big question how this rich data can be used efficiently for different purposes and what's role of NSI in this.
- **Proposal: Our profession should aim at broader code, including advices for emerging data ecosystem roles:**
 - **General principles.**
 - **Guidance for practical cases:**
 - Secondary and tertiary use of data (use of admin data for statistics and for research).
 - Combining data from different sources (both admin data and privately held data).
 - Use of micro-data (or very granular statistics) in knowledge-based decision making.

