

Official statistics: Quo vadis?

by

Sæbø, Hans Viggo; Hoel, Marit

*"This is the Accepted peer-reviewed manuscript version of the following article, accepted for publication in **Statistical Journal of the IAOS**. It may contain minor differences from the journal's pdf-version.*

The definitive, peer-reviewed and edited version of this article is published in

Statistical Journal of the IAOS : 15 March 2023

The final authenticated version:

Sæbø, Hans Viggo and Hoel, Marit. 'Official Statistics: Quo Vadis? Statistical Journal of the IAOS, 39 (1), 39-45, 20231. <https://doi.org/10.3233/SJI-220100>

© Hans Viggo Sæbø; Marit Hoel, 2023



Statistisk sentralbyrå
Statistics Norway

Official statistics – Quo vadis?

Hans Viggo Sæbø* and Marit Hoel

Statistics Norway, Akersveien 26, 0177 Oslo, Norway

Abstract

The data revolution has resulted in discussions in the statistical community on the future of official statistics. Will official statistics survive as a brand, or will such statistics drown in the flow of data and statistics from new sources and actors, including misused statistics and fake news? The Covid-19 pandemic has been an additional driver for discussion. There is a need to maintain the quality of official statistics and highlight the value of such statistics for the users as a basis for – and supplement to – other statistics and information. It is at the same time important to implement new developments to improve and keep up the relevance of official statistics. Key pillars today are statistical legislation, quality frameworks and core values defining requirements for official statistics. Possibilities are linked to new statistics, use of new data sources and possible extended roles of the statistical institutes within coordination, collaboration, and data stewardship. The paper addresses these issues in the light of trends in official statistics since the UN Fundamental Principles of Official Statistics were formulated about 30 years ago. Quality challenges for statistics and dilemmas in defining the roles of statistical institutes are considered. The paper includes examples from Statistics Norway.

Keywords: Official statistics; data revolution; coordination; cooperation; data stewardship

* Hans Viggo Sæbø, Statistics Norway, Oslo, Norway. Tel.: +47 45004936; E-mail: hvs@ssb.no

1. The data revolution

The future of official statistics might be threatened by the data revolution characterized by the abundance of data, data providers and new technology. Davies [1] wrote an article in the Guardian with the title “How statistics lost their power – and why we should fear what comes next”. He claims that the ability of statistics to accurately represent the world is declining, and that big data controlled by private companies is taking over and even putting democracy in peril. This points at both the role of global companies beyond the reach of national statistical legislation and fake news. This is a warning, that is and should be taken seriously in the statistical community.

Even if the concept data revolution has been used since the 1960's, there are reasons to believe that we have a paradigm shift, such as the one experienced with the emergence of Internet in the 1990's. MacFeely [2] has discussed if the paradigm for official statistics has shifted, and concluded that it has, in fact several times over the last 30 years, and both from the supply side of data and the demand side. On the supply side this paradigm shift is linked to the more recent definition of data to comprise monitoring of all activities. Over some years, there has been an expansion in the use of secondary data sources for statistics, from administrative data to other and new data, often denoted big data. On the demand side MacFeely mentions the common understanding that official statistics are a public good, available for free for all users (see 3.1).

Among international initiatives addressing the challenges of official statistics in today's “datafied” society is the Krakow Working Group endorsed at the 2022 IAOS General Assembly¹.

The need for information and the supply of new and timely data to fight the Covid-19 pandemic have illustrated the data revolution and challenged official statistics. Radermacher [3] notices that the authority of official statistics seems to have lost influence. He mentions the lack of statistical expertise and influence in the collection of data, and a lack of statistical literacy among the general population. Ljones [4]

¹ See: <https://officialstatistics.com/news-blog/fast-evolving-landscape-official-statistics-how-respond-challenges-3>

notes that the key figures for the prevalence of and deaths from Covid-19 are usually sources from epidemiological institutes, and not from the statistical institutes, who, however, may deliver background data. We can talk about an “infodemic”, a concept introduced during the SARS epidemic in 2003 to characterize the “information pandemic”. The possibilities for misinformation and fake news represent one side of the data flow. The other side is that many of the data generated are useful. Key statistics or indicators are the number or rates of registered infected persons, hospitalisations, and people in intensive care, deaths from the pandemic and vaccinations. In many countries, only some of these statistics (e.g., the number of deaths) have normally been classified as official. The sudden need for timely information about the pandemic has generated new statistics, though with quality problems. Decisions affecting societies severely are taken based on these statistics. This particularly concerns the number of registered infected persons – a measure that depends heavily on the amount of and representativeness of testing. Accuracy is therefore poor, but timeliness is very good. However, such figures have probably been good enough for decision making, at least in the first phase of the pandemic. This is a good example of important statistics normally produced outside the system of official statistics.

It is important to both be able to understand the numbers correctly and communicate them in a relevant way. This has been a challenge both for the health authorities, the media and thus for the public. Factors that may lead to misunderstandings when interpreting statistics and data on the pandemic comprise the uncertainty of statistics based on few observations, the effects of missing data (e.g., difference between infected persons and what has been registered), comparisons between geographical areas not based on rates, and difference in definitions (e.g., causes of deaths). It is important to check for relevant background variables (such as affected persons' age) and not to confuse concurrence (correlation) with causality. National Statistical Institutes (NSIs) in general contribute to correct use of their own statistics, to a lesser extent other statistics.

In Norway, health authorities are responsible for the statistics describing the pandemic, while Statistics Norway produces statistics on how the pandemic and the measures against it affect the society.

2. Data and statistics

Data is not the same as statistics, though statistics are also data. Statistics are numerical information relating to an aggregate of data on units or observations. There is a classic way of ordering statistics above data on the road to knowledge, i.e., statistics is closer to decisions than data. However, today the concept of data is widened and dominates the public discourse. The age of statistics is being replaced by the data era as expressed by Radermacher [5]. The emergence of data science as a discipline could be mentioned in this context. The programme of the World Data Forum in 2021 can illustrate this. The word data appears 121 times in contributed paper's headings, statistics only 7 times.

Statisticians and data scientists cooperate and participate in the same international fora. In their communication, it is important to use a clear and professional language to avoid misunderstandings. Open data are an advantage for use and reuse of data in the society. What is sometimes not communicated in their discussions is that official statistics today are open almost by definition. The UN Fundamental Principles of Official Statistics [6] and all quality frameworks for such statistics emphasize easy access for all.

The request for more open data may indicate a need for more relevant and disaggregated statistics. Access to source data used in the production of official statistics can normally not be open to all because of necessary confidentiality rules. They may be accessible, with specific restrictions, for research or other specified purposes. A lot of work is going on to improve access to more data by anonymization and advanced technical solutions.

3. Main trends in official statistics

Some trends in the development of official statistics over the last 30 years are described in this chapter. Some of the developments are mentioned by Sæbø [7] as examples of developments in quality work in statistics during the period of the European conferences on official statistics (Q-conferences) from 2001 to 2016.

3.1. Official statistics as a public good

The understanding that official statistics are a public good, available for free for all users, is a relatively new development. The UN Fundamental Principles of Official

Statistics (UNFPOS) [6] first adopted by the UN Statistical Commission in 1994 and reaffirmed in 2014, states that official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy, and the public with data about the economic, demographic, social and environmental situation. UNFPOS is a basis for the more recent quality frameworks such as the European Statistics Code of Practice (ES CoP) [8] and the UN National Quality Assurance Framework (UN NQAF) [9]. Official statistics as a public good is also included in new statistical legislation.

The quality frameworks for official statistics provide coherent and holistic systems for statistical quality management, covering the statistical production process from the statistical system and the institutional environment through the production processes, to the resulting output. They demonstrate that quality is a multidimensional concept, which was not so obvious in the last century where the quality of statistics often was regarded as accuracy alone.

3.2. Requirements for official statistics

Historically, there has been no internationally agreed definition of official statistics. It is a national responsibility to define the scope of a country's national statistical system, and hence defining and delimiting its official statistics. However, in many countries the label "official" has pointed at quality statistics produced by central public institutions, foremost the NSI. A modern definition describing requirements for official statistics can be found in the manual of UN NQAF [9]. Key points in this definition are official statistics of public interest, disseminated as a public good, and in compliance with requirements of UNFPOS and quality frameworks.

In the European Union and the European Economic Area including Norway, the European Statistics Code of Practice (ES CoP) is the basis for the quality requirements for official statistics. These requirements are reflected in new statistical legislation both in the European Union and nationally.

In Norway, the requirements for official statistics are given in the Statistics Act [10]. Such statistics need to be of general interest and fulfil the quality requirements from the ES CoP. This separates official statistics from sectorial management information.

Professional independence and impartiality are pillars in the Norwegian Statistics Act. It is not easy to measure professional independence and impartiality, but there are some schemes or measures that should be in place to fulfil these principles. The producers themselves shall decide how their statistics are to be produced and published, and when they should be disseminated. In ES CoP this is formulated by indicators that support independence and impartiality:

- Statistical release dates and times are pre-announced – in other words, equal and simultaneous access for everyone. A good practice is a release calendar where the publishing date is fixed at least 3 months in advance.
- Information on data sources, methods and procedures used is publicly available.

Official statistics must be clearly visible and easily accessible on the websites of the responsible institutions.

3.3. WWW

Another shift was caused by the introduction of the Internet, which has led to great improvements in accessibility of statistics. Several NSIs including Statistics Norway reacted quickly to the new possibilities and launched their first websites early in 1995. In the annual plan for Statistics Norway written during the autumn 1994 this was not even mentioned! About 10 years later, the Internet was the main channel for disseminating statistics, as news, tables and publications, in addition to databases where users could specify their own tables and download statistics. Today, statistics can be transferred in different formats and by machine-to-machine transfer through the application programming interface (API). It has improved transfer of data between systems (interoperability), and this has had great significance for statistics as open data. The NSIs' use of social media for spreading or referring to statistics is common.

3.4. Use of secondary data

During the last 30 years there has also been a shift in data sources for statistics, first from statistical sources (data collected for the purpose of statistics) to administrative sources (data registers developed for public administration), then to other sources including what is denoted as new sources including big data. While the first European

Q-conference in 2001 treated quality of administrative registers only in a session on business registers and macroeconomics, there were 5 sessions devoted to this in Q2014 [7]. In Q2022 in Vilnius almost all sessions treating data collection focused on utilising new data sources, in addition to administrative data. New data sources dominate the discussion on development in different working groups or task forces in the international statistical community.

In Norway, to a large extent, official statistics are based on administrative data systems or registers. Statistics Norway uses more than 100 such data sources from more than 30 public institutions as a basis for its production of statistics. Statistics Norway has agreements of cooperation with these institutions, and structured quality reports exist for all registers used for production of official statistics. Recent examples on the use of new data sources are mentioned in chapter 5.3.

3.5. From NSIs to statistical systems

NSIs are not sole producers of official statistics, but the NSIs normally have a coordination role that comprises quality assurance. The European Statistical System currently consists of Eurostat, 31 NSIs and 285 other national statistical authorities². In the present round of European peer reviews coordination is a main issue.

According to the Norwegian Statistics Act [10], Statistics Norway shall coordinate all development, production and dissemination of official statistics in Norway, and produce an annual public report to the Ministry of Finance on the quality of official statistics. The Ministry has appointed a Committee for Official Statistics, led by Statistics Norway. The members mainly represent authorities who are responsible for official statistics or hold administrative data systems that are important for official statistics. The Committee shall contribute to the quality of official statistics and an effective national statistical system.

Based on the Statistics Act, Norway has a national programme that defines and delimits official statistics. The programme is drawn up by Statistics Norway, in consultation with the Committee for Official Statistics. The first programme is valid for the period 2021 – 2023. Statistics Norway and 10 other public authorities have the

² According to Eurostat website <https://ec.europa.eu/eurostat/web/european-statistical-system>

responsibility for official statistics. The members of the Committee exchange experiences and develop competence on statistical matters, also through a subgroup that discusses methodological issues.

A new programme covering the period 2024 – 2027 is expected to be adopted by the Government in 2023. For more details about the coordination and quality assurance through the Norwegian programme for official statistics, see Sæbø and Andersen [11].

4. Challenges

Some challenges have already been considered in chapter 1 on the data revolution. This includes the competition from new producers of data and statistics, and the difficulties for the public to differentiate the continuous data flow from official statistics with specific quality requirements.

Statistics from new producers based on new data can often be timelier and produced at a lower cost than traditional official statistics, though possibly less accurate since there may be methodological challenges linked to coverage and representability. User surveys and focus groups often indicate that timeliness is the main quality challenge of official statistics today given that it is relevant. Low frequency of published statistics is a related challenge.

Official statistics traditionally change slowly to ensure coherence and comparability. Rapid changes in user needs can therefore be a challenge. But to stay relevant official statistics must be able to adapt quickly. There is often a need for improved granularity, and this must be solved without compromising statistical confidentiality. Access to privately held data such as big data may be a challenge for NSIs. Such access depends on proper (normally new) statistical legislation.

Funding of official statistics is a challenge in many countries. New developments are often dependent on additional funding. Many NSIs have a considerable part of their funding from other sources than the normal government funding. These sources may be public institutions if not directly from the market. Some such funding is valuable by ensuring user orientation of official statistics. However, such funding is vulnerable. In addition, the balance between earmarked appropriations and the professional

independence might be an issue dependent on the scope of the funding, see Sæbø and Holmberg [12].

5. The way forward

Issues linked to how to meet the challenges to official statistics are considered below.

5.1. Core values

There has been a discussion in the international statistical community about core values for official statistics, following the 69th plenary session of the Conference of European Statisticians (CES) in 2021 [13]. The aim of selecting and communicating core values is to promote trust in official statistics. A limited number of core values should be used to emphasize the essence of the quality standards of such statistics, and be suitable to differentiate official statistics from other statistics and data.

However, some values should also reflect what we want official statistics to be.

A set of core values was endorsed by the CES session in 2022 [14], following a proposal from a task team:

- Relevant
- Impartial
- Transparent
- Professionally independent
- Respects confidentiality
- Collaborative

Official statistics must be relevant. The values covering impartiality, transparency, professional independence and protection of privacy are largely specific for official statistics, while collaborative points at the direction producers of official statistics should go.

All the core values are anchored in newer statistical legislation and quality frameworks.

5.2. Quality improvements

The NSIs must continuously strive to respond to new user needs and improve the relevance of official statistics. This can imply that existing statistics are produced and disseminated with higher periodicity or granularity, or that new statistics are developed to illuminate new areas. Improved timeliness is particularly important, which calls for utilizing new data sources. The Covid-19 pandemic has led to several new statistics, and improvements of existing statistics. Statistics Norway has for example started to publish weekly statistics of deaths and daily statistics of bankruptcy. These statistics might be regarded as experimental when first published, in the sense that they did not have full maturity in terms of other quality dimensions than timeliness (accuracy may for example be affected by time lag). More use of experimental statistics should be considered by the NSIs.

The visibility of official statistics should in general be improved. Such statistics are associated with the NSIs, and with many other producers this need has become more pertinent. In Norway this has been addressed in connection with the programme for official statistics.

5.3. Innovation and development

Larger improvements require innovation. It is not enough to do things right; one also must do the right things. The examples in chapter 3 on trends demonstrate that new technology and developments outside the statistical community may offer possibilities for major steps forward. It is important that an NSI reserves sufficient resources for innovation and experimenting.

The Norwegian Statistics Act grants access to all privately held data for use in official statistics. Still, it is important for Statistics Norway to cooperate with the companies holding data. Statistics Norway has during the past few years been in dialog with different owners of privately held data. One example has been the acquirement of purchase receipt data and debit card data for use in the household budget survey. Other applications of new data comprise the use of scanner data for the consumer price index (CPI), electricity statistics based on data from an electricity data hub with information from all metering points, including “smart meters”. Statistics Norway has started a collaboration with Nordic colleagues regarding the use of Mobile Network Operator (MNO) position data.

5.4. Coordination and cooperation

Most NSIs are responsible for coordinating the national statistical system. This requires comprehensive collaboration with other producers of official statistics. This function has been strengthened in new legislation and quality frameworks for statistics such as the ES CoP.

However, collaboration with other producers of statistics should go beyond European and national official statistics. This applies to both new statistics that may be regarded as official in the future, and other statistics. A good climate for collaboration implies that other producers and non-official statistics are recognized. In addition, collaboration with data holders is crucial. Collaboration with academic institutions and research institutes is often a precondition for successfully exploiting new data sources. International cooperation is important in this context.

Cooperation with fact-checkers and the open data movement are other examples of possibilities that should be exploited for the benefit of official statistics and their users. Cooperation is also central when considering extended roles for the NSIs.

5.5. New roles - data stewardship?

There is a discussion in the international statistical community on the role of NSIs as *data stewards*. Several working groups and meetings have had this on the agenda during the last years, and the discussions continue. A group of countries, led by Estonia, prepared a paper on this topic for the Conference of European Statistics (CES) in 2019 [15]. This paper provides an overview of issues linked to the role of NSIs in public data governance and concludes that there is a need to find out what countries are currently doing or plan to do, aiming at considering the development of generic guidelines on the role of NSIs in the new data ecosystem.

Case studies presented in a background paper for the UN Statistical Commission in March 2021 [16], revealed a range of approaches to data stewardship with different levels of involvement from the NSIs. A working group has been active since then. At the same time there is a task force on the same topic organized by the UN Economic Commission for Europe (UNECE).

In brief, data stewardship can be defined as the responsibility to manage a data ecosystem, in cooperation with data providers, users (of data and statistics) and other stakeholders including possible other data stewards. The objective is to improve the use of data and statistics in the society. A data ecosystem can be defined as a system in which several actors interact with each other to produce, exchange and utilize data, statistics and analyses.

There are several case studies, among them also one carried out by Eurostat in 2021. While few NSIs have the full responsibility for all public data as a national data steward, many hold this role for data used for official national statistics, including access to administrative data, sharing of data, privacy protection and coordination of the national statistical system. Almost all European NSIs collaborate extensively with other government bodies on issues related to data and information management processes. This may be prescribed in the statistical legislation, but it varies between countries.

This is in line with the experiences and plans for Statistics Norway. One example is the “a-melding” (electronic dialogue with employers) based on a machine-to-machine solution. Statistics Norway collaborated with other public authorities to develop the solution which has simplified the employers’ communication with the different authorities.

To take a more active role in such collaboration, quality assurance, data sharing, teaching, and explaining statistical literacy in the society is natural for the NSIs. Most of these extensions of the role of an NSI are relevant regardless of the legislative basis. However, legalisation and organisation of the public sector may delimit *how far* the role can be extended.

6. Conclusions

The paper considers and recommends the following measures to ensure that official statistics will continue to be fit for purpose and matter in a world of data:

- Stick to the core values of official statistics based on new statistical legislation and frameworks for quality assurance of official statistics.
- Promote both continuous quality improvements (e.g., timeliness and granularity of statistics), and innovation.
- Effective coordination of the national statistical system.

- Collaboration as a key linking the core values, innovation, coordination, and new roles of the NSIs.

A collaborative NSI open to other actors, but still true to the core values of official statistics, is a preferable approach to ensure or improve trust by extending its role in the direction of data stewardship. More expanded roles as data stewards may comprise extended collaboration, quality assurance, standardisation, data sharing and promoting data and statistics literacy in society.

Funding of official statistics may be uncertain. Collaboration is probably the key also to sufficient and predictable funding!

References

- [1] Davies, W., How statistics lost their power – and why we should fear what comes next, The Guardian 19 January 2017. Available at <https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy>
- [2] MacFeely, S., In search of the data revolution: Has the official statistics paradigm shifted? Statistical Journal of the IAOS 36, 2020, 1075-1094. Available at <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji200662>
- [3] Radermacher, W. J., Governing-by-the numbers – Resume after one and a half years, Statistical Journal of the IAOS 37, 2021, 701 – 711. Available at: <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji210819>
- [4] Ljones, O., Pandemic and official statistics; some comments on recent COVID-19 experiences, Statistical Journal of the IAOS 36, 2020, 285 – 290. Available at: <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji200672>
- [5] Radermacher, W. J., New and Emerging Methods. Standardisation and Statistics. The Survey Statistician, 2021, Vol 84, 24-31. Available at: http://isi-iass.org/home/wp-content/uploads/Survey_Statistician_2021_July_N84_04.pdf

- [6] UNFPOS, Fundamental Principles of official Statistics. Resolution adopted by the General Assembly on 29 January 2014. Available at: <https://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf>
- [7] Sæbø, H.V., Quality in Statistics – from Q2001 to 2016, Statistika - Statistics and Economy Journal - No. 3/2016: 72 - 79. Available from: <https://www.czso.cz/csu/czso/statistika-statistics-and-economy-journal-no-32016>
- [8] Eurostat, European Statistics Code of Practice. Revised edition 2017. Available from <https://ec.europa.eu/eurostat/web/quality/european-quality-standards/european-statistics-code-of-practice>
- [9] UN NQAF, National Quality Assurance Frameworks (NQAF) Manual for Official Statistics (of 2019). Available at: <https://unstats.un.org/unsd/methodology/dataquality/un-nqaf-manual/>
- [10] Act relating to official statistics and Statistics Norway (the Statistics Act). Available at: <https://www.ssb.no/en/omssb/ssbs-virksomhet/styringsdokumenter/statistikkloven>
- [11] Sæbø, H. V. and Andersen, M., Coordination and Quality Assurance through a Programme for Official Statistics: The Norwegian Case, Statistical Journal of the IAOS 37, 2021, 361 – 369. Available at: <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji200738>
- [12] Sæbø, H.V. and Holmberg, A., Beyond code of practice: New quality challenges in official statistics, Statistical Journal of the IAOS 35, 2019; 171 – 178. Available at <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji180463>
- [13] UNECE, 69th plenary session of the Conference of European Statisticians. 23 - 25 June 2021. Papers are available from: <https://unece.org/statistics/events/CES2021>
- [14] UNECE, Core values of official statistics. Published 09/06/2022. Available from: <https://unece.org/statistics/documents/2022/06/working-documents/core-values-official-statistics>

- [15] UNECE, The role of national statistical systems in the new data ecosystem
Paris 26 – 28 June 2019. Available at:
https://unece.org/fileadmin/DAM/stats/documents/ece/ces/2019/ECE_CES_2019_25-1906075E.pdf
- [16] UN, Approaches to data stewardship. Background paper for the 52 session of the Statistical Commission 1 – 3 and 5 March 2021. Available at:
<https://unstats.un.org/unsd/statcom/52nd-session/documents/BG-3a-DataStewardship-E.pdf>