



# Economic and Social Council

Distr.: General  
11 April 2019

Original: English

---

## Economic Commission for Europe

### Conference of European Statisticians

#### 67<sup>th</sup> plenary session

Paris, 26-28 June 2019

Item 6 of the provisional agenda

#### **Emerging role of national statistical offices as offices for statistics and data**

### **The role of national statistical systems in the new data ecosystem**

**Note by Estonia with contributions from Canada, Ireland, Italy, the Netherlands, Poland and the UNECE Secretariat**

#### *Summary*

The Bureau of the Conference of European Statisticians asked a group of countries, led by Estonia, to prepare a paper on this topic. The Bureau discussed the first draft at its meeting in February 2019 and made several comments and suggestions. The authors have incorporated this feedback in the current version of the paper, which is now presented to the 2019 Conference of European Statisticians as a contribution to the seminar on “Emerging role of NSOs as offices for statistics and data”.

The Conference is invited to express views on the way forward.



## I. Introduction

### A. Objective and motivation of the paper

1. The objective of this document is to address the challenges facing national statistical systems (NSS) and national statistical offices (NSO) regarding the development of national data ecosystems. These data ecosystems typically comprise many different sorts of data, statistical, administrative, geospatial and other new sources including “big data”. Data ecosystems are driven by new possibilities for linking these diverse data sources to create valuable information assets for policy makers and the general public.
2. This document aims to address WHAT are the strategic questions concerning the new role of NSOs in the data ecosystem, WHERE should NSOs exercise leadership, and HOW should NSOs respond to these challenges in order to fulfil the expectations of different user groups, from policy makers to citizens. The authors recognize that in different countries, national statistical systems currently have different roles and capabilities, and there are different expectations by policy makers and society about what the new role of NSOs should be. The differences between countries in terms of general economic development and use of information technology have a substantial impact on how NSOs might see their role.
3. This document was prepared in response to a request from the Conference of European Statisticians (CES) Bureau meeting on 14–15 October 2018. It is intended for discussion at the CES seminar on the emerging role of NSOs as offices for statistics and data, in June 2019.

### B. Ongoing discussion about the new role of national statistical offices

4. Discussion about the role of NSS and NSO in the new data ecosystems started in the CES Bureau in February 2018. The issues raised in this paper also take into account the outcomes of the relevant meetings of CES, Organization for Economic Cooperation and Development (OECD) and the European Union.
5. In November 2014, the Independent Expert Advisory Group to the United Nations presented the paper “A World That Counts. Mobilising the Data Revolution for Sustainable Development”<sup>1</sup>. This document described the current situation as an ongoing data revolution, which needed to be used for sustainable development. It recommended that the United Nations should develop a comprehensive strategy and a roadmap towards a new “Global Consensus on Data”. This roadmap should include the development and adoption of legal, technical, geospatial and statistical standards. The document highlighted the following issues: “Openness and exchange of data and metadata, including interoperability of data and information systems; demographic and geospatial information, including “geographic semantic” management and exchange; global exchange of information on illicit financial flows; open data and digital rights management and licensing.”<sup>2</sup>
6. OECD has prepared a paper “Which strategies for NSOs in the digital era? Towards ‘Smart Data’ Strategies”<sup>3</sup>. This paper calls for the creation of a smart data framework – something contrasting to the hype surrounding big data. It also addresses the new data ecosystem. This paper sees the role of NSOs as follows: “In the context of the National Statistical System: NSOs are in a natural position to orchestrate a national Smart Data Strategy.”

---

<sup>1</sup> [A World That Counts. Mobilising the Data Revolution for Sustainable Development](#). Independent Expert Advisory Group, United Nations. November 2014.

<sup>2</sup> [A World that Counts](#) p. 21

<sup>3</sup> [Which strategies for NSOs in the digital era? Towards ‘smart data’ strategies](#). 15th meeting of the OECD Committee on Statistics and Statistical Policy, June 2018.

7. Peter Struijs and Sofie de Boer of Statistics Netherlands wrote in their paper “Big data strategies for official statistics” that both “big data” and “data ecosystem”<sup>4</sup> are buzzwords.

8. At its October 2018 meeting, the CES Bureau discussed a paper from Canada and the United Kingdom on strategic partnership framework for official statistics, and a paper from the United Nations Economic Commission for Europe (UNECE) Secretariat on the role of national statistical systems in the new data ecosystem.<sup>5</sup>

## II. Changes in the environment and new data ecosystem

### A. Data revolution and datafication of society

9. The term “data revolution” was used in a report “A World That Counts”, prepared in 2014 by the UN Secretary General’s Independent Expert Advisory Group. It is defined as follows: “A true data revolution would draw on existing and new sources of data to fully integrate statistics into decision making, promote open access to, and use of, data and ensure increased support for statistical systems” and “Most people are in broad agreement that the ‘data revolution’ refers to the transformative actions needed to respond to the demands of a complex development agenda, improvements in how data is produced and used; closing data gaps to prevent discrimination; building capacity and data literacy in ‘small data’ and big data analytics; modernizing systems of data collection; liberating data to promote transparency and accountability; and developing new targets and indicators.”<sup>6</sup>

10. Digitalisation of society’s functions (or environment) changes how the society behaves. We act according to and trust digital forms more and more compared to paper documents; our traces are increasingly digital and may be a source for statistics.

11. The datafication of society refers to the fact that individuals, organizations and non-living objects leave a rapidly increasing amount of digital traces. Sensors are everywhere, the Internet of Things (IoT) is rapidly growing. In fact, nowadays it seems that no movement, no action or transaction, no change can take place without somehow, somewhere, creating data.<sup>7</sup>

### B. Digital infrastructure and data ecosystem

12. The greatest driver of change in the data ecosystem has been the data revolution of the last several years with the increasing digitization of information and the emergence of the concept of “big data”. With the data revolution come the inherent issues of data access rules, data quality issues and privacy rights. One of the results of the data revolution has been the drive towards the management of data as a strategic asset to increase return on investment (ROI). This has resulted in the creation of statistical output by other organizations that are leveraging the availability of new data sources, tools and techniques to create output. The data revolution has also resulted in a change in stakeholders’ expectations – the demand for more open data, available faster, and accessed through web portals, APIs and sharing platforms. The demand for access to data is driven by the fast-paced change in technology, including artificial intelligence (AI) and machine learning, and the increasing data literacy skills of the data analytics and data science communities.<sup>8</sup>

13. Why “smart data” rather than “big data”? As observed in the projects analyzed, value is not necessarily about the data being “big”. Most common way to characterize “big data” is by 3Vs (volume, variety and velocity), sometimes added two more Vs (veracity and value).

<sup>4</sup> Peter Struijs, Sofie de Broe. *Big Data Strategies for Official Statistics*. 2018. 11 p.

<sup>5</sup> *Role of National Statistical Systems in the New Data Ecosystem*. CES Bureau meeting, October 2018.

*Modern Partnership Framework for Official Statistics*. CES Bureau meeting, October 2018

<sup>6</sup> <http://www.undatarevolution.org/data-revolution/>

<sup>7</sup> Peter Struijs, Sofie de Broe. *Big Data Strategies for Official Statistics*. p.2

<sup>8</sup> *Modern Partnership Framework for Official Statistics*. § 5

Value results from taking creative angles on data, sometimes combining existing (small and big) data sources in new ways, as well as tapping into unconventional sources and forging new methods and algorithms. Value also arises from combining the more traditional statistical approaches with the newest data science techniques – rather than opposing them. “smart data”, from that perspective, can be seen as a semantic advocacy for a value proposition that combines and transcends both, bringing together the statistician and the data scientist.<sup>9</sup>

14. The data revolution has been driving change in the data ecosystems for several years, with the increasing digitization of data and the emergence of “big data”. As a result, data has become a strategic asset. Many organizations outside statistics create statistical output using new data sources, tools and techniques. This has also changed stakeholders’ expectations towards the demand for more open data, easily accessible and available faster.<sup>10</sup>

### **III. Strategic questions**

#### **A. What will be the role of national statistical offices and official statistics in society?**

15. Issues raised by previous papers:

- To meet the new demands, the current official statistics business model needs to overcome several gaps:
  - Data sourcing gap – the ability to combine data from traditional sources with the new sources. Access to new data becomes a strategic question that needs a global data sourcing strategy
  - Platforms gap – new platforms are needed to manage the increasing amounts of data that are often hosted remotely (on a cloud). This requires adapting financial models, privacy policies, using emerging AI techniques, etc.
  - Skills gap – new expertise is needed (in data science, modelling, exploration, analytics, AI, etc.).<sup>11</sup>
- Work at the international level is very important. The international statistical community could widen its role to become the international statistical and data community. Partnerships with global organizations can benefit from global coordination.<sup>12</sup>

16. The position of NSOs may be addressed in different topics: legislation, changing functions, the international dimension, and the needs of users from policy-makers to citizens.

#### **B. What are the traditional functions that national statistical offices have to maintain? How the new role might change values and traditions?**

17. Issues raised by previous papers:

- As Struijs and De Broe put it, the value added of official statistics includes:
  - High quality standards of official statistics
  - Full transparency of methods and assumptions

---

<sup>9</sup> Which Strategies for NSOs in the Digital Era? Towards ‘Smart Data’ Strategies. § 12.

<sup>10</sup> Role of National Statistical Systems in the New Data Ecosystem. § 8.

<sup>11</sup> Which Strategies for NSOs in the Digital Era? Towards ‘Smart Data’ Strategies. § 7; Role of National Statistical Systems in the New Data Ecosystem. § 16.

<sup>12</sup> Role of National Statistical Systems in the New Data Ecosystem. § 6.

- Professional independence or submission to independent professional judgement
  - Information becomes public to everyone at the same moment
  - Coherence and comparability.<sup>13</sup>
- Historically, statistical organizations have built their products in vertically organized, fully controlled lines of production based on surveys (households, firms) and reporting from other institutions (for example, NSOs collecting data from line ministries, agencies and local governments; or central banks collecting data from financial institutions). For years, NSOs have strived to (a) increase efficiency of the reporting mechanisms, and (b) gain access to administrative sources to complement these data and possibly reduce the burden, especially on businesses.<sup>14</sup>
18. Official statistics has certain operational traditions, including topics like statistical quality, neutrality, transparency of methods and assumptions, coherence and comparability of statistics, etc. When developing their new role, what is the strategy for NSOs to ensure that these core characteristics will remain important?
19. What would be a viable and sustainable business model for NSOs of the future? In particular, how do we secure the financial resources needed for the realization of our ambitions? What should our market approach and ambitions be?
20. The new role will definitely change the way NSOs work. Is it possible even now to set some new principles or should we depend on the experience of some countries as case studies?

### **C. How will national statistical offices respond to competition from other organisations, and what strategic partnerships are needed?**

21. Issues raised by previous papers:
- The goal of a modern partnership framework is to think about NSS first, and to build on the comparative advantage of partners who share a common goal and vision. Modernized partnership agreements must be flexible, adaptable and be grounded upon a cost effective, efficient, quality driven agenda.<sup>15</sup>
  - The strength of NSOs in partnership arrangements are sound data governance and data stewardship; however, given the fast-paced change in data availability and technology, NSOs must look to strategic partnerships with data providers, technology providers, academics and researchers and media outlets to continue to respond to the ever-changing needs of stakeholders.<sup>16</sup>
  - NSOs might see other institutions dealing with statistics and data analysis as competitors or as partners. The first institution which is competing with NSOs is academia. Of course, there are many joint ventures or research programmes, but still the question of the role of both in the future must be addressed. Another type of organization that NSOs are competing with are the data analysis companies.
22. The issue of partnerships is becoming increasingly important. Partners will include other public service bodies, geospatial and mapping agencies, academic institutions, research facilities and think-tanks; data scientists and engineers; data protection organizations; private sector organizations that may be able to provide training, expertise and software solutions; open source communities, etc.

<sup>13</sup> Peter Struijs, Sofie de Broe. Big Data Strategies for Official Statistics. P. 4.

<sup>14</sup> Which Strategies for NSOs in the Digital Era? Towards 'Smart Data' Strategies. § 15

<sup>15</sup> Modern Partnership Framework for Official Statistics. § 7.

<sup>16</sup> Modern Partnership Framework for Official Statistics. § 14.

23. When trying to solve problems via partnerships, NSOs will act naturally like coordination bodies. This raises the question of data stewardship, sharing of data and its protected and ethical use.

#### **D. How to maintain high trust in delivering quick insights to changes in society?**

24. Issues raised by previous papers:

- Will NSOs of the future function as clearing houses for clean, processed and standardized data sets, made discoverable through open data catalogues and inventories? Will this future be service oriented rather than data oriented, including the provision of data, including master and reference data as a service, analysis as a service and simulation as a service?<sup>17</sup>
- With the digitalisation of our societies, and people's ability to obtain real-time information (reliable or otherwise) on virtually every aspect of life, expectations have also evolved regarding statistics. At least four characteristics mark this new demand:
  - Minimum "time-to-market" is underlying much of the demand – the readiness to accept delays in the provision of official statistics has declined
  - Granularity – beyond timeliness, there is also an expectation that data are granular, for example locally relevant ("how is my community doing?"), that they distinguish different socio-economic groups and that they cover a broad array of social, economic and environmental subject matters
  - A quest for "trusted quality" – confronted with an ever increasing (and overwhelming) data supply, there is an expectation from users that NSOs make sense of the noise associated with "big data", cut through sometimes contradictory information and add value and context to the orphan data that may be dropped in social networks, the media or other new sources
  - Expectations have also risen in regard to a wider range of new data services and products, to which experts (such as academia, NGOs, businesses) or perhaps citizens at large can contribute.<sup>18</sup>

25. Moving NSOs to the next stages include:

- Becoming more user centric, reaching the entire society, providing data in new ways and countering fake news
- Playing a larger role in governance of data, advising the government about strategic issues related to statistics and data, providing expertise for developing government data hubs
- Playing a bigger role in ensuring the quality of information used in decision-making.

26. When providing data services, increased timeliness, close to real-time, is likely to be increasingly important. Data services should also take into account the need for increased granularity.

### **IV. Possible new roles for national statistical offices: general considerations**

27. This section contains some general considerations for NSOs to take into account when considering their new roles in national data ecosystems. They may eventually serve as a basis for developing more detailed guidance. These considerations address NSOs, governments and society as the whole, particularly those who are keen to lead this "data revolution". Special attention is paid to NSOs and the international statistical community as target group,

---

<sup>17</sup> Modern Partnership Framework for Official Statistics. § 33.

<sup>18</sup> Which Strategies for NSOs in the Digital Era? Towards 'Smart Data' Strategies. § 5.

which has both an institutional position and the necessary competencies to advise on data governance principles for other parties.

## A. The role of NSOs in public data governance

28. The legal framework for official statistics should be enhanced to include the new roles of NSOs. The scope of NSS is widening as new data sources become available and society's expectations are changing regarding what might come out of data and statistics.

**Consideration 1:** NSOs could start discussions within the national governments in order to establish a legal system enabling use of data. NSOs should promote the benefits of following the Conference of European Statisticians' *Guidance on Modernizing Statistical Legislation*, particularly regarding:

- Open data, linked open data and official statistics
- Data exchange among producers of official statistics
- Evolving population and housing census and the legal aspects
- Collaboration with central banks
- Integrating geospatial data and statistics
- Statistical offices and government data management.

29. NSOs should seek a leading role in data governance across government. Data governance comprises the following elements:

- Data architecture
- Development of common technical data standards
- Metadata management
- Findability and access to data
- Data quality management
- Information security and privacy
- Information life-cycle management.

**Consideration 2:** NSOs could initiate discussions within national governments on how they could be involved in the elaboration of national data strategies, either in an advisory or leadership role.

**Consideration 3:** NSOs should work closely with the international community in sharing experience in data governance.

30. Data stewardship involves all data governance activities. Some of them are managerial, some are more technical, and some might also include analytical skills. Data stewardship responsibilities and skills include but are not limited to the following activities:

- Supporting high quality and optimized use of data
- Facilitating access to data
- Promoting expertise, skills and data literacy
- Promoting common standards, frameworks and data policies
- Elaborating data strategies, including sharing and collaboration aspects, etc.

**Consideration 4:** NSOs could encourage governments to appoint a Chief Data Officer or other high responsible official, and institutions to appoint their data stewards. Alternatively, a Data Governance Board could be established.

**Consideration 5:** NSOs should participate in promoting data literacy and numeracy, especially concerning statistics.

**Consideration 6:** NSOs should work with other government bodies to develop a system for training data stewards for the public sector. Training could be addressed to business and technical data stewardship as well as take into account the data domain and managerial aspect of stewardship.

## B. National data ecosystems and data services

31. To support national data ecosystems, data architecture principles elaborated in the statistical community, such as the Common Statistical Data Architecture and the Generic Statistical Information Model could be broadened in their scope to be implemented as standards for a national data ecosystems architecture.

**Consideration 7:** NSOs could work with other government bodies to move towards managing and governing data in line with common data architecture principles.

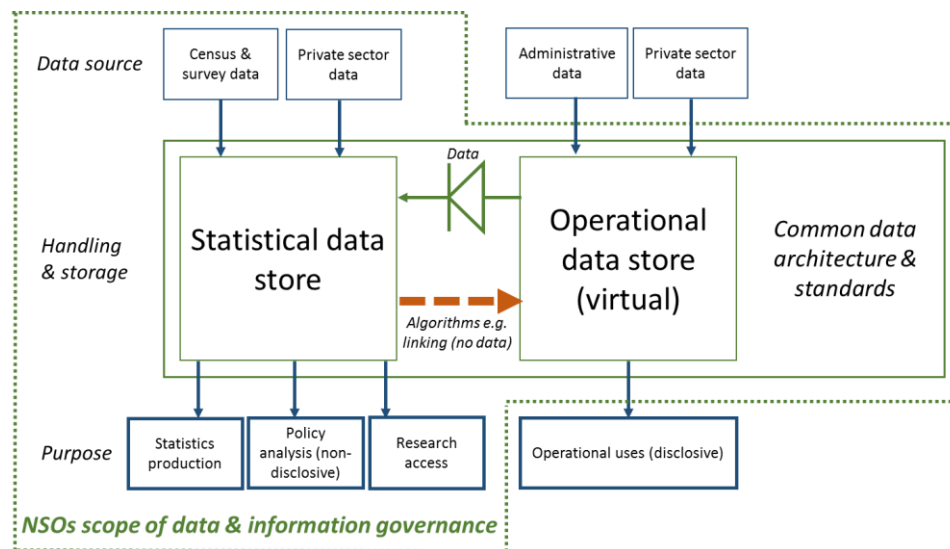
32. Data architectures in NSOs to support data governance and data services need to be agreed to manage different data sources. Figure 1 presents a high-level architecture for data governance. The scope of data governance is defined as a data lake with two zones (stores). In the operational data store, there are individual and aggregate data from non-statistical sources. The statistical data store contains data that are subject to statistical confidentiality requirements. Data can be transferred from the operational data store to the statistical data store, but not vice versa.

**Consideration 8:** NSOs should review their data, application and technical infrastructure to be sure that data are highly secured and protected, whilst facilitating data integration where needed for statistical purposes and data services.

**Consideration 9:** NSOs should communicate that statistical confidentiality will remain in place in the statistical data store.

Figure 1

### High level architecture for data governance using data lake metaphor



33. Data quality includes capturing and documenting metadata (including definitions, rules, logical models, code sets and classifications), as well as identifying data custodians and applying a data quality framework.

**Consideration 10:** NSOs should encourage other government bodies to review their data quality, especially for administrative sources, and take steps for their continuous improvement.

**Consideration 11:** NSOs could participate in developing and implementing data quality standards in different application domains.



34. Users expect a certain level of quality from NSOs, therefore some sort of certification is needed for published statistics, data sets and data services provided by NSOs and other data service providers.

**Consideration 12:** NSOs could be the certification bodies for data held by other organizations that are part of the national statistical system. This certification should be based on the same standards as those applied to the data held by the NSO.

**Consideration 13:** NSOs could participate in open data initiatives and support these with their expertise of data governance and data services.

### C. The next generation of statistics and analysis

35. The data revolution enables increased timeliness and granularity in data analysis. This will facilitate more detailed information on socio-economic and sustainable development indicators.

**Consideration 14:** NSOs should work closely with customers to deliver needed indicators as close to real-time as possible and drop indicators that are not needed.

**Consideration 15:** NSOs should develop analytical capabilities, more specifically in the areas of data literacy, data science and data engineering, to support effective data management.

36. Data science and AI-enabled statistics will be increasingly important.

**Consideration 16:** NSOs should implement data science methodology and tools in the production of statistics and delivery of data services.

**Consideration 17:** NSOs would gain from stronger partnerships with the AI community, both in the academia and business sectors.

### V. Implementation case studies

37. To support NSOs to take a pro-active role in defining and developing national data ecosystems in line with the above considerations it would be useful to share implementation experiences. An additional recommendation is to set up a platform for this, and to develop the necessary template(s) to hold the information in a structured way.

38. To support NSOs it is also recommendable to draw up common data strategy roadmap. Ireland, Canada, and Estonia are already implementing their respective strategies: on Irish Public Service Data Strategy, Canadian Data Strategy Roadmap for the Public Service, Estonian Data Governance Strategy. Overall, the strategies are very similar, and they represent work-in-progress that set direction and provide targeted end-state. They are closely aligned on most themes, including:

- Vision, how data must be treated
- Privacy and transparency, how data should be understood and used
- Governance and standards, covering roles and responsibilities
- Analytics and data science
- Infrastructure
- Management of digital collection and providing data services.

### VI. The way forward

39. The paper was discussed at the CES Bureau meeting in February 2019, and some recommendations were made on the way forward.

40. As a result, the authors of the paper propose a series of next steps as follows:

- A light survey of CES members, to find out what countries are currently doing or plan to do
  - Based on the answers to the survey, selected countries could be invited to share in more detail information about their approaches and lessons learned in implementing their data governance strategy. This information could be shared in the form of case studies
  - CES could consider the possibility, based on the information in the case studies, to develop generic guidance on the role of NSOs in the new data ecosystem.
41. CES is invited to:
- Comment on the issues and considerations presented in the paper
  - Comment on the proposed way forward.

## VII. Sources

- Role of National Statistical Systems in the New Data Ecosystem. Paper for the CES Bureau meeting, October 2018. [https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/October/04-Redefining\\_national\\_statistical\\_systems.pdf](https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/October/04-Redefining_national_statistical_systems.pdf)
- Modern Partnership Framework for Official Statistics. Paper for the CES Bureau meeting, October 2018. [https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/October/08-Framework\\_for\\_establishing\\_partnerships\\_approved.pdf](https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/October/08-Framework_for_establishing_partnerships_approved.pdf)
- Which Strategies for NSOs in the Digital Era? Towards ‘Smart Data’ Strategies. Paper for the 15th meeting of the OECD Committee on Statistics and Statistical Policy, June 2018. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP\(2018\)7&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP(2018)7&docLanguage=En)
- Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service (Government of Canada, September 2018) [https://www.canada.ca/content/dam/pco-bcp/documents/clk/Data\\_Strategy\\_Roadmap\\_ENG.pdf](https://www.canada.ca/content/dam/pco-bcp/documents/clk/Data_Strategy_Roadmap_ENG.pdf)
- Peter Struijs, Sofie de Broe. Big Data Strategies for Official Statistics. 2018. <http://www.dgins2018.ro/wp-content/uploads/2018/10/20-NL-DGINS-Session-2-paper-Struijs-Big-Data-Strategies.pdf>
- Development Co-operation Report 2017. Data for Development (OECD, October 2017) [https://www.oecd-ilibrary.org/development/development-co-operation-report-2017\\_dcr-2017-en](https://www.oecd-ilibrary.org/development/development-co-operation-report-2017_dcr-2017-en)
- A World that Counts. Mobilising the data revolution for sustainable development. Independent Expert Advisory Group, United Nations. November 2014. <http://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>
- Data ecosystems for sustainable development. An assessment of six pilot countries. UNDP, September 2017 <http://www.undp.org/content/dam/undp/library/Sustainable%20Development/Data%20Ecosystems%20for%20Sustainable%20Development.pdf>
- UNECE Conference of European Statisticians "Guidance on modernising statistical legislation" [https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2018/CES\\_6\\_Comm\\_on\\_elements\\_of\\_statistical\\_legislation\\_Guidance\\_for\\_consultation\\_for\\_upload.pdf](https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2018/CES_6_Comm_on_elements_of_statistical_legislation_Guidance_for_consultation_for_upload.pdf)  
[Note – link to be updated when new publication is available]