

Statistical Commission
Fifty-fourth session
28 February – 3 March 2023
Item 4(c) of the provisional agenda
Items for discussion and decision: Data Stewardship

Background document
Available in English only

**Update on the workstreams of the Working Group on Data Stewardship
(WGDS)**

Prepared by Working Group on Data Stewardship

Table of Contents

Data Governance: workstream 1 (WS1)	2
Equity and inclusion: workstream 2 (WS2):	11
Data stewardship and the city data agenda: workstream 4 (WS4)	19
Annex I: Organisational membership of each workstream	21

This background document provides some supplementary reading on three of the five workstreams of the Working Group on Data Stewardship (WGDS), covering workstream 1: data governance, workstream 2: equity and inclusion and workstream 4: links with the city data agenda.

Data Governance: workstream 1 (WS1)

Lead(s): Statistics Poland/ World Privacy Forum (see Annex I for other members)

1. Ideas regarding potential approaches for data governance, data stewardship, and data use by National Statistical Offices (NSOs) as well as other public sector institutions has crystallized into new thought over the past several years. This maturation has opened fresh strategic avenues for NSOs to consider as they seek to provide responsible leadership as modern data institutions and data stewards. A set of varied governmental and multi-stakeholder responses have grown from pandemic-era governance efforts, created in an endeavor to map an ethical pathway through complex data ecosystems and their governance. This has provided NSOs with opportunities for policy development and leadership.
2. During the Covid-19 pandemic, expanded use of personal and non-personal data held by the private sector surfaced as a prominent trend, one that put significant tension on how NSOs and other government-sector institutions could balance the need for data use for the public good, along with the concomitant need to govern data ethically.¹ The frictions resulting from a lack of legislative guardrails regarding government use of personal data held by the private sector did not begin in 2020. However, the pandemic exacerbated the already existing frictions and brought public awareness of the difficulties to the point that work toward resolution was necessary.² One such effort has resulted in a new and pivotal set of multilateral principles regarding government access to and use of personal data held by the private sector in the law enforcement context.³ These principles have significant implications for NSOs.
3. Also, during the pandemic, governments — which were experiencing an acute need for faster, better data analysis — raced to craft comprehensive national artificial intelligence strategies and frameworks to facilitate the development of better analytical engines with which to address national crises.⁴ In response to this and

¹ *Data Stewardship and the Role of National Statistical Offices in the Changing Data Landscape*, High Level Event, United Nations Statistical Commission 52nd Session (2021) Side event. <https://unstats.un.org/unsd/statcom/52nd-session/side-events/20210210-1M-data-stewardship-and-the-role-of-NSOs-in-the-changing-data-landscape/> . See also: *UN/DESA Policy Brief #89: Strengthening data governance for effective use of open data and big data analytics for combatting COVID-19*, United Nations Department of Economic and Social Affairs, <https://www.un.org/development/desa/dpad/publication/un-desa-policy-brief-89-strengthening-data-governance-for-effective-use-of-open-data-and-big-data-analytics-for-combating-covid-19/> .

² *Data for international health emergencies: Governance, operations, and skills, Statement of the Science Academies of the G-7*, The Royal Society, 31 March 2021. <https://royalsociety.org/-/media/about-us/international/g-science-statements/G7-data-for-international-health-emergencies-31-03-2021.pdf?la=en-GB&hash=E6AE132F624E4C01EB10BFC25CB83F1C> .

³ *Declaration on Government Access to Personal Data Held by Private Sector Entities*, OECD, December 2022. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0487>

⁴ OECD AI Observatory, Live Repository of National AI Strategies. OECD. <https://oecd.ai/en/dashboards/overview>.

other factors, a fresh body of global work regarding ethical principles for artificial intelligence was completed during the pandemic. Some of this work has shaped new ethical recommendations around health data and artificial intelligence,⁵ some of this work has created new normative soft law instruments around the ethics of artificial intelligence more generally.⁶ The recent multilateral work on the ethics of artificial intelligence provides new thinking for NSOs as they redefine their roles and ethical leadership within broader government AI work. While NSOs already have a body of well-established ethical norms regarding their statistical work,⁷ the new ethical norms that have emerged regarding the use of public sector data and artificial intelligence can potentially expand NSOs' ethical canon and facilitate definitional work on what data stewardship means for NSOs. It is through examining ethical considerations and principles that NSOs can most clearly elucidate their roles as data stewards.

4. The activity around creating ethical data guidance during the pandemic is taking place against the background of ideas around data governance that for the first time in modern history are normative nearly globally. These norms, created in part by the near- global adoption of Europe's General Data Protection Regulation (GDPR) legislative model, also created a problematic vacuum for NSOs by virtue of routinely exempting them from much or part of their purview, as discussed in the WS1 2022 background report and in other fora.⁸
5. It is pertinent to articulate how data governance norms interact with NSOs. Throughout the world, formal legislation embodying data governance norms have historically shifted in slow-moving regional waves, beginning most recently in the 1970s with the enactment of the first modern data protection law in Hesse, Germany,⁹ followed by the Fair Credit Reporting Act in the U.S. in 1974,¹⁰ followed by six decades of legislative activity pertaining to data governance.¹¹ Today 158 jurisdictions and counting have passed modern versions of data governance legislation.¹² Many if not most jurisdictions have at this point passed either a version of the governance model established by Europe's General Data Protection Regulation (GDPR), or a legislative model influenced by it. This can be seen in

⁵ *Ethics and governance of artificial intelligence for health*, World Health Organization, 28 June 2021. <https://www.who.int/publications/i/item/9789240029200> This report developed six consensus principles as well as recommendations for the public and private sectors.

⁶ *Recommendation on the Ethics of Artificial Intelligence*, UNESCO, 23 November 2021. <https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence> .

⁷ *Fundamental Principles of Official Statistics* (A / RES/ 68/ 261 from 29 January 2014), United Nations Statistical Commission. <https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx> .

⁸ *Data stewardship and the role of National Statistical Offices in the changing data landscape*, UN Statistical Commission High Level Event, UN Statistics, February 2021. <https://unstats.un.org/unsd/statcom/52nd-session/side-events/20210210-1M-data-stewardship-and-the-role-of-NSOs-in-the-changing-data-landscape/>.

⁹ *Hessisches Datenschutzgesetz (Data Protection Act)*, 30 September 1970. <https://starweb.hessen.de/cache/GVBL/1970/00041.pdf> .

¹⁰ *Fair Credit Reporting Act of 1974*, 15 U.S.C. §1681. https://www.ftc.gov/system/files/documents/statutes/fair-credit-reporting-act/545a_fair-credit-reporting-act-0918.pdf .

¹¹ *Data Protection and Privacy Legislation Worldwide*, UNCTAD, as of 2023. <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide> .

¹² Count derived by authors, based on multiple data sources. See: Graham Greenleaf, *Now 157 Countries: Twelve Data Privacy Laws in 2021/22*, Privacy Laws & Business International Report 1, 3-8, UNSW Law Research, 15 March 15, 2022. <https://ssrn.com/abstract=4137418> . The count of 158 includes the addition of a bill passed recently in Indonesia. India and Bangladesh have proposed data governance and privacy bills.

models from Africa (Mauritius, Ghana, South Africa, and others) from India (the draft Privacy Bill of 2022) from Asia (Japan, Korea, Mongolia, and others) from Latin America and the Caribbean (Brazil, Mexico, Ecuador, and others).

6. The data governance norms set forth in Europe's GDPR were built atop an older European data protection law, Directive 95/46/EC, the Data Protection Directive originally passed in 1995.¹³ After lengthy deliberation, the GDPR, with its robust set of modern norms, was adopted in 2016 and went into effect in 2018. Thereafter the norms enshrined in the GDPR quickly swept through the business world as de-facto standards for the treatment and processing of data. These norms focused on fair and lawful processing, a right to transparency, a right to be informed, a right to access, a right to erasure, and a right to restrict processing, among other rights.
7. Something that was not well-understood by the public prior to the pandemic was that large and important tranches of data were typically exempted from these rights granted under GDPR and GDPR-like legislation.¹⁴ Notably, public health data in national health emergencies as well as government statistical data and data relating to national security are routinely exempted from data governance legislation to varying degrees.¹⁵ These exemptions and others have something in common: the use of data is exempted in certain circumstances from regulatory guardrails so that data can be more easily utilized for the benefit of society. During the pandemic, data governance guardrails were further loosened to facilitate greater use of private sector data, again for public benefit.
8. In 2022, our workstream noted that while the pandemic had increased private sector data-sharing activity, the use of private sector data was coming under increasing legal and public scrutiny.¹⁶ Early surveys of NSOs regarding this topic indicated at the time that although the pandemic initially increased the statistical uses of private sector data under the auspices of the health emergency, legal barriers were still significant enough that expanded uses were not seen as likely to continue at the same rate after the emergency use authorizations were lifted.¹⁷
9. It is possible that had the Covid-19 pandemic not occurred, that the legal and normative barriers regarding government use of private sector data would have remained in place and would likely have evolved much more slowly, over time, taking at least a decade or more prior to advancing toward measurable change. However, the pandemic did occur, and it has had a pronounced effect on data governance norms in that there has been a marked shift toward normalizing increased data use,

¹³ *European Parliament and Council Directive 95/46/EC* of 24 October 1995, Protection of personal data. <https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:31995L0046>

¹⁴ *General Data Protection Regulation, Chapter 9, Art. 89*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679>. European member states may have additional derogations regarding statistical work. See: *Derogations Tracker*, GDPR Resource Center, Lathan & Watkins. <https://gdpr.lw.com/Home/Derogations> .

¹⁵ *The History of the General Data Protection Regulation*, European Data Protection Supervisor. https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en.

¹⁶ *In-depth review on collaboration with private sector data providers*, Poland (lead), Canada, Mexico, UK, Eurostat and IMF, pre-publication copy. ECE/CES/2022/.

¹⁷ *In-depth review on collaboration with private sector data providers*, Poland (lead), Canada, Mexico, UK, Eurostat and IMF, pre-publication copy. ECE/CES/2022/.

particularly regarding government use of data, including personal data, held by the private sector.

10. At the same time as these developments have been occurring, there has also been a movement toward a formal reassessment of the GDPR, at least in part. In June 2022, the European Data Protection Supervisor, Wojciech Wiewiórowski, foreshadowed in a now famous editorial:

“I’m glad that what was once taboo — merely acknowledging there may be structural issues behind the malfunctioning of the GDPR — is now not only being internalized, but ideas of how to address them, including potential legislative initiatives, are being shared and floated by more voices and stakeholders. We need genuine debate on whether current data protection legislation — proposed 10 years ago, adopted six years ago, and in application for the last four years — is actually serving people in the way it was designed to. Whether it protects everyone equally and sufficiently; whether the intention behind it has been fulfilled.”¹⁸

11. As of 2023, the European Commission has published an initiative to further specify rules relating to the GDPR regarding enforcement.¹⁹ This step by the Commission is notable coming so soon after enactment for such an influential piece of legislation, which typically moves at a much slower pace. Together, the development of soft law around artificial intelligence as well as government sector use of private sector data and a partial re-thinking of the GDPR has opened a new conversation around modern data uses.

The issue of government access to personal data held by the private sector as a change agent

12. Government access to and use of private sector data, even when seen as more permissible during the pandemic, has been deeply controversial.²⁰ In the NSO context, as governments have broadened access to private sector data stores, complex ethical and legal issues have come into play, particularly if data stores involve personal data. Privacy laws do not provide NSOs a complete map through this complexity. Some ideas derived from privacy and data governance laws can help, particularly those around appropriate data processing, security, and handling. However today’s data governance legislation usually also includes core privacy and data governance tenets such as a right of access, a right to erasure, and in some jurisdictions and sectors, a right to restrict disclosures, among other rights.²¹ While

¹⁸ Wojciech Wiewiórowski, *We still need to talk about data protection*, Opinion. Politico, 9 June 2022. <https://www.politico.eu/article/eu-data-protection-gdpr-brussels-regulation-supervision/>

¹⁹ *Proposal for a Regulation further specifying procedural rules relating to the enforcement of the General Data Protection Regulation*, European Commission, Proposed for Q2 2023. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13745-Further-specifying-procedural-rules-relating-to-the-enforcement-of-the-General-Data-Protection-Regulation_en

²⁰ *In-depth review on collaboration with private sector data providers*, Poland (lead), Canada, Mexico, UK, Eurostat and IMF, pre-publication copy. ECE/CES/2022/.

²¹ The Right to Erasure is prominent in most GDPR-influenced regulations. The Right to Restrict Disclosure is granted in some sectoral regulations, such as the health privacy regulation HIPAA in the U.S. See: *Right to Request a Restriction*, U.S. Department of Health and Human Services. <https://www.hhs.gov/hipaa/for->

these are important rights, they also create a direct conflict with some of the activities of NSOs, in that they make the analytical work of NSOs particularly difficult, if not impossible in some instances. For these reasons, NSOs are typically granted meaningful exemptions from aspects of data governance and privacy laws in most jurisdictions. The exemptions may apply unevenly to accessing data held by the private sector for a variety of reasons, which can differ based on jurisdictional rules and context. However, even with exemptions, ethical and other questions abound. The issue of access to personal data held by the private sector is messy for NSOs through and through, and there is not currently a consensus resolution to this issue.

13. As discussed earlier, the newer tranche of ethical guidelines can provide a potential pathway for NSOs. Elements of another potential path toward resolution can also be glimpsed by observing a government sector that is typically exempted from data governance and protection regulation, which is law enforcement and national security. During the pandemic, beginning in December 2020, the OECD's Committee on Digital and Economic Policy (CDEP) created a working group to negotiate consensus principles for "Trusted Government Access to Private Sector Data." The principles were to apply specifically to access to personal data held by the private sector, in law enforcement and national security contexts. This launched a formal, 2-year negotiation process among OECD member countries and the European Union. The OECD process was admittedly fraught, in that negotiations were often tense and suffered notable pauses and setbacks.²² The negotiations occurred primarily between OECD member countries, with some inputs from civil society and business stakeholders.

14. At the OECD Ministerial in December 2022, the member governments launched the completed principles in the form of a Ministerial Declaration.²³ While the OECD principles are not an OECD Recommendation and are therefore not soft law, they are potentially normative as a significant and formal negotiated agreement amongst OECD member countries. The OECD principles are as follows:

- Legal basis,
- Legitimate aims,
- Approvals,
- Data handling,
- Transparency,
- Oversight, and
- Redress.

Each principle is built out in the Declaration with explanations and recommendations.

15. For NSOs, the existence of these negotiated consensus principles regarding government access to personal data held by the private sector - even in a different government context — raises the possibility that NSOs could potentially negotiate a

professionals/faq/right-to-request-a-restriction/index.html . See also: *Article 17, Right to Erasure*, GDPR. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679> .

²² Theodore Christakis, Kenneth Propp, Peter Swire. *Towards OECD Principles for Government Access to Data*, Lawfare, 20 December 2021. <https://www.lawfareblog.com/towards-oecd-principles-government-access-data> .

²³ *Declaration on Government Access to Personal Data Held by Private Sector Entities*, OECD/LEGAL/0487, OECD. Adopted on 13/12/2022. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0487> .

global consensus around ethical principles regarding NSO access to and use of personal information held by the private sector. NSOs have different needs and would need to negotiate quite different principles for statistical purposes.

16. Asked what would facilitate ethical and legal collaboration and use of privately held data for statistical purposes, an in-depth survey of 34 NSOs from 32 countries in 2022 (28 UNECE member countries and 4 countries outside UNECE) listed trust, company corporate social responsibility commitments, communication, mutual understanding of interests between the parties, a flexible approach to data formats or sharing technologies, open communication around data quality and methodology, the NSO's organizational capabilities, an enabling legal framework, and openness for cost competition as important components.

Additional consensus work on principles and implementation models for consideration

17. A great deal of work has been done in adjacent areas that could be relevant for NSOs. The following ethical principles and implementation models have the potential to be assistive to NSOs as they seek to define potential data stewardship roles and responsibilities.

A. Ethical principles relating to artificial intelligence in conceptualizing NSO data stewardship

18. NSOs are well aware of the *Fundamental Principles of Official Statistics*. In considering what defines data stewardship in the NSO context, consideration of additional ethical principles could be of assistance, particularly recent consensus work in the area of artificial intelligence.
19. The World Health Organization crafted, in 2021, a series of ethical principles regarding artificial intelligence, *Ethics and Governance of Artificial Intelligence for Health*. These principles are “aware” of modern digitalization issues as well as the impact of the pandemic and are particularly well-crafted. Some of these principles have excellent potential for adaptation in an NSO context in thinking about data stewardship. The document discusses applications of AI for health, laws, policies, and principles that apply to the use of AI, key ethical principles for the use of AI for health, as well as elements of a framework for governance of AI for health, among other topics.²⁴ The full report contains potentially helpful material for NSOs, particularly the six consensus principles the WHO adopted:
 - Protect autonomy,
 - Promote human well-being, human safety, and the public interest,
 - Ensure transparency, explainability and intelligibility,
 - Foster responsibility and accountability,
 - Ensure inclusiveness and equity,

²⁴ *Ethics and governance of artificial intelligence for health*, World Health Organization, 28 June 2021. <https://www.who.int/publications/i/item/9789240029200> This report developed six consensus principles as well as recommendations for the public and private sectors.

- Promote artificial intelligence that is responsive and sustainable.

The WHO materials are rich, and grapple with multiple research complexities.

20. UNESCO's *Recommendation on the Ethics of Artificial Intelligence* was adopted in November 2021.²⁵ It is soft law, and as such, it should be considered carefully by NSOs with particular regard to artificial intelligence ethical principles, which could form one or more aspects of data stewardship pillars for NSOs. The UNESCO Recommendation is lengthy. This is a straightforward listing of the principles:

- Proportionality and Do No Harm
- Safety and security
- Fairness and non-discrimination
- Right to privacy, and data protection
- Human oversight and determination
- Transparency and explainability
- Responsibility and accountability
- Awareness and literacy
- Multi-stakeholder and adaptive governance and collaboration

21. Both sets of ethical principles, from WHO and UNESCO, have much to offer NSOs in terms of procedure, models of approach and overall content. Not all principles will work for NSOs, but they provide a richly nuanced modern viewpoint.

B. Contextual implementation of ethical principles, and high-functioning models to assess

22. Data stewardship responsibilities for NSOs will differ considerably based on regional and national contexts, along with legal, economic, and digital contexts, among others. Inclusiveness of the wide variety of NSOs in their varying contexts is an important consideration in designing implementation of data stewardship principles. There are several possible models to work from.

Chemical Safety Regulation Models

23. Most countries have adopted some form of chemical safety regulations. The administration and adaptation of chemical safety policies, which are crafted at the national and subnational levels and are built according to a common framework, use the same definitions, and are also harmonized globally while respecting jurisdictional contexts. It is a complex harmonization that has taken years to fully develop.

24. Chemical safety regulations are typically written as omnibus legislative proposals at the national level and present an umbrella under which many types of chemicals can be individually regulated.²⁶ The country-level legal frameworks are then harmonized

²⁵ *Recommendation on the Ethics of Artificial Intelligence*, UNESCO, 23 November 2021. <https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence> .

²⁶ Pam Dixon, *Regulating and harmonizing biometric ecosystems: addressing the full spectrum of risks using global safety models and controls*, Turkish Policy Quarterly, Volume 20, Issue 4. March 2022.

by two multilateral institutions, the World Health Organization and the United Nations. The UN has a program called the **Globally Harmonized System of Classification and Labelling of Chemicals (GHS)**, which is regularly updated.²⁷ The idea of the UN GHS is to bring a global, standardized approach to chemical safety across all jurisdictions. Labeling is to be the same, the level or grade of the risk is to be defined similarly, and risk mitigation strategies would be similarly harmonized internationally. The UN GHS plan is part of the implementation of the UN Sustainable Development Goals (SDGs).

25. For NSOs, the model of UN chemical safety harmonization is a salient model to study. If agreement amongst the NSOs could be achieved regarding core definitions and applications of data stewardship activities, for example, then specifics as to implementation could be contextualized as necessary in each jurisdiction. The UN could facilitate harmonization of core aspects of the stewardship model, such as definitions, as mentioned.

Human Subject Research Models

26. Human subject research has generated significant ethical codes and research standards. Of interest for NSOs regarding human subject research models is how ethical codes, such as the Nuremberg Code and the Declaration of Helsinki, can be articulated into implementable regulations and practices at the country level. There is a significant literature regarding how the ethical codes have been translated into national legislation and implementable regulations across jurisdictions. This literature will be of interest to NSOs and interested parties that seek to understand in very practical terms how to take ethical guidelines and create working implementations from that at the country level.
27. There are more than 1,000 standards including legislation, regulations, ethical guidelines, and other guidance regarding human subject research protections across more than 130 countries. Amidst this preponderance of work, there are a handful of extraordinary beacons which have guided much of the modern ethical standards and research implementations in this area.
28. In the area of ethical codes and standards, the Nuremberg Code,²⁸ the Declaration of Helsinki,²⁹ and the Belmont Report³⁰ are significant.
29. In the area of research standards, the Council for International Organization of Medical Sciences (CIOMS): International Ethical Guidelines for Biomedical Research

<http://turkishpolicy.com/article/1094/regulating-and-harmonizing-biometric-ecosystems-addressing-the-full-spectrum-of-risks-using-global-safety-models-and-controls> .

²⁷ GHS, Rev. 8, (2019). United Nations. <https://unece.org/ghs-rev8-2019> .

²⁸ *The Nuremberg Code*, August 1947. Re-Published in the British Medical Journal, No. 7070 Volume 313, 7 December 1996.

https://media.tghn.org/medialibrary/2011/04/BMJ_No_7070_Volume_313_The_Nuremberg_Code.pdf .

²⁹ *The Declaration of Helsinki: Ethical principles for medical principles involving medical subjects*, Adopted 1964, most recent amendment 2013. World Medical Association. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> .

³⁰ *The Belmont Report*, 1976. <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>.

Involving Humans,³¹ and the International Conference for Harmonization - Integrated Addendum to the Guideline for Good Clinical Practice,³² are important models. The prominent research standards models cited here have been updated over the years; as CIOMS has noted: “After several years, ethical issues emerged for which the 1993 CIOMS Guidelines had no specific provisions,” a sentiment that many researchers and statisticians can identify with.

30. Even though human subject research models stem from a different topical area than NSO statistical research, the models provide practical and useful exemplars of how practitioners, administrators, and legislators have brought important ethical principles to implementation in specific research and country contexts.

³¹ *International ethical guidelines for health-related research involving humans, in collaboration with WHO*, 2009. https://cioms.ch/working_groups/bioethics/ .

³² *ICH Harmonized Guideline - Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice. E6 (R2)* 9 November 2016. https://database.ich.org/sites/default/files/E6_R2_Addendum.pdf .

Equity and inclusion: workstream 2 (WS2):

Lead(s): Departamento Administrativo Nacional de Estadística (DANE-Colombia); Global Partnership for Sustainable Development Data (GPSDD) (see Annex I for other members)

31. The Equity and Inclusion workstream has as its main deliverable one briefing document created through the compilation of case studies which highlight the main existing trends on how National Statistical Offices (NSOs) foster equity and inclusion through their steward role, but also raise possible questions for further research on the matter. Taking this into account, the workstream will engage in the following activities:
- Identify specific research questions to shed light on thematic and sectorial dimensions of the equity and inclusion agenda (indigenous communities, gender, intersectional approach), to support different NSOs understanding how they can build their own approach as data stewards while achieving relevance by meeting users' expectations as a cornerstone of statistical quality assurance. These research questions will guide the case studies.
 - Systematize learning from different national and international approaches that have been implemented to foster equity and inclusion, strengthen NSOs role as stewards, and foster the empowerment of and collaboration with vulnerable communities through better governance schemes.
 - Following the thematic and sectorial dimensions, the workstream will develop guidance on how NSOs can promote equity and inclusion along the data value chain to strengthen their role as stewards, recognizing the realities under which this promotion must be undertaken, and the specific expectations users have regarding the NSOs' role and their own.
32. This document contains guiding questions to draft the case studies, setting a common conceptual framework and aligning members' expectations for the final deliverable. Nonetheless, we do not intend to prevent members from portraying their particular situation and contexts. While the following conceptual framework and the guiding questions are meant to facilitate comparability among cases, we also encourage members to shed light on this new and challenging topic through their unique experiences.
33. Furthermore, this document also contains the most relevant outputs developed by this workstream. Since the consolidation of this workstream, members have drafted two main documents to advance on the deliverables referenced in the Annex II³³ of the report –these deliverables are also mentioned in this document. The first one is a briefing text on the relationship between equity and inclusion to data stewardship. The second document is the suggested terms of reference to guide the countries' case studies drafting process. Both documents are found below.

³³ E/CN.3/2023/20, https://unstats.un.org/UNSDWebsite/statcom/session_54/documents/2023-20-DataStewardship-E.pdf.

Common Basis for the Case Studies

Data Stewardship

34. According to the work developed by workstream 5 on the “overall conceptual framework on data stewardship”³⁴, despite being context specific, both in terms of definition and application, Data Stewardship can be understood in terms of its defining elements:
- Data stewardship includes governing the data ecosystem to improve the use and reuse of data for the public good.
 - Data stewardship is the ethical and responsible creation, collection, management, and use of data.
 - Data stewards, including NSOs, must facilitate coordination and cooperation between data providers and users.
 - Data stewards ensure data protection, standardization, and quality to build and maintain trust.
35. By systematizing these elements, workstream 5 concluded that data stewardship can be defined as follows:

“These [the common elements] include a common understanding of the data assets that are being considered. These assets can be categorized as people (producers of statistics, data collectors, analyzers, and users), technology (technical infrastructure), and processes (governance, laws, policies, and procedures) within a country’s data ecosystem. Effective data stewardship ensures that these elements work harmoniously to increase trust in (and the value, use, and impact of) data for the public good. Within this framework, the data steward provides oversight and guidance, reduces risk, and increases collaboration across the system”.³⁵

Data stewardship and equity and inclusion

36. A data steward has as its main responsibility to guarantee that people, technology, and processes within data ecosystems work harmoniously. In this sense, data stewardship is defined as a tool, a way to approach data governance to ensure we can face the advent of new data users and producers amid the fourth industrial revolution. However, as any other tool, its meaning has to do with the purpose we give it, rather than the instrument itself.
37. The ultimate purpose of data, and therefore of data stewardship, is to improve peoples’ lives. Nonetheless, as any other source of power, not every segment of the population has been able to benefit from data and, as data user and producer

³⁴ Preliminary results from the Working Group on Data Stewardship (Background document to the 53rd session of the UN Statistical Commission:

https://unstats.un.org/UNSDWebsite/statcom/session_53/documents/BG-3b-DSWG-E.pdf.

³⁵ Preliminary results from the Working Group on Data Stewardship (Background document to the 53rd session of the UN Statistical Commission (Page 36)

communities expand, the risk of deepening unequitable power structures increases. In this context, equity and inclusion is nothing more than the idea that we need to unlock the value of data for all, which means concentrating on those marginalized communities that have not traditionally benefited from statistical information.

38. In the context of equity and inclusion, data stewardship involves facilitating data ecosystems in a way that enables marginalized communities to benefit from data. The data steward is in charge of setting up incentives, procedures (formal and informal), ethical standards, technology infrastructure, and rules, among others, to empower marginalized populations. Nevertheless, this way of giving meaning to Data Stewardship comes with great uncertainty: the settings in which NSOs can facilitate those elements is infinite.
39. Inclusive data system features can vary widely both in terms of kind and extent. The Data Values white paper “Reimagining Data and Power: A roadmap for putting values at the heart of data” argues that the way in which data is designed and produced has implications for how people, especially those who are marginalized, are represented and included in data processes and in related decision making. When people have agency in the design, production, and use of data, they can actively engage and influence what and how data is collected and analyzed. Inclusive approaches can maximize benefits, expand agency, and redistribute power, but they must be undertaken systematically so that inclusion becomes embedded across data systems. The paper lays out three features of inclusive data systems: representation, co-creation, and review that can be employed separately or together to foster data agency.
40. Representation means making people visible through data, producing granular information in terms of their income, gender, age, race, ethnicity, migratory status, disability, location and so forth. From the official statistics perspective this could even mean to include these communities for the first time in official statistics, as well as identifying best practices to ensure that they remain included in spite of changes in the administration.
41. Another way to approach equity and inclusion could be through co-creation. Besides making people visible through statistical information and granting access to available data, a Steward could also include those marginalized communities into the discussion about statistical concepts, legal frameworks regarding data and its management, decision-making processes around data, and defining information priorities. Furthermore, an equitable approach could be the inclusion of these populations throughout the statical cycle, defining the information to be collected, being part of the collection and processing phases and finally participating in the dissemination of results.
42. An alternative could also be to promote review scenarios, so communities could provide feedback and contribute to how data is generated, processed, and used in relation to their own experience and needs. This could mean the creation of working arrangements such as committees or task forces that bring together the communities and experts from different disciplines to assess data gaps or biases.
43. Overall, the options available to promote equitable and inclusive data ecosystems, as the very concept of data stewardship, need to be context specific. Each of these

approaches come with advantages and limitations and have to be assessed carefully by stewards considering the needs of the populations they are trying to empower and the pertinence of the options (both in terms of kind and extent) or a combination of them.

Case studies guiding questions

44. To draft the case study of your choice, we encourage you to refer to the following guiding questions, as they might help you determine the content and narrative of the document.
- How did the actions in your case study respond to the three main objectives: identify barriers and enablers to the widespread use of data; fostering that the use of data ends in more inclusive and equitable policies and practices; and integrating marginalized communities into data production at one or more points in the data value chain.
 - Is the problem clearly defined in terms of the assets (people, technology, and processes) and the working arrangements among them?
 - How would you characterize the relation between the data steward with marginalized communities? Is there a legal mandate for your NSO to work with these communities? What are the main takeaways you have from building the relationships with these communities?
 - Are the marginalized communities involved identified and their interests defined in terms of their needs and expectations? The identification of these needs is done by the data steward or another stakeholder in the administration?
 - Is the problem-solving process and rationale explicitly stated and organized by a common criterion (chronological, type of intervention, etc.)?
 - Is the solution to the problem thoroughly explained and defined in terms of its features (representation, co-creation, review)?
 - Are the advantages and limitations of this solution clearly stated in the document, both in terms of its pertinence and plausibility?
 - What were the challenges faced by the NSO in implementing the solution?
 - Are the results and impacts of the intervention or action stated in the text?
 - Are there any lessons learned from this solution that you would like to share?

Open data and the case study

45. During the 53rd United Nations Statistical Commission³⁶, it was decided that the Working Group on Data Stewardship (WGDS) will incorporate the Working Group on Open Data (WGOD) outputs to build on their advancements, as both the open data

³⁶ <https://unstats.un.org/UNSDWebsite/statcom/documents/53>.

principle and the data stewardship concept are closely related. To fulfill that commitment, we ask you to include in the case study, if relevant, any information regarding the following elements provided by the WGod outputs and how they contribute to the promotion of equity and inclusion:

- Show transparency regarding data sets that cannot be released publicly: Under an open by default policy, a “negative list” of data sets which are excluded from open access is issued, rather than a list of specific data sets that are open for public access. If certain data cannot be published, or cannot be published within a specified time period, this should be clearly stated in the dissemination policy of the NSO and announced online at the location where one would generally expect those data to be found.
- Establish a transparent data request process: users may request the release of data not currently available on public databases. NSOs should therefore establish a system with clear and easy-to-understand instructions for requesting access to specific data sets. It is acceptable to charge for the marginal cost of preparing special data requests. If a request is denied, an explanation should be provided, as well as the details associated with any possible appeal process.
- Establish clear procedures for microdata access and publication: because reidentification of individuals can remain a risk, NSOs should develop clear guidance on how microdata is to be published and accessed by users and how they may be used. When disclosure risks have been adequately addressed, microdata should be published as public use files with licenses (often more restrictive in this context) and accountability mechanisms in place to prevent misuse and disclosure of confidential information.
- Build external users’ capacity to use and understand open data: dedicate resources and collaborate with other organizations towards building the data literacy of users and raising public awareness regarding the open data policies of the national statistical office.
- Addressing the concerns of statistics producers: there are many reasons why producers of statistics have been hesitant to release data with a public domain dedication or an open data license as part of a default open data model. This hesitation is associated with many of the challenges those NSOs generally face when opening data that have been discussed at length by third parties to the public. Solutions to these concerns should be explored to deepen the Open by Default Principle. Some of the concerns in this area are prevention of data misuse or misinterpretation; protection of the reputation or integrity of the national statistical office (moral rights); protection of individual privacy; fear of losing revenue.

The role of data stewards to guarantee equity and inclusion

46. The 2030 Agenda has set as one of its main principles the idea that no one can be left behind in the world’s development process. As it is clearly stated by the UN Sustainable Development Group, “Leave no one behind” consists not only to reach the poorest of the poor, but also requires combating inequalities and discrimination, as they jeopardize people’s agency. To comply with this commitment, the

international community have adapted their work and efforts to ensure that data collection, processing and dissemination are reflecting everyone's realities, particularly for the most vulnerable –and usually underrepresented. The quality and pertinence of the insights we deliver to help the most vulnerable depend on our ability to provide and foster inclusion and equity in data collection, production, and dissemination.

47. Despite this pledge, the fulfillment of the principle is directly linked with our ability to overcome the barriers that prevent the most vulnerable to benefit from data. As stated in the World Bank's World Development Report "Data for Better Lives Report", data have a dark side. Data access is becoming increasingly asymmetrical, creating environments that may incentive perverse private and public practices. Market forces are likely to reward data agglomeration, which could lead to abuse of power market, anti-competitive practices, or even discrimination. On the other hand, government data agglomeration may help amass and maintain political power, discourage dissent, and even discriminate against some population segments.
48. In this context, the different members of the National Statistical Systems (NSSs), and particularly the NSOs, are expected to operationalize data equity and inclusion within their governance and planning arrangements. Yet, how can NSOs guarantee that everyone is on board on this path, in a context of increasingly less control over data collection, processing, and dissemination? Although there is not a unique answer to this question, a consensus has been rising: NSOs must act as data stewards.
49. Nowadays, anyone can be a data producer: not because anyone can produce data properly, but because a quality data producer can come from anywhere. With the unfolding of the Fourth Industrial Revolution, new technologies have permitted firms, universities, public institutions –national and subnational–, and Civil Society Organizations (CSOs) to collect and produce data to an extent that was only possible for NSOs in the past.
50. In that sense, apart from adapting technologies such as Artificial Intelligence or the Internet of Things, NSOs must shift their role as exclusively data owners and processors, broadening their functions as stewards. As the head of the NSS, NSOs can guarantee the implementation of core values and principles in their statistics, but also oversee and promote that data provided by other stakeholders comply with pre-established standards of equity, inclusion, quality, relevance, impartiality, misuse prevention, and confidentiality. In other words, to keep our promise of leaving no one behind we need to become guardians that prevent that any data producer crosses to the data's dark side.
51. Moreover, as data stewards, NSOs have the responsibility to improve better access to data. Through the massification of data access, a wider variety of users have found value in statistics. In the past, the value of collected data was solely concentrated in the data producers and some of their partners –particularly in NSOs and public institutions. As we enlarge data sources and improve processing methodologies, the insights derived from data are becoming endless. Academics and policymakers are finding innovative ways to rely on data, by developing new processing techniques to contribute to decision-making processes. Satellite images are being used to predict multidimensional poverty or oversee illegal mining; national censuses are being used to geolocate and prioritize the most vulnerable amidst the

pandemic, and machine learning models are being put in place to correct ethnic groups' inclusion bias in sample-based statistical operations. Moreover, the private sector is currently relying on official statistics and Industry 4.0 technologies to produce new products and services. Data are becoming a public good and, as such, we must foster their responsible widespread usage and production.

52. Bearing this in mind, data stewards should focus on three tasks to foster equity and inclusion. First, NSOs must identify barriers and enablers to widespread the use of data by different stakeholders in society. We must focus on issues such as data literacy and the development of an ethical framework that allows us to navigate an increasingly decentralized data ecosystem. Second, it is paramount for data stewards that the increased use of data results in inclusive and equitable policies and practices. NSOs, in coordination with decision-makers and public institutions, need to address the limitations of different types of data to inform public policy. Finally, NSOs need to define their approach to foster the inclusion of different communities along the data value chain, generating mainstreaming disaggregated data for both policymakers and the communities themselves. To reach a world without poverty and discrimination, the development path must be built and traveled together, particularly with the most vulnerable. Equity and inclusion also mean to involve everyone in the discussion, and, to this respect, that includes them participating in the collection, processing, and dissemination of the data used to improve their future and everyone's else.
53. This workstream will contribute to the overall goals of the WGDS, by discussing the trends described above and further refining the main actions to be taken by NSOs as data stewards in the promotion of equity and inclusion. As a first step, the workstream group will compile examples from NSO members of the WGDS and from the broader community by leveraging the Data Values Project led by the Global Partnership for Sustainable Development Data (GPSDD).

Terms of Reference

54. This document contains the guidelines for workstream members to draft case studies on how to address the equity and inclusion perspective as data stewards. This information will allow the workstream to systematize the different approaches that can be taken by other countries to foster equity and inclusion throughout the entire statistical cycle. Bearing this in mind, and to facilitate the information analysis, member countries are encouraged to comply with the following suggestions:
55. The document should clearly justify why the presented experience is related to the data stewardship approach and how it is related to equity and inclusion perspective. The intervention group –ethnic groups, vulnerable communities, women, etc. – needs to be clearly defined.
56. The case studies should identify:
 - a. The barriers and/or enablers to encourage the widespread use of data by the intervention group or the participation of the intervention group in one or several steps of the data value chain,

- b. How was the intervention expected to result in the inclusion of the intervention group, and
 - c. What actions were undertaken throughout the different statistical phases – planning, collection, processing, and dissemination.
57. To integrate the relevant elements of the WGOD, if appropriate, cases studies should examine:
- a. How open and accessible is data, including microdata, related to marginalized groups?
 - b. Is it clear how to access data that is not open and is there a transparent process for making such requests?
 - c. Whether capacity to understand and use the relevant data was improved?
 - d. Did data openness facilitate community-level engagement?
58. When possible, country experiences should also describe the data governance schemes that allowed NSOs to lead and execute the interventions to foster equity and inclusion.
59. The case studies should describe, when possible, the different issues countries had to tackle to execute the actions contemplated by the interventions.
60. Case studies should explicitly state unresolved barriers and challenges they have identified or foreseen in the promotion of equity and inclusion in official statistics.

Data stewardship and the city data agenda: workstream 4 (WS4)

Lead(s): Sustainable Development Solutions Network (SDSN) TReNDS (see Annex I for other members)

61. Led by the Sustainable Development Solutions Network (SDSN) TReNDS, this workstream links the Working Group on Data Stewardship (WGDS) and those engaged in the smart cities initiative with the aim of fostering knowledge exchange, finding areas of mutual interest and suggesting remedies in these areas to improve and connect data stewardship at the national and cities level.
62. Data can be a powerful tool for helping city governments improve the efficiency of their operations, save money, provide better services, and enhance citizen engagement. Yet, to fully leverage the value of data to achieve these objectives requires city governments to improve their data stewardship and data analytics capabilities. As such, the experiences of National Statistical Offices (NSOs) are important for advancing data stewardship and supporting whole of government data strategies within cities.

Summary of Key Achievements to Date

63. This workstream has engaged in several activities to date. The well-attended webinar in March 2022 entitled “Enabling a data-driven culture between the national and city level”, provided an inclusive platform to delve into the recommendations this workstream provided to the Commission during previous sessions. As follow up to the webinar, the workstream has concentrated its efforts in two areas. The first area explores the potential for the Sustainable Development Goals (SDGs) to act as a use case for data innovations within cities. Consistent with the OECD’s findings that at least 105 of the 169 SDG targets will not be reached without proper engagement and coordination with local and regional governments. With cities and sub-national entities having conducted more than 50 Voluntary Local Reviews in 2021 / 2022, the workstream is considering city use cases that could showcase innovations in data production, dissemination, and use under the SDG framework.
64. Second, City Chief Data Officers (CCDOs) have long recognized academic institutions as an important resource and collaborations between university-based data labs and city data units are common, combining local knowledge with specialized technical skills to co-develop solutions. City data partnerships with private sector entities are also on the upswing. As such, the workstream is identifying and reviewing case studies that could inform how these partnerships are established and what makes them effective. These lessons could inform NSOs strategies aimed at developing sustainable and productive collaborations with academia and the private sector to improve their access to non-traditional forms of data and incorporate new analytics into their service offerings.

Strategic Developments and Plans for 2023

65. The forthcoming SDSN TReNDS report “Smart Cities: Establishing Multi-Stakeholder Collaborations to Leverage the Full Value of Data ” is intended for city and national government officials looking to establish stronger partnerships across sectors to advance evidence-informed decision-making. The ‘smart city approach’ to

collaboration points to some common recommendations that national governments might consider when establishing and maintaining projects across the whole of government to maximize the value of their data. This report will present case studies that illuminate how these partnerships are established and what makes them effective, with a particular focus on city-university, city-private sector, and city-national government partnerships. SDSN TRenDS will consult with key stakeholders to gather partnership insights and best practices at both the planning and implementation stages of smart city initiatives. These lessons could inform government-wide strategies aimed at developing sustainable and productive collaborations with academia and the private sector to improve their access to non-traditional forms of data and incorporate new analytics into their service offerings.

Annex I: Organisational membership of each workstream

(As of 31 January 2023)

No	Workstream/ role	Lead(s)	Members
	Secretariat	United Nations Statistics Division (UNSD-DESA)	
1	Governance and legal frameworks	Statistics Poland/ World Privacy Forum	INSTAT-Albania; Statistics Indonesia (BPS); Statistics Lithuania; NSO-Malawi; Statistics Netherlands (CBS); Statistics New Zealand; Statistics South Africa (Stats SA); Instituto Nacional de Estadística (INE-Spain); Statistics Sweden; Swiss Federal Statistical Office (SFSO); United Kingdom Statistics Authority (UKSA); UNESCAP; UNECA
2	Equity and inclusion	Departamento Administrativo Nacional de Estadística (DANE-Colombia); Global Partnership for Sustainable Development Data (GPSDD)	Statistics Canada; INEC-Ecuador; Statistics Indonesia (BPS); Statistics New Zealand; United Kingdom Statistics Authority (UKSA)
3	Sharing and collaboration	DANE-Colombia; Instituto Nacional de Estadística y Censos (INDEC-Argentina)	Australian Bureau of Statistics (ABS); Instituto Nacional de Estadísticas (INE-Chile); Central Statistical Bureau of Latvia; Statistics Netherlands (CBS); Statistics New Zealand; United Kingdom Statistics Authority (UKSA)
4	Data stewardship and the city data agenda	Sustainable Development Solutions Network (SDSN) TRenDS	Centro de Investigación y Docencia Económicas (CIDE), Mexico; City of Los Angeles; Open Data Watch (ODW); University of the Philippines (UP); William and Flora Hewlett Foundation; United Kingdom Statistics Authority (UKSA); United Nations Statistics Division (UNSD-DESA)
5	Overall conceptual	Open Data Watch	Australian Bureau of Statistics

No	Workstream/ role	Lead(s)	Members
	framework on data stewardship	(ODW); Statistics Poland	(ABS); DANE-Colombia; Federal Statistical Office of Germany (Destatis); Statistics Netherlands (CBS); Statistics New Zealand; Statistics Norway; Swiss Federal Statistical Office (SFSO); United Kingdom Statistics Authority (UKSA); UNECE