Statistical Commission
Fiftieth session
5 – 8 March 2019
Item 3(b) of the provisional agenda
Items for discussion and decision: Fundamental Principles of Official Statistics

Background document
Available in English only

Supplementing the United Nations Fundamental Principles of Official Statistics: Implementation Guidelines

Prepared by the United Nations Statistics Division

Supplementing the United Nations Fundamental Principles of Official Statistics: Implementation Guidelines

Introduction

In the forty-eighth session of the United Nations Statistical Commission it was agreed that a Friends of the Chair (FoC FPOS) group be set up to provide additional material that could supplement the existing *United Nations Fundamental Principles of Official Statistics: Implementation Guidelines*.

This paper provides ideas and guidance for consideration as chapters to the implementation guidelines, organized in two sections. Section 1 looks at ideas for how the implementation of the principles could be evaluated, including a discussion about compliance and evaluation approaches along with good practice examples (Attachment 1). Section 2 looks at opportunities and issues using non-traditional and non-conventional data sources, mapping the United Nations Fundamental Principles of Official Statistics to them and providing examples of where these type of data sources are being developed and/or used in the production of official statistics (Attachment 2).

At the end of each section, there are proposals for the UNSC to consider as part of the discussion of the formal paper (E/CN.3/2019/5).

Background

As the international statistical community reflects on the past 25 years since the adoption of the UNFPOS in 1994, it is worth revisiting their history and acknowledging how foundational they continue to be for National Statistical Offices and National Statistical Systems.

In 1991, a year marking the end of a long period of time where the countries of Central Europe were striving towards market-oriented democracies, the Conference of European Statisticians (CES) developed and adopted the Fundamental Principles of Official Statistics (FPOS)[1], which were subsequently approved in 1992 at the ministerial level by the United Nations Economic Commission for Europe (UNECE).[2] The document provided strong support for official statistics in Central European countries and aimed to restore or strengthen their public function as impartial, independent and relevant for their country's information system. In addition, through strongly worded provisions regarding the confidentiality of information, the FPOS made it possible to restore and deepen the confidence in official statistics. Another innovative aspect of the FPOS was to articulate mechanisms for achieving methodological compliance for national official statistics with international professional and scientific standards, which has in turn made for an effective information exchange on a global scale.

Considering the critical importance of official statistics for the national and international development agenda, it was quickly realized that the FPOS were of global significance, and the UN Statistical Commission (UNSC) adopted them at its special session of 11-15 April 1994 – with a revised preamble – as the United Nations Fundamental Principles of Official Statistics (UNFPOS), reaching a unique milestone in the history of official statistics.

The relevance of the UNFPOS remains as important as ever, which was discussed and acknowledged by the UNSC at its 42nd session in 2011. Moreover, the UNSC did not consider it necessary to revise the set of 10 principles themselves. However, in view of the world's technological advancement, data revolution and new challenges for official statistics emerging, it was decided to revise and update the preamble, which was adopted in its new shape in 2013, at the 44th session of the UNSC. At the same

[1] (CES/702)
[2] C(47)

session, the UNSC recommended to the Economic and Social Council (ECOSOC) the adoption of a draft resolution on the UNFPOS. The Council endorsed the principles in its resolution 2013/21 of 24 July 2013 where it also recommended the resolution to the General Assembly (GA) for endorsement. At the sixty-eighth session of the General Assembly, the UNFPOS were endorsed in the GA resolution 68/261 of 29 January 2014.

The UNFPOS form a solid basis for all ethical and quality-related conceptual documents throughout the world. The European Statistics Code of Practice (CoP), The OECD Recommendation for Good Statistical Practice or the Principles governing International Statistical Activities can be named here as some of many examples of the incorporation of the UNFPOS and their adaptation into various statistical quality frameworks[3]. In addition, the UNFPOS constitute the main quality annex in the *United Nations Handbook for Statistical Organizations* and in the *United Nations National Quality Assurance Framework Manual*.

Following the successful collaboration and bearing in mind the need for more practical guidance in their usage, the UNSC established the Friends of the Chair Group on the Fundamental Principles of Official Statistics (FOC FPOS) dedicated to developing a practical guide for the implementation of the Principles.[4] The document, enriched by countries' examples of good practices in UNFPOS application, was adopted at the 46th session of the UNSC in March 2015.

From their onset it should have been realized that the UNFPOS are not only relevant to the roles that governments and official statisticians play in creating and maintaining trusted official data and statistics. The role of the media, respondents (both public and private), academics, donor agencies and many others are equally relevant to the implementation of the UNFPOS. While governments are expected to provide legal frameworks and adequate resources, the roles of others (ie to respond truthfully, produce useful data and statistics (either in a traditional or non-traditional manner), provide development funds and support) are critical to maintaining the trust and confidence in the national statistical systems and most certainly affect the quality and integrity of the official statistical information they deliver.

Section 1: Evaluating the Implementation of the Fundamental Principles of Official Statistics

A. Introduction

This first section is conceived a supplementary chapter for *United Nations Fundamental Principles of Official Statistics: Implementation Guidelines*. It aims to support the use of the UNFPOS; provide criteria to evaluate actual and/or perceived non-compliance; and utilise information from the 2018 United Nations Global Survey along with selected country experiences to provide insights into the implementation challenges and successes of the UNFPOS. The chapter concludes with good practice country examples relating to the *United Nations Fundamental Principles of Official Statistics* (UNFPOS); and suggests a set of actions to assist countries in their transition towards compliance with the UNFPOS.

The *UNFPOS Implementation Guidelines* were published in 2015[5]. They break down the ten Fundamental Principles into various dimensions, and for each, describe activities that are taken into account in order to improve implementation of each principle. These are complemented by suggested good practices and case studies.

[3] An inventory of quality documents stemming from the Fundamental Principles of the Official Statistics is to be found in the Annex 1.
[4] UNSC decision 42/111
[5] See https://unstats.un.org/unsd/dnss/gp/Implementation_Guidelines_FINAL_without_edit.pdf

Coupled with any set of principles and implementation guidelines is the desire for an understanding of compliance. In 2015, the attention of the United Nations Statistical Commission initially focused on country compliance with the UNFPOS; and subsequently explored ways to effectively address perceived UNFPOS non-compliance[6]. The discussion at that time noted the absence of definitions of compliance and non-compliance, leaving these definitions open to interpretation across different national and regional contexts.

This chapter attempts to address this gap by describing non-compliance; by providing criteria for non-compliance and an evaluation process in Part B. Part B also presents three models for determining non-compliance, namely: (i) a self-assessment model; (ii) a peer review evaluation process; and (iii) an accreditation model.

The intent of the *Implementation Guidelines* is to remain a 'living document', enabling the UNFPOS implementation experiences and learnings of countries to contribute to the evolution of implementation approaches over time, particularly building on good practice examples. The *Guidelines* specifically seek the ongoing submission of country implementation experiences to update and enhance the current set of practical implementation guidelines and experiences. This chapter provides recent experiences in Part C.

The information presented in Part C has been supplemented with responses from the 2018 United Nations Global Survey. Key insights from the *Global Survey* include:

- There is a need to fill knowledge and understanding gaps among the broader national statistical system regarding the UNFPOS
- A sound and up-to-date statistical legal framework is necessary to support the implementation of UNFPOS in all aspects of the statistical production process
- It is crucial to keep users informed about the methods, quality and proper use of official statistics, engaging with them in order to promote data literacy to avoid the erroneous interpretation and misuse of statistics.
- The use of innovative sources of data, such as big data and citizen-generated data, for the production of official statistics, requires the investment of resources and efforts to ensure that internationally agreed standards, methods and classifications are systematically implemented by all data providers
- It is critically important to enable statistical offices to access administrative and other sources data while honoring data confidentiality
- In most countries, the sharing of resources and the implementation of joint data collection and dissemination activities could help improve coordination across entities that are part of the national statistical
- The level of implementation of international statistical standards is very heterogeneous across countries and regions.
- It is necessary to strengthen coordination mechanisms among partners and donors that support statistical capacity building, and to make sure that such support is in alignment with FPOS, particularly when it involves line ministries in the recipient country.
- With respect to open data and interoperability, a the majority of countries still provides data downloads in proprietary formats for data analysis software, while the dissemination in open, machine-readable formats.

Part D provides a set of suggested actions to assist countries to reach compliance with the UNFPOS.

[6] https://unstats.un.org/unsd/statcom/doc15/2015-18-FP-E.pdf

B. An evaluation guide

There are significant benefits to implementing, and achieving compliance with, the UNFPOS to produce high quality, reliable and impartial official statistics.

Official statistics are a recognised part of the knowledge infrastructure for a country (and region) underpinning decisions by governments, businesses, communities, citizens and the international community; influencing investment choices; monitoring progress; and informing policy development and measuring its success. Official statistics are also integral to good governance and meeting citizens' expectations of authoritative and trustworthy information about their country, informing debate and decisions on issues of national importance and monitoring the outcomes of those decisions.

When official statistics are objective, reliable, impartial and accessible, users (including Ministers) and observers (including international investors and the media) can have confidence in the production and release of the statistics and focus the debate on what the results mean and how they can be used to support opportunities for growth, progress and innovation.

Compliance with the UNFPOS is currently the most effective mechanism available to secure the benefits of official statistics and manage the risks of non-compliance. The risks associated with non-compliance are significant and can manifest as a loss of trust and confidence in the National Statistical System and/or National Statistical Office; a lack of professional and ethical behaviour in the production of official statistics; a lack of coordination at the local and national level; and, finally, a lack of co-operation at the regional and international level.

If these risks are not addressed, they can seriously undermine the role of official statistics in supporting sustainable development; peace and security; as well as for the flow of knowledge and trade among States and peoples of an increasingly connected world.[7]

B.1 Criteria for evaluation

The *UNFPOS Implementation guidelines*, particularly the activities to improve the implementation of each principle, provide a sound basis for developing a set of generic evaluation criteria. This is because the UNFPOS and related activities, good practices and case studies are generally accepted, having been developed by the global statistical community. At the same time, they are sufficiently wide-ranging to be usable in very different national contexts.

Evaluation criteria are generally interpreted to mean benchmarks or yardsticks against which results can be compared. These criteria could be developed by drawing on the questions in the global surveys, and from the good practice activities in the Implementation Guidelines.

Similarly, evaluation criteria could draw on the risks from non-compliance for each principle in the *UNFPOS Implementation Guidelines* to frame a set of outcomes. A suggested approach is provided in Table 1 along with potential sources of evidence for assessment. The two approaches (whether it is via the questions from the Global Survey or via the risks from non-compliance) create similar criteria which could be further refined into one set.

[7] https://unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf

Table 1. Examples of possible criteria to evaluate the implementation of the UNFPOS, based on the 2012 questionnaire and the 2015 Guideliens

From 2012 questionnaire	2.6 Is the National Statistical Office free from political interference in relation to methodology and survey design?
From 2015 Guidelines	3.1 Statistics are compiled on the basis of common standards determined only by statistical considerations. (p18) 3.2 Choice of sources and statistical methods as well as decisions about the dissemination of statistics are only made by statistical considerations 1.1 There must be clear provisions in the law to ensure scientific standards (p23)
Examples of possible evaluation criteria	1. Legal framework requiring use of appropriate statistical methods. 2. Policy setting out how statistical methods are to be decided, with corresponding implementation process. 3. Absence of public concern on political interference on statistical methods. 4. Methods and quality measures are publicly available. 5. Methods are subject to peer review, for which the process and outcomes are public when this is appropriate.

Table 2. Evaluation criteria based on the risks from non-compliance and sources of evidence

Criteria based on the risks from the implementation guide	**Sources of evidence for evaluation**
The production of official statistics is supported through legal mandate and/or policy settings.	• Enacted legislation • Published policy settings • Reviews of the legal and regulatory framework
Official statistics are recognised nationally and internationally as being produced on an impartial basis	• Trust and use -type survey results • Media reports -national and international which recognise impartiality • Self and/or Independent assessments • Country reviews eg OECD, IMF
Official statistics policies and practices are acknowledged as being guided by professional principles and guidelines.	• Compliance with code of practices • Media reports – national and international which acknowledge professionalism • Self and/or Independent assessments
Official statistics policies and practices are acknowledged as being transparent.	• Published material which outlines official statistics policies and practices • Reports on the monitoring of official statistics policies and practices • Compliance with code of practices • Media reports – national and international which acknowledge transparency • Independent assessment • Country reviews (e.g., by OECD, IMF)
Official statistics policies and practices are recognised as being guided by ethical standards.	• Published material which outlines ethical standards and their application • Records of staff capability, training and monitoring • Compliance with code of practices • Records of breaches and the consequences • Media reports – national and international which recognise ethical behaviour • Independent assessment • Country reviews (e.g., by OECD, IMF)

Using the criteria suggested, countries could then be asked as part of the *Global Survey* to outline their evidence base which would then be assessed as strong, weak or an absence of evidence (see Table 2).

Table 3. Areas of compliance and the types of evidence base

Areas of Compliance \ Evidence Base	Strong	Weak	None
The production of official statistics is protected through legal mandate and/or policy settings.	Full Compliance	At risk of non-compliance	Non-compliance
Official statistics are recognised nationally and internationally as being produced on an impartial basis	Full Compliance	At risk of non-compliance	Non-compliance
Official statistics policies and practices are acknowledged as being guided by professional principles and guidelines.	Full Compliance	At risk of non-compliance	Non-compliance
Official statistics policies and practices are acknowledged as being transparent.	Full Compliance	At risk of non-compliance	Non-compliance
Official statistics policies and practices are recognised as being guided by ethical standards.	Full Compliance	At risk of non-compliance	Non-compliance

A strong evidence base would support compliance, a weak evidence base would suggest a risk of non-compliance and an absence of evidence would mean non-compliance. The results for each region and/or country would be published and subject to public scrutiny and debate.

Using such an approach would allow countries whose evidence base was weak or absent to continue working on the relevant activities towards future compliance and help mitigate against the consequences of non-compliance. The approach also acknowledges areas where there appears to be progress for countries towards full compliance.

Developing what constitutes a strong and weak evidence base will require further development.

Clearly, some countries might be nervous about their results being published. However, open and public scrutiny of country results seems to be the best way to meet the remit of UNSC, for four reasons:

1. It would provide the widest possible access to the results.
2. It would enable countries to assess their own compliance with the Fundamental Principles and identify both strengths and weaknesses.
3. It would help to identify common themes and shared issues across the international statistical community (which might be addressed collectively through UNSC).
4. It would showcase useful practices that could strengthen compliance.

Countries could also be invited to provide a voluntary supplementary statement setting out the processes and practices that they have put in place to work towards full implementation. This might include, for example, local codes of practice and the processes that support their adoption. There

would an added advantage of sharing good practice as widely as possible and building a set of global resources that can be used by everyone to support implementation.

B.2 Towards and evaluation process

Our remit from UNSC 2017 required us to '*re-evaluate the mechanisms focused on how to best address perceived non-compliance with the Fundamental Principles*.' To do this we need to be clearer on what the pre-cursors might be for non-compliance and how to address them as National Statistical Organisations, National Statistical Systems and as an international statistical community.

A potential framework[8] that looks at the way the statistical institutions can be undermined has been applied to the case study for Argentina and provides a pragmatic place to start when identifying areas of risk and the potential precursors for non-compliance before they become major issues for the National Statistical System (NSS) and/or National Statistical Office (NSO). It may also assist with the prioritising and sequencing of implementation activities depending on the risk profile identified. It is envisaged this type of analysis could be done as part of wider strategic risk management assessment or strategy undertaken by or for an NSO and/or NSS.

[8] Seltzer, W, "Politics and Statistics: Independence, Dependence or Interaction?", Department for Economic and Social Information and Policy Analysis, United Nations, New York, United States of America, 1994.

Table 4. Suggested criteria for identifying non-compliance

Relevant Criteria	Area of risk for non-compliance	Examples of pre-cursors to non-compliance
The production of official statistics is protected through legal mandate and/or policy settings.	Mission of the statistical service	• The primary aim of the statistical system does not appear to be aligned to the enduring statistical needs and priorities of the country. • The legal mandate and/or policy settings are in doubt or are becoming outdated. • There are growing doubts about the credibility of the institution(s) from a broad range of users and stakeholders. • Institutional arrangements may compromise the ability of the NSO and/or NSS to operate professionally and impartially. • International organisations are raising concerns in advance of escalating to more formal mechanisms (e.g., IMF Statement of Concern and Declaration of Censure)
The production of official statistics is protected through legal mandate and/or policy settings.	Financial resources and controls	• Resources appear to increasingly compromise the ability of the NSO and/or NSS to operate sustainably and independently. • There is growing concern that decisions about the sources of official statistics are being compromised through budgetary control and/or are difficult to justify on statistical grounds.
Official statistics policies and practices are acknowledged as being guided by professional principles and guidelines.	Staff	• The integrity and professional independence of staff is beginning to lose support and/or be actively undermined. • There are growing numbers of staff being appointed who are not sufficiently qualified and/or being appointed on merit • Staff turnover rates are reaching levels which could undermine the overall health and capability of the NSO and/or NSS to perform credibly.
Official statistics are recognised nationally and internationally as being produced on an impartial basis	Suppression or changes in data series	• Suppression or changes in data series are (increasingly) being requested based on non-statistical grounds and are difficult to justify. • There is growing concern being expressed about the range and/or credibility of a country's official statistics across a broad range of users and stakeholders (including international organisations and the wider statistical community).

Official statistics are recognised nationally and internationally as being produced on an impartial basis	Definitions, concepts and methodology	• More challenges are being made to existing and new definitions, concepts and methodologies which depart from internationally accepted norms and cannot be justified on statistical grounds. • Changes are becoming increasingly ad-hoc and lacking in transparency for users and stakeholders
Official statistics policies and practices are acknowledged as being guided by professional principles and guidelines.	Terms and nomenclature	• Decisions about terms and nomenclature increasingly appear to undermine the independence and credibility of the Government Statistician and the wider National Statistical System. • Decisions about terms and nomenclature increasingly appear to lack regional and international comparability and coordination, departing from widely-accepted norms and approaches. • Changes are becoming more ad-hoc and lacking in transparency
Official statistics are recognised nationally and internationally as being produced on an impartial basis	Altering specific numbers	• Unusual patterns are emerging in the coherence of a country's official statistics and cannot be easily explained. • There is growing concern about the credibility of a country's official statistics from a broad range of users and stakeholders (including international organisations and the wider statistical community).
Official statistics policies and practices are acknowledged as being transparent.	Extent and timing of release data	• The extent and timing of official statistics releases is becoming more ad-hoc and aligned with political considerations. • There is growing concern expressed from a broad range of users and stakeholders (including international organisation and the wider statistical community) about the release schedule and its perceived manipulations. • Concern about ongoing unjustifiable pre-release access, or lack of transparency about pre-release access, is expressed by users, media or other commentators.
Official statistics policies and practices are recognised as being guided by ethical standards.	Threats to confidentiality	• There is potential for significant harm to and/or legal action from data suppliers and respondents due to their lack of trust and confidence in official statistics. • There are systemic and on-going confidentiality and privacy issues which are undermining the trust and confidence of data suppliers and respondents. • Survey response rates are approaching levels that don't support the credible production of official statistics.

Official statistics policies and practices are recognised as being guided by ethical standards.	Use of the NSO for political analysis or other work	• Statistical resources including staff are perceived to support political activities including the direct contribution, display and/or distribution of political material (e.g., posters, flags, pictures and leaflets).
Official statistics are recognised nationally and internationally as being produced on an impartial basis	Active campaigns to discredit statistical service outputs, methods or staff	• There are sustained efforts which appear designed to discredit statistical service outputs, methods or staff • There is escalating concern across a broad range of users and stakeholders including international organisation and the wider statistical community about activities which appear designed to discredit statistical service outputs, methods or staff.

Clarifying the precursors for non-compliance, while useful in itself, is most effective when coupled with an evaluation process that applies the criteria to a national statistical system and allows countries to consider their areas of risk. At least three evaluation options exist: (i) A self-assessment model; (ii) Peer review; and (iii) Accreditation.

(i) *A self-assessment tool*

The FOC FPOS group's 2015 discussion of ways to promote compliance and detect non-compliance[9] recognised the important role of self-assessment in bringing compliance challenges into the open and promoting good practice. It is our view that self-assessment, primarily through the regular global review of FPOS should form the basis for evaluating the level of compliance and risks of non-compliance in the first instance.

Global reviews were conducted in 2003 and 2012 using self-completion questionnaires. The questionnaires generally match the activities in the *Implementation Guidelines*, together with a self-assessment for each principle of the extent to which the country feels the principle is implemented. Most of the questions include follow up questions requiring textual responses explaining the answers. A new round of surveys, the 2018 United Nations Global Survey, has been conducted using a similar questionnaire approach.

The 2012 survey showed that countries considered the principles to be well implemented, and that this had improved over the previous ten years. Nevertheless, issues continue to emerge, principally about the independence of statistical agencies, which has driven the desire to be clearer about the risks from non-compliance and their identification.

(ii) *A peer review evaluation process*

An alternative to self-assessment, also put forward for consideration by the FOC FPOS group, would be some form of country-specific peer review programme. However, it is important to acknowledge that this would be complex and expensive to implement, especially if the intention was to have the same reach as the voluntary compliance process outlined above. In particular, some countries might be unable or unwilling to participate and it would be a significant challenge to achieve consistency.

While a universal peer review is impractical, a more manageable option might be for countries (or groups of countries) to invite independent peer review on a voluntary basis, in between the self-

[9] See: https://unstats.un.org/unsd/statcom/doc15/2015-18-FP-E.pdf

assessment process outlined above. Such processes have been used effectively in the European Union[10], for example, and could be very helpful in supporting countries to understand how they might improve its practices to build compliance and in showcasing good practice in specific cases. In line with the voluntary compliance procedure, it is our view that the outcomes of such reviews should be made public. For the process to work effectively, there would need to be an open and transparent process to appoint peer review teams and agreed terms of reference for the remit of such reviews.

(iii) *Accreditation*

An additional tool option by FOC FPOS for further improving the implementation of the UNFPOS might be some form of certification or accreditation, conferred on producers that exemplify sustained and consistent application of some or all of the UNFPOS.

Certification would usually be based on the assessment of a third party. It might be delivered through the voluntary peer review outlined above. Once again, practical challenges, especially time and cost, would be significant but recognising that the benefits from compliance are also significant. The process would require, even more than the other tools mentioned, a proper assessment frame and indicators.

C. Good practice examples

The part provides two recent evaluation experiences of the implementation of the UNFPOS. The first example is a Self-Assessment approach based on a regional code of practice and the second is a case study approach by Argentina based on the framework by Seltzer (1994).

C.1 Self-assessment

A Self-Assessment utilising the ASEAN Code of Practice

The ASEAN Code of Practice (CoP) was based on the *UN Fundamental Principles of Official Statistics* (UNSC 1994) and adopted by all the ASEAN Member States in 2012. The objective of the CoP was to strengthen the ASEAN institutional framework and governance relating to statistics. The CoP is a set of indicators to assess the progress toward ensuring trust, accountability and the highest professional standards in the development, production, dissemination and communication of ASEAN statistics. The committee agreed to implement eight key principles in the CoP and to conduct self-assessments based on the indicators in the CoP.

The CoP comprises:

A. Institutional Environment

1. Mandate for Data Collection;
2. Professionalism & Integrity;
3. Confidentiality;
4. Accountability;
5. Statistical Cooperation & Coordination;

[10] https://ec.europa.eu/eurostat/web/quality/peer-reviews

B. Statistical Process

6. Cost Effectiveness;
7. Reduced Respondent Burden;

C. Statistical Output

8. Commitment to Quality (Relevance, Reliability, Timeliness, Comparability & Accessibility)

The self-assessment was conducted by the committee and the working group of the ASEAN Community Statistical System (ACSS) comprises ACSS Sub-Committee on Planning and Coordination (ACSS-SCPC), Working Group on International Merchandise Trade Statistics (WGIMTS), Working Group on Statistics of International Trade in Services (WGSITS), Working Group on Data Sharing, Dissemination and Communication of Statistics (WGDSA) and Working Group on Foreign Direct Investment Statistics (WGFDIS).

The tool consists of two modules as follows:

- **Module 1:** Assesses the functioning of the national statistical system (NSS) as a whole, with a special focus on institutional, legal and organisational aspects. To be implemented for CoP Section A. The assessment was conducted by the ACSS-SCPC in 2015 and 2017.
- **Module 2**: Assesses the performance of the statistical system in a specific sector (such as trade in goods, foreign direct investment, trade in services,…). In this module, there is a special complement to assess the quality of specific sector indicators. To be implemented for CoP Section B and C). The assessment was conducted by WGIMST, WGSITS, WGFDIS and WGDSA in 2016 and 2018.

Self-assessment on Institutional Environment (Section A)

The first self-assessment exercise on the Institutional Environment (Section A) dimension of the CoP was conducted by the ACSS-SCPC in 2015.

The 2015 assessment was presented at the 5th Session of the ACSS committee in 2015. In general, the assessment showed a very good application of the 5 Key Principles (KP) of the CoP *Section A* (Institutional Environment), with the highest records by KP 1 – *Mandate for Data Collection* (88% of AMS was fully complied with quality requirements).

Meanwhile, the second assessment was conducted in 2017 and the results were presented in the 7th Session of the ACSS Committee in 2017. As a whole, the results shows improvement in all principles in the region compared to 2015 but there are still some gaps to be narrowed within AMS.

Self-assessment on Statistical Process (Section B) and Statistical Output (Section C)

In pursuit of the ACSS Vision of "a responsive ASEAN Community Statistical System providing high quality statistics", the implementation of the first self-assessment of the ACSS Code of Practice (CoP) focusing on sections B and C at the AMSs occurred in 2016 and followed by second assessment in 2018.

The assessment was conducted in 2016 by three working groups of ACSS in generating official statistics for IMTS, SITS, FDIS, while the statistical dissemination is included in the assessments in 2018. The result was presented at the 6th Session of the ACSS Committee in 2016 and 8th Session in 2018 respectively.

Based on the assessment results, the committee proposes that some important elements need to be considered in producing official statistics where AMS needs to improve in the following areas:

- The adequacy of staffs and financial resources;
- Capacity building for staffs through training, participating in international seminar etc;
- Explore administrative sources to reduce respondent burden;
- Institutional Framework to be better in incorporating administrative sources;
- Lack of revision studies on regular basis in some AMSs;
- Processes in investigating discrepancies, including mirror analysis;
- The process to regularly assess the accuracy of data.
- Dissemination by utilising appropriate information and communication technology to meet users' needs; and
- Prices of the statistical products and services, where appropriate, were not clearly disclosed in some AMS.

Based on overall self-assessment results, the recommendations of the ASEANStats are to identify new measures for assessment considering the following elements:

- The general orientation stemming from these feedbacks is that quantitative indicators should not be systematically required in the context of this self-assessment questionnaire.
- In many cases, detailed qualitative information could be more meaningful.
- In WGSITS, some quantitative indicators are kept when they appear to be directly related to the assessment measure proposed and to bring a more precise and meaningful information about the level of satisfaction of this measure.

ASEANStats encouraged the regular conduct of such assessment to ensure compliance of key principles of CoP Section A, B and C by updating the questionnaire by using the 'Snapshot tool'.

Detailed information relating to these assessments is available in Attachment A.

A case study of self-assessment in Argentina

Between 2007 and 2015, National Institute of Statistics and Censuses (INDEC) in Argentina was subject to political intervention to produce and release statistical data in line with the official discourse.[11] This meant that many of INDEC's technical reports were implausible to the population. Initially, the debate was mainly around the Consumer Price Index (CPI), but later distrust grew to most INDEC technical reports.

In December 2015, the new National Government declared an administrative emergency of the National Statistical System. As a result, INDEC followed a consolidation path based on five strategic pillars, due to the large deficits of the Institute: institutional transformation, strengthening statistical capacity, improving coordination of the National Statistical System, boosting dissemination and friendly access to statistics and developing international relations.

To both guarantee its projected institutional development path and support the effective implementation of the UNFPOS, INDEC developed a project for a new Statistical Law. Among the most relevant characteristics of the draft law that can be mentioned are the following: a) organizational and financial independence; b) strengthening the role of head of the NSS; c) a new policy for appointment of the Director-General of the Institute, through parliamentary agreement and with a fixed term and public contest for the appointment of staff; d) the use of new data sources, such as administrative records for official statistics; and e) the formal and explicit adoption of the UNFPOS and the recommendations on statistical good practices of the organisations to which Argentina adheres (i.e., the OECD).

[11]Inter-American Development Bank, "Who wants to know? The Political Economy of Statistical Capacity in Latin America", Washington DC, United States of America, 2018.

INDEC continues to re-build the statistical infrastructure of Argentina. A detailed description of actions and outcomes is available in Attachment B.

C.2 Peer review

Peer Review of Myanmar's National Statistical System

A peer review mission for the National Statistical System (NSS) of Myanmar was carried out in Nay Pyi Taw from 26 June to 1 July 2016. The peer review team comprised three members: one senior official from each of the National Statistics Offices (NSOs) of Malaysia and the Philippines; and a senior official of the Statistics Division of ASEAN Secretariat. Secretariat support of the review was provided by the Partnership in Statistics for Development in the 21st Century (PARIS21) Secretariat through a Consultant, and the partnership's Regional Programme Coordinator for Asia-Pacific and Small Island Developing States. The team evaluated a number of documents supplied by Myanmar's Central Statistical Organization (CSO); met with a number of Myanmar NSS stakeholders; met with officials from various government agencies who are producers and users of statistics; as well as representatives from development partners (DPs). The evaluation report produced by the Review Team identified areas of good practice, as well as areas for improvement. The report is available at the Paris21 website[12].

D. Actions to support a transition towards compliance

This chapter has discussed non-compliance criteria; assessment processes – including good practice examples; and now provides actions to assist countries transition their official statistical systems towards compliance.

Promoting the benefits of compliance

Compliance with the UNFPOS has real, practical benefits, especially around reducing risk for decision makers and enhancing public trust in official statistics. Highlighting and sharing good practice and highlighting the risks around non-compliance is a vitally important way to support member states in developing and maintaining compliance with the UNFPOS.

PROPOSAL 1: A central repository of case studies should be maintained to promote and preserve best practice in complying fully with the Fundamental Principles. This repository could include examples of good practice as well as highlighting the issues arising from non-compliance through case studies after the fact.

Publishing the outcomes of FPOS evaluation assessments

PROPOSAL 2: Self-assessment for measuring compliance with the Fundamental Principles should be maintained. Additional support for this process should be provided, through the development of centrally agreed criteria and a common guidance framework against which countries can assess compliance. The proposals in Section B set out a starting point. We propose that a working group should be set up to take this forward, finalising the criteria and developing supporting guidance.

The existing review of compliance via questionnaire already provides a mechanism for taking forward regular self-assessment of compliance. The results of self-assessment should be published by the country in question and by ISC.

Encouraging on-going assessment of compliance

[12] http://www.paris21.org/sites/default/files/2017-09/Myanmar-Peer-Review-Final.pdf

Although there will be practical challenges around time and cost, a mechanism to invite review from an independent panel of member states would provide a very powerful additional tool for enhancing trust and transparency and highlighting best practice. The OECD approach for accession countries and the existing peer review framework in place in Eurostat are helpful examples that might be built upon.

PROPOSAL 3: The working group set up under proposal 2 should move to consider how an invited, independent peer review might work, and consider any additional mechanisms, once criteria and guidance for self-assessment are in place.

The criteria developed to support self-assessment could feed through into independent peer review, building on the topics covered in the FPOS compliance questionnaire. While self-assessment should address all of the UNFPOS, an independent peer review process could be more flexible, and focus on a subset of the criteria, determined on a case by case basis. Once again, the outputs of peer review should be published and fully transparent. We recommend that UNSC considers this option alongside self-assessment as an additional means of supporting the UNFPOS.

Part 2: Mapping of the United Nations Fundamental Principles of Official Statistics against non-conventional and non-traditional data sources

A. Background

In modern statistical systems, a lawful right of access to administrative and secondary data sources cannot be neglected. Without that right, national statistical systems would not be able to meet the growing demands of users.[13] Going further, an important element for an efficient functioning of national statistical systems are infrastructures that make it possible to share data and to interlink them, particularly those stemming from different sources and stakeholders.

Background for this situation can be found in today's reality, the so-called "data revolution". Indeed, the data industry has been expanding very fast with data providers proliferating, thereby challenging the traditional data producers, National Statistical Office (NSOs), to redefine themselves and their data production systems. In addition, the value of pace in getting the information has been outgrowing the value of quality thereof.

Users frequently turn to sources other than official statistics, unconscious of the quality limitations of what they are receiving. Policy makers, businesses, citizens want not only to be informed, they want to be informed quickly and in an easily accessible way. While having progressively acquired and adopted new technologies, methods and standards, NSOs cannot remain solely providers of good-quality official statistics. Their roles are evolving towards the ones of story-tellers, data stewards, data integrators and finally, quality assessors and providers of standards and secure data architecture, especially in the era of fake news and post-truths.[14]

Public statistics are currently facing new needs resulting from phenomena changing previously known social and economic structures. Advancing globalization and digitization is forcing official statisticians to cover new areas, change existing surveys and finally explore the potential hidden in new data sources. Moreover, the data revolution is also impacting social attitudes to data sharing, privacy and confidentiality. Official statistics has been striving to regulate access to private sector data for statistical purposes[15], however the social perception may differ in this case from the attitude taken to statistical access for public sector administrative data. Therefore, the existence of the reference principles describing universal key values of the official statistics is of utmost importance.

The second global and vital challenge of today is the 2030 Agenda for Sustainable Development (2030 Agenda) where NSOs fulfil the task of supporting data needs, as "[n]o country is at the point of being able to report on 100% of the indicators, not even the most sophisticated among us. The 2030 Agenda has brought a sense of humility and camaraderie between national statistical offices. It is an opportunity

[13] See: Statistical Capacity Building for Sustainable Development: "While national statistical legislation should define and legislate for all aspects of a statistical system, there are three absolutely necessary components: the professional independence of that system and in particular of the head of that system; The NSO and other statistical agencies of the NSS must be legally empowered to collect primary information for statistical purposes from persons, households, businesses and other institutions and have legal access to all appropriate administrative microdata held by national and regional public administrations for the purposes of compiling aggregate statistical information; and the safeguarding of confidentiality must be clearly set out in law". In : *Developing the fundamental pillars necessary for modern national statistical systems*/Statistical Journal of the IAOS 33 (2017) 895–909 895 DOI 10.3233/SJI-160331 IOS Pres.

[14] Anil Arora, opening speech from the High Level Workshop on Strategic Partnerships, Geneva, 11 April 2018.

[15] To this end, a B2G Data Sharing Expert Group has been created at the European level as an informal body whose main task is to assist the European Commission in assessing issues connected with business-to-government data sharing. It is managed by Commissions' Directorate-General for Communication Networks, Content and Technology (DG CNECT) The governing principles for B2G data sharing were outlined in 2018 by the Commission in its Communication 'Towards a common European data space'. Those are accompanied by details legal and practical considerations set out in Guidance on private sector data sharing.

to bring producers and consumers together around an important cause and demonstrate the critical role of NSOs and the relevance of our science and craft". [16]

The trends outlined above indicate the necessity of introducing innovative methods and techniques for the production of official statistics, using and combining many (new) data sources, including those called non-traditional or non-conventional. The non-traditional sources cover big data (mobile phone data, social media, etc.), citizen generated data (CGD), and a range of other so-called new data sources. In each of their cases, we deal with external data handlers, and in the case of big data, most often with private sector companies.

As coordinators of the National Statistical Systems (NSSs), NSOs need to improve their own capacity in adopting technologically advanced ways of including the new data in their production processes, which requires meeting several conditions. These are related to obtaining the consent of the data owners for their use, while complying with applicable law (e.g. protection of personal data), durability of the data source, data compatibility with the official statistical production process or its change and adaptation, ensuring methodological regime and data quality. In addition, NSOs are faced with the challenge of developing modern systems that incorporate standards, quality and interoperability in the data production processes.

Statistical surveys and administrative registers are and will remain important sources to produce official statistics, as it is confirmed, among many others, by examples from the United Kingdom or New Zealand[17]. Yet, the enhancement and enrichment of the traditional data gathering methods by big data analytics appears inevitable.

Bearing this in mind, opening up privately-held data of general interest for re-use by national statistical systems is undoubtedly beneficial for society. However, the lack of uniform legal basis determining conditions under which privately-held data can be made available for official statistics hinder the success of this venture. The arguments advanced by data owners against granting access to their data for statistical offices concern mainly property rights, confidentiality and trade secrets.

In view of the above, the need to reconsider the meaning of the UNFPOS resurfaces. Official statistics must meet the test of practical utility, regarding its indispensable character for the information system of a democratic society. Statisticians need to follow strictly ethics in the collection, processing, storage and presentation of statistical data, along the idea of transparency of methods used.[18]

---

[16] Anil Arora, opening speech from the High-Level Workshop on Strategic Partnerships, Geneva, 11 April 2018

[17] The United Kingdom uses a range of sources to produce official statistics. While the key sources are surveys and administrative datasets, the UK Office for National Statistics (ONS) has been conducting some preliminary experiments with new data sources including web scraped price information, geo-located Twitter traces, smart energy meter data and aerial photography. The summing up of the mentioned works is available at:
https://www.ons.gov.uk/aboutus/whatwedo/programmesandprojects/theonsbigdataproject
More detail is given below:

a) The use of online property databases to classify housing types, particularly caravan sites, with a view to improving our on-the-ground intelligence for census enumeration. See https://www.ons.gov.uk/methodology/methodologicalpublications/generalmethodology/onsworkingpaperseries/onsmethodologyworkingpaperseriesno11identifyingcaravanhomesinzoopladatajune2017
b) Web scraping price data from retailers with a view to generating more timely pricing information to feed into price indices.
c) Preliminary trials with anonymous, aggregated mobile telephone data to see whether we can estimate commuting flows
d) The ONS Data Science Campus is doing wider research around how big data can serve the public good – see https://datasciencecampus.ons.gov.uk/projects/

New Zealand uses a range of sources to produce official statistics, most are based on government administrative data sources and surveys. However, data from credit card providers, data from electronic payments switch companies (all electronic card transactions), scanner data for consumer electronics and web-scraped prices are also being used. A trial to use utilities data has been done, being however an *ad hoc* example to understand population movements during the Canterbury earthquakes. Statistics New Zealand is continuing to explore a range of different sources to support the production of official statistics.

[18] Dominik Rozkrut, keynote speech from the Conference of European Statistics Stakeholders, Bamberg, 18 October 2018.

Including non-traditional and non-conventional data sources in the statistical production constitutes another challenge for the applicability, timeliness and relevance of the UNFPOS. As it is almost impossible for policy makers to differentiate between official and non-official statistics, a challenge to communicate quality is being faced by NSOs.[19] Stemming directly from the Principles, the professional independence of statistical offices and their longstanding experience in dealing with personal and confidential data is widely appreciated and perceived as providing extra assurance with regard to the further use of the data owned by private companies.

The following mapping of the UNFPOS against the mentioned sources aims at proving that those features remain unchanged. For the purposes of the document, notions "new data sources", "non-conventional data sources", "non-traditional data sources" and "big data" are used interchangeably.

Although, no uniform definition has yet been agreed upon for big data, it is broadly accepted that the notion refers to data sets of increasing volume, velocity and variety; the 3 V's.[20] In the paper "An Assessment of big data for official statistics in the Caribbean", a fourth V is added to describe big data, namely their "veracity", conformity to facts.[21] Another feature of those data sources is their lack of consistent structure, meaning the lack of pre-defined data model and/or the fact that they do not fit well into conventional relational databases. Being interesting for official statistics, this type of data represents considerable challenge about its inclusion into the statistical production due to ethical concerns, law and ownership-related issues.

Following the classification of the High-Level Group for the Modernization of Official Statistics (HLG-MOS), established in 2010 under the umbrella of CES, the large data sources comprise:

- Administrative data (arising from the administration of a program, be it governmental or not), e.g. electronic medical records, hospital visits, insurance records, bank records, food banks, etc.
- Commercial or transactional sources (arising from the transaction between two entities), e.g. credit card transactions, on-line transactions (including from mobile devices), etc.
- From sensors, e.g. satellite imaging, road sensors, climate sensors, etc.
- From tracking devices, e.g. tracking data from mobile telephones, GPS, etc.
- Behavioural, e.g. online searches (about a product, a service or any other type of information), online page view, etc.
- Opinion, e.g. comments on social media, etc.[22]

Another type of classification, developed by the UNECE Task Team on Big Data[23], provides a division into human-sourced information (social networks, blogs and comments, personal documents, pictures, videos, internet searches, mobile data content, sms, user-generated maps and e-mail) process-mediated data (data produced by public agencies, medical records, data produced by businesses, commercial transactions, banking/stock records, e-commerce, credit cards) and machine-generated data, so called *internet of things*, i.e. data from sensors, home automation, weather/pollution sensors, traffic sensors, mobile sensors, mobile phone location, security/surveillance videos/images, mobile locations, and satellite images.

Irrespective of the classification adopted, it is evident that the use of new data sources in official statistics represents a set of considerable challenges falling into at least one of the following categories:

[19] Sylvie Michaud, TA2.14 Non-official Data: Challenges and Opportunities for NSOs, Canadian Experience, UN World Data Forum, Dubai 2018.
[20] *What does "big data" mean for official statistics?,* United Nations Economic Commission for Europe, Conference Of European Statisticians, 10 March 2013What does big data mean for official statistics?, UNECE, CES, 10 March 2013
[21] Abdulkadri Abdullahi, Evans Alecia, Ash Tanisha, *An assessment of big data for official statistics in the Caribbean: Challenges and opportunities,* January 2016
[22] *What does "big data" mean for official statistics?,* UNECE, CES, 10 March 2013
[23] https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data

- Legislative, i.e., with respect to the access and use of data;
- Privacy, i.e., managing public trust and acceptance of data re-use and its link to other sources;
- Financial, i.e., potential costs of sourcing data vs. benefits;
- Management, e.g., policies and directives about the management and protection of the data;
- Methodological, i.e., data quality and suitability of statistical methods;
- Technological, i.e., issues related to information technology.[24]

In light of the above, UNFPOS emerge as the main reference and a solid basis for considering ethical issues related to the inclusion of the new data sources in the official statistics. In 2017 Europe's *Ethical Guidelines*[25] have been published, aiming to draw the attention of statistical authorities to possible issues of professional ethics that can appear with the use of big data in the production of official statistics. Based on the results of the projects carried out by NSOs and research, investigating different angles of the use of big data in official statistics, the *Guidelines* recommend an approach that is also compliant with the statistical code of conduct. For the purposes of the present mapping exercise, the *Guidelines'* approach of examining the possible issues at three main stages of the statistical production process is applied.

Principle 5 is considered to be the most applicable of the UNFPOS to the new data sources[26].

> *"Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents".*

It sets up the framework for new data sources to be considered as a natural supplementation to statistical production. Provided that non-conventional sources ensure the quality of statistical output are cost-efficient and minimize the reporting burden for the data providers, the tendency is to increase their use in the official statistics. Building on that, the examples of using sensor data, mobile phone data, and satellite images can be given across countries.

At all main stages of the statistical production process – acquisition, processing and dissemination – questions of an ethical nature concerning the cornerstone values of official statistics are naturally raised. While acquiring data, the problem of data ownership emerges. Big data are mostly collected by private companies, which are usually not bound by the law in most countries to make data available for official statistics. Therefore, the provision of data is often based on strategic partnerships established with private companies to gain access to data. Here the main ethical reference should be Principle 1 called the impartiality principle and ensuring the right for information.

> *"(...) official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information".*

This principle stresses the need to avoid any pressure from businesses to put their interest above the public interest. Secondly, the selection of partners should be done with full transparency and be preceded by thorough research. From practical examples, given by the members of the Group, strengthening the Research Divisions was the recurrent one. Another form of ensuring that the use of new data sources is consistent with Principle 1 is a preference given to a multi-mode approach while collecting the

[24] *What does "big data" mean for official statistics?,* UNECE, CES, 10 March 2013
[25] Services concerning ethical, communicational, skills issues and methodological cooperation related to the use of Big Data in European statistics, Task 1. Ethical Review, Deliverable 1.3. Report on ethical guidelines June 2017 https://ec.europa.eu/eurostat/cros/system/files/draft_ethical_guidelines_final.pdf.
[26] See F. Perucci, *Implementing the Fundamental Principles in a transforming statistical system*, 15th IAOS Conference, Abu Dhabi, UAE, 6-8 December 2016

information. Validation of information and constant crosschecks have also been given as examples thereof.

Relevance, impartiality, equal access, interaction with users and planning being most important features covered by this principle, a risk of competition between national statistical offices and other stakeholders to access new data sources has been mentioned as a factor to be avoided.

Therefore, according to the Principle 6 – the confidentiality principle – NSOs need to assure big data providers that their data will be used exclusively for the purposes of official statistics and there is minimal risk of harm to their business. The confidentiality principle affords a distinct advantage for national statistical offices compared with other data providers. In order to work consistently to this principle, NSOs should be informed by data providers whether their customers are aware that the data about them can be delivered to statistical authorities.

Other examples of actions to be undertaken in respect of the applicability of UNFPOS to the use of new data sources, at the stage of data collection, are actions of legislative nature. Within Principle 6, all sorts of adherence to national, international and regional codes have been quoted as examples. In light of new data sources, strict adherence to confidentiality has been described as the most important factor, due to the possibility of inadvertent disclosure of confidential information deriving from a new data source. A case of social media use has been evoked as an example of risk and difficulty in obtaining users' consent to the usage data containing information about them.

At the stage of data processing, the UNFPOS appear as a remedy for the major threat, which is the number of quality-related issues that may be compromised while integrating big data with the official statistics. Examples: risk of bias and manipulation within big datasets, no guarantee in stability and continuity of data structure, lack of scientific proof while using statistical models or imputation techniques for the data processing.

The key reason being that big data in general are not designed for statistical purposes and therefore they do not comply with statistical definitions, standards and methods.[27] According to Principle 3 *("To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics"),* it is of utmost importance to use new data sources according to verifiable and internationally comparable, transparent standards and procedures.

Members of the FOC FPOS Group also highlighted the increasing importance of metadata and paradata, as well as the need for establishing and agreeing upon new quality attributes for new data sources, if they differ from the existing ones. This would emphasize the professionalism of statistical offices, which is one of their major assets that makes official statistics a reliable source of information.

During big data processing a risk of revealing personal information is also pertinent, as well as improper use of those data that can damage the reputation of the official statistics. Therefore again, Principle 1 (on relevance, equal access and impartiality), Principle 6 (on confidentiality) and Principle 2 ("*To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.")*, provide for the importance of legislation, standards and ethics to prevent loss of trust - may be widely applicable.

At the stage of dissemination of statistical information, faced with complex techniques needed for the production of statistical output while using new data sources, it is vital to: inform the users about methods and procedures used to produce statistics (Principle 3), adhere to international standards ensuring quality and comparability (Principle 9 – *"The use by statistical agencies in each country of*

[27] Report on ethical guidelines, p. 5. https://ec.europa.eu/eurostat/cros/system/files/draft_ethical_guidelines_final.pdf

international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels".), as well as to communicate and educate users, preventing an erroneous interpretation of statistical information (Principle 4 "*The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics").* Based upon that, new data sources should be duly described, and applied methods and models should be documented to support an independent assessment of data processing and statistical results.

Different types of new data sources present different questions of ethical nature – according to a range data characteristics, e.g. access to the data, content of personal information, quality issues in terms of suitability for the purpose of official statistics, the clarity of the methods to be applied in order to get statistical output etc.[28] Further elaboration is made using the types of new data sources used in official statistics and their linkage to respective UNFPOS which constitute a set of recommendations on how to cover the mentioned issues in full compliance with the existing ethics. The *Table 5* included at the end of the paper represents a matrix, highlighting the most relevant principles to consider for each of the new data sources. Country examples for each of the types of new data sources used in official statistics are listed in Attachment 2.

B. Mapping against different types of non-traditional and non-conventional data sources

B.1 Mobile phone data

Mobile phone operators' systems generate a very large amount of data on the use of mobile communication, including location information. These data are mostly used for business and marketing purposes. However, the location data can be used for generating statistics about space-time movement of phones, needed for instance to supplement the tourism statistics.

The ethical issues at stake with this type of data concern the privacy of the data subjects (this non-conventional data source contains sensitive personal information), as well as the professional independence of national statistical offices that might be compromised while creating partnerships with mobile data providers in view of their equal treatment.

Principle 1 can be applied here, providing for ensuring the lack of biases connected with data collection in order to best meet the users' needs, to stay relevant and impartial. When it comes to further stages of statistical production process, a major issue that can be identified during data processing is the lack of suitability of mobile phone data for statistical purposes. At the dissemination stage, because of differences in the definitions, the methodology concerning this type of data source can be complex. Here, Principle 3 is widely applicable, which requires a clear, transparent and understandable presentation of both official statistics and the corresponding metadata.

B.2 Data from smart electricity consumption meters

These data can be interesting for official statistics because they provide information on energy consumption, which can be beneficial for statistics on household consumption expenditure, consumer price indexes, environment statistics or statistics on energy consumption. The inclusion of this non-traditional data source into official statistics is well justified because of a considerable reduction in the burden on respondents, therefore the full application of Principle 5 can be observed here. A possible difficulty in securing access to a sufficient level of detail can be addressed using appropriate legal provisions, however the coverage of this data source should also be explored to assess if the statistics produced are representative and relevant. The smart meters' data are not intended for statistical purposes therefore the risk of discontinuity of data source may also be present

[28] Report on ethical guidelines, p. 6.

B.3 Satellite imagery data

NSOs are exploring the possibility of using data from satellite imagery in official statistics. They are expected to decrease burden on respondents, to improve timeliness, as well as to reduce survey costs. They can also contribute to providing more disaggregated data. Satellite data are used mainly to complement agriculture statistics. The increasingly widespread use of this kind of data is linked with relatively few concerns of ethical nature.

Data are mostly publicly available and they do not carry privacy concerns (or to a very limited extent). The major issue can be quality of statistical output that is directly related to the quality of satellite images, as well as methodology which, following Principle 3, should be clearly explained to the users. According to Principle 1, NSOs need to perform thorough research on the methods used to compile statistics based on satellite images to guarantee the quality of their final products. At this point we can observe the relevance of Principle 10

> "*Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries*"

Principle 10 provides for the joint efforts of international statistical and research communities targeted at finding common methodological solutions, sharing infrastructure, saving resources and taking advantages of synergies.

B.4 Social media data

Social Media can be used in at least three ways: As a subject of Official Statistics, e.g. use of Social media; to disseminate Official Statistics, thus reaching out to all kinds of users; and lastly as a source to compile Official Statistics. In what follows social media as a source (social media data) will be dealt with.

Disseminated via the internet, social media data still represent an area for further exploration for official statistics. Adopting the form of messages, images, video or searches, these data are voluntarily submitted by users on the web. In several countries research is conducted to use social media to measure the level of well-being of societies (studies on happiness explore sentiment analysis).
The ethical concerns associated with these types of non-conventional data sources are related to privacy issues (there is a recurrent question whether the users of social networks should be notified that the information they post, although being public, will be used by statistical agencies) or to lack of access to data which sometimes has to be purchased from private owners. In this case an ethical question arises as to whether statistical offices should pay for data sources that are going to be transformed into public official statistics.

Processing this type of data source depends critically on the methodology adeptly to ensure a sufficient level of quality in the resulting official statistics. Therefore, strengthening research divisions in statistical offices is important if these type of data sources are to be used effectively.

Social media are also a vulnerable source of data when it comes to bias and manipulation. While disseminating statistics based on social media, it is important to accompany them with proper metadata and to describe them in an understandable way.

Again, while using this type of non-traditional data source to produce statistics, Principles 1, 2, 3, 6 and 9 could be applied here, to ensure relevance, to prevent loss of trust, to make them subject to a comprehensive, comparable methodologies and standards, and to ensure users that their privacy is protected respectively.

The use of social media data should also be studied from the perspective of the fake news proliferation. Following Holan (2016), fake news can be defined as "invented material that has been cleverly manipulated so as to come across as reliable, journalistic reporting that may easily be spread online to a large audience that is willing to believe the stories and spread the message"[29].

In the context of social media, there is a danger of collecting data from users and making it available for advertisers who use it mainly to target advertisements. Another example of manipulating the truth is the existence of fake accounts, so called "bots" who may affect the factual image, hindering the quality of social media data to be used for statistical purposes.

Against this background, the ethical reference provided by the UNFPOS seems to be one of the most effective, as they promote official statistics rather than the raw social media data, thereby valuing more highly the production of reliable, comparable and high-quality data – those which meet international standards[30]. Elaborating on the arguments expressed during the conference "Truth in numbers: the role of data in a world of fact, fiction and everything in between"[31], firstly national statistical offices should be provided with the resources and infrastructure they need to support their roles as both a standard setter and co-ordinator across the national statistical system, and secondly it should be ensured by governments that there is no political interference in their national statistical systems, resulting in greater trust by citizens and reflecting national statistical offices as being "guardians of the facts"[32].

B.5 Web-scraped data

Web scraping refers to a technique used for extracting data from websites. Official statistics uses web-scraping techniques for example to collect prices of different goods from the internet and to use this type of data source as a supplement to the Consumer Price Index (CPI). If online prices replaced prices collected in a traditional way, costs of statistics could be considerably minimized. Similarly, to other non-traditional data sources, web-scraped data may raise legal issues (terms of use of websites differ across countries), they are not designed for statistical purposes, therefore methodological issues are also at stake. Once these data are made available, the statistics partially based on web-scraped data should – as in previously described cases – be accompanied by proper metadata.

B.6 Road traffic sensors data and passengers tracking sensors data

Vehicle detection loops, installed in pavements, or lasers can detect vehicles passing or arriving at a certain point, e.g. approaching a traffic light or in motorway traffic. or approaching a bus terminal. Normally, the data are stored in a central data warehouse of the responsible authority, e.g. the national transport agency.[33] Their use in official statistics is rather common. They serve to draw a picture of the number of vehicles, the speed with which they move, along with surveys for estimating commuting time. Passenger information collected by the sensors can be used to determine bus routes as well as monitor ridership to determine if service to certain areas should be increased or decreased. While using this type of non-conventional data source, the major issues are linked to quality. There are usually no privacy-related concerns. Therefore, the Principles that could be applied here refer to common methodologies, standards and definitions, as well to the proper explanation of statistics for the users.

[29] Holan, A.D. (2016). "2016: Lie of the Year: Fake News," *Politifact*,December 13, 2016. Verified April 7, 2017: http://www.politifact.com/truth-o-meter/article/2016/dec/13/2016-lie-year-fake-news/ in Vincent F. Hendricks, Mads Vestergaard, *Reality Lost. Markets of Attention, Misinformation and Manipulation*, p.63., https://link.springer.com/content/pdf/10.1007%2F978-3-030-00813-0.pdf

[30] V. Hendricks, J. Jütting, *Trust in numbers? Why we need strong official statistics in an attention seeking society*, http://www.paris21.org/news-center/news/trust-numbers-why-we-need-strong-official-statistics-attention-seeking-society

[31] 4 April 2018, Bern, Switzerland, http://www.paris21.org/news-center/events/conference-truth-numbers-role-data-world-fact-fiction-and-everything-between

[32] V. Hendricks, J. Jütting, *Trust in numbers? Why we need strong official statistics in an attention seeking society*

[33] Report on ethical guidelines, p. 10.

B.7 Scanner and cash register data, e.g. from supermarkets

While purchasing a product in a supermarket, a record is being created by the scanner. These transaction records collected from many sellers are of a great interest for statisticians. Used for the purposes of the Consumer Price Index (CPI) or complementing other domains of official statistics, such as business statistics or house expenditure, this non-traditional data source may outweigh the traditional collection methods mainly because of its lower cost.

Ethical issues in this respect concern i.e. establishing partnerships with private companies and can be linked with Principle 1 – providing for relevance of statistics and professional independence of statisticians. Another question related to scanner data concern methodological challenges (proper classifications of products, choosing appropriate index methodologies) – therefore the observance of Principles 3 and 9 can be considered a solution in the mentioned cases.

B.8 CCTV data (Security / Surveillance videos), e.g. for citizen security purposes

Several countries / cities have decided to install camera's for citizen security purposes. While these may provide more comprehensive coverage of certain types of crime in certain areas, one has to be on the alert for both privacy issues (camera's capturing everybody: criminals and non-criminals alike) and quality issues (proper distribution of the cameras, so as to have complete coverage is a must).

C. Conclusion

Specific features of non-traditional data sources result in challenges for NSOs to comply with such principles as professional independence, access to sources, mandate for data collection, adequacy of resources, impartiality, objectivity and clarity of the methods used.

However, the technological development of today's world makes it impossible not to include these sources into official statistical production processes – the outcome of reinforces their complementarity to traditional data sources. As much as the NSOs should strive to build partnerships with non-official data producers and owners to make those data part of official statistical outputs, the cornerstone values of official statistics, such as quality, standards, and professional independence should never be neglected, nor compromised.

The timelessness and relevance of the UNFPOS make them useful for the new technological reality. The proliferation of new data sources is a good opportunity to test and reinterpret the UNFPOS, and thereby, to confirm their universal character.

The mapping exercise that has been performed by the FOC FPOS confirmed an extensive applicability of the Principles to all aspects proper for new data sources. Moreover, owing to the fact that they constitute a basis for all data and statistical quality-related codes, charters, and frameworks throughout the world, it can be argued that they can respond to current challenges, be it data revolution or the 2030 Agenda, with NSOs playing the leading and coordination role in the monitoring process.

Table 5 Mapping of the Fundamental Principles of Official Statistics against non-traditional and non-conventional data sources									
New data sources	**Principle 1**	**Principle 2**	**Principle 3**	**Principle 4**	**Principle 5**	**Principle 6**	**Principle 7**	**Principle 9**	**Principle 10**
DATA from	Impartiality principle, right for information (...) official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information.	To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.	To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.	The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.	Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents	Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.	The laws, regulation and measures under which the statistical systems operate are to be made public.	The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.	Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.
Mobile phones	✓		✓	✓	✓	✓	✓		✓
Smart electricity consumption meters	✓		✓		✓	✓	✓		✓
Satellite imagery			✓		✓				✓
Social media	✓	✓	✓		✓	✓	✓	✓	✓
Web scraping activities			✓		✓				✓
Road traffic sensors			✓		✓				✓
Scanner and cash registers			✓	✓	✓	✓	✓	✓	✓
CCTV (security/ surveillance videos)			✓	✓	✓	✓	✓	✓	✓

Attachment 1: Examples of quality documents based on the UN Fundamental Principles of Official Statistics[34].

- The Fundamental Principles of Official Statistics https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx
- UNFPOS Implementation Guidelines https://unstats.un.org/unsd/dnss/gp/impguide.aspx
- European Statistics Code of Practice https://ec.europa.eu/eurostat/web/products-catalogues/-/KS-02-18-142
- Quality Assurance Framework of the European Statistical System (currently being reviewed) https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646
- Recommendation of the OECD council on good statistical practice http://www.oecd.org/statistics/good-practice-toolkit/Brochure-Good-Stat-Practices.pdf
- IMF Standards for Data Dissemination http://www.ins.tn/sites/default/files/methode/pdf/sdds_en_0.pdf
- ISI Declaration on Professional Ethics https://isi-web.org/index.php/news-from-isi/296-declarationprofessionalethics-2010uk
- United Kingdom Code of Practice for Statistics which sets out the detailed principles and practices the UK statistics commits to, under three pillars of Trustworthiness, Quality and Value https://www.statisticsauthority.gov.uk/code-of-practice/
- CCSA Principles Governing International Statistical Activities https://unstats.un.org/unsd/ccsa/principles_stat_activities/
- Statistics Canada Quality Assurance Framework https://www150.statcan.gc.ca/n1/pub/12-586-x/12-586-x2017001-eng.htm
- CARICOM'S Statistics Code of Practice https://caricom.org/documents/13410-caricom_statistics_code_of_practice.pdf
- New Zealand's publication providing guidance to the National Statistical System about the highest priority official statistics (Tier 1). https://www.stats.govt.nz/assets/Uploads/Principles-and-protocols-for-producers-of-tier-1-stats/principles-and-protocols-for-producers-of-tier-1-stats.pdf

[34] An extensive inventory of nationally and internationally developed data quality references is to be found at https://unstats.un.org/unsd/methodology/dataquality/quality-references/

Attachment 2: Examples of non-conventional and non-traditional types of data in official statistics

1. Use of mobile phone data in official statistics
    - New Zealand used cell phone data to understand short-term population movements after the 2010 and 2011 Canterbury earthquakes. This was a special case for access to the data due to the natural disaster.[35]
    - UK made an estimation of commuter flows[36]
    - France has been conducting experimental research work on particular subjects (the evolution of the resident population throughout the day, mobility of people within local areas…) with the view of learning about the possibilities of using the CDR data for statistical studies. The experimental research work enables INSEE to be more familiar with the structure of the data and eventually to propose an aggregated format that would be interesting for statistical purposes.[37]

2. Use of smart meters data in official statistics
    - Estonia uses smart electricity meters data to produce electricity consumption statistics. A study on the feasibility of using smart meter data for producing information on electricity consumption was done with the view of using an opportunity to obtain reporting burden reduction on businesses. The output of the study is final energy consumption by economic activity, by region and monthly, quarterly and annual aggregation for businesses and final energy consumption by household characteristics as they are contained in household registers (size of dwelling, number of rooms and persons, etc.) by region and monthly, quarterly and annual aggregation. Another objective of the study was is to identify vacant dwellings and to verify real places of residence of households and related persons[38].
      In the UK, the case was examined experimentally, using area-level summaries. The interests are twofold – validation of address data by cross referencing against aggregated counts of meter numbers, and exploring the ability to identify areas with lots of vacant properties to save on the cost of field visits.[39]

[35] http://archive.stats.govt.nz/tools_and_services/earthquake-info-portal/using-cellphone-data-report.aspx

[36] the results of which can be found at: https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsusingmobilephonedatatoestimatecommutingflows and https://www.ons.gov.uk/methodology/methodologicalpublications/generalmethodology/onsworkingpaperseries/onsmethodologyworkingpaperseriesnumber13comparingthedensityofmobilephonecelltowerswithpopulationestimates

[37] For further details see the *Position paper on access to privately held data which are of public interest. Opening up new data sources for a new generation of official statistics – in light of the growing European Digital Single Market and the revision of the Public Sector Information Directive*, November 2017, p. 14 https://ec.europa.eu/eurostat/web/ess/-/access-to-privately-held-data. More information is available at : https://www.conference-service.com/NTTS2017/documents/agenda/data/abstracts/abstract_88.html and https://www.conference-service.com/NTTS2017/documents/agenda/data/abstracts/abstract_97.html

[38] See ESS position paper p. 18.

[39] Work to date is available at: https://www.ons.gov.uk/file?uri=/aboutus/whatwedo/programmesandprojects/theonsbigdataproject/analysinglowelectricityconsumptionusingdeccdata_tcm77-418326.pdf https://www.ons.gov.uk/file?uri=/aboutus/whatwedo/programmesandprojects/theonsbigdataproject/onssmartmetercensusfeasibilitystudyfinalreportseptember2014v14_tcm77-408965.pdf https://www.ons.gov.uk/file?uri=/aboutus/whatwedo/programmesandprojects/theonsbigdataproject/comparingcountsofelectricitymetersandaddressesbypostcodeinenglandandwales.pdf

3. Use of satellite imagery data in official statistics

- Most statistical agencies use spatial data, if not satellite data, to map census boundaries. This is a pretty common use case. Many countries are also using satellite data in combination with other sources to produce SEEA Land Accounts:
- Australian Bureau of Statistics produced an experimental land account including estimates of the value of land.[40]
- Statistics Canada has been developing Land Accounts to track land cover change since the 1990s. The current version is also used as the basis for experimental ecosystem accounting: https://www150.statcan.gc.ca/n1/pub/16-201-x/2013000/appendix-appendice1-eng.htm.
- Several countries in the ESCAP region are working on land accounts, but have not yet published: Fiji, Indonesia, Nepal, and Vanuatu.
- Statistics Poland is actively involved in activities aimed at including innovative statistical methods and techniques, covering the use of Big Data, in statistical production. The works are carried out as part of international projects (including the European Space Agency project), national ones (e.g. „Gospostrateg", implemented by the National Center for Research and Development), as well as by means of ongoing negotiations with data owners (among others Google, Play (mobile phone operator), the Main Directorate of National Roads and Motorways). Areas in which solutions for advanced data analysis are developed, for the purposes of a real-time monitoring of the economy, which can be used to study selected aspects related to the single market policy and competitiveness, include among others agricultural research based on satellite imagery conducted by the European Space Agency. The use of satellite imagery technologies will allow to improve the identification and monitoring of agricultural crops and to develop methods for assessing the impact of extreme events such as: flood, drought, frosts, hail, etc. on the state of crops during the growing season, which cause direct repercussions in the economy (e.g. for prices of agricultural products, imports and exports, etc.);

4. Use of social media data in official statistics

- Statistics Netherlands uses social media messages to produce an alternative consumer confidence indicator. The research question was: Can we replicate the consumer confidence index by only using social media data, while reducing production time? A model has been built based on fitting characteristics derived from Facebook and Twitter messages.
- Istat produces an experimental social mood on economy index. The index was started by only using Twitter as a source, but further social media might be taken into consideration in the future.
- INEGI Mexico is working on the following Twitter-centered projects:
  - Mood of twitterers in Mexico – automatically measure and report the mood of twitterers, collecting more than 300 million twits. Process of sentiment classification, text normalization, machine learn algorithm, positivity quotient
  - Domestic tourism.
  - Mental health.
  - Mobility in Mexico City.
  - New agglomerations.
  - Consumer confidence.
  - Insecurity.
- Department of Statistics Malaysia (DOSM) uses online social media i.e. Twitter & Facebook for the statistical communication and dissemination of official statistics. Public Maturity Assessment on Official Statistics (PMAOS) is a tool used by DOSM to analyse sentiment analysis based on public opinion and reaction on the media. Data is presented in 3 different sentiment values – positive, negative and neutral. The analysis provides valuable input for DOSM to enhance its product and services.

[40] http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/1301.0Main+Features2822012

5. Use of web-scraped data in official statistics

- Statistics Poland is carrying out works on development of methodology and implementation of modern systems for measuring the change index in retail prices – regarded as a measure of inflation – using and linking new data sources, among others data from information systems of retail chains (scanned data), data obtained directly from their owners (private companies), data on retail prices from the Internet (including using web scraping techniques), administrative data. As one of the key macroeconomic indicators, the index of changes in retail prices is used in analyzes and on-going monitoring of economic phenomena and the condition of the economy;
- Statistics Poland is carrying works on development of a methodology and tools for collecting data on the labor market in a real time. An analysis of the possibilities of acquiring data from corporate portals on which job vacancies are posted has been made, using web scraping techniques. Job offers are a key indicator in the analysis of the employment market, used among others by the European Central Bank and the European Commission as a measure of the demand on the labor market and the business cycle.
- Statistics Poland obtains information from websites using web scraping and text mining techniques, for the purpose of building a database of enterprises. Currently, entrepreneurs show very high activity on the Internet, their actions include looking for employees, selling products or sharing information about their new investments, hence the above research approach allows for a thorough real-time analysis of the situation on the enterprise market, including information on the possible flows of people, goods, services and cash.
- Statistics Portugal has been using web scraping since 2015 to collect prices for the Consumer Price Index (CPI) with the view of quality gains, such as cost and time reductions, and improved coverage and frequency. A technical solution was developed and tested for data extraction, storage, processing and analysis for an existent commercial website, comparing prices collected in a traditional manner. A traditional price collected from the outlets of the targeted company has been substituted with the same data found by the web scraper prototype.The example of Statistics Portugal is also the one of a successful strategic partnership with a private company.[41]
- UK Office for National Statistics (ONS) has done some early work on using web scraped information about consumer prices to derive experimental statistics [42]
- New Zealand has issued a number of papers on using web-scraped data for CPI measurement in New Zealand.[43]
- The Australian Bureau of Statistics (ABS) have collected retail prices for the Consumer Price Index (CPI) using web scrapers since May 2016. Web scrapers are currently programmed and maintained by ABS staff using Microsoft excel scripts (Visual Basic for Applications), collecting approximately 500,000 prices per week across over 50 retailers. The ABS has adopted a phased approach to implementing web scraped prices for production purposes. The first phase was implemented in June quarter 2017 where the ABS implemented web scraped data into the CPI, using an average calculated price of a product over a given period – essentially replacing field collected data with a web scraped price (ABS, 2017). That is, in the same way we currently price a specific product (whether in the field, online or via some other method), we choose a specific product in the web scraped time series to use as the basis for a price. With respect to the next phase of implementation, the ABS continues development work on text mining techniques (to form broader product definitions, especially for clothing) and price index methods (both bilateral and multilateral index methods) that maximise and automate the use of web scraped data.
- Department of Statistics Malaysia (DOSM) adopts web scraping techniques and crawls online

[41] See: ESS position paper, p. 19.

[42] See https://www.ons.gov.uk/file?uri=/aboutus/whatwedo/programmesandprojects/theonsbigdataproject/researchindicesusingwebscrapeddatamay2016final002.pdf

[43] http://archive.stats.govt.nz/tools_and_services/newsletters/price-index-news/jan-18.aspx

price data from selected retailer's websites. The data is still being analysed and will be used as input for the computation Price Indices. The initiative is known as Price Intelligence.

6. Use of road traffic data in official statistics

- Statistics Poland is carrying research in the field of transport using data from sensors, such as the Automatic Ship Identification System (AIS) and Electronic Toll Collection System (viaTOLL), using stream processing which allows for data processing in a parallel and independent manner and for scaling the solution depending on the needs. In addition, a mechanism has been developed to enable quick analyses of real-time data from sensors together with historical data and data deriving from statistical surveys. Conducting transport analyses, using the above-mentioned data, allows to obtain new information on the volume of traffic, transport work, the volume of pollution emitted by transport, necessary to shape and monitor transport policy, including the selection of the most effective transport investments.
- The Department of Public Transportation (DPT) in Bermuda has tracked the number of passengers on public buses via a sensor fitted on each bus in the past. Upon the bus arriving at the central bus terminal, data was automatically downloaded once the bus sensor detected the terminal sensor. More than three million passenger journeys are typically recorded each year[44], indicating the large volume and frequency of events that were tracked automatically. This is a prime example of how the use of big data can be more accurate and efficient than the use of survey data which are subject to sampling errors and non-sampling errors, in addition to being costly and time consuming.

7. Use of scanner data in official statistics

- Statistics Poland is carrying out works on development of methodology and implementation of modern systems for measuring the change index in retail prices - regarded as a measure of inflation - using and linking new data sources, among others scanner data from information systems of retail chain.
- Statistics New Zealand uses scanner data for consumer electronics in our CPI measurement and continues to explore more opportunities.[45]
- Statistics Netherlands introduced supermarket scanner data into the CPI in June 2002 and in January 2010 they expanded the use of scanner data for the compilation of the Dutch CPI. The six chains for which scanner data were utilized at that time had a market share of around 50% and accounted for slightly more than 5% of the CPI-weight. In January 2013, scanner data fully replaced traditional price collection for supermarkets in the Dutch CPI.
- INSEE France has an objective of substituting scanner data owned by retailers for data collected in shops for industrial food products, household cleaning products and non-durable health and beauty products sold in food stores. They will be used to calculate the CPI in 2020. Since 2010, experimentations with 4 voluntary retailers representing 30 % of the market have been conducted[46].
- The Australian Bureau of Statistics (ABS) has used scanner (transactions) data in the production of the CPI since March quarter 2014. The ABS adopted a phased approach for implementation of scanner data, with the initial implementation replacing point-in-time prices for a sample of products (previously collected by field collectors) with a unit value (from scanner data). This implementation increased the coverage of scanner data to approximately 25% of the weight of the Australian CPI across several commodity classes including food, tobacco, and household non-durable products. From the December quarter 2017, the ABS

[44] Bermuda Government Department of Statistics (2014): *Bermuda Digest of Statistics*, p 77.
[45] http://archive.stats.govt.nz/browse_for_stats/economic_indicators/CPI_inflation/cpi-price-change-scanner-data.aspx
[46] See : ESS Position paper, p. 9. More information is available at https://www.INSEE.fr/en/statistiques/2912652

implemented a new methodology (known as multilateral index methods) to compile components of the Australian CPI using scanner data. This second phase of implementation overcome issues with traditional bilateral index methods and utilise a census of products available in these big datasets; uses expenditure data to weight products; and reduces data collection costs (ABS, 2016; ABS, 2017). In addition to its primary use in the Australian CPI, scanner data is utilised in other areas for both production and confrontation, including Retail Trade, National Accounts and Health surveys.

8. Use of CCTV data / Security videos

- The Bermuda Police Service has approximately 150 CCTV cameras across Bermuda which are monitored to assist with crime detection, 19 of which have license plate recognition capabilities. Incidents captured on the cameras are investigated before determining whether a crime has been committed and recorded as a statistic. Therefore, the big data captured by CCTV cameras indirectly leads to more comprehensive recording of crime statistics.

Attachment B – A Case Study of Argentina

1. Introduction

The National Institute of Statistics and Censuses (INDEC) is a technical deconcentrated body within the scope of the Argentine Ministry of Treasury. INDEC was created in 1968 by Law 17622 and leads all the official statistical activities carried out throughout the country. The legal framework was later expanded by executive orders 3110 (1970) and 1831 (1993) and INDEC Provision 176 (1999). The Director-General is at the level of a Secretary of State.

The current legal framework sets forth, among other matters, that it is INDEC's responsibility to implement a statistical policy for Argentina, to give structure and lead the National Statistical System (NSS), to design statistical methodologies, to organise and run statistical infrastructure operations and to produce basic indicators and social, economic, demographic and geographic data.

The NSS is legally centralised and functionally decentralised. This means that INDEC has regulatory authority over the central statistical services of the National Government and the peripheral services, such as the Provincial Statistical Offices. For some statistical operations, decentralisation is especially appropriate due to the specific topic or location.

During almost 50 years, INDEC has been regarded as an institution that promotes and complies with the highest professional standards in statistics within the region. Additionally, throughout its history, the Institute participated in numerous international initiatives to exchange statistical experience.

Until 2007, government administrations had respected INDEC's independence in methodologies, statistical processes and release of indicators and technical reports. However, independence had always been *de facto*, that is, not fully guaranteed by a legal framework, as in a *de jure* system.

2. Political intervention at INDEC (2007-2015)

Between 2007 and 2015, INDEC was politically intervened with a strong pressure by the government administration to produce and release statistical data in line with the official discourse ([47]). This meant that many of INDEC's technical reports were implausible to the population. Initially, the debate was mainly around the Consumer Price Index (CPI), but later distrust grew to most INDEC technical reports.

The political intervention caused significant institutional deterioration, including the replacement of several directors and the arbitrary displacement of a great number of technicians from their regular tasks, which in many cases meant resignation or dismissal.

The discretionary changes in methodology and procedures and, especially, the misleading results of statistical series, resulted in lack of credibility and mistrust in the information produced by INDEC, which caused strong scepticism among users and the general population, who began to use alternative indicators. As a consequence, INDEC also lost legitimacy within the Argentine government, which meant that it could not maintain its leadership over the National Statistical System.

There was great institutional weakness due mainly to the lack of a regulatory update to include modern standards of good statistical practice. INDEC's political intervention also failed to formalise operational and administrative processes. Additionally, both the building infrastructure and the infrastructure for technology and communication were neglected and at risk.

[47] Inter-American Development Bank, "Who wants to know? The Political Economy of Statistical Capacity in Latin America", Washington DC, United States of America, 2018.

INDEC's deterioration was also perceived by the international community. For example, the International Monetary Fund (IMF) publicly questioned the credibility of CPI and Gross Domestic Product (GDP) and, issued a Statement of Concern and Declaration of Censure in 2012 and 2013, respectively, in accordance with the Fund's legal framework.

3. INDEC's political intervention and the modes of undermining National Statistical Offices

INDEC's political intervention period can be placed within statistical theory. In fact, according to William Seltzer ([48]), there are 11 modes of institutionally undermining the work of a National Statistical Office (NSO). Many of these modes can be present simultaneously in an NSO when politically intervened.

However, not all threats to an NSO's integrity are from their own governments; they can also be from the media and powerful private or regional interests. In any case, most issues arise when government authorities either initiate or actively support these processes.

In the case of INDEC's political intervention, the scope of each mode of institutional undermining, as described by Seltzer (1994), was the following:

1. **Mission of the statistical service.** A statistical system whose primary aim is to provide data that the current political leadership want is running the risk of losing institutional credibility. Since INDEC only produced and released data consistent with the political discourse of the time, and not with the socio-economic reality, it only served a political purpose. In fact, one of the terms used during the political intervention process was "militant statistics", that is, statistics that had to support and sustain a political agenda. This meant a radical stray of the mission of the statistical service from that established in the regulation in force.
2. **Financial resources and controls.** Without resources, the Fundamental Principles of Official Statistics ([49]) cannot be applied. The lack of budget during the intervention period in INDEC was very clear, given the critical situation founded, by late 2015, in all infrastructure dimensions (building, IT and also statistical[50]).
3. **Staff.** A method of threatening the integrity and professional independence of a statistical service is to attack staff members, usually in the form of dismissing them or forcing their resignation. INDEC's political intervention either displaced, dismissed or forced the resignation of numerous specialised officials and replaced them with political activists. Additionally, job instability rose significantly due to flexible and politically driven hiring methods. Labour casualisation, compared to other periods, was significant.
4. **Suppression and change of statistical series.** Between 2007 and 2015 there were many interventions of this type, such as changes in indicator publication and processes, as in the case of the Consumer Price Index, Gross Domestic Product and Trade Balance. However, the most prominent case was INDEC's 2014 decision to indefinitely discontinue poverty and indigence by income indicators[51]. The index was becoming so low, due mainly to distortions in the basket used

[48] Seltzer, William, "Politics and Statistics: Independence, Dependence or Interaction?", Department for Economic and Social Information and Policy Analysis, United Nations, New York, United States of America, 1994.
[49]United Nations, "Fundamental Principles of Official Statistics", A/RES/68/261, New York, United States of America, 2014.
[50] The main statistical operations were outdated. Last valid National Agricultural Census (2002), National Economical Census (2004), and Income and Expenditure Household Survey (2004).
[51] https://www.lanacion.com.ar/1684626-sin-dar-explicaciones-el-indec-oculto-ayer-los-datos-de-pobreza-e-indigencia-de-2013

as deflator[52], that it was reaching near zero values, a fact that was mocked by public opinion. The claim of the then Minister of Economy was that continuing to measure these phenomena would "stigmatise" the poor[53].

5. **Definitions, concepts and methodology.** The definitions and concepts used in the collection, compilation and dissemination of statistics are the rocks on which the statistical contour of a country are built, although they are largely unseen. For this reason, they are sometimes the target of political manipulation, as in the case of INDEC's political intervention, by means of ad hoc methodological changes in the production of several statistical products.
6. **Terms and nomenclature.** Often the issues of nomenclature are so sensitive that they are decided at the political level. At INDEC, political influence was linked to the terms and nomenclature of the CPI, labour force and foreign trade.
7. **Altering specific numbers.** Statistical falsification runs counter to the very heart of the professional ethos of an NSO. Because of the interrelatedness of statistics, fabrication can easily be detected or is expensive to produce. This is one of the less used and more inconvenient modes of undermining an NSO. The suspicion of exhaustive manipulation of statistics during INDEC's political intervention period and the inconsistencies in the set of indices and the early detection, are in line with Seltzer's theory.
8. **Extent and timing of the release of data.** Manipulation as to the timing of the release of data quickly raises questions about manipulation of the data themselves, undermining the credibility of all government statistics. During INDEC's intervention, the dissemination schedule was flexible and unclear. Additionally, not all indicator databases were released.
9. **Threats to data confidentiality.** At the beginning of the political intervention, authorities pressured (and later displaced) INDEC officials to expose Consumer Price Index respondents. This breach of statistical confidentiality resulted in a court process against the authorities, which is still in progress[54]. These actions are a serious threat to data confidentiality, an essential element to ensure the accurateness of statistical programmes and to respect respondents' rights.
10. **Use of an NSO for political analysis or other political work.** The leaders of INDEC's political intervention publicly acclaimed the contribution of the statistical service to the agenda, noting the concept of "militant statistics". INDEC headquarters and staff were used for partisan meetings and campaigns, and the building was filled with the ruling party's posters, flags, pictures and leaflets.
11. **Active campaign to discredit statistical service outputs, methods or staff.** The political intervention persecuted the displaced officials and everyone who spoke against the process (such as independent analysts) with smear campaigns (even using INDEC's website[55]), in court, with fines[56] and with acts of verbal and physical abuse[57]. Many of these acts were taken to court, and verdicts finally favoured the victims, although some of them are still in process.

The following Table summarizes each aforementioned mode of institutional undermining:

[52] https://www.lanacion.com.ar/1740493-sin-datos-oficiales-subio-la-pobreza-en-el-ultimo-ano
[53] https://www.perfil.com/noticias/economia/kicillof-no-tengo-el-numero-de-pobres-es-una-medida-estigmatizante-20150326-0015.phtml
[54] https://www.infobae.com/politica/2018/08/27/pidieron-elevar-a-juicio-la-causa-contra-guillermo-moreno-por-la-manipulacion-del-indec/
[55] https://www.indec.gov.ar/nuevaweb/cuadros/novedades/nota_ta_28_09_11.pdf
[56] https://www.clarin.com/politica/Melconian-multa_a_consultoras-Guillermo_Moreno_0_rJxSNZN6wQg.html
[57] See "INDEC: Historia íntima de una estafa" [INDEC: An intimate story of fraud]. Gustavo Noriega. Buenos Aires. 2012.

Table 1. Modes of institutional undermining: INDEC under political intervention (2007-2015).

	Mode of undermining	Examples of the political intervention at INDEC (2007-2015)
1	Mission of the statistical service	Concept of "militant statistics"
2	Financial resources and controls	Lack of funding for statistics
3	Staff	Replacement of staff with political supporters
4	Supression or changes in series	Supression of poverty and indigence statistics (2013). Alteration of the CPI. Unexpected change of GDP base (March 2014)
5	Definitions, concepts and methodology	Ad hoc methodological changes
6	Terms and nomenclature	CPI. Labour market. Foreign trade
7	Altering specific numbers	Alteration of statistical series (CPI, poverty, GDP...)
8	Extent and timing of release data	Unclear dissemination schedule. Bases not published
9	Threats to confidentiality	Political pressure to reveal informants
10	Use of NSO for politics	Defence and promotion of militant statistics
11	Campaigns to discredit	Persecution of former officials and analysts

4. Reconstruction and statistical development (December 2015 - present).

Shortly after the new INDEC management took over, in December 2015, the new National Government declared administrative emergency of the National Statistical System and granted the Director-General broad powers to restore leadership of the Institute and make changes to the organisation structure.

Restructuring INDEC meant appointing new authorities to lead the rebuilding of their respective areas by incorporating qualified staff. As a result of this process, the official statistical service of the Republic of Argentina, led by INDEC, was eventually completely re-established, both in quantity and in quality. In 2016, the statistical emergency period concluded successfully.

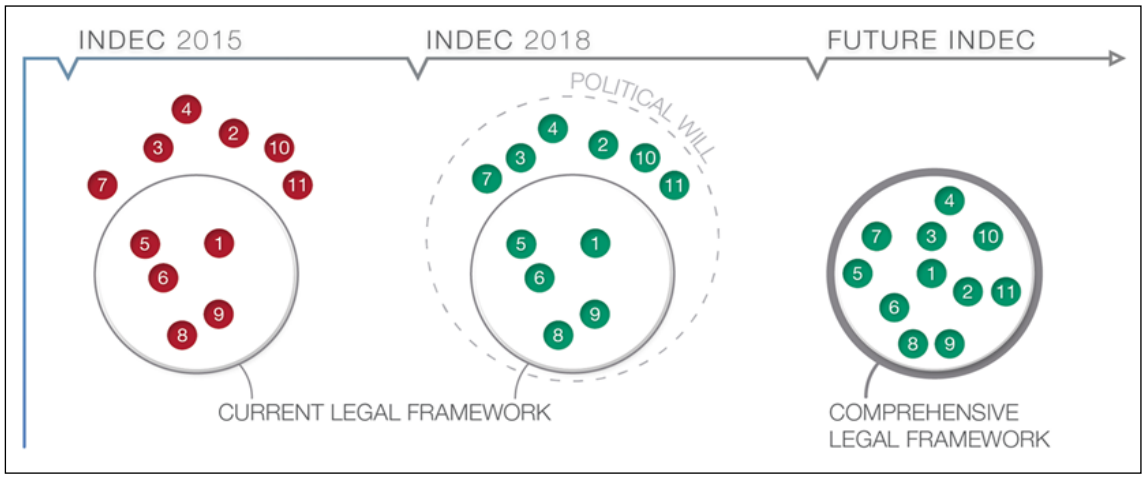
Later, INDEC followed a consolidation path based on five strategic pillars, due to the large deficits of the Institute: institutional transformation, strengthening statistical capacity, improving coordination of the National Statistical System, boosting dissemination and friendly access to statistics and developing international relations.

In connection with the institutional transformation, some adjustments to the organisational structure of the Institute (to promote a more efficient functioning) were introduced. These adjustments included, for example, selecting a new Management Team, designing a career development plan, modernizing its ICT infrastructure and even modifying INDEC's organization chart in order to, among other improvements, enable the creation of six new regional offices which will be located across the country's statistical regions.

In order to both guarantee its projected institutional development path and support the effective implementation of the Fundamental Principles of Official Statistics, INDEC presented a project for a new Statistical Law. Among the most relevant characteristics of the draft law that can be mentioned are the following: a) organizational and financial independence; b) strengthening the role of head of the NSS; c) a new policy for appointment of the Director-General of the Institute, through parliamentary agreement and with a fixed term and public contest for the appointment of staff; d) the use of new data sources, such as administrative records for official statistics; and e) the formal and explicit adoption of the United Nations Fundamental Principles of Official Statistics and the recommendations on statistical good practices of the organisations to which Argentina adheres (i.e., the OECD).

In connection with was recently mentioned, the following graph helps to explain the proposed changes, and how these legal adjustments could improve INDEC's institutional framework:

Graph 1. INDEC's recent and projected institutional evolution.



Note: The numbers depict the situation of the eleven modes of institutional undermining (as explained before). The ones whose situation was effective during the political intervention period (2007-2015) are depicted in red. On the contrary, the modes that were reversed due to the changes applied during the reconstruction and statistical development of INDEC (end of 2015-as of today) are depicted in green.

In connection with the strengthening of its statistical capacity, INDEC designed, implemented and projected several statistical operations that will allow the Institute to update and improve its fundamental statistical core, so as to enable INDEC to enhance the quality of its future production of statistical data and, also, to fill the statistical gaps wherever these were existent as of December 2015. This process includes statistical operations such as a National Household Expenditure Survey (2017-18), a National Mining Census (2016-17) and an Agricultural National Census (2018), as well as methodological improvements related to the Consumer Price Index and National Accounts.

INDEC also introduced several fundamental reforms in connection with its policies associated with the dissemination and friendly access to its statistics. In this sense, the Institute's web site was modernized, the presence of INDEC in the leading social media platforms was initiated and/or enhanced, statistical education programs were executed and relations with the media was restored and improved.

As for the international integration, INDEC resumed, as part of its institutional development policy, an active participation in the international area by strengthening its multilateral and bilateral relations in order to enhance and promote its modernization process and the building of new statistical capacities. The Institute's reinsertion in the international environment also means that the Directorate-General is committed to participating in the annual meetings of the United Nations Statistical Commission, the Statistical Conference of the Americas of ECLAC and the Specialized Meeting on Statistics of MERCOSUR. This international reinsertion also means that all INDEC teams participate in international technical meetings that allow them to remain updated on specific topics and to develop the Institute's statistical capacity.

Along INDEC's transformation and modernisation path, there have been several challenges that conditioned both the speed and the effective implementation of the work programme and reforms. This means that there are windows of opportunity that must be properly identified. There are three factors associated to consolidating the reconstruction of INDEC's statistical capacity:

1. **Public opinion.** It is imperative to have the support of society, to sustain and underpin the implementation of these changes.

2. **Political willingness and the role of decision makers.** INDEC's experience is evidence that a fundamental factor for the reconstruction of a crisis-ridden NSO is the will of the political power to revert the situation. In the Argentine case, the support of the Executive was particularly significant to INDEC's functional and technical independence and to effect the necessary changes.

3. **International reinsertion.** In Argentina, the international community's exposure and report was an important eye-opener. Similarly, the support of the statistical community was fundamental to the reconstruction of the Argentine National Statistical System, both as a guide and in terms of cooperation.

5. Conclusions

Considering the facts presented, the following conclusions may be drawn from the recent Argentine experience:

1. If there is a strong institutional framework to safeguard the functional and technical independence of an NSO, it is less likely that there will be a political intervention such as affected INDEC between 2007 and 2015.

2. It is beneficial to keep on promoting and disseminating the Fundamental Principles of Official Statistics and the recommendations on good practice.

3. The international statistical community is very important both in the detection of practices contrary to the Fundamental Principles of Official Statistics and in the process of rebuilding and strengthening the NSOs that have undergone important institutional crises.

4. It is necessary to have specific mechanisms and action protocols on how the international statistical community might act in cases of evident political interference that could distort the technical and functional independence of an NSO.

5. It is favourable to promote awareness of political leaders and other relevant social stakeholders (academia, media, private sector and others) on the institutional reality of NSOs and their statistical systems and, especially, on the consequences of having low-quality statistics and intentional distortion.

References

[1] Resolution adopted by the Economic and Social Council on 24 July 2013on the recommendation of the Statistical Commission (E/2013/24)2013/21. Fundamental Principles of Official Statistics

[2] United Nations Fundamental Principles of Official Statistics Implementation guidelines (Final draft, subject to editing, January 2015)

[3] Code of Principles in statistical cooperation of the Policy Group of Statistical Cooperation, Doc PGSC/2007/06

[4] UN Economic and Social Council, Resolutions. Supplement No. 1, Economic and Social Council official records, 50th session, 11–13 January and 26 April–21 May 1971. New York: UN, 1971.

[5] UN Economic and Social Council, Report of the Special Session of the Statistical Commission, New York, 11–15 April 1994, E/1994/29

[6] Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics amended by: regulation (EU) 2015/759 of the European Parliament and of the Council of 29 April 215

[7] Eurostat, European Statistics Code of Practice for the National Statistical Authorities and Eurostat (EU statistical authority). Adopted by the European Statistical System Committee 16th November 2017

[8] What does "big data" mean for official statistics?, United Nations Economic Commission for Europe, Conference Of European Statisticians, 10 March 2013What does big data mean for official statistics?, UNECE, CES, 10 March 2013

[9] Services concerning ethical, communicational, skills issues and methodological cooperation related to the use of Big Data in European statistics, Task 1. Ethical Review, Deliverable 1.3. Report on ethical guidelines June 2017.

[10] Abdulkadri Abdullahi, Evans Alecia, Ash Tanisha, An assessment of big data for official statistics in the Caribbean: Challenges and opportunities, January 2016

[11] Arora, Anil, Why are partnerships so important in the current context?, keynote speech from High-level Seminar on Strategic Partnerships in Official Statistics, Geneva, 11 April 2018

[12] Hendricks, Vincent.F., Jütting, Johannes, Trust in numbers? Why we need strong official statistics in an attention seeking society, http://www.paris21.org/news-center/news/trust-numbers-why-we-need-strong-official-statistics-attention-seeking-society

[13] Hendricks, Vincent F., Vestergaard, Mads, Reality Lost. Markets of Attention, Misinformation and Manipulation,
p.63., https://link.springer.com/content/pdf/10.1007%2F978-3-030-00813-0.pdf

[14] Michaud, Sylvie TA2.14 Non-official Data: Challenges and Opportunities for NSOs, Canadian Experience, UN World Data Forum, Dubai 2018.

[15] Perucci, Francesca Implementing the Fundamental Principles in a transforming statistical system, 15th IAOS Conference, Abu Dhabi, UAE, 6-8 December 2016

[16] Rozkrut, Dominik, Capacity development for international cooperation in a modern statistics office in the face of new challenges – business as usual or value added? , keynote speech from the Conference of European Statistics Stakeholders, Bamberg, 18 October 2018.

[17] Statistical Capacity Building for Sustainable Development in : Developing the fundamental pillars necessary for modern national statistical systems/Statistical Journal of the IAOS 33 (2017) 895–909 895 DOI 10.3233/SJI-160331 IOS Pres.

[18] Position paper on access to privately held data which are of public interest. Opening up new data sources for a new generation of official statistics – in light of the growing European Digital Single Market and the revision of the Public Sector Information Directive, November 2017 https://ec.europa.eu/eurostat/web/ess/-/access-to-privately-held-data

[19] Report by the UN Big Data Global Working Group Satellite Imagery and Geo-Spatial Data Task Team: https://unstats.un.org/bigdata/taskteams/satellite/

[20] Australian Bureau of Statistics 1352.0.55.144 - Research Paper: Methodological Approaches for Utilising Satellite Imagery to Estimate Official Crop Area Statistics (Methodology Advisory Committee) http://abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1352.0.55.144Sep%202014?OpenDocument

[21] Australian Bureau of Statistics (ABS). Making greater use of transactions data to compile the consumer price index. Information Paper 6401.0.60.003, Canberra: ABS; 2016.

[22] Australian Bureau of Statistics (ABS). Feature Article: The Australian CPI: A Contemporary Measure of Household Inflation. Cat. no. 6401.0, Canberra: ABS; 2017.

[23] Australian Bureau of Statistics (ABS). An implementation plan to maximise the use of transactions data in the CPI", Information paper 6401.0.60.004, Canberra: ABS; 2017.