



United Nations
Global Working Group on Big Data for Official Statistics
Task Team on Access and Partnerships

Deliverable 2: Principles for Access to Big Data Sources

Version 15 October 2015, for discussion at the 2015 Global Conference on Big Data for Official Statistics, Abu Dhabi, 20-22 October, 2015

This document contains a draft of a set of principles for access to Big Data sources together with a preamble listing a number of considerations. The principles and preamble are the core of deliverable 2 of the Task Team on Access and Partnerships (TTAP). The document also contains an introduction to the deliverable, explaining its aims and the process applied, and providing some background information.

Introduction

Reference to the deliverable

According to its Terms of Reference¹ (TOR), the objectives of the TTAP are “to facilitate access to Big Data sources for official statistics and facilitate forming partnerships with other public and private organisations in order to work with Big Data. This should be done in a way that reflects a mutual understanding with partners of what is reasonable to expect from each other, by respecting each other’s position, role, aims, business model, social responsibilities, limitations and possibilities.”

The TOR also states: “Access to Big Data sources and forging partnerships with other public and private organisations in order to work with Big Data is becoming ever more important to national statistical systems (NSS) for fulfilling their mission in society. The NSS should collaborate rather than compete with the private sector, in order to advance the potential of official statistics. At the same time, the NSS should remain impartial and independent, and invest in communicating the advantages of exploiting the wealth of available digital data to the benefit of the people. Building public trust will be the key to success.”

The TOR defines four deliverables. The second one is described as follows:

A set of principles setting conditions for access to Big Data sources. These principles cover rights and obligations of data providers, including globally operating data providers, as well as rights and obligations of national and international statistical organisations. The principles also cover the issues of data confidentiality and privacy protection and must link to the Fundamental Principles of Official Statistics of the UN Economic and Social Council.

In this document the TTAP proposes such a set of principles of access to Big Data sources, together with a preamble providing the context and main considerations for these principles. Hereafter these principles will be referred to as the Principles of Access. The Annex to this deliverable contains the Fundamental Principles of Official Statistics (FPOS) as adopted by the UN Social and Economic Council (ECOSOC) and later by the UN General Assembly².

Body of adoption

The TOR mentions that the Principles of Access are meant “to be formally adopted by an appropriate international body, as a reference for all national and global Big Data providers and national and international statistical organisations”. No decision has been taken yet concerning the international body to discuss and ultimately adopt the Principles of Access. This depends on the Statistical Commission, to which the Global Working Group reports. There could be a role for e.g. UN Global Compact, ITU, ECOSOC and/or the UN General Assembly. ECOSOC is the body that adopted the FPOS – which are also fundamental to the Principles of Access – before they were submitted for adoption to the General Assembly. For the Principles of Access, the same form as the FPOS has been chosen.

¹ Version 3 April 2015, available on [Trello](#).

² Resolution 68/261, adopted by the General Assembly on 29 January 2014.

The process, past and future

The TTAP carries out its tasks in accordance with a Work Plan³, which contains a list of the participating countries, organisations and team members. The participating parties represent various types of stakeholders at the national and international level, although the private sector was not represented. The team, which started its activities on April 1, 2015, has met nine times by video conference to discuss the contents and progress of the deliverables. The deliverables have been tuned to each other, so the Principles of Access are in line with the other deliverables. After discussing the Principles of Access within the TTAP, the other task teams of the Global Working Group on Big Data for Official Statistics were asked to comment on the document, thereby widening the involvement of stakeholders. These comments have been taken into account in this deliverable. The Global Conference on Big Data for Official Statistics, in Abu Dhabi in October, will provide the opportunity to widen the involvement of stakeholders even further, which is essential to improving the draft Principles of Access and gaining support for the Principles of Access to be eventually submitted to the appropriate international body for discussion and adoption.

Selection of the principles

The proposal lists eight Principles of Access:

- principle 1: on the social responsibility to provide data
- principle 2: on the adherence to the FPOS
- principle 3: on the need for a level playing field for businesses
- principle 4: on acknowledgement of business interests
- principle 5: on the need to provide adequate data, but no more than needed
- principle 6: on cost benefit proportionality
- principle 7: on non-discrimination of statistical agencies
- principle 8: on transparency

These principles were selected in order to cover all issues that need to be addressed in order to facilitate access to Big Data sources in practice. The first principle asserts the responsibility of businesses to provide data access; the second refers to the FPOS. However, businesses can only reasonably be asked to cooperate if their competitors are treated likewise (third principle), their legitimate business interests are recognized (fourth principle), adequate data is collected, but no more than needed (fifth principle) and the costs are reasonable compared to the public benefit that is expected (sixth principle). For statistical agencies, it is important, especially when dealing with Big Data providers that operate internationally, that similar data requests are treated similarly. The seventh principle is aimed at this. Finally, one of the main hurdles in practice for getting data access is gaining public trust and support. Transparency (eighth principle) is probably the most effective way of convincing the public that access to Big Data sources for official statistics is justified.

³ The latest version is of 12 May 2015, available on [Trello](#).

Selection of the considerations

The preamble to the Principles for Access lists a number of considerations. They are meant to underpin the eight Principles for Access and provide relevant context.

An example is the use of mobile phone location data for disaster relief, such as after the recent earthquake in Nepal. In response to the Ebola outbreak, GSMA guidelines⁴ have been developed on the protection of privacy in the use of mobile phone data. The first consideration refers to this.

Another example is the public discussion related to the revelations about wholesale data collection by the NSA and apparent breaches of confidentiality by other government bodies and institutions, and by certain private sector actors. Such practices are a threat to the accessibility of Big Data sources for official statistics. This piece of context is provided in the fifth consideration, which links this to a call for transparency.

Finally, the Principles for Access are very general, since they apply to all types of Big Data sources. Therefore, they may have to be worked out for different types of Big Data sources, and for certain branches, since businesses that are in a position to provide Big Data may be members of branch organizations. The need for elaboration according to type of Big Data sources and possibly branch is mentioned in the last consideration.

The last consideration also mentions the distinction between the use of Big Data for actual production of statistics and the preceding exploration phase. The principles refer to the use for production purposes, as is mentioned in the endorsement statement.

Preamble to the Principles for Access to Big Data Sources

The [*name of the appropriate international body*]

Taking notice of the high and urgent need for access to Big Data sources for the production of official statistics, such as indicators for the Sustainable Development Goals and statistics on phenomena related to modern society, and the social responsibility already shown by private organizations to provide access to Big Data sources, free of charge, for purposes such as disaster relief and the fight against epidemics,

Bearing in mind that in using such data the Fundamental Principles of Official Statistics, as endorsed by the UN General Assembly⁵, unconditionally apply, and that the statistical community has pledged to adhere to the professional ethics, as stated in the Declaration on Professional Ethics, as adopted by the International Statistical Institute⁶, thereby creating the foundation for sharing Big Data,

⁴ <http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2014/11/GSMA-Guidelines-on-protecting-privacy-in-the-use-of-mobile-phone-data-for-responding-to-the-Ebola-outbreak-October-2014.pdf>

⁵ Resolution 68/261, adopted by the General Assembly on 29 January 2014.

⁶ This declaration was adopted by the Council of the International Statistical Institute in its session of 22 and 23 July 2010, in Reykjavik, Iceland.

Recognizing the legitimate interests of businesses, including respect for their business model and value proposition, and the need to guarantee a level playing field for businesses considering the burden created by requesting access to Big Data for official statistics, as well as the legitimate interest of organizations in charge of compiling official statistics to have equal access,

Stressing that the burden to businesses resulting from data requests for official statistics must be fair in proportion to their envisaged public benefits and that the data should be adequate and relevant in relation to the purposes for which they are requested,

Considering that legislation specifically aimed at accessing and using Big Data is at an early stage of development, that existing national and international legal frameworks fully apply but need interpretation in view of Big Data, especially concerning privacy, data ownership, reuse of data by third parties, and liability in case of breaches of confidentiality, and that there is thus a need for agreed principles,

Highlighting the need to create public trust by applying full transparency in the use of Big Data for official statistics, in particular in view of privacy concerns, given a number of well-publicized cases of likely abuse outside the realm of official statistics, and the need to provide clarity concerning the possible use for statistical purposes of personal data in customer contracts with businesses, for instance by referring to the Principles for Access to Big Data Sources set out below,

Acknowledging that Big Data sources are diverse in many respects, such as owning the data versus providing a platform for the data of others, that source and branch specific principles and operational rules and guidelines may be needed for dealing with access to Big Data sources, that such rules and guidelines should be based on the Principles for Access to Big Data Sources set out below, that before access is requested for the purpose of producing official statistics data exploration may be necessary in collaboration with the Big Data source, and that this requires the development of partnerships between businesses providing and statistical agencies using data,

Endorses the Principles for Access to Big Data Sources for the production of official statistics set out below.

Principles for Access to Big Data Sources

Principle 1. The role of national and international systems of official statistics is to provide relevant, high-quality information to society in an impartial way. This role is indispensable to the well-functioning of societies. To this end, data is needed from individuals, households, businesses and institutions as inputs to these systems. These members of society have the social responsibility to make the data that is needed available to the statistical agency concerned, free of charge.

Principle 2. The data needed for official statistics may only be collected and processed if the statistical agency concerned acts in full accordance with the Fundamental Principles of Official

Statistics⁷. These principles guarantee, among other things, the professional independence and accountability of the statistical agency, and the strictly confidential use of the data, exclusively for statistical purposes.

Principle 3. When data is collected from businesses for the purpose of producing official statistics, the fairness of the distribution of the burden across businesses has to be considered, in order to guarantee a level playing field.

Principle 4. Data requests for official statistics must acknowledge and take into account the role of data in the business model and value proposition of businesses, in particular if their data has market value. There must be a fair balance between public and business interests when data is requested and possible harm to business interests has to be kept as low as possible.

Principle 5. The data must be adequate and relevant in relation to the purposes for which it is requested from the business. No more data should be requested than needed for these purposes. Operational arrangements have to be agreed on between the business and the statistical agency concerned, taking into account business concerns and data adequacy for official statistics. The metadata must also be adequate.

Principle 6. The cost and effort of providing data, including possible pre-processing, must be reasonable compared to the expected public benefit of the official statistics envisaged.

Principle 7. When businesses operate internationally, they are expected to treat requests for data from national statistical systems in a non-discriminatory way, unless different treatment is justified by differences in the national legislative frameworks concerned, and provided that adherence to the Fundamental Principles of Official Statistics and the Principles for Access to Big Data Sources is guaranteed in theory as well as practice.

Principle 8. Full transparency about the provision of data to national and international statistical systems has to be exercised by businesses as well as statistical authorities in order to ensure public trust and to support a level playing field for businesses and an equal treatment of national statistical systems by businesses. In particular, businesses have to make sure that their customers are aware of the possible use of their data for the purpose of official statistics, in full adherence with the Fundamental Principles of Official Statistics.

⁷ <http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>

Annex to deliverable 2: Fundamental Principles of Official Statistics

The Economic and Social Council

Recalling recent resolutions⁸ of the General Assembly and the Economic and Social Council highlighting the fundamental importance of official statistics for the national and global development agenda,

Bearing in mind the critical role of high-quality official statistical information in analysis and informed policy decision-making in support of sustainable development, peace and security, as well as for mutual knowledge and trade among the States and peoples of an increasingly connected world, demanding openness and transparency,

Bearing in mind also that the essential trust of the public in the integrity of official statistical systems and confidence in statistics depend to a large extent on respect for the fundamental values and principles that are the basis of any society seeking to understand itself and respect the rights of its members, and in this context that professional independence and accountability of statistical agencies are crucial,

Stressing that, in order to be effective, the fundamental values and principles that govern statistical work have to be guaranteed by legal and institutional frameworks and be respected at all political levels and by all stakeholders in national statistical systems,

Endorses the Fundamental Principles of Official Statistics set out below, as adopted by the Statistical Commission in 1994⁹ and reaffirmed in 2013, and recommends them further to the General Assembly for endorsement.

Fundamental Principles of Official Statistics

Principle 1. Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, 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.

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.

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.

⁸ These include General Assembly resolution 64/267 on "World Statistics Day", and Economic and Social Council resolutions 2006/6 on strengthening statistical capacity and 2005/13 on the 2010 World Population and Housing Census Programme.

⁹ For the original preamble used on the occasion of the initial adoption of the Principles in 1994, see the report on the Statistical Commission on its special session (E/1994/29), chapter V, available from <http://unstats.un.org/unsd/statcom/94report.pdf>. More information on the Fundamental Principles and their history is available from <http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>.

Principle 4. The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

Principle 5. 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.

Principle 6. 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.

Principle 7. The laws, regulations and measures under which the statistical systems operate are to be made public.

Principle 8. Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.

Principle 9. 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.

Principle 10. Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.