BUILDING STATISTICAL CAPACITY

THE CHALLENGES

Graham Eele

PARIS21
Partnership in Statistics for Development in the 21st Century

Discussion Paper No. 7
July 2015

This report was commissioned by the OECD's PARIS 21 Secretariat as part of a research project for the Bill and Melinda Gates Foundation, titled 'Informing a Data Revolution'.

Disclaimer

The opinions expressed in this paper are those of the author and should not be attributed to the PARIS21 partnership or its members.

Building Statistical Capacity

The Challenges

By Graham Eele

CONTENTS

BUILDING STATISTICAL CAPACITY THE CHALLENGES

1. INTRODUCTION

The Informing a Data Revolution (IDR) project[1], funded by the Bill and Melinda Gates Foundation, was set up in 2014 to take stock of the situation in developing countries and to identify the problems they have in producing and using data in their national statistical systems. IDR has published a Road Map for a data revolution[2], launched in April 2015, setting out a broad programme of action to help developing countries meet data needs related both to monitoring progress towards the Sustainable Development Goals (SDGs), as well as to supporting national priorities. The Road Map makes a series of recommendations for strengthening national statistical capacity aligned along four main themes: governance and leadership; capacity building; principles and standards, and; technology and innovation.

This paper was prepared as part of the background work for the IDR project and looks at the challenges facing developing countries in building statistical capacity over the next 15 years.

2. INVESTING IN NATIONAL STATISTICAL SYSTEMS

2.1. WHY NATIONAL STATISTICS?

The main focus of the Road Map is on how the data revolution can best support and strengthen the capacity of national statistical systems, especially those in developing countries. In this paper we look at how these systems work and discuss the main challenges they face and are likely to face over the next 15 years. We also present some analysis of how the international statistical system works and how it interacts with what is happening in countries.

The need for nation states to collect and compile statistics about their population, their economy and their resources has been recognised for millennia. From the census described in the Bible[3] to the "political arithmetic" of John Graunt and Sir William Petty[4], the motivation of these efforts at compiling statistics was to support the raising of taxation and the management of the affairs of state. In most developing countries this process has continued from the early efforts of colonial governments to establish systems and processes to compile data on some aspects of government activity, to the more sophisticated and complex arrangements in place today[5].

Governments of almost all political persuasions have found it necessary to establish statistical units and departments to carry out and lead this work. In general this is because most official statistics – statistical information, collected, compiled and

disseminated by governments – have the characteristics of public goods[6]. Because it is generally not possible to reply on market provision, if governments wish to compile data for their own use and for the use of others, then they need to establish their own statistical systems. At the global and regional levels, where, for example, the data needed to monitor the SDGs will be compiled, then international coordination is required[7].

If better statistics are to be provided and if this provision is to be sustained, therefore, then the data revolution must focus on investing in the capacity of national statistical systems and in better international coordination. The focus will need to be more on the processes, the people and the systems that generate the data – the infrastructure of the statistical system – than on indicators per se. Capacity is needed to identify what data are required, to collect the source data, to process this to ensure consistency and reliability, to compile indicators, to make this information accessible, to document what has been done and to ensure consistency with other processes.

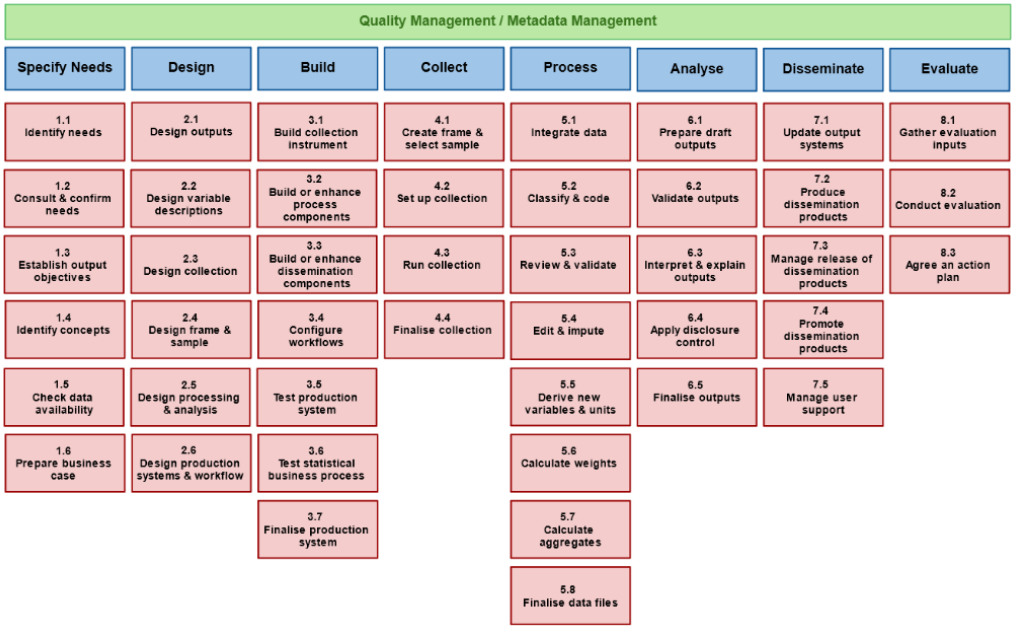
2.2. UNDERSTANDING THE STATISTICAL PROCESS

Developing, managing and improving statistical processes are not trivial tasks (see Box 1). Because data users are not able to assess the value and usefulness of any specific statistic or even a data set simply by examining the data themselves, special care is needed to ensure that methods and procedures are properly documented and that this information is available and accessible. If the data and the resulting statistics, indicators and analysis are to be of value then they must not only be of good quality, they must be seen to be so. It is essential that statistics are collected and compiled in line with agreed standards and principles so that users are able to trust the data and are able to determine to their own satisfaction that they are fit for purpose. Because the compilers of the data – the component parts of national statistical systems – will not know in advance who will use their data, or in what ways, it is very important that they have in place appropriate and robust mechanisms and processes to monitor data quality and to ensure the reliability of their data as far as possible.

The statistical process, if it is to be effective, must generate reliable and trustworthy data that can be used with confidence. The process is often complex and involves much more than simply collecting, compiling and disseminating data[8]. The value of data is greatly increased if it can be compared with information from other sources so that changes can be reliably identified and time series data built up to see how different phenomena change over time. This requires the use of standards and methods which ensure that like is being compared with like.

Box 1 - The Generic Statistics Business Process Model

The Generic Statistics Business Process Model (GSBPM) is an international standard describing and defining the processes needed to produce official statistics. The GSBPM can be used for harmonising statistical computing infrastructures, facilitating the sharing of software components, and providing a framework for assuring and improving the quality of the resulting data and statistics.



2.3. THE IMPORTANCE OF STANDARDS, ETHICS AND PRINCIPLES

The organisation and management of statistical processes also involves ethics and adherence to important principles[9]. Concerns such as professionalism, truthfulness and integrity underpin the statistical process in all fields of activity. In official statistics the need is even greater, adherence to these values and to professional ethics must be in place, but must also be clear to everyone with an interest[10]. To help ensure a uniform approach to these concerns, the UN has established the Fundamental Principles of Official Statistics (see Box 2), to promote the integrity, independence and objectivity of statistics compiled and published by Governments.

While the importance of established ethical and scientific principles is well established and widely accepted in the official statistics community, it will be important that these ideas are also applied more widely as the data revolution is rolled out. In particular, it will be essential to continue to protect the confidentiality of information about individuals and to promote procedures that ensure the integrity of the data. While the process of ensuring adherence to the fundamental principles is generally based on national legislation as well as international agreement[11], different mechanisms are likely to be needed for other types of data, especially “big data” derived from different business processes.

Box 2.2 Fundamental Principles of Official Statistics

In 1991, the Conference of European Statisticians developed and adopted a set of principles governing official statistics. In recognition of the need to ensure that national statistical systems are able to produce appropriate and reliable data that adhered to certain professional and scientific standards, the United Nations Statistical Commission adopted this very same set of principles in 1994. The principles have not been revised since and are still considered as relevant today as they had been in the past.

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.

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.

The definition, application and updating of standards is also an important aspect of official statistics at all levels. These may cover issues such as definitions, classifications, data collection methods, compilation procedures, and quality control processes. The application of standards has two main functions. The first is to ensure that data from different sources can be safely compared. The second is to inform data users that the data meet a defined technical standard and that confidence can be placed in the results. At the national level, the responsibility of setting and applying standards is usually defined by law and is typically the responsibility of the main national statistical agency. Internationally, the development and application of statistical standards is supported by agreement, rather than by a formal legal process. In the field of statistics, the United National Statistical Commission has the responsibility for setting and monitoring the use of standards. Again, an important concern of the data revolution is how agreement can best be reached on the definition and application of standards for new kinds of data.

2.4. WHAT MAKES AN EFFECTIVE NATIONAL STATISTICAL SYSTEM?

Ultimately, the effectiveness of any national statistical system is determined by the data it collects and the outputs it produces. Just producing data, however, is not sufficient, the indicators, data series and other outputs must meet the needs of users and must be provided in a form that supports their widespread use. The data outputs need to be of good quality, published within a time frame that means the data are still relevant and provided in a format that users are able to access and use. It is also important that statistical agencies are open and transparent about their methods and procedures and provide access to adequate metadata – detailed descriptions of the methods and procedures used to produce the data. Building and maintaining trust in the published statistics is an essential part of the work of all national statistical systems.

In order to document understand how national statistical systems operate and what they are able to do, PARIS21 has established a Metabase[12], which sets out how statistical systems operate in the following six main dimensions.

- **Access:** The public availability of data is the foundation of a better-informed society.
- **Innovations:** The ability of a statistical system to identify and use innovations is important in improving effectiveness and efficiency.
- **Timeliness:** Timely data helps decision makers react quickly and stay informed.
- **Soundness:** Making sure that statistics are produced using sound methodology builds trust in data and ensures transparency.
- **Institutions:** A healthy institutional environment is essential for statistical development.
- Use: Knowledge on the demand for data improves efficiency in data production.

The Metabase builds on and is closely associated with work already carried out by the World Bank through their recently re-launched Data on Statistical Capacity

website[13]. Since 2004, the World Bank has calculated and published a Statistical Capacity Indicator (SCI), which assess the extent to which a developing country's statistical system adheres to international statistical standards and methods, whether or not statistics are published in line with internationally recommendations and whether important statistical data are available in a timely fashion. The Indicator includes 25 indicators that monitor and "grade" a country's statistical capacity progress annually and form the basis for the calculation of the overall SCI score. By using the data in the PARIS21 Metabase and the Data on Statistical Capacity website, national statistical systems are able to monitor their own operations and development progress. At the same time, users are able to see what is currently being done and how well their statistical system meets international standards and good practice.

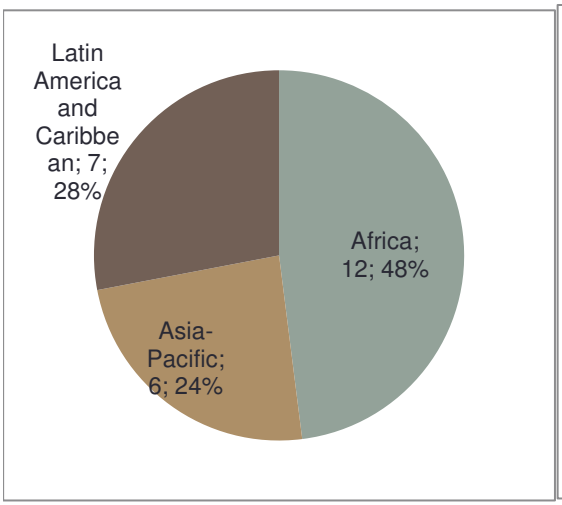
2.5. STRENGTHS AND WEAKNESSES OF NATIONAL STATISTICAL SYSTEMS

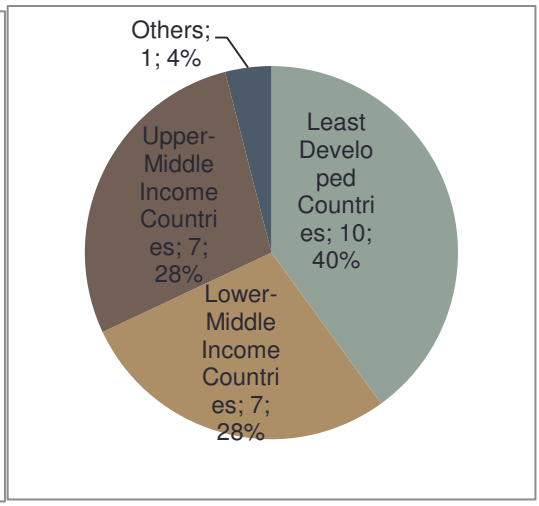
A cross-country study was carried out as part of the IDR project, covering 25 countries overall together with six more detailed case studies (see Box 3). Although there are some limitations to this study – including the non-random nature of the group of countries included and the fact that responses were only provided by national statistical agencies – nevertheless, it is felt that the results are both useful and illustrative of the facing National Statistical Offices (NSOs) in 2015.

Box 3 - Methodology

Questionnaires were sent to investigators from 43 National Statistical Offices. 27 questionnaires were returned, and most of the analysis is based on a sample of 25 NSOs.

The geographic and income level distributions of the sample are shown below. The 25 countries also include 16 IDA countries, 3 Small Island Developing States (SIDS), 5 fragile states and 6 landlocked countries.





The average score using the World Bank Statistical Capacity Indicator in 2013 is 71 compared with the average for all countries of 66.

Quintile	Countries in the analytic sample	
20% (0-50)	DR Congo*, Gabon	2
40% (51-63)	Burundi*, Ghana, Nepal	3
60% (64-72)	Bangladesh*, Cabo Verde*, Cambodia, Mali, Pakistan, Senegal, Vietnam, Trinidad and Tobago*	8
80% (73-80)	Bolivia, Costa Rica, Malawi, Mozambique, Nigeria, Dominican Republic, South Africa, Uganda	8
100% (81-100)	Colombia*, Mexico, Peru, Philippines*	4

Botswana and Tanzania did not follow pre-defined response categories, to be followed up. Recoding was done for Bangladesh and Malawi. Respondents who filled out the questionnaires include Director General, Deputy Director General or senior director of the corresponding NSO.

2.6. ISSUES ARISING FROM THE COUNTRY STUDIES:

Some of the key messages coming out the country study are as follows.

- **Coordinating national statistical systems for better statistical quality:** NSOs have recognized system-wide coordination and management as a key priority, particularly for improving data quality. Despite the fact that a number of formal processes and mechanisms have been put in place, there is still a lack of coordination of data processes among different data producers. A lack of system-wide coordination has impeded the implementation of statistical standards and quality measures, which is often limited to the NSOs themselves. To this end, NSOs expressed the need for skills development and technical assistance in effective, system-wide coordination and management.

- **Investing in people and skills development**: There is a strong call for financial investment in human resources. In particular, NSOs identified training and skills development in the design of statistical processes (statistical methodologies/ process/workflow), in data analysis and in strategic planning as the three main priority areas.

- **Improving data dissemination and use:** Among most statistical production processes, data dissemination and use is a key area for development in the next five years. Developing data dissemination policies and improving data documentation are recognized as a key priority. NSOs also hope to harness the potential of ICT and new technologies to further improve data access and use. In general, data dissemination and use is the area where there is the highest level of demand for technical assistance and innovations.

- **Harnessing the power of ICT:** The potential of ICT and innovation has been widely recognized by NSOs. There is a strong call for applying ICT to various statistical production processes from data collection, capture, processing, analysis, dissemination to archiving and storage. The development of ICT also presents challenges and many NSOs need more resources for investment and technical assistance to be able to use new tools and technologies effectively.

- **Strengthening statistical process design and management**: The need for improving statistical processes not only requires the NSOs to improve IT and statistical infrastructure, but also, to develop or redesign statistical processes and related management models and standards. The demand for process design is related to the increasing use of new data sources, especially registers or administrative sources and geospatial information. At the same time NSOs are looking for advice and support to manage the proliferation of ICT tools and technologies available for statistical processes, as well as the new and emerging demand for statistics. In particular, there is a demand for skills development and financial investment in statistical process design and management.

- **Aid in statistics is aligned with national priorities, but is not always delivered in line with a system-wide approach.** While most countries consider the current financial and technical aid for statistics well aligned with national priorities, many also pointed out that receiving aid in statistics imposes additional costs on the NSO and the national statistical system as a whole. In many countries, the delivery of financial aid in statistics is still not consistent with a system-wide approach and countries have difficulties in managing aid projects involving many donors.

The in-depth case studies provide more insights into these and some other issues and concerns and illustrate the inter-linkages between them. For example, all statistical systems face financial pressures, but these are often relatively greater in the smaller agencies in the lower income countries. Without well-qualified staff it is often difficult to take advantage of the opportunities provided by new technology and NSOs remain dependent n external technical assistance in many cases. Financial and human resource pressures also make it difficult for agencies to invest sufficiently in training. In the smaller statistical agencies it can be difficult to release keys staff for training especially overseas.

Other concerns include the difficulties of maintaining relations with data users and in improving access and the availability of statistics. In low-income countries both providers and users of statistics have limited capacity and while the potential for improvement is high, short-term resource constraints mean it is often difficult to make significant progress.

The problems that were identified in the cross-country study associated with the design and management of aid projects are also reflected in the case studies. Some countries find it difficult to participate actively in the international statistical system, they feel that there is rarely sufficient support for training and human resource development and they are concerned that funding from donors is too limited and unpredictable.

3. STRENGTHENING THE INTERNATIONAL STATISTICAL SYSTEM

3.1. THE ROLES OF THE INTERNATIONAL STATISTICAL SYSTEM

The international statistical system interacts with national statistical agencies and systems in a number of different ways. Over time, a complex and diverse international statistical architecture has evolved, developing and implementing standards and statistical methods, providing support to and coordination for national statistical systems, compiling consistent international data sets on a wide variety of topics, but especially the indicators needed to monitor the Millennium Development

Goals and mobilising and delivering financial and technical assistance to developing countries.

All of these tasks are important and will continue to be needed in the post-2015 period. Statistical standards cover classifications, quality frameworks, and data exchange protocols, and increasingly, technical standards for describing statistical processes and sharing software applications. Such standards aim to "industrialise" statistical systems to make re-use and implementation more efficient. Standards are the necessary foundation for modernisation of statistical processes.

The compilation of international data sets has been an important activity for a considerable period of time, but the emphasis on the need for consistency and, consequently, the need for much more complete metadata, became much clearer following the Millennium Summit and the launch of the MDGs. The need to make the most effective use of fairly limited resources, as well as the recognition that effective coordination was going to be essential if data needs were to be met, has led to the establishment of coordination mechanisms, especially with the UN system, including the Inter-Agency and Expert Group (IAEG) on MDG Indicators**.** In the 2015 and beyond, the launch of the Sustainable Development Goals will increase attention on this topic, especially as the number of indicators is expected to increase substantially over the MDGs. Even without the need to monitor global development goals, however, the need for internationally consistent data sets is expected to increase as trade and other international interactions grow. The aim must be to increase the quality of the indicators in these data sets and to ensure that as much as possible they reflect reality on the ground.

Since 2000, attention has also focused on the capacity of national statistical systems to provide the data needed to lead and to monitor development. New initiatives have been put in place within the international community to provide new resources to invest in capacity, to provide guidance and support and to ensure that countries had access to new tools and technology. The Marrakech Action Plan for Statistics, updated through the Busan Action Plan in 2011, provided a framework for prioritising assistance and, most importantly, for helping countries to set their own targets goals through the preparation of national strategies for the development of statistics.

3.2. THE RELATIONSHIP BETWEEN NATIONAL AND INTERNATIONAL SYSTEMS

To a large extent the development of the international statistical system has not been planned. Organisations and roles have evolved as needs have become apparent. In practice, then its structure and governance is diverse and, at first sight, somewhat overlapping and confusing. All the main international agencies have their own governance processes and the main responsibility of the agencies and their staff is to their own boards and management structures. It is also the case that the international system is chronically short of money for recurrent costs and there are perennial problems associated with the financing of global public goods.

Looked at from the point of view of developing countries, the current structure of the international system has some strengths, but also some important weaknesses. To some extent, it is possible for countries to get access to advice and support and if the first agency approached is unable to help then there are others that also have the expertise. There is however, the potential for overlapping initiatives and dealing with a multiplicity of agencies does impose real costs on under-resources and under-staffed statistical systems. Many countries find it difficult to participate actively in international and regional governance processes and the views of low income countries can easily be overlooked when work programmes and priorities are being decided.

For many countries the international statistical system is the main source of new knowledge and new technology. Small statistical agencies have neither the resources nor the expertise to innovate and rely on information about new technology and new ways of doing things coming from outside. One concern with this process, however, is that of duplication, especially in the area of information technology. Different agencies come up with potential solutions and new applications independently and these are then disseminated to countries through a variety of means including technical assistance programmes. Countries haver little or no input into the design of these products and are often obliged to take them, if they want to have access to the other parts of the package such as training and technical advice.

The duplication of both actors and products in different areas of statistics indicate that effective coordination remains a serious challenge. A number of different processes to strengthen coordination and to promote accountability have been put in place in recent years. But while coordination has been widely espoused, in practice it has often proved more difficult to put into effect.

4. Conclusions and Recommendations

While the Data Revolution undoubtedly provides many opportunities for making use of new types of data and new methods of analysis and dissemination, it is also true that continued investment in the capacity of national statistical systems in developing countries remains a priority. A substantial proportion of the data needed to support development programmes and to monitor progress towards the SDGs will continue to be generated by these systems and by the surveys, censuses and other data collection programmes they manage. New partnerships will be needed and new kinds of data should be identified and brought into the ambit of official statistics, but this will require both careful management and new skills and expertise.

Considerable progress has been made in building and strengthening statistical capacity since the launch of the Millennium Development Goals in 2000, but remains still to be done. If the challenge of the SDGs of eliminating absolute poverty by 2030, leaving no one behind is to be met, then existing data processes must be improved supplemented by new data from the private sector, academia and civil society. It will be essential, however to ensure that these new data sources are based on the same fundamental principles that have underpinned official statistics – the data must be objective and methods must be transparent and open to rigorous scrutiny.

Another challenge is not only to collect and compile data, but crucially to turn this into useful and useable information. This too will require new tools and new skills as well as much more effective use of new technology. In an increasingly connected world, there should be no reason why everyone should not have access to information to make decisions and to hold policy makers to account.

Finally we need to make sure that the international statistical system works more effectively, both to support efforts at the country level, but also to compile the data needed to assess progress globally. It will be important, for example, to develop and strengthen data standards allowing data from different sources to be more easily compared and used. There is also a need to achieve a significant increase in global data literacy, accessibility and use, in support of the post-2015 development agenda.

REFERENCES

---

[1] URL: http://datarevolution.paris21.org/

[2] URL: http://datarevolution.paris21.org/sites/default/files/Road_map_for_a_Country_led_Data_Revolution_web.pdf

[3] Luke 2, 1-7

[4] K Pearson, "The History of Statistics in the 17th and 18th Centuries", Charles Griffin, London 1978 p 2

[5] Eele, G. "The Organisation and Management of Statistical Services in Africa: Why does it Fail?" World Development, 16 (9), 1988.

[6] Tendulkar, S. "Demand for Better Statistics and Use of Data" PARIS21 http://paris21.org/sites/default/files/datause-Tendulkar-Consortium09.pdf

[7] Round, J. "Assessing the demand and supply of statistics in the developing world" PARIS21 Working Paper no.2, 2014

http://www.paris21.org/sites/default/files/PARIS21-DiscussionPaper4-Demand.pdf

[8] Center for Global Development "Delivering on the Data Revolution in Sub-Saharan Africa", http://www.cgdev.org/publication/delivering-data-revolution-sub-saharan-africa-0?callout=1-4

[9] For example, see the International Statistical Institute's Declaration on Professional Ethics published on World Statistics Day 20 October 2010, http://www.isi-web.org/about-isi/professional-ethics/43-about/about/296-declarationprofessionalethics-2010uk

[10] Seltzer, W. "Official Statistics and Statistical Ethics: Selected Issues", International Statistical Institute, 55th Session, 2005, http://unstats.un.org/unsd/WS%20%202005%20ISI%20paper.pdf

[11] https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx

[12] http://datarevolution.paris21.org/metabase

[13] http://datatopics.worldbank.org/statisticalcapacity/

PARIS21

www.PARIS21.org

Discussion Paper No. 7

July 2015