Brussels Rural Development Briefings
A series of meetings on ACP-EU
policy development issues

Data: the next revolution for agriculture in ACP countries

http://brusselsbriefings.net

This Reader was prepared by
Chris Addison, Senior Expert, Knowledge Management, CTA
Isolina Boto, Manager, CTA Brussels Office
Lebo Mofolo, Independent Expert

Briefing[1] n. 40

Data: the next revolution for agriculture In ACP countries

Brussels, 15 February 2015

This Reader was prepared by

Chris Addison, Senior Expert on Knowledge Management, CTA

Isolina Boto, Manager, CTA Brussels Office

Lebo Mofolo, Independent Expert

With the assistance of Yentyl Williams, Research Assistant, CTA Brussels Office

The information in this document was compiled as background reading material for the 40th Brussels Development Briefing on "Data Revolution". The Reader and most of the resources are available at http://brusselsbriefings.net

Table of Content

1. Context: The Data revolution: from data collection to real-time digital data

Despite the key role of agriculture as economic sector, serious weaknesses persist in the measurement of agricultural outcomes and in our understanding of the factors hampering agricultural growth among smallholders. While governments and donors target agriculture for large-scale investments with ambitious goals of raising agricultural productivity, little is done to ensure that accurate statistics are produced to monitor agricultural development.[2]

Agricultural statistics continue to suffer from poor quality, lack of relevance, insufficient funding and little use in national policy dialogues and the poorest countries – for which agriculture is a critical source of livelihood – often have the poorest data in quality and scope, being least able to direct their limited resources into improving the quality of their statistics to be able to inform policies (African Development Bank).

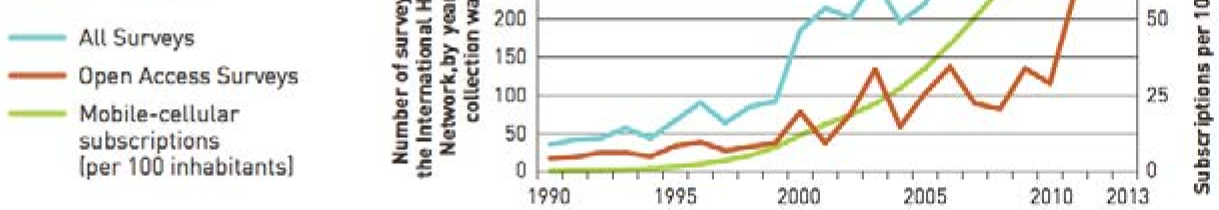
Today, however, the increased volume, velocity and variety of data used across the economy, and more importantly their greater social and economic value, signal a shift towards a data-driven socioeconomic model. In this model, data are a core asset that can create a significant competitive advantage and drive innovation, sustainable growth and development.[3]

Data has become a key asset for the economy, as important as human and financial resources. Whether it is geographical information, statistics, weather data, research data, transport data, energy consumption data, or health data, the need to make sense of "Big data" is leading to innovations in technology, development of new tools and new skills.

More data was created in 2013 than in all the preceding years of human history combined, and every minute the world generates enough data to fill more than 360,000 standard DVDs.[4]

These data hold the potential—as yet largely untapped— to allow decision makers to track development progress, improve social protection, and understand where existing policies and programmes require adjustment.[5] Whereas in previous generations, a relatively small volume of analog data was produced and made available through a limited number of channels, today a massive amount of data is regularly being generated and flowing from various

THE GROWTH OF DATA: TRENDS IN DATA AVAILABILITY, DATA OPENNESS AND MOBILE PHONE USE



Source: * International Household Survey Network (http://catalog.ihsn.org/index.php/catalog). For a detailed analysis of global trends in Survey data availability, see, e.g., Demombynes and Sandefur (2014), "Costing a Data Revolution." Center for Global Development, Working Paper 383.

** World Bank (http://data.worldbank.org/indicator/IT.CEL.SETS.P2), based on data from the International Telecommunication Union (ITU), World Telecommunication/ICT Indicators database sources, through different channels, every minute in today's Digital Age.

The "traditional data" (official statistics and survey data) will continue to generate relevant information, but the phenomenon of big data - where information comes from different sources ranging from connected devices to sensors and GPS - offers enormous potential to develop innovative products and services. Turning Big Data—call logs, mobile-banking transactions, online user-generated content such as blog posts and Tweets, online searches, satellite images, etc.—into actionable information is key.

The data revolution is not restricted to the industrialized world. The spread of mobile phone technology to the hands of billions of individuals may be the single most significant innovation that has affected developing countries in the past decade. Across the developing world, mobile phones are used daily to transfer money, buy and sell goods, and communicate information including test results, stock levels and prices of commodities. The numbers of real-time information streams and people using social media are growing rapidly in developing countries as well.

Some concepts[6]

Real-time digital data

The growing role of 'crowdsourcing'[7] and other "participatory sensing" efforts bringing together communities of practice and like-minded individuals through the use of mobile phones and other platforms including Internet, hand-held radio, and geospatial technologies etc. and advances in computing and data science now make it possible to process and analyze Big Data in real time. However, due to its size and often complex and unstructured nature, Big Data presents several analytical challenges that demand continually updated tools and expertise.

Data Revolution

The opportunity to improve the data that is essential for decision-making, accountability and solving development challenges. Coined in May 2013 in the report of the High-Level Panel of Eminent Persons on the post-2015 Development Agenda, the "data revolution" is an explosion in the volume of data, the speed with which data are produced, the number of producers of data, the dissemination of data, and the range of things on which there is data, coming from new technologies such as mobile phones and the "internet of things", and from other sources, such as qualitative data, citizen-generated data and perceptions data;

Big Data

BD is a popular phrase used to describe a massive volume of both structured and unstructured data that is so large that it's difficult to process with traditional database and software techniques. The characteristics which broadly distinguish

Big Data are also called the "Five Vs": Volume, Velocity, Variety, Veracity and Value.

Volume refers to the vast amounts of data generated every second. Just think of all the emails, twitter messages, photos, video clips, sensor data etc. we produce and share every second. We are not talking Terabytes but Zettabytes or Brontobytes. On Facebook alone we send 10 billion messages per day, click the "like' button 4.5 billion times and upload 350 million new pictures each and every day. If we take all the data generated in the world between the beginning of time and 2008, the same amount of data will soon be generated every minute! This increasingly makes data sets too large to store and analyse using traditional database technology. With big data technology we can now store and use these data sets with the help of distributed systems, where parts of the data is stored in different locations and brought together by software.

Velocity refers to the speed at which new data is generated and the speed at which data moves around. Just think of social media messages going viral in seconds, the speed at which credit card transactions are checked for fraudulent activities, or the milliseconds it takes trading systems to analyse social media networks to pick up signals that trigger decisions to buy or sell shares. Big data technology allows us now to analyse the data while it is being generated, without ever putting it into databases.

Variety refers to the different types of data we can now use. In the past we focused on structured data that neatly fits into tables or relational databases, such as financial data (e.g. sales by product or region).

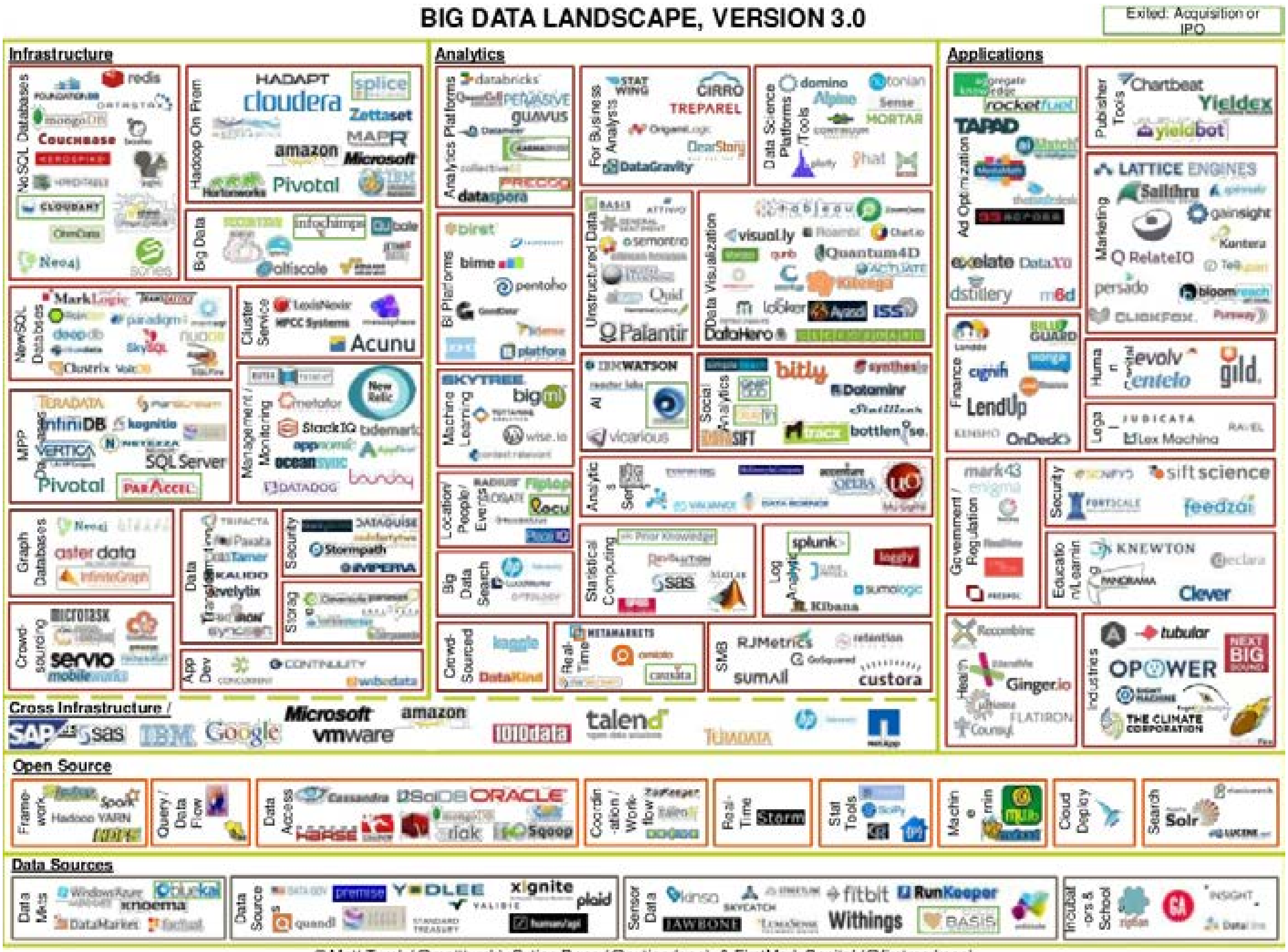
In fact, 80% of the world's data is now unstructured, and therefore can't easily be put into tables (think of photos, video sequences or social media updates). With big data technology we can now harness differed types of data (structured and unstructured) including messages, social media conversations, photos, sensor data, video or voice recordings and bring them together with more traditional, structured data.

Veracity refers to the messiness or trustworthiness of the data. With many forms of big data, quality and accuracy are less controllable (just think of Twitter posts with hash tags, abbreviations, typos and colloquial speech as well as the reliability and accuracy of content) but big data and analytics technology now allows us to work with these type of data. The volumes often make up for the lack of quality or accuracy.

Value: Then there is another V to take into account when looking at Big Data: Value! It is all well and good having access to big data but unless we can turn it into value it is useless. So you can safely argue that 'value' is the most important V of Big Data. It is important that businesses make a business case for any attempt to collect and leverage big data. It is so easy to fall into the buzz trap and embark on big data initiatives without a clear understanding of costs and benefits.

Some refer to "Seven Vs" and add visualization(way of presenting the data in a manner that's readable and accessible) and variability (refers to data whose meaning is constantly changing).

This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos posted online, transaction records of online purchases, and from cell phone GPS signals to name a few.



© Matt Turck (@mattturck), Sutian Dong (@sutiandong) & FirstMark Capital (@firstmarkcap)

Big Data for Development sources generally share some or all of these features:[8]

(1) Digitally generated – i.e. the data are created digitally (as opposed to being digitised manually), and can be stored using a series of ones and zeros, and thus can be manipulated by computers.

(2) Passively produced – a by product of our daily lives or interaction with digital services

(3) Automatically collected – i.e. there is a system in place that extracts and stores the relevant data as it is generated

(4) Geographically or temporally trackable – e.g. mobile phone location data or call duration time.

(5) Continuously analysed – i.e. information is relevant to human well-being and development and can be analysed in real-time

In some cases, big data is defined by the capacity to analyse a variety of mostly unstructured data sets from sources as diverse as web logs, social media, mobile communications, sensors and financial transactions. This requires the capability to link data sets; this can be essential as information is highly context-dependent and may not be of value out of the right context. It also requires the capability to extract information from unstructured data, i.e. data that lack a predefined (explicit or implicit) model. Estimates suggest that the share of unstructured data in businesses could be as high as 80% to 85% and largely unexploited or underexploited. In the past, extracting value from unstructured data was labour-intensive. With big data analytics silos of unexploited data can be linked and analysed to extract potentially valuable information in an automated, cost-effective way.[9]

Delivering value from big data is a challenge. Overcoming this requires building capacity in three distinct areas, namely, scalable data management (processing, storage, resource management), (ii) data analysis (harnessing statistics generated), (iii) expertise in the flied i.e. sustainable agricultural development.

Big Data Analytics

It refers to tools and methodologies that aim to transform massive quantities of raw data into "data about the data"—for analytical purposes. They typically rely on powerful algorithms that are able to detect patterns, trends, and correlations over various time horizons in the data, but also on advanced visualization techniques as "sense-making tools." Once trained, algorithms can help make predictions that can be used to detect anomalies in the form of large deviations from the expected trends or relations in the data.

Open data / Open access to data

It refers to data that is free from copyright and can be shared in the public domain. More, and more open, data can help ensure that knowledge is shared, creating a world of informed and empowered citizens, capable of holding decision-makers accountable for their actions.

Appropriate sharing of data across the economy requires more robust frameworks. Many sources of third-party data do not yet consider sharing their data, and economic incentives may not be encouraging. More needs to be known about pricing and licensing models, but also about ownership and control mechanisms, including intellectual property rights (IPR) regimes. Many governments wish to recover costs, partly for budgetary reasons and partly on the grounds that those who benefit should pay. However, the calculation of benefits can be problematic.

The public sector has nevertheless led the way in opening up its data to the wider economy through various "open data" initiatives. The **OECD** (2008) Council Recommendation for Enhanced Access and More Effective Use of Public Sector Information, describes a set of principles and guidelines for access to and use of PSI.

At **EU level**, a study estimates the total market for public sector information in 2008 at 28 billion across the EU. The same study indicates that the overall economic gains from further opening up public sector information by allowing easy access are around 40 billion a year for the EU27. The total direct and indirect economic gains from PSI applications and use across the whole EU27 economy would be in the order of 140 billion annually.[10]

The Commission has supported open data through its funding programmes, in particular the Framework Programmes for Research and Development, the

Competitiveness and Innovation Programme, and the ISA programme. The projects cover a range of research and application areas and types of organisations.

The **Bill and Melinda Gates Foundation** advocates for more public funded research to be made open data and has adopted an open data policy this year.

The World Bank is also starting an open data project called the Partnership for Open Data, together with the Open Data Institute and the Open Knowledge Foundation. The aim of the Partnership for Open Data idea is to have a broad coalition of these institutions so that we can coordinate better — not contradict each other — share knowledge and pool resources.

Open data in Agricultural Statistics: Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA)

The LSMS-ISA project is committed to providing timely and open access to data. The micro-data produced under the project is fully documented and publicly available within twelve months of the completion of each survey round. Data are disseminated by each country's statistics office as well as the LSMS-ISA website. Additionally, to increase the accessibility of data to policymakers and stakeholders, the LSMS-ISA supports the development of modules of the ADePT software (www.worldbank.org/adept) to produce standardized tables and graphs on agriculture and livestock.

Case Study in Improving Data Collection and Statistical Methodology for Agriculture Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA)[18]

The Living Standards Measurement Study – Integrated Surveys on

Relevance of Big Data dimensions in developmental issues

Dimension	Explanation	Situation in developing countries
Relative advantage	• Perceived benefits of a technology over previous technologies and the extent to which it is better than the idea it supersedes.	• Compared to traditional surveys (e.g. household, labor market, living standard) Big Data involves relatively shorter time to collect and analyze. • BD initiatives are likely to reduce the costs significantly compared to traditional surveys. • SM (e.g. Twitter's API) is more appropriate to detect crisis than traditional techniques such as interviews and surveys (Madsen, 2013).
Compatibility	• The degree to which a technology and the tasks it performs are perceived as being consistent with the existing values, beliefs, past experiences, and needs of potential adopters.	• The initiatives such as Kenya's ODP are likely to be incompatible with institutions filled with secrecy and corruption.
Complexity	• The level of difficulty of installing and using a technology (variety and uncertainty increase complexity).	• There is a shortage of skills for data-related manpower. People with the lack of appropriate technical skills and qualifications may find BD techniques more complex.
Observability	• The degree to which the features and benefits of a technology are visible, noticeable, and understandable to self/others, the results can be described to non-users.	• BD can help document quality standards of agricultural products.
Trialability	• The ability to experiment or try (on a limited basis) before formally adopting.	• Organizations can try with small data sets to see the appropriateness and effectiveness of BD projects.

Source: Kshetri, N. 2014. The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns.

Agriculture (LSMS-ISA) project is an innovative household survey program established with a grant from the Bill and Melinda Gates Foundation and implemented by the Living Standards Measurement Study (LSMS) within the Development Research Group at the World Bank. Recognizing that existing agricultural data in the region suffer from inconsistent investment, institutional and sectoral isolation, and methodological weakness, the LSMS-ISA project collaborates with the national statistics offices of its partner countries in Sub-Saharan Africa to design and implement systems of multi-topic, nationally representative panel household surveys with a strong focus on agriculture. The primary objectives of the project are to improve the availability and quality of smallholder agriculture data within a multi-topic framework and to foster innovation and efficiency in empirical research on the links between agriculture and welfare outcomes in the region.

The four basic tenets of the project are:

(i) Integration of the surveys into existing national statistical systems to ensure project sustainability;

(ii) A multi-topic framework featuring household, agriculture and community questionnaires that enable meaningful cross-sectoral policy analyses;

(iii) Analytical capacity building for local statistical office staff via hands-on training and technical workshops;

(iv) Active dissemination of new data, tools and best practices to benefit researchers and survey practitioners.

The LSMS-ISA project works across four fronts:

- **Collecting** household survey data with focus on agriculture in 8 SSA countries (Malawi, Tanzania, Uganda, Nigeria, Niger, Ethiopia, Burkina Faso and Mali)
- **Improving methodologies** in data collection, producing best practice guidelines & research
- **Documenting & disseminating** micro data, policy research
- **Building capacity** in national institutions

2. Data Collection and Agricultural Statistics

2.1 Instruments and methods of data collection

In the absence of robust statistical systems, the household survey is the foundation on which most development data is built but population census are limited in numbers and frequency (*only around 12 of the 49 countries in sub-Saharan Africa* have held a census in the past 10 years and *21 African countries* haven't had a survey in the past seven years). Furthermore, methodologies differ which makes comparing countries or combining data from different countries very difficult. A *recent experiment by World Bank researchers in Tanzania,* comparing results from the different methods, found that estimates of how many people in the country are hungry varied from just under 20% to nearly 70%, depending on the method chosen. Unreliability of data might well be in some cases purposely ignored as not to lose funding for programmes or question its findings. Better data might challenge some comfortable and familiar myths about development.

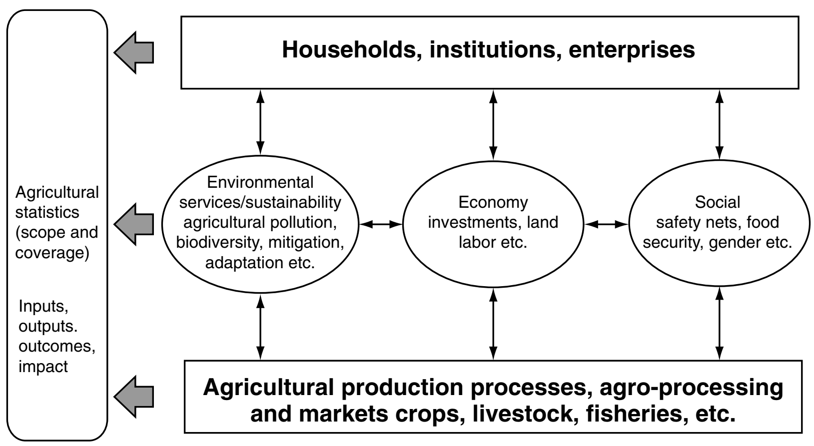
Data collection in the field of nutrition11 today is the same as it was in the 1990s or even in the 1980s. We have nutrition surveys every five years, and micronutrients surveys even less frequently. Mobile technologies—cheap android tablets, internet, mobile phones should be used. Simple surveys could be done every year using them. The data revolution is a call for better data, but also a different way of thinking about data in terms of costs and benefits and also in terms of supporting effective actions and suspending bad actions due to better data.

The lack of up-to-date research on survey methodologies has led to serious gaps in the existing knowledge base, limiting the identification and promotion of effective policies. Statistical systems for agriculture also lack integration, limiting the utility of the data for examining linkages between agriculture and key issues such as poverty or nutrition, as well as linkages between socioeconomic variables and environmental conditions. In order to better inform agricultural policies, multisectoral and integrated approaches are needed.

Traditionally, agricultural statistics have been collected outside of the National Statistical System (NSS), with little oversight by the National Statistical Office responsible for the enforcement of statistical standards. Notably, accounting methods used by some African countries to calculate essential statistics such as GDP may be up to five decades old or more (Bhushan, 2012). Numerous Latin America and Caribbean countries have been found by the World Bank to be lacking a strategic vision to manage their national statistics systems (The World Bank, 2010).

With new technologies becoming increasingly available at relatively low costs, more rigorous research is needed to create and promote improved, cost-effective standards in agricultural statistics. The digital revolution can assist in offering more efficient and cost-effective ways to capture the complexity of agriculture. The progress made to date with technologies such as Global Positioning Systems (GPS), satellite imagery, Computer Assisted Personal Interviewing (CAPI10) and mobile phones offers great potential.

The conceptual framework for agricultural statistics



Source: World Bank/UN/FAO/

There are a broad number of instruments and methods available for agricultural data collection, and the most appropriate depends on a number of factors, which should be appropriately integrated into the data collection methodology. The Global Strategy to Improve Agricultural and Rural Standards recognises three dimensions of agriculture which should be taken into account in the development and implementation of sound national, regional and international agricultural statistics, namely the economic, social and environmental. This type of analysis is motivated by the complex role of agriculture which necessitates that statistics gathered in this field are relevant for various disciplines and areas of intervention, including environmental conservation, education, promotion of gender equality, trade promotion and so on.[12]

Key factors to be considered are:

(a) The **frequency** of the type of statistical gathering undertaken. If it's a census which takes place every decade, then the data collection methods may vary from a survey carried out annually or a monthly estimate. Another challenge inherent in agricultural data collection is the seasonality of farming and even rural populations, which demands specific intervals for certain types of data collection. Statistical data may be gathered: Annually, Seasonally (e.g. six monthly); Quarterly; Monthly; Weekly; Daily; on Ad-hoc basis.

(b) The **responsible institution** for carrying out the data collection will also play a role in influencing the methodology, whether it be a government ministry or a subcontracted private institution or perhaps an international organisation. Each type of institution will have certain capacities that make it stronger or weaker at different methods of data collection. For agriculture, the key institutions may include: National Statistics Office; Ministry of Agriculture; Other Line ministries; Central Bank; Commodity board; Producers' association, and; Customers/ Revenue Authority.

(c) The **scope and coverage** of the data collection means that the method used should be appropriate for the temporal, geographic and thematic needs. If the data is to be collected over a day, week or month, and in a farm, village or administrative district, then the tools and methodology will likely vary. Additionally, the collection may extend beyond traditional farming and include fisheries and aquaculture, agroforestry and other agricultural activities.

(d) The **source** of the data cannot be overlooked in the methodological process, as a particular source of data can provide results that are more appropriate, reliable, replicable, comprehensive, cost effective etc. than another. Data can come from, inter alia, censuses, sample surveys, administrative records, estimates/ forecasts, special studies; or expert opinion/ assessments.

Examples of data sources:

a. **Administrative data**. Governmental interventions such as subsidies, regulation, and legislation often require agricultural holders to report production information. Land ownership and cadastral surveys provide useful information for constructing registers. Food inspections, animal health inspections, and trade data provide input to the utilization accounts.
b. **Remotely sensed data**. These include vegetative indices that show overall crop conditions and information about changes in land cover and use. The survey framework should include the need to provide ground truth data if remote sensing information is to be used to estimate cropland areas.
c. **Agribusinesses** are the source of utilization data and prices.
d. **Expert judgment and windshield surveys** can be used to collect data from experts whose judgments inform evaluations of agricultural conditions. For instance, the Sourcebook (World Bank 2008b) refers to a procedure in which experts travel a specified route on a periodic basis and record the condition of crops, which provide an input into crop yield forecasts.
e. **Community surveys.** *The World Programme for the Census of Agriculture (FAO 2005b)* provides an overview of data that can be collected at the village level. These data include information about the infrastructure and services available to households and agricultural holdings, occurrences of food shortages, frequency of natural disasters, etc.

Different statistical methods and practices have their advantages and disadvantages, merits and demerits. These may include the reliability, replicability or sustainability, affordability, objectivity, relevance, accuracy, comparability, accessibility etc. of different approaches. However, evaluating different statistical methods and practices according to these criteria is not only difficult, but raises challenges in that very often different experts, academics, practitioners, governments, institutions may not all agree as to the final outcome of such an evaluation.

Challenges in data collection

In their analysis of agricultural statistics is Sub-Sahara Africa, Carletto, Jolliffee and Banerjee evaluated various case studies and statistical methods and practices to demonstrate the challenges inherent in agricultural data collection and statistical evaluation in Africa.

Sources of Agricultural Data:

(a) *Routine data system based on systems of resident or local extension officers exists in virtually all countries. Extension officers collect different type of data on a regular basis at a geographically granular level, including information on land usage, crop forecasting and production. One major drawback of current routine data systems is the high degree of arbitrariness and subjectivity in the data collection protocols. The lack of statistical rigor is likely to lead to systematic biases, while the lack of documentation on the actually data collection procedures often hinders any attempt of comparison and reconciliation with other data sources.*

(b) *Agricultural census which, based on FAO guidelines, countries are recommended to implement every ten years. However, because of the high costs of full enumeration and the rather limited amount of information collected, agricultural censuses are less and less common, leading many experts to even question the feasibility of a decennial census in African countries.*

(c) *Surveys are the third source of data on agriculture. Farm surveys have been, and still are, the backbone of agricultural statistics in Africa, with great variation in terms of content, frequency and quality. While in principle indispensable to get a good depiction of the agricultural sector based on sound statistical principle, there are at least two major drawbacks. One the one hand, the statistical properties of many farm surveys have been repeatedly questioned. On the other hand, by focusing almost exclusively on measuring agriculture, they fail to provide enough information to understand it. In fact, even among the most remote and poorest rural household, agriculture does not exists in a vacuum and diversification in terms of income sources at both the household and individual level is the norm, not the exception (Davis et al, 2009). Shifting the attention from the farm to the household as a unit of analysis, and expanding the thematic coverage of the survey, would partly solve the problem. Integrated household surveys do that, but not without cost. First, the breath of the data collected may result in significant trade-offs in terms of depth. Also, the timing must be adjusted to the agricultural season, and the added requirements of collecting information on highly seasonal and fast changing processes. But sampling is probably the most challenging aspect because of the differences in sampling requirement and optimal domains of inference of the different variables.*

Measurement of Productivity

The measurement of agricultural productivity, which for many countries is a core component of agricultural statistics, is used by the authors to demonstrate the complexity, and advantages and disadvantages of various statistics gathering methods and practices and the impact of variables on statistics collection. As a point of departure, measurement of agricultural productivity is defined as "land productivity, or yields, equivalent to the amount or value of crop harvested (the numerator) over cultivated land (the denominator)."

(a) Amount or value of crop harvested: statistical measurement challenges

- The type of crop being harvested can play a significant role in the ability to collect accurate data on crop quantities.
- Farmers may be better at recalling harvests quantities or revenues of high-value marketed crops like rice versus cash crops like cassava and bananas
- Crops can vary in state, which is often overlooked in surveys, so maize may be on the cob, in grain, in flour etc.
- Non-standard units are used to measure crops, so there may be a heap or individual pieces of cassava, but a bunch of bananas, which then make it difficult to convert these into standard units such kilograms or tons.

- Consistency in the construction and application of accurate conversion factors of non-standard units is low.
- Smallholders, particularly the poorest among them, consume a large share of their production, which is then not factored into the production data.
- The lack of farm gate prices or record keeping by farmers of far gate prices even when crop is sold, means that more general market prices or unit values averages over a geographic zone have to be used, even though these may vary widely throughout the year.
- Crop cutting, which is considered the gold standard of quantifying production, is more appropriate for cereal crops than say root crops and as its burdensome financially and time-wise, it may not be used in large farm and household surveys, or else it is used selectively.
- Intercropping, a common practice in developing countries, also makes crop production estimates challenging.

(b) Size of Cultivated Land: statistical measurement challenges

- Traversing (compass and rope/tape), considered by FAO to be the gold standard, is often not used as it is time-consuming and costly, particularly in the case of large national household surveys.
- Satellite imagery to delineate land parcel boundaries can be accurate, but this accuracy is highly diminished in the case of tree dense areas and areas with regular cloud cover. It can be costly for long term and large scale surveys which would require acquisition and processing of high resolution imagery.
- "Eye estimates" by extension officers, which involves the highly subjective measurement of plot size, often chosen arbitrarily, can result in very inaccurate and potentially biased estimates.
- Self-reporting by farmers is equally inaccurate or imprecise, as there are many factors which contribute to errors in self-reporting. Topography or geography of the parcel, especially parcel slope, is one of these. Furthermore, it is common for farmers to use non-standard units.
- GPS technology has become increasingly affordable, accurate and user friendly, but is not without its own limitations. GPS-based coordinates are known to have measurement errors arising from satellite position, signal propagation and receivers. Furthermore, it cannot be given as certain that all household plots will be measured, even if GPS technology makes this feasible.

Beyond the particulars of agricultural statistical methods and practices for gathering data on a single variable (e.g. productivity), there may be broader issues with other national data sources or data gathering methods outside the field of agriculture which then make it difficult for accurate agricultural statistics to be established. In their Africa Country Assessment of Agricultural Statistical Systems, the African Development Bank found that there is a discrepancy in the quality of the statistical methods and practices on the continent used for the production of agriculture and rural statistics.

2.2 Agricultural Statistics, Big Data and Open Data

Big Data sources offer huge volumes of data which require storage and processing that exceed the capacity of traditional statistical production tools. Using Big Data would necessitate moving away from exclusive dependence on statistical methods that cannot handle huge volumes of information. Instead, a more diverse set of tools should be adopted. This can be addressed through the use of data mining and machine learning algorithms having the required computational efficiency (Bondi, 2000). Besides, due to the enormous heterogeneity of Big Data sources as regards formats, contents, storage, etc., methods to produce statistical information should be developed *ad hoc* for each case and the Generic Statistical Business Process Model might not be applicable. In the beginning, the Big Data source should be explored using data mining procedures to learn about the unknown data structure and decide on the possible outcomes, and how to combine these data with traditional statistical production procedures.[13]

The role of "Big Data" and "Open Data" in the context of agricultural statistics in ACP countries is an emerging but growing one. Whereas the bulk of the current programmes concerning the improvement of national statistic systems in developing countries, particularly in Africa, are focused on the more traditional, orthodox data collection techniques, new technology has made inroads in other areas of activity such that the information produced by technology in these areas has become increasingly important as a source of statistical data. An example of this would be the use of mobile telephony in the context of banking services, such as the mobile money transfer system M-Pesa which has gained widespread use in a number of African countries. Other technologies are applied in sectors such as health, education, environment or even security and disaster prevention, which all contribute to the growth of information that could potentially be used by the national statistics office and other bodies responsible for statistics in the field of agriculture.

Big data also raises important questions about the potential role of the private sector in national statistic systems, as the private sector continues to be both a source of and generator of big data.

Characteristics of a technology influencing its diffusion: the case of BD

Relative advantage	Perceived benefits of a technology over previous technologies and the extent to which it is better than the idea it supersedes.	- Compared to traditional surveys (e.g. household, labor market, living standard) Big Data involves relatively shorter time to collect and analyze. - BD initiatives are likely to reduce the costs significantly compared to traditional surveys. - SM (e.g. Twitter's API) is more appropriate to detect crisis than traditional techniques such as interviews and surveys (Madsen, 2013).
Compatibility	The degree to which a technology and the tasks it performs are perceived as being consistent with the existing values, beliefs, past experiences, and needs of potential adopters.	The initiatives such as Kenya's Open Data Portal are likely to be incompatible with institutions filled with secrecy and corruption.
Complexity	The level of difficulty of installing and using a technology (variety and uncertainty increase complexity).	There is a shortage of skills for data-related manpower. People with the lack of appropriate technical skills and qualifications may find BD techniques more complex.
Observability	The degree to which the features and benefits of a technology are visible, noticeable, and understandable to self/others, the results can be described to nonusers	BD can help document quality standards of agricultural products.
Trialability	The ability to experiment or try (on a limited basis) before formally adopting.	Organizations can try with small data sets to see the appropriateness and effectiveness of BD projects.

Source: Kshetri, N. 2013

Even as the global debate continues on the issue of protecting rights in the access to and application of big data, particularly privacy rights, in developing countries with large rural populations and capacity constraints, governments and civil society have not addressed the issue of public policy or regulation of big data with the same energy as in developed countries. Filling this governance vacuum over the use of big data will certainly become an issue that more stakeholders in ACP countries will have to pay attention to, particularly with a view to ensuring that the use of big data complements the overall sustainable development goals and reduction of poverty which remain a priority in most of the ACP states.

Two examples of the way in which big data can complement traditional national statistical systems in developing countries, both generally and specifically for agriculture, are with improved transparency and accuracy.

Open Access Data in developing countries: challenges and opportunities

In terms of transparency, developing country governments, national statistical bodies or other institutions infrequently publish the data gathered as part of censuses or other large scale surveys. Open access to data remains a big challenge because of various factors. Some reasons for the lack of openly accessible data related to statistical figures in develop-

COUNTRY	READINESS SUBINDEX	IMPLEMENTATION SUBINDEX	IMPACT SUBINDEX	ODB OVERALL
Africa	25.90	14.73	5.72	14.29
Kenya	49.70	45.88	21.55	43.06
Morocco	36.46	27.84	16.59	27.24
Mauritius	35.71	30.59	0.00	26.08
Rwanda	36.71	27.84	0.00	24.27
Ghana	39.51	23.53	0.00	21.60
Tunisia	63.52	10.98	26.46	21.02
South Africa	35.39	18.43	10.31	19.20
Botswana	12.16	21.57	0.00	16.08
Uganda	23.99	13.33	23.07	16.15
Tanzania	20.43	17.65	0.00	14.51
Malawi	28.24	11.76	16.52	14.47
Ethiopia	15.45	10.59	0.00	8.70
Burkina Faso	17.63	8.24	0.00	7.35
Benin	11.60	9.41	0.00	7.28
Namibia	11.57	9.02	0.00	7.00
Senegal	28.57	4.71	0.00	6.46
Cameroon	7.11	6.67	5.56	5.65
Zimbabwe	15.20	5.88	0.00	5.30
Zambia	11.84	5.10	0.00	4.23
Nigeria	36.90	0.00	0.00	4.35
Mali	6.15	0.39	0.00	0.00

Note: ODB is Open Data Barometer.

Source: ODI (2013).

Rankings of selected African countries on open data readiness, implementation, and impact

ing countries may be due to government secrecy, lack of complete data, inability to host data and provide it in accessible formats (technical or resource limitations).

The *World Wide Web Foundation* and the *Open Data Barometer Report by the Open Data Institute* show that in terms of indicators relating to right to information laws, civil society demand for data, and open government initiatives, Africa is lagging behind Europe and the Americas but outperforming the Middle East and Central Asia. Their indexes do not give much weight to the statistical products of NSOs, however, focusing instead on government budgets and commercially useful information, such as maps and transportation timetables. With encouragement from donors and other partners, NSOs in Africa could and should take the initiative in making their data widely available.

Beyond these aggregate measures, there is the simple reality that many citizen- and donor-funded household surveys remain unavailable— as reports or microdata— to other public ministries or to the public at large, limiting their utility for affecting policymaking or holding government accountability. The catalog of the International Household Survey Network (IHSN) indicates that only 56 percent of the microdata from household surveys conducted between 2000 and 2014 are available to the public. For example, in principle, the microdata from the 2005–06 Kenya Integrated Household Budget Survey, the country's most recent multipurpose consumption survey, are available on request from the Kenya National Bureau of Statistics.

Center for Global Development and the African Population and Health Research Centre. 2014. Delivering on the Data Revolution in Sub-Sahara Africa.

Kenya's Open Data Portal

Among developing economies, Kenya is probably the most spectacular example in making data available to the public and facilitating the use of BD. In 2011, Kenya launched an Open Data Portal (ODP) with the help of the World Bank. The project received support at the highest levels of the government. The data in the ODP includes a full digital edition of the 2009 census, government expenditure for 12 years, household income surveys, and data about the location of schools and health facilities.

Increasingly, public and private bodies worldwide face the challenge that, with an ever more information and technology savvy population, and the growing accessibility of tools to access, read or analyse data, there is a stronger call by civil society and even individuals for data to be made available openly. Some countries have met this challenge by making data related to government or public studies, policies, and initiatives available on open access platforms. Big data also forces the private sector to account for business decisions and practices; as more information about products, goods and services become available, consumers are empowered to make demands on producers and services providers to change certain practices. Whereas in both the public and the private sector cases, traditional statistics based on orthodox data gathering methods may have provided a more general picture of a

certain matter, big data and open data have been applied to add clarity to the situation, by providing the type of detail which may not be feasible with orthodox data collection methods

The potential role of open data for agriculture was a driving force in the decision behind the G-8 Group's decision to hold a conference on Open Data in Agriculture in 2013. According to the Group, "at the 2012 G-8 Summit, G-8 leaders committed to the New Alliance for Food Security and Nutrition, the next phase of a shared commitment to achieving global food security. As part of this commitment, they agreed to "share relevant agricultural data available from G-8 countries with African partners and convene an international conference on Open Data for Agriculture, to develop options for the establishment of a global platform to make reliable agricultural and related information available to African farmers, researchers and policymakers, taking into account existing agricultural data systems." An outcome of the conference was the publication of action plans by the of G-8 members which make agricultural data streams available to users in Africa and worldwide.[14]

Big Data in developing countries: challenges and opportunities

A large volume of big data is generated in real time, and provides information that can be fine-tuned to improve accuracy beyond what is feasible with traditional data gathering methods used in agricultural statistics. For example, statistics on weather patterns used by NSO or other public bodies, which rely on satellite and other traditional tools, may not be sufficiently accurate to the date, time and location that a farmer may need to make a decision on planting, harvesting or irrigating, or an insurer to verify a claim on weather related crop damage. Big data, whether through specific applications or instruments related to farming, or even via social media tools such as Twitter or Facebook, can generate a range of information that can provide data useful in painting a more accurate picture of the needs of consumers, from the point of view of businesses, or help smallholders make more informed decisions. The entire agricultural value chain is therefore able to benefit from the availability of big data, whether at the start or at the end of the chain. As more data becomes available, more complex networks within the value chain are made possible through the increase of accessible, detailed and disaggregated data.

Data collection through mobile phones

In the late 1990s, cellphones came to Africa and revolutionized the continent. Suddenly, Africans could communicate — with each other and with the world. In less than 10 years, there was a new revolution, only now it was no longer about making phone calls; it was all about the apps. A whole universe opened up with e-health, e-government, e-learning and, best of all, mobile banking. Fast forward another decade to today, and we are on the verge of another revolution. It's no longer just about an individual being able to make calls or what an individual can do with his phone. Far from being just an accessory, mobile phones are starting to be used to collect data in an increasing number of disciplines. Today, it's all about the aggregate. With cellphones, it's now possible to gather individual African's views, hopes and habits, and aggregate them to create accurate and reliable data in near real time. This is actionable information that can drive investment, create jobs and fuel development.

New agricultural cell phone and tablet applications are also being developed for Africa, where there are now more mobile phones than in the U.S. or Europe and where there has been a 20-fold increase in Internet bandwidth since 2008. These apps are connecting farmers to financing, market information, agricultural expertise and sometimes simply to each other so they can share information and best practices.

Data collection through mobile phones includes **financial services and digital payments, education, health**, **agriculture (**payments for agricultural products, input purchases and subsidies, availability of storage), early detection of droughts or disasters in remote are as allowing timely needed humanitarian assistance.

Sources such as online or mobile financial transactions, social media traffic, and GPS coordinates now generate over 2.5 quintillion bytes of so-called, big data every day[15].

Participatory Geographical Information Systems

PGIS combines a range of geo-spatial information management tools and methods such as sketch maps**, Participatory 3D Models** (P3DM), aerial photographs, satellite imagery, Global Positioning Systems (GPS) and Geographic Information Systems (GIS) to represent peoples' spatial knowledge in the forms of virtual or physical, 2 or 3 dimensional maps used as interactive vehicles for spatial learning, discussion, information exchange, analysis, decision making and advocacy.

PGIS practice is geared towards community empowerment through measured, demand-driven, user-friendly and integrated applications of geo-spatial technologies. GIS-based maps and spatial analysis become major conduits in the process. A good PGIS practice is embedded into long-lasting spatial decision-making processes, is flexible, adapts to different socio-cultural and bio-physical environments, depends on multidisciplinary facilitation and skills and builds essentially on visual language. The practice integrates several tools and

methods whilst often relying on the combination of 'expert' skills with socially differentiated local knowledge. It promotes interactive participation of stakeholders in generating and managing spatial information and it uses information about specific landscapes to facilitate broadly-based decision making processes that support effective communication and community advocacy.

Online advertisement

Data generated when consumers use the Internet can create value and give firms opportunities to improve their operations and market their products more effectively.

The collection of data is not limited to the firm's website. By using service providers such as social networking sites and advertising networks, firms can also collect data generated elsewhere. Such data are increasingly available through data markets and can be combined with data from sources such as census data, real estate records, vehicle registration and so forth. These enhanced user profiles are then sold to advertisers looking for consumers with particular profiles in order to improve behavioural targeting.
Overall, the revenue generated by online advertisement has grown much faster, especially in the last five years, than traditional advertising channels did in their first 15 years. In the first quarter of 2012, online advertising revenues of the top 500 advertisers in the United States, for example, reached USD 8.4 billion, according to the latest IAB Internet Advertising Report (BusinessWire, 2012). This is USD 1.1 billion (15%) more than in the first quarter of 2011. In 2011, AdWords generated more than USD 20 million a month on average from the top 20 websites. This was largely due to the increasing ability to target potential customers and measure results. However, the added value is not limited to advertisement revenue. There are also benefits for consumers. According to McKinsey (2010), consumers in the United States and Europe received EUR 100 billion in value in 2010 from advertising-supported web services. This is three times more than current revenue from advertising and suggests that the consumer value created is greater than advertising revenues would indicate. OCDE

Traditional sectors such as retail are changing: firms like Tesco, the UK supermarket chain, exploit huge data flows generated through their fidelity card programmes. The Tesco programme now counts more than 100 market baskets a second and 6 million transactions a day, and it very effectively transformed Tesco from a local, downmarket "pile 'em high, sell 'em cheap" retailer to a multinational, customer-oriented one with broad appeal across social groups.[16]

2.3 Priorities in Agricultural Data collection for developing countries

The need to improve agricultural data collection and statistics has been recognised by developing countries, international institutions and the private sector alike, hence the various programmes, initiatives and projects which aim to achieve this. As the challenges in this field are numerous and varied, as a starting point, it has been recognised that establishing priorities for agricultural data collection should be a central component of the drive to improve statistics in agriculture in developing countries. The priorities reflect the areas where different stakeholders, including governments, international and regional development and agriculture institutions, academia/researchers and other experts perceive the greatest challenges to lie.

A Stakeholder Conference, held in Tunis in 2010 to discuss the research component of the Implementation Plan for Africa of the Global Strategy to Improve Rural and Agricultural Statistics, undertook Stakeholder Survey Questionnaire which requested the stakeholders to rank the criteria for prioritising a range of research topics and propose a ranking of a pre-defined list of research topics based on their expert knowledge and experience.[17]

The results of the priority ranking for the pre-defined list of research topics saw an unequivocal affirmation of the fact that even the most basic agricultural statistical information, namely improving the estimation of crop, yield and production, was still needed. In terms of decreasing priority, the overall rankings were as follows:

Priority level 1
- Improvement of estimation of crop area, yield and production

Priority level 2
- Use of GPS in the production of agricultural statistics
- Development of master sampling frames

Priority level 3
- Methodology for the compilation of food security statistics
- Methods for estimating crop area, yield and production of mixed and/or repeated cropping
- Methods for estimating yield of root crops, edible forest products, etc.

Priority level 4
- Development of an integrated survey programme

Priority level 5
- Linking area frames with list frames
- Estimation of food stocks
- Estimation of farm gate prices

Priority level 6
- Reconciliation of census data with survey data

Priority level 7
- Use of remote sensing
- Determination of user's information needs for decision making

Priority level 8
- Use of small area estimation methods for improving agricultural statistics

Priority level 9
- Estimation of informal cross border trade data
- Use of administrative data for improving agricultural statistics

It is further interesting to note that use of GPS was also recognised as a top priority for research into for improving African agricultural statistics, particularly for the African government stakeholders, but that the use of administrative data in agricultural statistics did not receive a high priority ranking. According to Keita (2011), although this approach "has been receiving great attention in Europe in the last decades, [it] did not have a high level of priority for the African agricultural statisticians as well as for all stakeholders, probably because they believe that the quality of administrative data is poor."

3. Initiatives and partnerships to improve data collection and use

3.1 Global initiatives

The initiative to develop the **Global strategy to improve agricultural and rural statistics** came as a response to the declining quantity and quality of agricultural statistics. The global strategy will also address the emerging data requirements posed by the Millennium Development Goals (MDGs), mainly on biofuels, global warming, the environment and food security. The purpose of the global strategy is to provide the vision for national and international statistical systems to produce the basic data and information to guide the decision making required for the 21st century. The global strategy is based on a thorough assessment of data user needs and what is currently available.

UN secretary-general Ban Ki Moon named his independent expert advisory group, 24 experts tasked with providing recommendations on how best to use data to deliver the sustainable development goals. Those recommendations were published in a report entitled **A world that counts**.

The **UN's Global Pulse Initiative**, is a United Nations innovation initiative of the Secretary-General launched in 2009, forges partnerships between UN agencies, governments, academia and the private sector to accelerate discovery, development and adoption of Big Data and real-time analytics for sustainable development and humanitarian action. The three-fold project strategy includes: 1. **Research & Development**: Conducting research to discover new proxy indicators for tracking development progress and emerging vulnerabilities in real time, and assembling a toolkit of technologies for analyzing real-time data. 2. **Big Data partnerships**: Forging partnerships with companies, organizations, researchers and academic institutions that have the data, technology and analytical expertise needed for Big Data for Development projects and advocacy. 3. **Pulse Lab network**: Establishing an integrated network of country-level innovation centers that bring together government experts, UN agencies, academia and the private sector to prototype and pilot approaches at country level and support successful adoption.

The **Global Open Data for Agriculture and Nutrition initiative (GODAN)** aims to demonstrate the value of Open Data in Agriculture and convince more governments to encourage its availability. A **timeline** of the history so far was presented together with **deliverables.**

EU initiatives on Big Data and Open Data

Estimates suggest that better exploitation of data could significantly increase efficiency, with billions of savings for the public sector. According to MGI (2011), full use of big data in Europe's 23 largest governments might reduce administrative costs by 15% to 20%, creating the equivalent of EUR 150 billion to EUR 300 billion in new value, and accelerating annual productivity growth by 0.5 percentage points over the next ten years. The main benefits would be greater operational efficiency (due to greater transparency), increased tax collection (due to customised services, for example), and fewer frauds and errors (due to automated data analytics). Similar studies of the United Kingdom show that the public sector could save GBP 2 billion in fraud detection and generate GBP 4 billion through better performance management by using big data analytics (Cebr, 2012). These estimates do not include the full benefits for policy making to be realised from real-time data and statistics.

In July 2014, the European Commission outlined a new strategy on Big Data, supporting and accelerating the transition towards a data-driven economy in Europe which will stimulate research and innovation on data while leading to more business opportunities and an increased availability of knowledge and capital, in particular for SMEs[19].

The agriculture research projects focus on how big data can increase food safety and productivity, and farmers' incomes. Using big data in a smart way changes the practice of farming for the better: using sensors to provide for efficient use of natural resources; accessing real time data on farm machinery; being informed about weather patterns, topography and crop performance.

Two main projects have been carried out. The e-Agri provides data crop monitoring for developing countries and the Miniturbine project provides data on clean energy supply in agriculture.

The e-Agri project aims to support the uptake of European ICT research results in developing countries. It has the objective of establishing an advanced European e-agriculture service through crop monitoring in Morocco and China[20]. In Kenya, activities shall focus on capacity building and raising the interest of local stakeholders on European e-agriculture practices. This project builds on approaches based on the European Information and Communication Technologies including space-based Earth Observation (EO) geographical information systems and agro-meteorological modelling. It also draws on the work developed by European research institutions (including VITO, Alterra, JRC and University of Milan) in the area of agricultural monitoring for the Common Agricultural policy. Together, this project intends to transfer, adapt and locally apply e-agriculture practices to benefit developing country farmers and policy-makers. It also enhances European crop production forecasting technology on a global scale, which strengthens capacity for addressing the global challenge of food security and sustainable agricultural growth. The project works in close collaboration with other European food security projects focusing on Africa, namely GMFS or AGRICAB and more.

EU initiatives on open data/PSI

The 2003 Directive on the re-use of public sector information[21] set out the general legislative framework at European level. The general re-use policy is complemented by legislative or policy initiatives in specific sectors. Examples are:

• the Access to Environmental Information and INSPIRE Directives[22], aimed at the widest possible dissemination of **environmental information** and the harmonisation of key datasets;

• the Commission Communication on Marine Knowledge 2020[23], aiming amongst other things to make **marine data** easier and less costly to use;

• the initiatives within the 2008 **Action Plan[24] for the Deployment of Intelligent Transport Systems (ITS)**, looking at, amongst other things, access for private service providers to travel and real-time traffic information;

• the Commission's policy on open access to **scientific information[25]**, which includes a pilot for open access to publications resulting from projects funded by the European Union and a pan-European, participatory **e-Infrastructure of Open Access Repositories**; The JRC publications repository is also relevant in this context.

3.2 Public Private Partnerships in support of Big Data

The public sector cannot fully exploit Big Data without leadership from the private sector which includes mobile phone carriers, credit card companies and social media networking sites and manages enormous data sets that hold rich insights. Companies analyse this data to support decision-making or provide market intelligence. More recently, public sector institutions have begun leveraging similar techniques to generate actionable insights for policymakers.

Governments – ideally working with private institutions, civil society and academia – should set and enforce legal frameworks to guarantee data privacy and security of data for individuals, and ensure its quality and independence. They can also can balance public and private interests and create systems that foster incentives without creating unacceptable inequalities.[26]

There are a few emerging examples[27] of multi-stakeholder agreements on data issues that provide a baseline set of insights to build upon.

These include:

- HP Earth Insights is a joint project between the technology company Hewlett-Packard and Conservation International dedicated to monitor tropical forests to enable proactive responses to environmental threats and supporting the protection of hundreds of threatened and endangered species (harnessing-big-data-drive-environmental-progress). The scientists' largest challenge was collecting, managing and analyzing the biodiversity and climate data they've captured at 16 sites across four continents. "Scientists had to manually collect and analyze this data from tropical forests – often taking weeks, months or more to analyze information – making it difficult to identify new patterns and intervene to protect biodiversity. HP's analytics platform is able to manage volumes of data on biodiversity and climate measurements like precipitation, temperature, humidity and solar radiation nine times faster than before – and improving accuracy of data analysis.

- Orange Data for Development Challenges. The Orange company, one of the world's leading telecommunications operators in Africa has sponsored and provided Call Detail Records (CDR) data for two competitions among data researchers. The first one was held in 2013 using CDRs from Cote d'Ivoire. The second was held in 2014 using CDRs for Senegal. The challenge rewards the "best work on improving algorithms for anonymisation, data mining, and data visualisation and cross-matching" (Data for Development challenges). Researchers have been able to use CDRs to estimate mobility patterns, epidemiological trends, measures of well-being, and other development issues.

- Weather companies working with weather data agencies both to share their data for public use, and to help host and disseminate the data.

Current practice in developed countries has shown that private sector companies can use open government data to the benefit of ordinary citizens while making profits. These data flows –from public to private sector to citizens- are starting to become reality in developing countries as well. An important reason for the reluctance of private companies to share data is risks, particularly, legal risks arising from actual or perceived invasion of individual privacy.

The recent Ebola crisis in western Africa is a good example, where several NGOs and data researchers requested government intervention in persuading private companies to release data that could be put to help to address or diffuse that crisis. But the response from both private companies and governments was not as forthcoming as was expected. Privacy was probably the biggest obstacle. In an effort to address this, the ebola crisis prompted GSMA (the GSM Association of mobile operators and related companies operating in 220 countries) to release GSMA Guidelines on the Protection of Privacy in the Use of Mobile Phone Data for Responding to the Ebola Outbreak.

EU & the data public private partnership

The EU Data public private partnership recognises the global potential of Big data as an enabler to achieve competitiveness, growth and jobs due to its ability to impact on a large scale. Launched on 1 January 2015, the EU data PPP in partnership with the Big Data Value Association has given itself the challenge of harnessing big data as a veritable economic asset, which also has the potential to bring veritable social gains. The partnership worth 2.5billion is part of the big data master plan to strengthen the EU's data sector; currently, only 2 of 20 top companies worldwide working on Big data are European.

4. Issues and risks around new data

There are important issues specifically relevant to Big Data for Development such as:

Governance

Although much of the publicly available online data has potential utility for development purposes, private sector companies may be reluctant to share data due to concerns about competitiveness and their customers' privacy. It is therefore critical to ensure a **legal framework** that defines rules for privacy-preserving analysis and protect the competitiveness of the private sector companies willing to share data. Suitable legal frameworks, ethical guidelines and technological solutions for protected data sharing are at the centre of efforts to leverage Big Data for development. There are big amounts of data generated by individuals which are of great importance for a collective use but mechanisms must be developed to ensure adequate user privacy and security.

Business models must be created to provide the appropriate incentives for private-sector actors to share and use data for the benefit of society. Such models already exist in the Internet environment. Companies in search and social networking profit from products they offer at no charge to end users because the usage data these products generate is valuable to other ecosystem actors. Similar models could be created in the mobile data sphere, and the data generated through them could maximise the impact of scarce public sector resources by indicating where resources are most needed.[28]

Privacy

There is a need to guarantee the right of individuals to control what information related to them may be disclosed, which has implications for all areas of work, from data acquisition and storage to retention, use and presentation. In many cases, the production of data itself raises concerns, as people may be unaware of the data they are generating on a daily basis and understanding how it may be used.

Cybersecurity risks

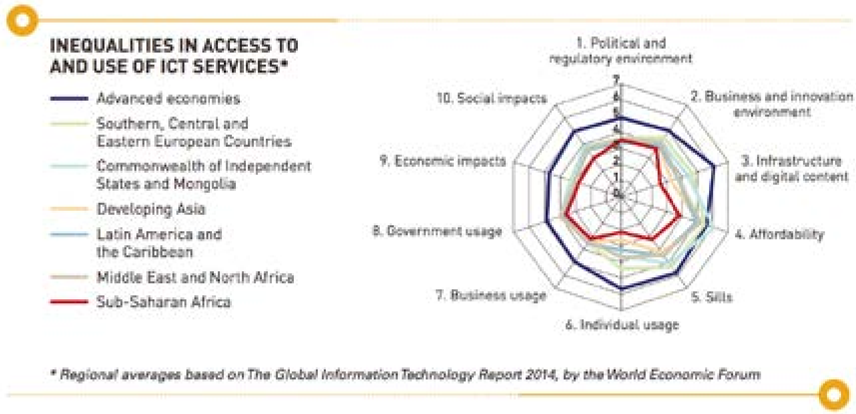
As the volume and value of data stored increases so does the risk of data breaches. According to company surveys, reported thefts of electronic data surpassed losses of physical property as the major crime problem for global companies for the first time in 2010 (Masters and Menn, 2010; Kroll, 2012). This demonstrates the increasing corporate value of intangible assets, such as data, as compared to tangible assets.[29] These are typical cyberespionage incidents often targeting a sector's key organisations or key competitors to steal data or different forms of intellectual property and to reduce these organisations' competitive advantage. As data usage today requires information systems and networks to be more open, organisations are obliged to adapt their security policy to the more open and dynamic environment in which data are widely exchanged and used.

Digital divide and capacity development

Major gaps are already opening up between the data haves and have-nots. Without action, a whole new inequality frontier will open up, splitting the world between those who know, and those who do not. Many people are excluded from the new world of data and information by language, poverty, lack of education, lack of technology infrastructure, remoteness or prejudice and discrimination.[30] According to McKinsey, African countries spend about 1.1% of GDP on investment in and use of internet services, less than a third of what, on average, is spent by richer countries – meaning that the gap in internet availabil-

Source: Gregor Aisch. A graphics editor at The New York Times.

Access, too, is often restricted behind technical and/or legal barriers, or restricted by governments or companies that fear too much transparency, all of which prevent or limit effective use of data.



* *Regional averages based on The Global Information Technology Report 2014, by the World Economic Forum*

ity and use is growing every year, as some regions accelerate ahead.[31]

The graph below shows how advanced economies are ahead of the rest of the world on almost every indicator of access to, use of, and impact of the use of digital technologies.

The Scale of the Digital Divide
Mobile communication and the rise of the internet have contributed in ways unforeseen to easing communication and access to information in Africa. Sadly, access is still not affordable for most people and even though mobile subscription rates are at 60%, internet penetration is only 16%. Limited internet access and general access to ICTs in rural areas impacts on participation in policy processes (Munyua, 2007; NEPAD, 2013).

The available international bandwidth to Africa has increased from 100 Gbps in 2008 to 1.5 Tbps in 2013, yet per capita, Africa still lags behind the rest of the world: the average bandwidth per person is 2 Kbps compared to more than 90 Kbps in Europe. It is clear that with the current pace of development, Africa will continue to lag behind other regions and the so-called ‘digital divide’ will continue to increase. Furthermore, most investments in ICTs are made in urban areas where returns are higher, resulting in a rural–urban digital divide that manifests at national level. This rural–urban divide may be more detrimental to national development than the divide between countries or regions. The digital divide also contributes to some of the challenges of information asymmetry that present at national and regional levels.

Nyirenda-Jere TPR and Kazembe JA. 2014. Improving Policymaking for Agricultural and Rural Development in Africa: The role of ICTs and Knowledge Management. IIED, IDS, ODI.

A pool of qualified personnel with skills in data management and analytics (data science) is essential for the success of a “smarter” data-driven economy (OECD, 2012i). Demand for highly specialised skills is expected to intensify as data analytics proliferate, and a shortage of data scientists is likely in the near future. MGI (2011), for example, estimates that the demand for deep analytical positions in the United States could exceed supply by 140 000 to 190 000 positions by 2018. This does not include the need for an additional 1.5 million managers and analysts who can use big data knowledgeably.

The European Union economies have reported a huge skill shortage for data-related manpower. The EU Competition Commissioner Neelie Kroes noted that of the top 20 global BD companies, 17 are from the US and 2 from Europe. It is argued that the highest performance computers are unaffordable even to a member of the EU (Kroes, 2013).[32]

5. Implications of Big Data for Agricultural Growth

Big data can optimise the value chain and manufacturing production to more efficient use of labour. The promise of big data lies in innovation-related areas: creation of new data-intensive products; use of data to optimise or automate production or delivery processes (data-driven processes for "smart" grids, "smart" logistics and transport); use of data to improve marketing, for instance by providing targeted advertisements, data-driven marketing and design; use of data for new organisational and management approaches; use of data to enhance research and development (data-driven R&D).

Furthermore, data collected from distribution networks allow utility providers to identify losses and leakages during the distribution of energy and other resources. By deploying smart water sensors in combination with data analytics, Aguas Antofagasta, a water utility in Chile, was able to identify water leaks throughout their distribution networks and reduce total water losses from 30% to 23% over the past five years, thereby saving some 800 million litres of water a year. As in the case of public-sector data, opening smart meter data to the market has led to a new industry that provides innovative goods and services based on these data which have contributed to green growth and created a significant number of green jobs.

Big data in agriculture: the leading role of private-sector

The relationship between the private sector and big data in agriculture, particularly in developing countries, is probably consequent to the fact that the private sector has been the source of many of the key technological innovations, and subsequent diffusion of these innovations. Multinationals have played a strong role in the diffusion of new technologies which then generate big data, and the activities of enterprises, including exports, outsourcing, and licensing have led to the rapid spread of new technologies in far flung countries, including those which may not necessarily be priority countries in the development agenda. According to Kshetri, "Transnational Corporations (TNCs) such as Monsanto and Syngenta are likely to drive the international technology transfer of BD in the agricultural sector".[33]

In Africa, it is it private sector which has been most attuned to the potential of big data. According to a report by IBM, up to 40 percent of business in Kenya and Nigeria are in the planning stages of a big data project, which compares positively with the global average of 51 percent, and a further twelve percent of firms in these two countries already have big data projects, which is very close to the 13 percent global mean. The report further revealed an unforeseen impact of big data, namely that it favours small firms, as 43 percent of small firms polled said they were in the planning stage for big data projects, compared with 24 percent of medium companies.[34]

Wolfgang von Loeper, a former farmer, launched in 2014 with IBM **MySmart-Farm app** in South Africa. It leverages agriculture themed data, including climate, soil information and disease, to help farmers make the best decisions about irrigation, pest control and fertilization. Detailed information about irrigation helps farmers to irrigate only when required, and avoid wasting a precious resource.

Furthermore, developing countries are increasingly becoming the source of the technologies which are leading to the boom in big data, often through collaborative relationships and in response to the needs of sectors that have often been neglected in the technological boom. An example is the AgriLife platform which was developed by Kenya-based IT company MobiPay and was launched in late 2012. Mercy Corps then supported the expansion of AgriLife to Uganda and helped build relationships with other service providers and integrate them into the platform, so they can reach rural clients more effectively.[35]

Will Big Data benefit small farmers?

Big data has small players who are especially important as generators of new applications, tools and other means of using technology in innovative ways. This is no more visible than in the area of mobile phones and other portable technology, which have spurred the demand of accessible, affordable, and resource light applications that allow users to enhance the functionality and utility of their devices, not only as means of communication or entertainment, but also as instruments in the work place, in disaster zones, in the school or laboratory.

A big challenge for big data in agriculture, particularly in developing countries, is accessibility and in some cases, affordability. Farmers, particularly smallholder farmers, are the poorest and most vulnerable stakeholders in the agricultural supply chain, and as such are at a particular disadvantage in terms being able to garner the benefits of the big data revolution.

Large growers can afford specialised machineries, which is not the case for small-farmers. The conditions that stimulated the growth of BD in the US farming industry (widespread adoption of mechanized tractors, genetically modified seeds, computers and tablets for farming activities) are less prevalent in developing countries. A main concern is that BD collection efforts will only benefit big and well-educated farmers.[36]

Big Data is needed most by farmers who least can afford it. Major corporations are investing heavily in Big Data for agriculture, and start-ups in the space are proliferating, supported by the increasing availability of venture capital. But all this market-driven activity does little to help poor, developing areas such as sub-Saharan Africa, where productivity is very low by U.S. standards and where virtually all of the world's population growth is predicted to take place in the coming decades.

As Big Data increasingly plays a role in developed countries to increase productivity, the gap between developed and developing countries can become even greater and the perceived limited application of new technologies in farming in developing countries also decreases the attractiveness of the sector for investors and for youth as a source of future employment.

5.1 Precision agriculture: towards e-farming?

Precision agriculture is closely associated with technology and its application to large-scale farms in developed countries. Technology allows today to have GPS-equipped sensors on tractors enable to measure and respond to soil variability across vast tracts of land and dispense the right amounts of fertiliser and water exactly where it's needed.

Higher volume of data on farming activities is available than in the past. For instance, data such as farmers' credit history and the amounts of seeds and pesticides used was not available in the pre Big Data-environment. As to the data speed, near-real-time data and information on farmers' needs and capabilities are available. This means that financial institutions, produce buyers, and other relevant actors can fulfill farmers' needs more quickly than in the past. Regarding the variety, most data currently used in farming-related activities is structured data. Such data can be combined with unstructured data. For instance, farmers can upload pictures and videos related to a problem they are facing, which can be analysed by experts to offer customized advice.

Some argue that BD is the source of the next revolution in farming (Jacob Bunge, 2014). An overview of the deployment of BD in industrialized countries would be helpful for how the condition can be improved in developing countries. On this front, precision agriculture or precision farming has been a key trend in industrialized countries. Data collected on soil conditions, seeding rates, crop yields, and other variables from farmers' tractors, combines, and drones is combined with detailed records on historic weather patterns, topography, and crop performance collected by the providers of prescriptive-planting technology. The data is crunched by algorithms and human experts and turned into customized useful advice and is sent directly to farmers and their machines, instructing them as to the optimum amount of pesticides, herbicides, fertilizer, and other applications.

The cloud-based platform AgriLife, which is accessible via mobile phone. It is used for collecting data and analyzing farmers' production capability and history. In order to ensure fast, easy and efficient availability of resources and services to distant, rural farmers, the platform also acts as an integration point for financial institutions, mobile network operators, produce buyers, and their agents (Yeoman, 2013).

The data analysis provides a better understanding of small farmers' needs and production capability. Service providers can tailor their offerings such as crop insurance, input payments, and savings accounts based on the data (Yeoman, 2013). Uganda's Farmers Centre (FACE) was an early adopter of AgriLife. FACE started uploading information on its 10,000 farmer clients, who travel long distances to purchase inputs and aggregate their produce at FACE warehouses for processing/sale.

Before using AgriLife, FACE collected information by paper-based questionnaires. Small farmers' transaction data would help them build a credit history, which is used by value-chain actors to provide credit and other resources such as seeds, fertilizers, and pest-control chemical agents (Yeoman, 2013).

As of September 2013, AgriLife facilitated over US$ 2 million in revolving credit lines to about 120,000 small farmers in Kenya and Uganda. The AgriLife platform is also being used in Zimbabwe, Zambia, and Senegal (G-Analytix, 2013).

Prior researchers have recognized that developing world-based farmers face difficulties in meeting the quality and safety standards set by the developed world (Oluoch-Kosura, 2010). In this regard, the availability of easily accessible data that include digital records of farming activities such as the amounts of seeds and pesticides will obviously play a major role in documenting quality standards of agricultural products.

Many farmers who have implemented data-driven prescriptive planting based on the analysis of nutrients in soil and other factors have reported a significant increase in productivity.

TNCs, which are often producers, processors, or traders of agricultural products or sellers of inputs or machinery, engage in a contracting system in which they assume a variety of responsibilities including providing technical assistance and marketing to developing world-based small farmers. TNCs such as Monsanto and Syngenta, which have become a driving force behind the utilization of BD in the industrialized world, are thus likely to act as a key channel in the international technology transfer of BD. A related point is that international technology transfer in BD is likely to have differential effects across different categories of crops. For instance, foreign companies are more active in newly emerging export crops, which are integrated into the international supply chain. Traditional cash crops such as coffee, cotton, tea, and tobacco are thus more likely to realize the need to adopt various aspects of BD (Hoeffler, 2006).

Many tractors and combines are guided by global **positioning system satellites**. Devised for industrialised farms, precision agriculture now has the potential to increase the yields of smallholder farmers and the technology being more widely accessible. For example, a new handheld device known as the **GreenSeeker**, which is based on the relationship between the light reflectance in the red and near infrared spectrum of a plant, and the nitrogen status of that plant can be used to measure the health and nitrogen status of plants, enabling farmers to make more precise assessments of fertiliser requirements. Better use of nitrogen fertiliser not only increases profitability but also reduces groundwater pollution. GreenSeeker is used in Mexico and being evaluated in Ethiopia and South aia (CIMMYT). The GreenSeeker costs about $500 (£297), making it relatively affordable though still expensive for many small-scale farmers.

In 2013, Monsanto acquired the weather-data-mining firm Climate Corp. Likewise, the agricultural cooperative Land O'Lakes bought satellite imaging specialist Geosys. In the same vein, in order to provide real-time climate and market information to its data service users, DuPont announced collaboration with the weather-and-market analysis firm DTN/The Progressive Farmer. In 2013, Deere agreed to send data from its tractors, combines, and other machinery to the computer servers of DuPont and Dow (Bunge, 2014).

Nutrient management is another area where BD may be relevant. In Africa, outdated knowledge is pervasive and ubiquitous in recommendations for nutrient management.

This often leads to too much fertilizer in relation to potential crop demand and on a uniform basis irrespective of the type of land (Giller et al., 2011). A model-based and data-driven approach is thus likely to reduce the costs of fertilizer and increase productivity.

A further area in which BD might have potential to facilitate agricultural and farming activities in developing countries relates to the availability of near-real-time data and information regarding farmers' needs and capabilities, which can be used by value chain partners to effectively serve the farmers.

With evidence that precision agriculture techniques can work, the challenge is creating appropriate enabling environments to encourage take-up.

5.2 Data for investment

One of the biggest impediments to investment is lack of information. With information, investor confidence grows. With confident investors, money flows and that's when real, sustainable development happens. Information is power, but ask any trade minister trying to promote his or her country without adequate data, and you will understand lack of information is powerlessness.

Every day, according to the coconut-milk-drinking nerds in Silicon Valley, the world generates 2.5 quintillion bytes of electronic data.

Big data is quickly being seen as the next big investment in agribusiness. In early November Monsanto $930 million cash purchased The Climate Corp., a San Francisco-based tech company, which processes 52.43 million megabytes of weather information daily including eight years of soil, moisture and precipitation records for each of the 29 million farm fields in the U.S. Using this data the company uses an algorithm that divides the country into approximately half a million plots and then generates 10,000 daily weather scenarios for each of them. This information is used to creatme custom insurance policies for corn, soybeans and wheat farmers. Then when data shows adverse weather conditions causing loss of crops an insured farmer receives a check – no claims, forms, or negotiating. The Climate Corp. charges approximately $40 per acre to insure crops and Monsanto expects the company will yield $20 million in the coming years.

Climate Corp.'s whiz-bang crop insurance scheme stands on two pillars. First, it lifts buckets of free weather and yield information from the National Weather Service and the U.S. Department of Agriculture and, second, Monsanto's 2012 purchase of Precision Planting, an Illinois firm that specializes in on-the-go seed selection and placement.

Together it's Big Data meets Big Seed meets Big Iron, and the pairings will drive your tractor, select and place your seed (from personalized varieties) by the foot or meter and fertilize, irrigate and insure the crop while you're monitoring it all from your kitchen or farm office.

5.3 Successful applications of relevance to agriculture

The retailer **Tesco** combines weather information with its own sales history to predict demand for certain items and prevent waste. Adidas is able to immediately redirect supply trucks to stores running low on a particular style of shoe, based on its real-time monitoring.

In one well-known **study**, researchers followed the position of 1.9 million active mobile phone SIM cards in Haiti around the time of the 2010 earthquake to track the displaced population.

At the same time, Big Data allows novel ways to monitor a population's characteristics in real time. For example, **researchers from Belgium** used anonymised data on how much airtime credit mobile phone users in Cote d'Ivoire purchased to estimate the relative income of individuals, and the diversity and inequality of income. This analysis paints a nuanced picture of shifts in the distribution of wealth.

Data for climate change. Through the **Regional Centre for Mapping of Resources for Development (RCMRD)** in Nairobi, Kenya launched a satellite tracking system that can collect real-time data from 75 percent of Africa's land area. Capable of capturing images with a 250-metre resolution, the Moderate Resolution Imaging Spectroradiometer (MODIS) monitors factors affecting the environment, like forest fires, in areas where human surveillance cannot reach without the aid of aerial photography. It enables the acquisition of direct data which can be processed into different products for a variety of applications, such as flood mapping, crop monitoring, fire assessment, water quality assessment and hailstorm prediction. The satellite receiving station in Nairobi collects data from several earth observation satellites, which it shares with the RCMRD's 15 member states in eastern and southern Africa.

IBM Research Africa in Nairobi begins their project in using their **supercomputer "Watson"** to address agriculture problems on the continent.

In 2011, **Kenya launched its new Open Data Portal**, which includes a full digital edition of the 2009 census, 12 years of detailed government expenditure data, government household income surveys, and the location of schools and health facilities. The portal provides unlimited data access on the web and through mobile phones to researchers, web and software developers, journalists, students, civil society and the general public.

Crowdsourcing. NGOs like Ushahidi are already using crowdsourcing to obtain, verify and disseminate real-time information about natural disasters and election monitoring, and Ushahidi is developing ways to filter and use the huge amounts of information being created with applications such as SwiftRiveriii.

GeoPoll, a tool for surveying populations through cellphones, using text messages, voice recordings and Web applications. Database of more than 150 million users which allows to collect real-time data that can help businesses, organizations and governments make better-informed decisions.

African Development Bank (AfDB) launched the **Africa Information Highway (AIH)**, which comprises two types of portals for each participating country: a statistical data portal and an open data portal. AIH is Africa's new "one-stop centre" for development data.

'**Data for Development Senegal'** is an innovation challenge open on ICT Big Data for the purposes of societal development. Following the **'D4D' in the Ivory Coast** in 2013, Sonatel and the Orange Group are making anonymous data, extracted from the mobile network in Senegal, available to international research laboratories, as well as data on hours of sunshine.

Data for insurance: the automatic weather stations can show insurance companies, governments and farmers how much rain is received over a given period of time much more accurately than the weather stations. The SERVIR platform, set up in 2008, integrates satellite observation and predictive models with other geographic information to track and forecast ecological changes, and respond to natural disasters.

The **Mtrac programme in Uganda**, supported by UNICEF, the WHO and USAID, uses SMS surveys completed by health workers to alert public health officials to outbreaks of malaria, and lets them know how much medicine is on hand at health facilities, so they can anticipate and resolve any shortages. Now the Ugandan government is collecting data from thousands of health facilities, capturing and analysing results within 48 hours at a total cost of less than US$150 per poll.

The **Rapid Family Tracing and Reunification** (RapidFTR) is an open source mobile application used to collect crucial information about children who have been separated from their families in disaster situations. Information is shared securely on a central database for family members looking for a missing child. RapidFTR uses the same type of security as mobile banking to ensure that family-tracing information, especially photos, is accessible only by authorised users, to protect these vulnerable children..[37]

Mobilising Data to Deal with an Epidemic : In the wake of Haiti's devastating 2010 earthquake, researchers at the Karolinska Institute and Columbia University demonstrated that mobile data patterns could be used to understand the movement of refugees and the consequent health risks posed by these movements. Researchers from the two organisations obtained data on the outflow of people from Port-au-Prince following the earthquake by tracking the movement of nearly two million SIM cards in the country. They were able to accurately analyse the destination of over 600,000 people displaced from Port-au-Prince, and they made this information available to government and humanitarian organisations dealing with the crisis. Later that year, a cholera outbreak struck the country and the same team used mobile data to track the movement of

people from affected zones. Aid organisations used this data to prepare for new outbreaks. The example from Haiti demonstrates how mobile data analysis could revolutionise disaster and emergency responses.

As part of the development of their **National Strategy for the Development of Statistics** (NSDS) completed in partnership with the Partnership in Statistics for Development in the 21st Century (PARIS21), **Rwanda** identified some simple, yet systematic, improvements that could dramatically help make better use of evidence for policy making. One innovation included moving up the publishing date of the Consumer Price Index by five days each month in response to needs from both policy makers and businesses. The release date of the Demographic and Health Survey and Living Conditions Survey was changed so that the information could be used in measuring Rwanda's first poverty reduction strategy and so the information could inform planning for the next one. These changes in data scheduling increased the usefulness of the data and allowed for better evidence-based decisions to be made.[38]

The U-report social monitoring platform established by UNICEF in Uganda has more than 240,000 young people reporting on issues that affect their communities. Early reporting of an infectious disease in banana production contributed to halting the spread of the disease, which could have cost the country $360 million per year if left unchecked.[39]

In Kenya, powered by solar energy, the **automatic weather stations** are fitted with a General Packet Radio Service (GPRS), which enables them to record rainfall data from farms within a radius of 20 km every 15 minutes (Center for Training and Research in Arid and Semi Arid Lands Development, CETRAD). Information gathered can show insurance companies, governments and farmers how much rain is received over a given period of time much more accurately than the weather stations. The SERVIR platform, set up in 2008, integrates satellite observation and predictive models with other geographic information to track and forecast ecological changes, and respond to natural disasters.

IMF and EAC team up to develop finance data
The East African Community (EAC) and the International Monetary Fund (IMF) have come together to help regional states in compiling finance statistics for their different governments aimed at assisting the EAC partner states meet the fiscal data requirements associated with the East Africa Monetary Union (EAMU) Protocol. The intervention is timely in facilitating production of robust statistical data required for the establishment of the regional monetary Union and transition to EAC single currency by 2024.

5.4 Using Big Data for Monitoring and Evaluation

While technology in itself is not suffcient without trained staff and adequate resources, using

ICT – mobilephones, tablets, applications and software – to collect data in the field, and to perform M&E in development projects, while also working closely with rural communities and taking their feedback offers new opportunities. Some tools include:

iFormBuilder: An iOS mobile data collection platform that features an application that requires no paper or connection and is available worldwide. This application is being used for data collection in over 110 countries and it allows real time data upload and offline data collection, while immediately sending any updates to a mobile workforce with server assignment. Catholic Relief Services used iFormBuilder to register and distribute vouchers to beneficiaries during a seed fair in Central African Republic, and they were able to save over a week of preliminary work and reduce staff by 50%.

Cropster: is an initiative that seeks to support sustainable agriculture by empowering farmers with access to key information and ensuring data transparency. It enables them to make informed decisions, and also supports people and communities at all levels of supply chain. This app offers an M&E tool that facilitates data collection as well as the exchange of information within producer groups, NGOs and commercial partners in Latin America. This tool provides decentralised monitoring, real-time information, and the ability for users to customise data and verify input.

EpiSurveyor: An award winning mobile app that lets users create an account, design forms, download them to their mobile phones, collect data and send it to a server. According to a World Bank report, in 2010, nine data collectors used EpiSurveyor to interview beneficiaries in 25 municipalities in a secondary survey (the first one, conducted in 2009, used paper and pen) in a World Bank Conditional Cash Transfer project in Guatemala. Digitisation cut the cost of an interview by 71%, increased the sample size from 200 to 700 beneficiaries, and reduced the individual interview time by 3.6%.

6. The way forward

As the volume, variety and velocity of data continues to increase, so too do the possibilities for how it can be applied to tackle global challenges such as agriculture. The era of big data is potentially transformational. But to achieve these benefits, we need a deeper understanding of several interdependent issues. We need to learn more about the dynamics of constantly changing data flows. New means and mechanisms for engaging individuals need to be developed. The approaches will need to be meaningful, pragmatic, adaptive and proportional. With so much uncertainty, the need for continuous experimentation, learning and sharing is paramount. Investing in small-scale pilots that bring together the private sector, regulators, civil society and local communities will provide the insights and local knowledge critical for long-term resilience and adaptation.[40] Efforts to improve cooperation between old and new data producers, ensure the engagement of data users, and develop global ethical, legal and statistical standards to improve data quality and protect people from abuses in a rapidly changing data ecosystem are needed.

Brynjolfsson et al.(2011) estimate that the output and productivity of firms that adopt data-driven decision making are 5% to 6% higher than would be expected from their other investments in and use of information technology. These firms also perform better in terms of asset utilisation, return on equity and market value. Growing investments in data management and analytics partly reflect the increasing economic role of data.

Towards "data-driven decision-making"

The availability of timely, relevant, and reliable data on the agriculture sector is necessary for effective planning, monitoring, and evaluation of agricultural and rural development policies and field interventions. However, at a time when more than ever, reliable data are needed on this sector, several studies point to a steady decline in the quality of agricultural statistics in many developing countries, particularly African countries[41]. Innovations in digital data, rapid data collection and analysis need to be used as tools to help decision-makers.

More data does not mean better decisions. Collecting and processing data and turning them into information, using data and making them open for others to use and re-use is an indispensable investment to get information for actionable data by policy-makers. Therefore, governments should empower and finance statistical independent offices to address the opportunities of data revolution and innovation and to adapt to the new world of data to collect, process, disseminate and use high-quality, open, disaggregated and geo-coded data, both quantitative and qualitative. Governments should also design and implement governance systems which regulates the use and management of data.

The IEAG report proposes ways to improve data for achieving and monitoring sustainable development. The report highlights two big global challenges for the current state of data: (i) the challenge of invisibility (gaps in what we know from data, and when we find out); (ii) The challenge of inequality (gaps between those who with and without information, and what they need to know make their own decisions) and calls for fostering and promoting innovation to fill data gaps; mobilising resources to overcome inequalities between developed and developing countries and between data-poor and data-rich people and promote leadership and coordination to enable the data revolution to play its full role in the realisation of sustainable development.

Development partners should also support data investment, especially for capacity building, infrastructure development in order to get qualitative data usable to demonstrate their impact.

They should also support the setting and enforcement of common standards for data collection, production, anonymisation, sharing and use to ensure that new data flows are safely and ethically transformed into global public goods, and maintain a system of quality control and audit for all systems and all data producers and users.

GLOSSARY[123]

Algorithm

A mathematical formula placed in software that performs an analysis on a set of data.

Analytics

Using software-based algorithms and statistics to derive meaning from data.

Analytics platform

Software or software and hardware that provides the tools and computational power needed to build and perform many different analytical queries.

Integrate and analyse data to uncover new insights, and help companies make better-informed decisions. There is a particular focus on this field on latency, and delivering insights to end users in the most timely manner possible.

Anonymization

The severing of links between people in a database and their records to prevent the discovery of the source of the records.

Automatic identification and capture (AIDC)

Any method of automatically identifying and collecting data on items, and then storing the data in a computer system. For example, a scanner might collect data about a product being shipped via an RFID chip.

Behavioral analytics

Using data about people's behavior to understand intent and predict future actions.

Big data

This term has been defined in many ways, but along similar lines. Doug Laney, then an analyst at the META Group, first defined big data in a 2001 report called "3-D Data Management: Controlling Data Volume, Velocity and Variety." Volume refers to the sheer size of the datasets. The McKinsey report, "Big Data: The Next Frontier for Innovation, Competition, and Productivity," expands on the volume aspect by saying that, "'Big data' refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze."

Business Intelligence (BI) Platforms

Used for integrating and analysing data specifically for businesses. BI Platforms analyse data from multiple sources to deliver services such as business intelligence reports, dashboards and visualizations

Cassandra

A popular choice of columnar database for use in big data applications. It is an open source database managed by The Apache Software Foundation.

Cell phone data

Cell phones generate a tremendous amount of data, and much of it is available for use with analytical applications.

Citizen Reporting or Crowd-sourced Data

Information actively produced or submitted by citizens through mobile phone-based surveys, hotlines, usergenerated maps, etc; While not passively produced, this is a key information source for verification and feedback

Cloud

A broad term that refers to any Internet-based application or service that is hosted remotely.

Cloud computing

Cloud computing is described as "a service model for computing services based on a set of computing resources that can be accessed in a flexible, elastic, on-demand way with low management effort (OECD). Due to economies of scale, cloud computing providers have much lower operating costs than companies running their own IT infrastructure, which they can pass on to their customers. Cloud computing has played a significant role in the increase in data storage and processing capacity. In particular, for small and medium-sized enterprises (SMEs), but also for governments that cannot, or do not want to, make heavy upfront investments in ICTs, cloud computing enables organisations to pay for supercomputing resources via a pay-as-you-go model.

Cognitive systems

Artificial systems that can perceive their environment, understand the situation and execute tasks efficiently even in challenging conditions.

Such systems, including robots, smart spaces and sensor networks, are robust, intuitive to use and capable of adapting to users and changing situations. They can work with various degrees of autonomy, in isolation or in cooperation with people. The next generation of such ICT systems and products with more intelligence has a huge potential impact on many sectors, from industrial production to personal assistance.

Collective Awareness Platforms for Sustainability and Social Innovation (CAPS)

ICT systems leveraging the emerging "network effect" by combining open online social media, distributed knowledge creation and data from real environments ("Internet of Things") in order to create awareness of problems and possible solutions requesting collective efforts, enabling new forms of social innovation. The Collective Awareness Platforms are expected to support environmentally aware, grassroots processes and practices to share knowledge, to achieve changes in lifestyle, production and consumption patterns, and to set up more participatory democratic processes.

Computer-generated data

Any data generated by a computer rather than a human–a log file for example.

Content management system (CMS)

Software that facilitates the management and publication of content on the Web.

Cookies

Cookies are short text files stored on the users' computer by a website. Cookies are normally used to provide a more personalised experience and to remember the users' profile without the need for a specific login. They can also be placed by third parties (such as advertising networks) in end users´ devices and be used to track users when surfing across different websites associated to that third party.

The EU privacy legal framework requires users' consent to store cookies in their terminal devices (computers, lap tops, smartphones) or gain access to information collected through cookies.

Creative Commons Licenses

Models released by Creative Commons, a U.S. non-profit corporation. Many of the licenses, notably all the original licenses, grant certain "baseline rights", such as the right to distribute the copyrighted work without changes, at no charge. Creative Commons licenses are currently available in over 70 different jurisdictions worldwide. Licenses for jurisdictions outside the United States are under the purview of Creative Commons International.

Crowdfunding

A collective effort by many individuals who network and pool their resources to support efforts initiated by other people or organizations, often over the internet. Individual projects and businesses are financed with small contributions from a large number of individuals; innovators, entrepreneurs and business owners utilise their social networks to raise capital.

Crowdsourcing

The act of submitting a task or problem to the public for completion or solution.

Cyber security

Commonly refers to the safeguards and actions available to protect the cyber domain, both in the civilian and military fields, from those threats that are associated with or that may harm its interdependent networks and information infrastructure. Cyber security strives to preserve the availability and integrity of the networks and infrastructure and the confidentiality of the information contained therein. The term cyber security also covers prevention and law enforcement measures to fight cybercrime.

Quality of official statistics.

Traditional definition of quality as statistical accuracy based on sampling theory has been complemented over the years by more user-oriented quality criteria such as relevance, timeliness, coherence or integrity.

Data

A quantitative or qualitative value. any and all of the digital materials that are collected and analyzed in the pursuit of scientific advances

Common types of data include sales figures, marketing research results, readings from monitoring equipment, user actions on a website, market growth projections, demographic information, and customer lists.

Data access

The act or method of viewing or retrieving stored data.

Data Access and Interoperability Task Force (DAITF)

A new international initiative for Data Initiated in Europe, DAITF aims to create an open forum for discussing and agreement on data- related standards, APIs, policy rules, and data interoperability mechanisms, from the very basic integration layers (AAI, PIDs, registries, etc.) to the semantic and regulatory levels. DAITF will integrate its bottom-up activities under the RDA (Research Data Alliance) governance structure.

Data aggregation

The act of collecting data from multiple sources for the purpose of reporting or analysis.

Data analytics

The application of software to derive information or meaning from data. The end result might be a report, an indication of status, or an action taken automatically based on the information received.

Data architecture and design

How enterprise data is structured. The actual structure or design varies depending on the eventual end result required. Data architecture has three stages or processes: conceptual representation of business entities. the logical representation of the relationships among those entities, and the physical construction of the system to support the functionality.

Database

A digital collection of data and the structure around which the data is organized. The data is typically entered into and accessed via a database management system (DBMS).

Database management system (DBMS)

Software that collects and provides access to data in a structured format.

Data center

A physical facility that houses a large number of servers and data storage devices. Data centers might belong to a single organization or sell their services to many organizations.

Data collection

Any process that captures any type of data.

Data Exhaust

Passively collected transactional data from people's use of digital services like mobile phones, purchases, web searches, etc., and/or operational metrics and other real-time data collected by UN agencies, NGOs and other aid organisations to monitor their projects and programmes (e.g. stock levels, school attendance); these digital services create networked sensors of human behaviour;

Ddata governance

A set of processes or rules that ensure the integrity of the data and that data management best practices are met.

Data integration

The process of combining data from different sources and presenting it in a single view.

Data integrity

The measure of trust an organization has in the accuracy, completeness, timeliness, and validity of the data.

Data management

According to the Data Management Association, data management incorporates the following practices needed to manage the full data lifecycle in an enterprise:

- data governance
- data architecture, analysis, and design
- database management
- data security management
- data quality management
- reference and master data management
- data warehousing and business intelligence management
- document, record, and content management
- metadata management
- contact data management

Data Management Association (DAMA)

A non-profit international organization for technical and business professionals "dedicated to advancing the concepts and practices of information and data management."

Data- and text-mining

Automated research technique in the digital environment for the purpose of discovering and extracting knowledge from unstructured data. Via keyword searches, lexical analysis tools and mining algorithms, text- and data mining enables researchers to structure the content of scientific (numerical) databases or written publications after their preferences and for their own research purposes. It allows also for retrieving and extracting relevant information and automatically gaining structured results without having to browse each and every scientific work found in a simple keyword search. It is growingly applied in many different areas, from market and business intelligence to science and engineering (bioinformatics, genomics, medicine, education).

Data model, data modeling

A data model defines the structure of the data for the purpose of communicating between functional and technical people to show data needed for business processes, or for communicating a plan to develop how data is stored and accessed among application development team members.

Data profiling

The process of collecting statistics and information about data in an existing source.

Data protection and Privacy

Data protection refers to personal data, gathered and processed in a safe and secure manner. Privacy is the prerogative of indviduals to be left alone, out of public view, and in control of the collection and sharing of information about themselves (informational privacy). The concepts of data protection and privacy therefore overlap, but do not coincide.The right to privacy is enshrined in the Universal Declaration of Human Rights (Article 12), the EU Charter of Fundamental Rights (art 7,8) as well as in the European Convention of Human Rights (Article 8).

Data quality

The measure of data to determine its worthiness for decision making, planning, or operations.

Data repository

The location of permanently stored data.

Data science

A recent term that has multiple definitions, but generally accepted as a discipline that incorporates statistics, data visualization, computer programming, data mining, machine learning, and database engineering to solve complex problems.

Data set

A collection of data, typically in tabular form.

Data structure

A specific way of storing and organizing data.

Data value chain

Underlying concept to describe the idea that data assets can be produced by private actors or by public authorities and exchanged on efficient markets like commodities and industrial parts (or made available for reuse as public goods) throughout the lifecycle of datasets (capture, curation, storage, search, sharing, transfer, analysis and visualization). These data are then aggregated as inputs for the production of value-added goods and services which may in turn be used as inputs in the production of other goods and services.

Data warehouse

A place to store data for the purpose of reporting and analysis.

Digital Science

Digital Science (also Open Digital Science) is about the way research is carried out, disseminated, deployed and transformed by digital tools, networks and media. These issues are often also covered by concepts such as e-science, e-infrastructures, open science, science2.0, web science, or internet science.

Grid computing

The performing of computing functions using resources from multiple distributed systems. Grid computing typically involves large files and are most often used for multiple applications. The systems that comprise a grid computing network do not have to be similar in design or in the same geographic location.

Hackathon (also known as a hack day, hackfest or codefest)

Event in which computer programmers and others involved in software development, including graphic designers, interface designers and project managers, collaborate intensively on software projects.]Occasionally, there is a hardware component as well.

Hackathons typically last between a day and a week. Some hackathons are intended simply for educational or social purposes, although in many cases the goal is to create usable software.

Hadoop

An open source software library project administered by the Apache Software Foundation. Apache defines Hadoop as "a framework that allows for the distributed processing of large data sets across clusters of computers using a simple programming model."

A whole ecosystem of technologies designed for the storing, processing and analysing of data. The core Hadoop technologies work on the principle of breaking up and distributing data into parts and analysing those parts concurrently, rather than tackling one monolithic block of data all in one go.

Information management

The practice of collecting, managing, and dsitributing information of all types–digital, paper-based, structured, unstructured.

Internet Corporation for Assigned Names and Numbers (ICANN)

Non-profit private sector US corporation formed in 1998 that helps with the global technical coordination of key Internet resources such as Top Level Domain Names and Internet address spaces. ICANN also develops policies for the introduction of new generic Top Level Domains and Top Level Domains using IDNs (internationalised domain names using non-English characters such as Arabic, Hindi and Chinese).

Internet of Services (IoS)

Refers to services which can be managed through IT and, being combined, lend themselves into value-added services. This is one to keep an eye on, especially as the service sector has become the biggest and fastest-growing business sector in the world.

Internet of Things (IoT)

Dynamic global network infrastructure with self-configuring capabilities based on standard and interoperable communication protocols where physical and virtual "things" have identities, physical attributes and virtual personalities and use intelligent interfaces and are seamlessly integrated into the information network. More on the internet of things.

Kafka

LinkedIn's open-source message system used to monitor activity events on the web.

liLnked data

As described by World Wide Web inventor Time Berners-Lee, "Cherry-picking common attributes or languages to identify connections or relationships between disparate sources of data."

Load balancing

The process of distributing workload across a computer network or computer cluster to optimize performance.

Location analytics

Location analytics brings mapping and map-driven analytics to enterprise business systems and data warehouses. It allows you to associate geospatial information with datasets.

Machine learning

The use of algorithms to allow a computer to analyze data for the purpose of "learning" what action to take when a specific pattern or event occurs.

Whereas the analytics platforms input processed data and output analytics/ dashboards/visualisations for end users, the input in machine learning is data the algorithm 'learns from', and the output depends on the use case. One of the most famous examples is IBM's super computer Watson, which has 'learned' to scan vast amounts of information to find specific answers, and can comb through 200 million pages of structured and unstructured data in minutes. The computer recently combed through recipes and taste combinations to create its own sauce.

Mashup

The process of combining different datasets within a single application to enhance output–for example, combining demographic data with real estate listings.

Massively Parallel Processing (MPP) Databases- MPP databases work by segmenting data across multiple nodes, and processing these segments of data in parallel, and uses SQL. Whereas Hadoop is usually run on cheaper clusters of commodity servers, most MPP databases run on expensive specialised hardware.

Master data management (MDM)

Master data is any non-transactional data that is critical to the operation of a business–for example, customer or supplier data, product information, or employee data. MDM is the process of managing that data to ensure consistency, quality, and availability.

Metadata

Any data used to describe other data–for example, a data file's size or date of creation.

NoSQL- Stands for Not Only SQL; also involved in processing large volumes of multi-structured data. Most NoSQL databases are most adept at handling discrete data stored among multi-structured data. Some NoSQL databases, like HBase, can work concurrently with Hadoop.

Official Statistics

Refers to figures that must be generated:
(a) with the purpose to serve the whole spectrum of the society;
(b) based on quality criteria and best practices;
(c) by statisticians with assured professional independence and objectivity.

Online Information

Web content such as news media and social media interactions (e.g. blogs, Twitter), news articles obituaries, e-commerce, job postings; this approach considers web usage and content as a sensor of human intent, sentiments, perceptions, and want.

Open access

Practice of providing on-line access to scientific information that is free of charge to the reader.

Open Data

Free and widely available data for consultation and reuse, including reuse for commercial purposes, with a view to increasing transparency and stimulating economic activity. Applies mostly, but is not strictly limited to government data.

Open Data Center Alliance (ODCA)

A consortium of global IT organizations whose goal is to speed the migration of cloud computing.

Open source software (OSS)

Software distributed freely with its code, allowing anyone to access, to study, to redistribute and to change it. It must be distributed under a license recognised by the Open Source Initiative or the Free Software Foundation (FSF).

Operational data store (ODS)

A location to gather and store data from multiple sources so that more operations can be performed on it before sending to the data warehouse for reporting.

Parallel data analysis

Breaking up an analytical problem into smaller components and running algorithms on each of those components at the same time. Parallel data analysis can occur within the same system or across multiple systems.

Petabyte

One million gigabytes or 1,024 terabytes.

Physical Sensors

Satellite or infrared imagery of changing landscapes, traffic patterns, light emissions, urban development and topographic changes, etc; this approach focuses on remote sensing of changes in human activity

Recommendation engine

An algorithm that analyzes a customer's purchases and actions on an e-commerce site and then uses that data to recommend complementary products.

Reference data

Data that describes an object and its properties. The object may be physical or virtual.

Satellite Broadband

Satellite Broadband is a high-speed bi-directional Internet connection made via communications satellites instead of a telephone landline or other terrestrials means. Today, satellite broadband is comparable to DSL broadband in terms of both perform

Sawzall

Google's procedural domain-specific programming language designed to process large volumes of log records.

Search

The process of locating specific data or content using a search tool.

Semantic Web

A project of the World Wide Web Consortium (W3C) to encourage the use of a standard format to include semantic content on websites. The goal is to enable computers and other devices to better process data.

Server

A physical or virtual computer that serves requests for a software application and delivers those requests over a network.

Smart grids

Upgraded electricity networks enhanced by two-way digital communication between supplier and consumer and intelligent metering and monitoring systems. Smart Grids can efficiently integrate the behaviour and actions of all users connected to it: in order to ensure an economically efficient, sustainable power system with low losses and high quality, security of supply and safety.

Social Media

A group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content. It employs mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content.

Software defined networking (SDN)

An approach to building computer networks that separates and abstracts elements of these systems. These elements are called the “control plane” and the “data plane”. SDN decouples the system that makes decisions about where traffic is sent (the control plane) from the underlying system that forwards traffic to the selected destination (the data plane). This technology simplifies networking and enables new applications, such as network virtualization in which the control plane is separated from the data plane and implemented in a software application. This architecture allows network administrators to have programmable central control of network traffic without requiring physical access to the network's hardware devices.

Storm

An open-source distributed computation system designed for processing multiple data streams in real time.

Terabyte

1,000 gigabytes.

Trade Related Aspects of Intellectual Property Rights (TRIPS)

The TRIPS Agreement which came into effect on 1 January 1995, is to date the most comprehensive multilateral agreement on trade-related intellectual property. TRIPS covers notably copyright and related rights (i.e. the rights of performers, producers of sound recordings and broadcasting organisations)

Transactional data

Data that changes unpredictably. Examples include accounts payable and receivable data, or data about product shipments.

Transparency

As more data becomes openly available, the idea of proprietary data as a competitive advantage is diminished.

Visualization Platforms- Specifically designed- as the name might suggest- for visualizing data; taking the raw data and presenting it in complex, multi-dimensional visual formats to illuminate the information.

Weather data

Real-time weather data is now widely available for organizations to use in a variety of ways. For example, a logistics company can monitor local weather conditions to optimize the transport of goods. A utility company can adjust energy distribution in real time.

Web 2.0

Refers to an incremental development of the technologies behind the World Wide Web, allowing the user to participate and contribute directly to the production of information, rather than being a mere passive receiver of it.

Whole Earth Model

An integrated data management system that allows geophysicists, engineers, and financial managers in the oil and gas industry evaluate the potential of oil and gas fields.

Sources: Eurostat, Data-informed, Global Pulse, Dataconomy, EC-Europa, GODAN.

ACRONYMS

ADP	Accelerated Data Program
AFCAS	African Commission for Agricultural Statistics
AfDB	African Development Bank
AFRISTAT	Observatoire Economique et Statistique d'Afrique Subsaharienne
AGROST	African Group on Statistical Training and Human Resources
ASCI	Agricultural Statistics Capacity Indicator
ASDCI	Agricultural Statistics Development Composite Indicator
ASSD	Africa Symposium on Statistical Development
AUC	African Union Commission
AQUASTAT FAO	Information System on Water and Aquaculture
BD	Big Data
DQAF	Data Quality Assessment Framework
FAOSTAT	Food and Agriculture Organization Corporate Statistical Database
CGD	Center for Global Development
COFOG	Classification of Functions of Government
CPC	Central Product Classification
CRM	Customer Relationship Management
ESGAB	European Statistical Governance Advisory Board
Eurostat	Statistical Office of the European Union
FAO	Food and Agricultural Organization of the United Nations
GDP	Gross Domestic Product
GPS	Global Positioning System
HPC	High-performance computing

ICAS-V	Fifth International Conference on Agricultural Statistics
IDA	International Development Association
IHSN	International Household Survey Network
ICP-Africa	International Comparison Program for Africa
IHSN	International Household Survey Network
IP	Internet Protocol
ISI	International Statistical Institute
ISIC	International Standard Industrial Classification of Economic Activities
IoT	Internet of Things
ITU	International Telecommunication Union
LCCS	Land Cover Classification System
MDG	Millennium Development Goals
M&E	Monitoring and Evaluation
MPPS	Multiple Probability Proportional to Size
NASS	National Agricultural Statistics Service
NIS	National Institute of Statistics
NOPEMA	Product classification of AFRISTAT member states
NSAS	National System of Agricultural Statistics
NSC	National Statistical Council
NSDS	National Strategy for the Development of Statistics
NSO	National Statistics Office
NSS	National Statistical System
ODI	Open Data Institute
ODP	Open Data Portal

OECD	Organization of Economic Cooperation and Development
OSS	Open source software
PARIS21	Partnership in Statistics for Development in the 21st Century
PDA	Personal Digital Assistant
PPPs	Purchasing Power Parities
PRESS	Partner Report on Support to Statistics
PSI	Public-sector data, public sector information
RRSF	Reference Regional Strategic Framework for Statistical Capacity Building in Africa
SDG	Sustainable Development Goals
SEEA	System of Integrated Environmental and Economic Accounting
SHaSA	Strategy for the Harmonization of Statistics in Africa
SNA	System of National Accounts
SPARS	Strategic Plan for Agricultural and Rural Statistics
STATCOM	United Nations Statistical Commission StatCom-Africa Statistical Commission for Africa
TFSCB	Trust Fund for Statistical Capacity Building
TQM	Total Quality Management
TWG	Technical Working Group
UN	United Nations
UNECA	United Nations Economic Commission for Africa
UNSC	United Nations Statistical Commission
UNSD	United Nations Statistics Division
VA	Value Added

RESOURCES AVAILABLE ONLINE (English and French)

En italique les documents disponibles en français

AFRICAN DEVELOPMENT BANK

African Development Working Group. 2014. Delivering on the Data Revolution in Sub-Saharan Africa Final Report of the Data for African Development Working Group. Center for Global Development & The African Population and Health Research Center.
http://www.cgdev.org/sites/default/files/CGD14-01%20complete%20for%20web%200710.pdf

African Development Bank. 2014. Country Assessment of Agricultural Statistical Systems in Africa Measuring the Capacity of African Countries to Produce Timely, Reliable and Sustainable Agricultural Statistics.
http://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/AfricaCountryAssessment_ASCI_Report_Final_Web_11_2014.pdf

African Development Bank. 2014. Implementation Update: Action Plan to Improve Statistics for Food Security, Sustainable Agriculture and Rural Development in Africa. Statistics Department, Chief Economist Complex Bulletin No. 5, April 2014.
http://www.afdb.org/fileadmin/uploads/afdb/Documents/Policy-Documents/April_2014_-_Implementation_Update-_Action_Plan_to_Improve_Statistics_for_Food_Security__Sustainable_Agriculture_and_Rural_Development_in_Africa.pdf

African Development Bank, African Union, Food and Agricultural Organization and UN Economic Commission for Africa. 2011. Improving Statistics for Food Security, Sustainable Agriculture, and Rural Development: An Action Plan for Africa 2011-2015.
http://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Improving%20Statistics%20for%20food%20security_Sustainable_Agriculture%20and%20Rural%20Development.pdf

African Development Bank, African Union and UN Economic Commission for Africa. 2010. Strategy for the Harmonization of Statistics in Africa SHaSA.
http://auc-statdivision.voila.net/pdfstat/SHaSA_strategy.pdf

CENTRE FOR GLOBAL DEVELOPMENT

Centre for Global Development. 2014. Delivering on the African Data Revolution
http://www.cgdev.org/sites/default/files/CGD14-01%20complete%20for%20web%200710.pdf

Sandefur, J. and Glassman, A. 2013. The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics. Center for Global Development. Washington DC.
http://www.cgdev.org/sites/default/files/political-economy-bad-data.pdf

Sandefur, J. 2012. Statistical independence and the missing macroeconomics of Africa. Centre for the Study of African Economies Newsletter 07. pp. 8-10. CSAE. Oxford University.
http://www.csae.ox.ac.uk/output/newsletters/pdfs/CSAE-newsletter-07-2012.pdf

EUROPEAN COMMISSION

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of Regions. Open data: An engine for innovation, growth and transparent governance. COM(2011) 882 final. 2011
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF

European Commission. 2010. Riding the Wave: How Europe can gain from the rising tide of scientific data. Final report by the High-level Expert Group on Scientific.
http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

ACP-EU TECHNICAL CENTER FOR AGRICULTURAL AND RURAL COOPERATION (CTA)

Andre Jellema, Wouter Meijninger and Chris Addison. Open Data and Smallholder Food and Nutritional Security.Alterra. CTA. 2015

Matthew McNaughton and David Soutar. Agricultural Open Data in the Caribbean. Institutional perceptions, key issues and opportunities. CTA,. 2015
Blog of the Agrihack Talent Initiative. 2014. An initiative of CTA. Facilitated by the ARDYIS Project.
http://hackathon.ict4ag.org

FOOD AND AGRICULTURE ORGANISATION

Food and Agriculture Organisation, United Nations and World Bank. Global Strategy to Improve

Agricultural and Rural Statistics Economic and Sector Work. Report No. 56719-Glb.2010 http://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/Ag_Statistics_Strategy_Final.pdf

Carletto G. 2010. Improving the Availability, Quality and Policy-Relevance of Agricultural Data: The Living Standards Measurement Study – Integrated Surveys on Agriculture. FAO. Prepared for the Wye City Group on Statistics on Rural Development and Agricultural Household Income.. http://www.fao.org/fileadmin/templates/ess/pages/rural/wye_city_group/2010/May/WYE_2010.2.1_Carletto.pdf

Keita, N and Carfagna, E. 2010. Overview of methodological issues for research to improve agricultural statistics in developing countries. Proceeding of ICAS-V, Fifth International Conference on Agricultural Statistics, Integrating Agriculture into National Statistical Systems Kampala, Uganda 13-15 October 2010. http://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/PDF/ICASV_3.4_101_Paper_Keita.pdf

Food and Agriculture Organisation, United Nations and World Bank. 2010.Global Strategy to Improve Agricultural and Rural Statistics Economic and Sector Work. Report No. 56719-Glb http://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/Ag_Statistics_Strategy_Final.pdf

Food and Agriculture Organization (FAO). 2006. The Challenge of Renewal, Report of the Independent External Evaluation of FAO. Submitted to the Council Committee for the Independent External Evaluation of FAO (CC-IEE) September 2007. ftp://ftp.fao.org/docrep/fao/meeting/012/k0827erev1.pdf

OECD

OECD. 2013. Exploring data-driven innovation as a new source of growth. Mapping the Policy Issues Raised by “Big Data”. OECD Publishing. http://www.intgovforum.org/cms/wks2014/uploads/proposal_background_paper/DSTI-ICCP%282012%299-FINAL.pdf

Jutting, Johannes. 2013. Strengthening National Statistical Systems to Monitor Global Goals. Organisation for Economic Co-operation and Development (OECD) and Partnership in Statistics for Development in the 21st Century (PARIS21).Element 5. Paper 1. www.oecd.org/dac/POST-2015%20P21.pdf

OECD. 2013. Building Blocks for Smart Networks. OECD Digital Economy Papers.No. 215. OECD Publishing. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP/CISP%282012%293/FINAL&docLanguage=En

OECD. 2012. ICT Applications for the Smart Grid: Opportunities and Policy Implications. OECD Digital Economy Papers. No. 190. OECD Publishing. http://www.scp-responder.eu/pdf/knowledge/other/ict/OECD%202012%20-%20ICT%20for%20smart%20grid.pdf

OECD. 2012. Summary report of the E-Leaders Meeting 2012: New ICT Solutions for Public Sector Agility. Mexico City. OECD Publishing. http://www.oecd.org/governance/eleaders/E-Leaders-Summary-2012.pdf

Reimsbach-Kounatze, C. 2012. The Proliferation of “Big Data” and Implications of Official Statistics and Statistical Agencies. A preliminary analysis. OECD Publishing. http://www.oecd-ilibrary.org/docserver/download/5js7t9wqzvg8.pdf?expires=1423739572&id=id&accname=guest&checksum=A7BD0D5C0B12A3D15CC95F99F45CBE62

OECD. 2009. Smart Sensor Networks: Technologies and Applications for Green Growth. Digital Economy Papers. No. 167. OECD Publishing. http://www.oecd-ilibrary.org/docserver/download/5kml6x0m5vkh.pdf?expires=1423739295&id=id&accname=guest&checksum=6D9E1E1E032563D6D854036F6A76BD64

OECD. 2008. Recommendation for Enhanced Access and More Effective Use of Public Sector Information. OECD Publishing. http://www.oecd.org/internet/ieconomy/40826024.pdf

OECD. 2005. Principles and Guidelines for Access to Research Data from Public Funding. OECD Publishing. http://www.oecd.org/sti/sci-tech/38500813.pdf

UNITED NATIONS

United Nations Secretary-General. 2014. Independent Expert Advisory Group on a Data Revolution for Sustainable Development. A World that Counts - Mobilising the Data Revolution for Sustainable Development. http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf

United Nationals Economic and Social Council. 2014. Economic Commission for Europe. Big data – an opportunity or a threat to official statistics? Conference of European Statisticians. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2014/32-Eurostat-Big_Data.pdf

United Nations. 2013.Statistics and Indicators for the Post-2015 Development Agenda. New York. http://www.un.org/en/development/desa/policy/untaskteam_undf/UNTT_MonitoringReport_WEB.pdf

United Nations. 2013. A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development. http://www.un.org/sg/management/pdf/HLP_P2015_Report.pdf

UNICEF Newsline. 2013. Rapid route to reuniting Congolese refugee children in Uganda with their families. http://www.unicef.org/infobycountry/uganda_70090.html

United Nations. 2012. The Secretary-General's High-Level Panel of Eminent Persons on the Post-2015 Development Agenda. New York. www.un.org/sg/management/hlppost2015.shtml

United Nations Global Pulse. 2012. Big Data for Development: Opportunities & Challenges. Global Pulse. http://unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf

UNECA. 2010. Regional Strategic Framework for Statistical Capacity Building in Africa. http://ecastats.uneca.org/acsweb/rrsf/Baselineinformation/Statisticalorganization/tabid/183language/en-US/Default.aspx

Kiregyera, B. 2008. Reforming National Statistical Systems in Africa: A Case for Making the National Statistics Offices Autonomous. Statistics Working Paper 2. United Nations Economic Commission for Africa. Addis Ababa http://repository.uneca.org/bitstream/handle/10855/15359/bib-61473_I.pdf?sequence=1

WORLD BANK

Beegle K, De Weerdt J, Friedman J and Gibson J. 2010. Methods of Household Consumption Measurement through Surveys: Experimental Results from Tanzania. Policy Research Working Paper Series 5501. World Bank, Washington DC. http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2010/12/13/000158349_20101213140710/Rendered/PDF/WPS5501.pdf

Croke, K. Dabalen, A. Demombybes, G. et al. 2012. Collecting High Frequency Panel Data in Africa Using Mobile Phone Interviews. Policy Research Working Paper Series 6097. World Bank, Washington D.C. https://openknowledge.worldbank.org/bitstream/handle/10986/9313/WPS6097.pdf?sequence=1

Demombynes, G. 2012. Opening Up Microdata Access in Africa. World Bank Development Impact Blog. http://blogs.worldbank.org/impactevaluations/opening-up-microdata-access-in-africa

Devarajan, S. (2011). Africa's statistical tragedy. Africa Can...end Poverty. Retrieved from http://blogs.worldbank.org/africacan/africa-s-statistical-tragedy

Carletto, G. Jolliffe, D. Banerjee, R. 2015. From Tragedy to Renaissance. Improving Agricultural Data for Better Policies. Policy Research Working Paper 7150. World Bank. http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2015/01/06/000158349_20150106100529/Rendered/PDF/WPS7150.pdf

World Bank and WHO. 2014. Global Civil Registration and Vital Statistics Scaling Up Investment Plan 2015-2024. Working Paper 88351. http://www.who.int/healthinfo/civil_registration/WB-WHO_ScalingUp_InvestmentPlan_2015_2024.pdf

Kumar, R. 2014. How Youth Saved Bananas in Uganda. The World Bank . http://blogs.worldbank.org/youthink/how-youth-saved-bananas-uganda

WORLD ECONOMIC FORUM

World Economic Forum. 2015. Data-Driven Development Pathways for Progress. http://www3.weforum.org/docs/WEFUSA_DataDrivenDevelopment_Report2015.pdf

World Economic Forum. 2012. Big Data, Big Impact: New Possibilities for International Development. http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf

OTHER RESOURCES

Alvarez, G. D. Tran, M. and Raina, R. 2011. MDGs: Sub-Saharan Africa: Overcoming Data Gaps and Ranking Progress. Department of Economics Discussion Paper. Fordham University. Bronx. NY. http://stage.web.fordham.edu/images/academics/graduate_schools/gsas/economics/dp2011_01_alvarez_tran_raina.pdf

Anderson, C. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. Wired magazine. http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory

Béné C, Chijere ADG, Allison EH and Snyder K. 2012. Summary report: Uganda and Malawi field pilots of proposed LSMS Fisheries Module. http://pubs.worldfishcenter.org/resource_centre/WF_3059.pdf

Bill and Melinda Gates Foundation. 2013. Using Mobile Data for Development. http://www.cartesian.com/wp_content/upload/Using-Mobile-Data-for-Development.pdf

Bollier, D. 2010. The Promise and Peril of Big Data, The Aspen Institute. Washington DC. http://citpsite.s3.amazonaws.com/events/big-data/Aspen-Big_Data.pdf

Brynjolfsson, E. Hitt, L. M. and Kim, H. H. 2011. Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? Available at SSRN http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486

Caribbean ICT Research Programme. 2014. MFisheries Background. MFisheries. http://www.cirp.org.tt/mfisheries/index.php/about/background

Carletto C, Jolliffe D and Banerjee R. 2013. The Emperor has no Data! Agricultural Statistics in Sub-Saharan Africa. Conference paper - African economic development: Measuring success and failure, Vancouver, Canada, 18-20 April 2013 http://mortenjerven.com/wp-content/uploads/2013/04/Panel-3-Carletto.pdf

Cavoukian, A. and Jonas, J. 2012. Privacy by Design in the Age of Big Data, Privacy by Design. https://privacybydesign.ca/content/uploads/2012/06/pbd-big_data.pdf

Centre for Economics and Business Research Ltd (Cebr). 2012. Data equity: Unlocking the value of big data. Report for SAS. http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf

Chang, F. Dean, J. Ghemawat, S. Hsieh, W. C. Wallach, D. A. Burrows, M. Chandra, T. Fikes, A. and Gruber, R. E. 2008. Bigtable: A distributed storage system for structured data. ACM Trans. Comput. Syst. 26, 2, Article 4. http://web.stanford.edu/class/cs347/reading/bigtable.pdf

Christiaensen L, Carletto D, and Barrett CB. 2013. No food security without good data or local analysts. Presentation at the Workshop Café, 1st Global Food Security Conference, 29 Sept-2 October 2013, Noordwijkerhout, Netherlands http://dyson.cornell.edu/faculty_sites/cbb2/presentations/Global%20Food%20Security%20Conf%20No%20food%20security%20without%20good%20data%20or%20local%20Sept%2030,final.pptx

Chui, M. Farrell, D. and Jackson, K. 2014. How government can promote open data. McKinsey&Company. http://www.mckinsey.com/insights/public_sector/how_government_can_promote_open_data

Cios, K. J. Pedrycz, W. Swiniarski, R. W. and Kurgan, L. A. 2010. Data Mining: A Knowledge Discovery Approach, Springer P. C. Incorp. 1st ed. http://dl.acm.org/citation.cfm?id=1204561

Cisco. 2012. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update. 2011–2016. White Paper. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

Dillon, B. 2011. Using Mobile Phones to Conduct Research in Developing Countries. Journal of International Development. Volume 24. Issue 4. http://www.edi-africa.com/docs/Dillon_MobilePhoneSurveys.pdf

Dumbill, E. 2012a. What is big data? An introduction to the big data landscape. O'Reilly Radar. http://radar.oreilly.com/2012/01/what-is-big-data.html

Dumbill, E. 2012b. Big data market survey: Hadoop solutions. O'Reilly Radar. http://radar.oreilly.com/2012/01/big-data-ecosystem.html

Edmunds, R. and Marchant, T. 2008. Official Statistics and Monitoring and Evaluation Systems in Developing Countries: Friends or Foes? Paris: Partnership in Statistics for Development in the 21st Century. http://paris21.org/sites/default/files/3638.pdf

Gartner. 2012. Gartner Says Worldwide Business Intelligence, Analytics and Performance Management Software Market Surpassed the $12 Billion Mark in 2011. Press release. www.gartner.com/it/page.jsp?id=1971516

Gartner. 2011. Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data. Press release. www.gartner.com/it/page.jsp?id=1731916

GSMA. The Mobile Economy 2014. . http://www.gsmamobileeconomy.com/GSMA_ME_Report_2014_R_NewCover.pdf

Global Open Data for Agriculture and Nutrition. 2014. "Statement of Purpose." http://www.godan.info/statement.html

ITU. 2012. Key ICT indicators for developed and developing countries and the world (totals and penetration rates).ITU. www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom.html

Jensen, R.T. 2010. Information, efficiency, and welfare in agricultural markets. Agricultural Economics 41, no. s1. 2010: 203-216. http://ageconsearch.umn.edu/bitstream/53206/2/ICTs_Agriculture_Jensen.pdf

Jerven, M. 2013. Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It. Ithaca. NY: Cornell University Press. http://www.cornellpress.cornell.edu/book/?GCOI=80140100939320

Jones, S. 2012. Why 'Big Data' is the fourth factor of production. Financial Times. http://www.ft.com/intl/cms/s/0/5086d700-504a-11e2-9b66-00144feab49a.html#axzz3RWiVn4fQ

G-8 International Conference on Open Data for Agriculture 2013, Action plans https://sites.google.com/site/g8opendataconference/action-plans

Hilbert M. Big Data for Development: From Information- to Knowledge Societies. University of California, Davis. January 15, 2013. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2205145\

Jacobs, S. 2014. Big Data Comes to Africa. This is Africa Global Perspective, 03/03/2014 retrieved 13 February 2015. http://www.thisisafricaonline.com/News/Big-Data-comes-to-Africa

Keita, N and Carfagna, E. 2010. Overview of methodological issues for research to improve agricultural statistics in developing countries. Proceeding of ICAS-V, Fifth International Conference on Agricultural Statistics, Integrating Agriculture into National Statistical Systems Kampala, Uganda 13-15 October 2010. http://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/PDF/ICASV_3.4_101_Paper_Keita.pdf

Kirkpatrick, R. 2014. Big data: big potential. Capacity4dev. http://capacity4dev.ec.europa.eu/article/big-data-big-potential#sthash.sFrWbRij.dpuf

Kshetri, N. 2014. The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns. http://bds.sagepub.com/content/spbds/1/2/2053951714564227.full.pdf

Kodila-Tedika, O. 2012. Africa's Statistical Tragedy: Best Statistics, Best Government Effectiveness. University of Kinshasa. http://mpra.ub.uni-muenchen.de/45719/8/MPRA_paper_45719.pdf

Letouzé, E. (2012). Big Data for Development: Opportunities and Challenges (White p). New York: United Nations Global Pulse. http://www.unglobalpulse.org/projects/BigDataforDevelopment

Long, J. and Brindley, W. 2013. Insights into the role of technology in addressing development challenges: The role of big data and analytics in the developing world. Accenture Development Partnerships. http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf

Loukides, M. 2010. What is data science? The future belongs to the companies and people that turn data into products. O'Reilly Radar. http://radar.oreilly.com/2010/06/what-is-data-science.html

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey & Company. http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation

Mittal, S. and Gaurav T. 2009. Role of mobile phone technology in improving small farm productivity. Agricultural Economics Research Review 22. http://ageconsearch.umn.edu/bitstream/57502/2/15-S-Mittal.pdf

McKinsey Global Institute (MGI). 2013. Lions go digital: The Internet's transformative potential in Africa. http://www.mckinsey.com/insights/high_tech_telecoms_internet/lions_go_digital_the_internets_transformative_potential_in_africa

McKinsey Global Institute(MGI). 2011. Big data: The next frontier for innovation, competition and productivity. McKinsey & Company. www.mckinsey.com/~/media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx

de Montjoye, Y-A. Kendall, J. and Kerry, C. F. 2014. Enabling Humanitarian Use of Mobile Phone Data. Brookings Institute. http://www.brookings.edu/~/media/research/files/papers/2014/11/12%20enabling%20humanitarian%20mobile%20phone%20data/brookingstechmobilephonedataweb.pdf

Nyirenda-Jere TPR and Kazembe JA. 2014. Improving Policymaking for Agricultural and Rural Development in Africa: The role of ICTs and Knowledge Management. IIED, IDS, ODI. http://pubs.iied.org/pdfs/14636IIED.pdf

Open Data Institute (ODI). 2013. Open Data Barometer: 2013 Global Report. London. www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf

Partnership in Statistics for Development in the 21st Century (PARIS21). 2012. Partner Report on Support to Statistics: PRESS. Paris. http://paris21.org/sites/default/files/PRESS2012-fullreport-WEB.pdf

Pentland, A. 2013. The Data Driven Society. Scientific American. http://www.nature.com/scientificamerican/journal/v309/n4/full/scientificamerican1013-78.html

Peres, W., & Hilbert, M. (2010). Information Societies in Latin America and the Caribbean Development of Technologies and Technologies for Development. Santiago: United Nations ECLAC. http://www.cepal.org/publicaciones/xml/3/43803/Libro_Cepal_98.pdf

Shuang, C. Fonteneau, F. Jutting, J. and Klasen, S. 2013. Towards a Post-2015 Framework That Counts: Developing National Statistical Capacity. PARIS21 Discussion Paper No. 1, Partnership in Statistics for Development in the 21st Century. http://www.paris21.org/sites/default/files/PARIS21-DiscussionPaper1-MDG.pdf

Taylor, K. C. 2011. A Review of Freedom of Information, Data Protection and Open Data in the Caribbean. Lecture, Internet Governance Forum. Nairobi. Kenya. http://www.scribd.com/doc/66799430/OpenDataCaribbean220911-V1

Thakkar, M. Floretta, J. Dhar, D. Wilmink, N. Sen, S. Niall Keleher, Michelle McConnaughay, Shaughnessy, L. 2013. Mobile Based Technology for Monitoring & Evaluation: A Reference Guide for Project Managers, M & E Specialists, Researchers and Donors. http://www.theclearinitiative.org/mobile-based-tech.pdf

Villars, R. Olofson. C. and Eastwood, M. 2012. Big Data: What It Is and Why You Should Care. White Paper. IDC. http://sites.amd.com/us/Documents/IDC_AMD_Big_Data_Whitepaper.pdf

Waverman, L. Meschi, M. and Fuss, M. 2005. The impact of telecoms on economic growth in developing countries. The Vodafone policy paper series 2. no. 03. 2005.10-24. http://www.ictregulationtoolkit.org/en/toolkit/docs/Document/3532

Woolfrey, L. 2013. Leveraging Data in African Countries: Curating Government Microdata for Research. DataFirst Technical Paper 22. University of Cape Town. http://blog.narotama.ac.id/wp-content/uploads/2014/11/DataFirst-Technical-Papers.pdf

Zinow, R. 2012. Big Data, Mobile and Cloud combined – the new paradigm shift?. SAP. Presentation at the ECD Conference. http://www.ecd-conference.de/wp-content/blogs.dir/30/files/2011/03/Zinow_SAP_1545_2.pdf

WEBSITES

AfDB

http://www.afdb.org/en

Open Data for Africa
www.opendataforafrica.org

Agrihack Talent
http://hackathon.ict4ag.org

Bill & Melinda Gates
http://www.gatesfoundation.org/

European Commission
https://ec.europa.eu/

https://ec.europa.eu/digital-agenda/en/towards-thriving-data-driven-economy

CIARD Ring
http://ring.ciard.net

CISCO
http://www.cisco.com/c/en/us/index.html

CGIAR Consortium Data Management System
http://www.cgiar.org/resources/open/data-managementsystem

ckan
www.ckan.org

Climate.Gov
https://www.climate.gov

CTA
http://www.cta.int/en

ICT Update: http://ictupdate.cta.int/

PGIS: http://pgis.cta.int/en/

European Commission DG Connect
http://ec.europa.eu/dgs/connect/en/content/dg-connect

Eurostat

http://ec.europa.eu/eurostat/help/new-eurostat-website

E-agri
http://www.e-agri.info/

Digital Agenda for Europe
http://ec.europa.eu/digital-agenda/

EU Open Data Portal
https://open-data.europa.eu/en/data

FAO Data
http://data.fao.org/

Finovaregio
http://www.finnovaregio.eu/

Global Viral Forecasting
www.gvfi.org

GODAN
http://www.godan.info/

GSMA Mobile Economy
http://www.gsmamobileeconomy.com

Harvest API
www.harvestapi.io

IBM Big data
http://www-01.ibm.com/software/data/bigdata/

International Telecommunication Union (ITU)
http://www.itu.int/en/Pages/default.aspx

Jamaica Agricultural Marketing Information System (JAMIS)
http://www.ja-mis.com/CompanionSite/home.aspx

Land Portal
http://landportal.info/

Open Definition
http://opendefinition.org/

OECD
http://www.oecd.org/

Open Knowledge Foundation
http://okfn.be / https://okfn.org

NOAA Caribbean Tagged Datasets on Data.Gov
http://catalog.data.gov/organization/noaa-gov?q=Caribbean

Regional Centres for Learning on Evaluation and Results

http://www.theclearinitiative.org/

The Caribbean Open Institute
http://data.caribbeanopeninstitute.org/

Trinidad and Tobago National Agricultural Market Information System (NAMIS)
http://www.namistt.com/

United Nations Global Pulse
www.unglobalpulse.org

United Nations data Revolution
http://www.undatarevolution.org/

http://data.un.org
Ushahidi
http://www.ushahidi.com/

Ushahidi SwiftRiver Platform
http://www.ushahidi.com/contribute/swiftriver/

World Economic Forum Mobile Financial Services Development Report
www.weforum.org\mfs

World Bank Open Data Portal
http://data.worldbank.org/indicator

ENDNOTES

1 The Reader and most of the resources are available at http://brusselsbriefings.net

2 Gero Carletto, Dean Jolliffe, Raka Banerjee. From Tragedy to Renaissance. Improving Agricultural Data for Better Policies. Policy Research Working Paper 7150. World Bank,. January 2015.

3 OCDE. Exploring data-driven innovation as a new source of growth. Mapping the Policy Issues Raised by "Big Data". June 2013

4 http://capacity4dev.ec.europa.eu/article/big-data-big-potential#sthash.sFrWbRij.dpuf. This includes tweets, public Facebook posts, geotags that locate where photos were taken and news stories. It can also include de-identified records of mobile phone activity, anonymised internet search data, consumers' purchasing trends, and soon perhaps, transcripts of talk-back radio. In the near future for those seeking to understand demographic change, official government statistics will be just one component of a much richer picture.
http://capacity4dev.ec.europa.eu/article/big-data-big-potential#sthash.sFrWbRij.dpuf

5 UN Global Pulse report. http://unglobalpulse.org/

6 A detailed glossary is provided as Annex.

7 The word "crowdsourcing" refers to the use of non-official actors ("the crowd") as (free) sources of information, knowledge and services, in reference and opposition to the commercial practice of outsourcing. "

8 UN Global Pulse report.

9 OCDE. Exploring data-driven innovation as a new source of growth. Mapping the Policy Issues Raised by "Big Data". June 2013

10 European Commission (2010), "A Digital Agenda for Europe", Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2010)245 final, 19 May 2010.

Review of recent studies on PSI re-use and related market developments, G. Vickery, August 2011.

11 ICTupdate issue 79 February 2015

12 'Global Strategy To Improve Agricultural And Rural Statistics Economic And Sector Work' United Nations, Food and Agriculture Organisation and World Bank. Report No. 56719-Glb

13 ibid

14 G-8 conference on Open Data for Agricture, April 29-30 2013. Action plans https://sites.google.com/site/g8opendataconference/action-plans

15 World Economic Forum. Big Data, Big Impact: New Possibilities for International Development. 2012

16 OCDE. Exploring data-driven innovation as a new source of growth. Mapping the Policy Issues Raised by "Big Data". June 2013

17 Keita, N and Carfagna, E. 2010. Overview of methodological issues for research to improve agricultural statistics in developing countries. Proceeding of ICAS-V, Fifth International Conference on Agricultural Statistics, Integrating Agriculture into National Statistical Systems Kampala, Uganda 13-15 October 2010.

18 LSMS-ISA "Improving Household Survey Data on Agriculture in Sub-Saharan Africa";

"No food security without good data or local analysts" Luc Christiaensen and Gero Carletto, World Bank, and Christopher B. Barrett, Cornell University. Workshop Café, 1st Global Food Security Conference, 29 Sept-2 October 2013, Noordwijkerhout, Netherlands

19 https://ec.europa.eu/digital-agenda/en/towards-thriving-data-driven-economy

20 Both Morocco and China already have mature ICT knowledge and earth observation expertise, while also having a large agriculture-based population. The fact that they already have these ICT knowledge bases allows for greater exploitation of European technologies. In comparison, Kenya lacks these technologies and therefore, the project application shall focus on capacity building.

21 OJ L 345/90, 31.12.2003

22 Directives 2003/4/EC, OJ L 41/26, 14.2.03 and 2007/2/EC, OJ L108/1, 14.3.2007.

23 COM(2010) 461 final.

24 COM(2008) 886 final/2.

25 Communication on scientific information in the digital age, COM(2007) 56, Communication on ICT infrastructures for e-Science COM(2009) 108 and related Council Conclusions.

26 United Nations Secretary-General, by the Independent Expert Advisory Group on a Data Revolution for Sustainable Development. A World that Counts- Mobilising the Data Revolution for Sustainable Development, November 2014

27 Amparo Ballivian, World Bank , William Hoffman, World Economic Forum. Public-Private Partnerships for Data. Issues Paper for Data Revolution Consultation. This issues paper has been prepared for online consultation and will be discussed at a meeting on financing key priorities of the data revolution in New York on January 22 2015. http://data.worldbank.org/sites/default/files/issue-paper-financing-the-data-revolution-ppps_0.pdf

28 World Economic Forum. Big Data, Big Impact: New Possibilities for International Development. 2012.

29 OCDE. Exploring data-driven innovation as a new source of growth. Mapping the Policy Issues Raised by "Big Data". June 2013

30 United Nations Secretary-General, by the Independent Expert Advisory Group on a Data Revolution for Sustainable Development. A World that Counts- Mobilising the Data Revolution for Sustainable Development, November 2014.

31 McKinsey Global Institute (November 2013). "Lions go digital: The Internet's transformative potential in Africa." http://www.mckinsey.com/insights/high_tech_telecoms_internet/lions_go_digital_the_internets_transformative_potential_in_africa

32 Kshetri, N. 2013

33 Kshetri, N. 2013.

34 'Big Data Comes to Africa', Sherellee Jacobs, This is Africa Global Perspective, 03/03/2014 http://www.thisisafricaonline.com/News/Big-Data-comes-to-Africa

35 ibid

36 Kshetri, N. 2013

37 UNICEF. http://www.unicef.org/infobycountry/uganda_70090.html

38 PARIS21. http://www.cgdev.org/blog/better-data-rwanda

39 Kumar, R. (2014) "How Youth Saved Bananas in Uganda". http://blogs.worldbank.org/youthink/how-youth-saved-bananas-uganda

40 World Economic Forum. Data-Driven Development Pathways for Progress. January 2015.

41 A report of the Independent External Evaluation of the Food and Agriculture Organisation concluded that "the quantity and quality of data coming from national official sources has been on a steady decline since the early 1980s, particularly in Africa" (FAO, 2006).

Brussels Rural Development Briefings

A series of meetings on ACP-EU development issues

Along with our partners in this joint initiative, the European Commission (DG DEVCO), the ACP Secretariat and ACP Group of Ambassadors, CONCORD and various media, we at CTA look forward to welcoming you at our next Brussels Development Briefing.

Never miss the latest daily news on key ACP-EU programmes and events in Brussels related to agriculture and rural development with our Blog brussels.cta.int

For more information Email: brussels.briefings@cta.int Tel: + 32 (0) 2 513 74 36

www.brusselsbriefings.net