12th World Telecommunication/ICT Indicators Symposium (Tbilisi, 2014)

Background Documents

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/1-E
18 November 2014

English

SOURCE: Ministry of Communication and Information Technology, Indonesia

TITLE: ICT Indicators and Prioritization Strategic Development on ICT Sectors

ICT INDICATORS AND PRIORITIZATION STRATEGIC DEVELOPMENT ON ICT SECTORS

Basuki Yusuf Iskandar[1] and Hedi Muhammad Idris[2]
ICT R&D and Human Resources Development Agency
Ministry of Communication and Information Technology, Republic of Indonesia
basu@kominfo.go.id and hedi002@kominfo.go.id

Introduction

Government of Indonesia and other stakeholders in the country deem a common belief that ICT can provide the foundation whereby a society can be built with equitable opportunity and knowledge. We also believe that it is a key determinant of the sustainable competitive advantages of the Nation. This concept is often left undeveloped for relevant ICT design and its policy instrumentalization. GOI seeks to present the composite ICT initiatives and qualitative as well as quantitative measurement to ensure that ICT utilization is addressed in productive ways. Ministry of Communication and Infromation Technology (MCIT) of Government of Indonesia (GOI) has annually conducted ICT Surveys. For example we captured telecommunication market share and built analysis based on ICT household indicators.

ICTs should be viewed as a strategic national resource which has implications on the sustainability of the state and nation. The consequence of this, ICT must be managed as a resource or treasury capital toward national competitiveness. As a resource, ICT is unique because it holds the potential benefits and at the same time it relates the potential risks that need to be mitigated while securing information security known comprehensively as the concept of the information resilience. In the context of global competition, ICT has determinedly become a strategic necessity and not merely a strategic choice anymore. ICT should be addressed strategically due to its enormous and long-term impact covering almost all development sectors correspondingly. So it is unquestionably imperative to set ICT as the blood vessels of this Nation, the content as the blood flow entirely of the Nation, the infrastructure and political will as the heart of the Nation. Otherwise, ICT development and competition will lead us to disadvantage position. Re-intrepretation from Trisakti principles from Indonesia founding father, Soekarno, with regard to ICT development,.ICT, therefore, is to be developed to create national souvergnity by inducing productive and competitive economy, a dynamic and national resilience, and the patriotic, innovative, and superior character nation by shielding our virtous cultural roots.

GOI is working hard to combine our policy work with research and analysis, providing concrete evidence of the outcome and benefits of ICT. As technology became more pervasive in

the workplace, education, entertainment, creative industry and daily living, MCIT's research and development focused on design of information and communication technologies, so that they are more accessible and usable by policymaker in various domains. MCIT have two sort of R&Ds. The first one is policy research aimed at proactive analysis, policy impact evaluation, and modelling. Another side we are also grappling applied R&D to produce prototype, innovative product and supporting the business making analysis of the product commercialization. Through this approach, MCIT provides national leadership by developing position papers, hosting meetings, and engaging in dialogue with researchers, practitioners, and policymakers in order to identify promising areas of research, development, and dissemination for the field and to advance evidence-based policy and practice, as well as contribute to the development of ICT-related protoype product and its industrial penetration.

The number of ICT users either individually or in any sector like household, business, education, etc., cause every country needs to observe, evaluate and monitor the development of ICT infrastructure, access and use. All the data and information are needed to build a proper policy to regulate access and ICT usage and so does infrastructure development. In this case, the description of the access and use of ICT becomes indispensable by the countries. ICT measurement standards are aiming to obtain a picture of the access progress , as well as the usage of ICT infrastructure in each country is also useful to benchmark position of ICT developments in one country against another country.

ICT Household Indicators

In 2013, the Ministry of Communications and Information Technology Republic of Indonesia conducted a survey of ICT access and usage in household sector. Survey of the use of ICT access in households nationwide conducted by the MCIT. Household population used as the sampling method is based on BPS (Indonesia National Statistic Agency) census. While the sample in this survey totaled 9,680 households with a 95% confidence level and a estimation margin of error 1% . Survey of ICT access and usage in households based on 12 core ICT indicators compiled by the ITU (International Telecommunication Union International) in 2012.

The sampling methods in of ICT access and usage in household sector survey can be seen in figure 1 below.

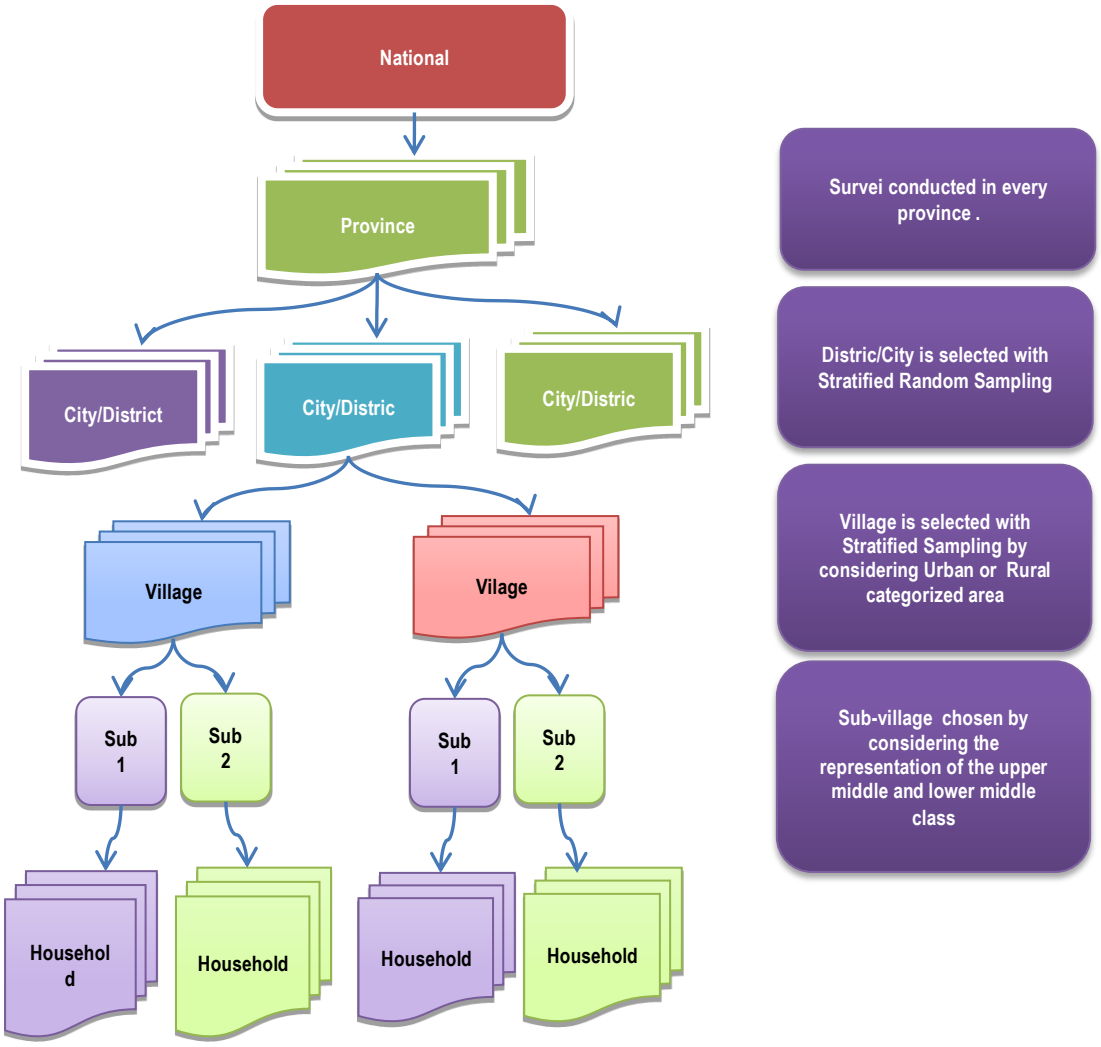


Figure 1. Sampling Methods

From the sampling frame, the distribution of the respondents in this survey is based on the region are: 32% of respondents from Sumatra, 28% of the area of Java, 18% from Sulawesi region, 9% from Bali and Nusa Tenggara region, 8% respondent from the Moluccas and Papua and 5% from Borneo.

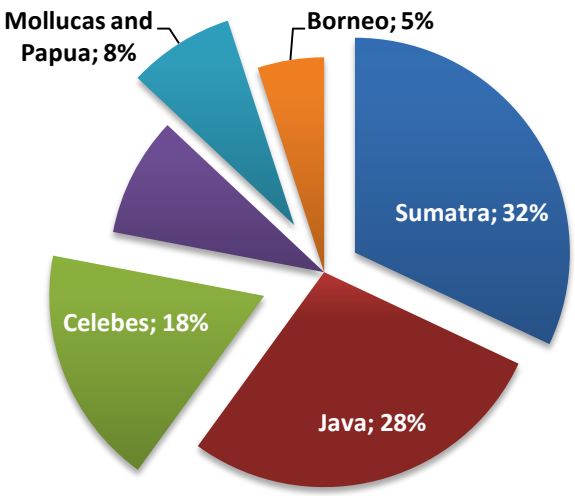


Figure 2. Sample Distribution by Region

HH1 Proportion of Households With a Radio

Radio is one of the simple and common ICT tools. Definition of radio here, is a conventional radio or radio integrated in the vehicle, an alarm clock, or mp3 player, but does not include radio on the computer and HP. In addition, owned radio must be in good condition and functioning properly. Based on survey results in Indonesia, 35% of the surveyed households have a radio, while 65% of households do not have a radio.

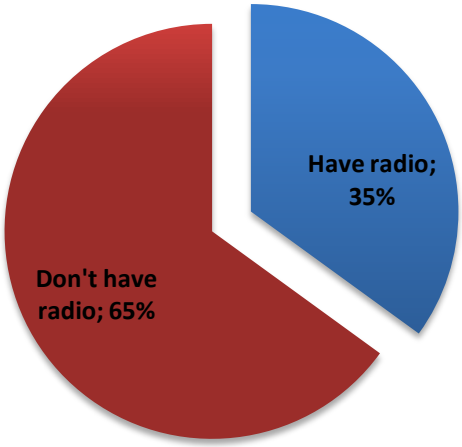


Figure 3. Proportion of households with a radio

HH2 Proportion of Households With a Television

Television is the most popular ICT devices. Currently, almost all of the home is certain to have a television. Based on the survey in Indonesia, there are 10% of households that do not have a television, and 90% of households have a television. Television sets are referred to in this survey is the conventional television and not the television that include in another device (mobile phone / mobile) or in a vehicle. In addition, household owned television must be in good condition and functioning properly.

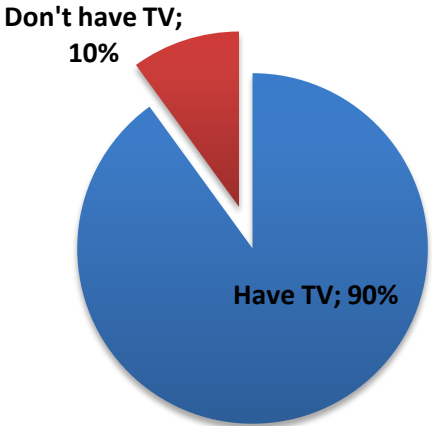


Figure 4. Proportion of households with a television

HH3 Proportion of Households With a Telephone

Fixed Telephone network in this survey is refer to a line connecting the subscriber's terminal equipment (telephone) to the public switched telephone network (PSTN) and has a dedicated port on the phone. Based on the survey, only 9.31% of households with telephone wires, while 90.69% of households do not have a fixed phone network. This data based on survey is greater than the number percentage of households with fixed telephone based on data collected by Village Indonesia National Statistic Agency PODES survey in 2012 with 8.6%.

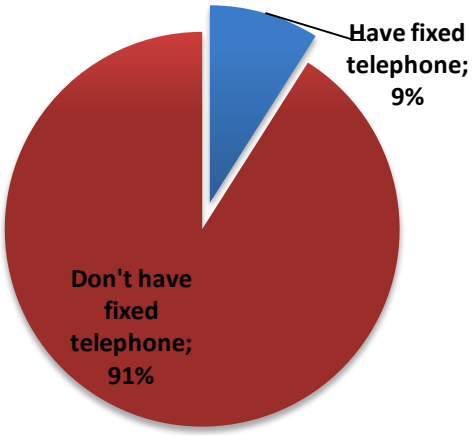


Figure 5. Proportion of households with telephone

HH4 Proportion of Households With a Computer

Nowadays a computer is a modern multi function ICT devices. First computer more useful to the calculating process and office activities, but development of computer technology also serves communication, learning media, games and others. In this survey, Computing devices are referred to in this survey is a desktop computer or laptop and not a mobile phone, PDA (Personal Digital Assistant) or TV. In addition, the computer also must be in good condition and used properly according to its function. Based on survey results, 27% of households having a computer and 73% of households do not own a computer.

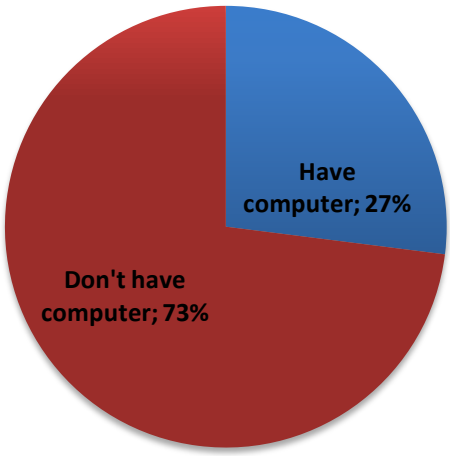


Figure 6. Proportion of households with a computer

HH5 Proportion of Individuals Who Used a Computer In the Past 3 Months

This survey also measures computers usage by individuals as a household member. The computer is defined as desktop computer or a laptop and does not include mobile phones, PDAs (Personal Digital Assistant). And also, computer must be in good condition and functioning. Computer usage in this survey refer to individual user (respondent) that use computer at least one time within the last 3 months, even though the respondent do not own computer by himself/herself. They can borrowed computer from family, friends and others. Based on survey results, 26% of respondents had used a computer for the last 3 months. While 74% of respondents do not use a computer.

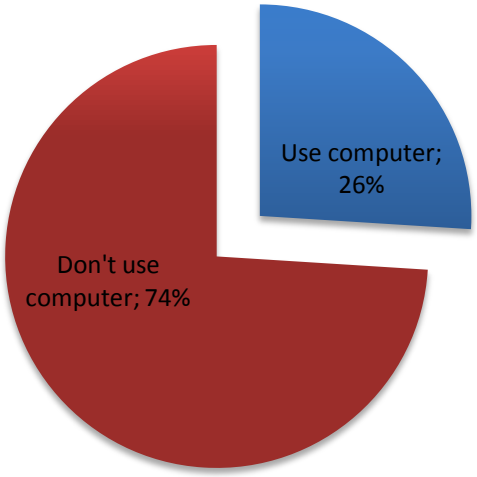


Figure 7. Proportion of individuals who used a computer in the past 3 months

HH6 Proportion of Households With Internet Access

The Internet is a worldwide public computer network, providing access to a number of communication services including the World Wide Web and run e-mail, news, entertainment, and data. Internet in this survey meant that the internet is accessible not only through a computer, but also can with cell phones, PDAs, game machines, digital television, etc.. Based on the survey results it appears that 21% of households have internet access, while 79% of households do not have internet access.

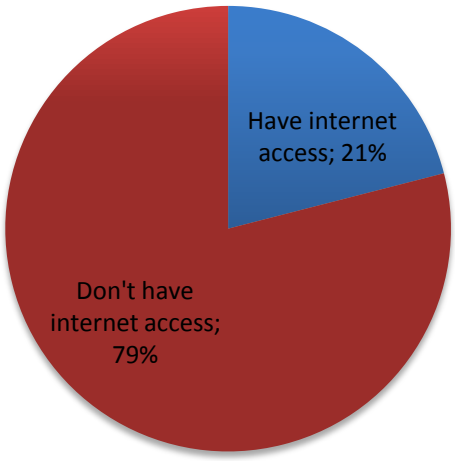


Figure 8. Proportion of households with internet access

HH7 Proportion of Individuals Who Used the Internet In The Past 3 Months

This survey also measures internet usage by individuals in Indonesia. Internet user refer to individual user (respondent) that use internet access at least once within the last 3 months, even though the respondent do not own internet network by himself/herself. They can use internet access from internet cafe, friend's home, office, school and others location. Based on survey results, 74% of respondents use internet access in the last 3 months. While 26% of respondents do not use internet.

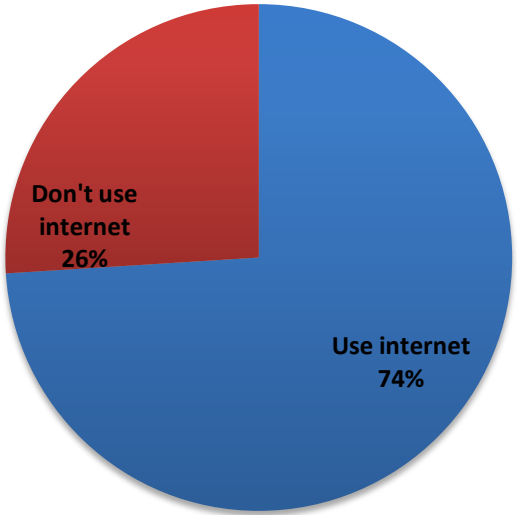


Figure 9. Proportion of individuals who used the Internet in the past 3 months

HH8 Location of Individual Use of the Internet In The Past Month

Each individual accessing the internet in different places. From the survey results, it can be seen the percentage of locations that are often used for accessing internet. Home has the highest percentage of location that respondents usually use the Internet access with 54.1%. Second highest percentage is to access internet from anywhere via HP at 47.9% i. While the internet cafe has the third highest percentage with 32.6%.

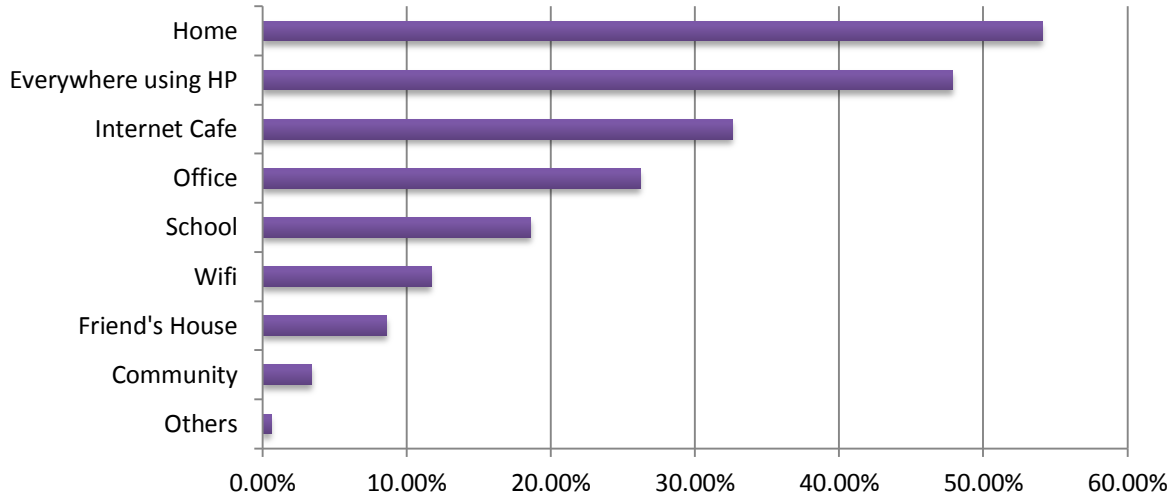


Figure 10. Location of individual use of the Internet in the past month

HH9 Internet Activities Undertaken By Individuals In The Past Month

In Indonesia, respondent have so many activities through internet. Opening a social networking site is the most commonly performed activities for respondents while using Internet access with a percentage of 67.9%. Then proceed with searching for information about goods or services by 42%, and then, learning activity is third highest activities by 40.1%.

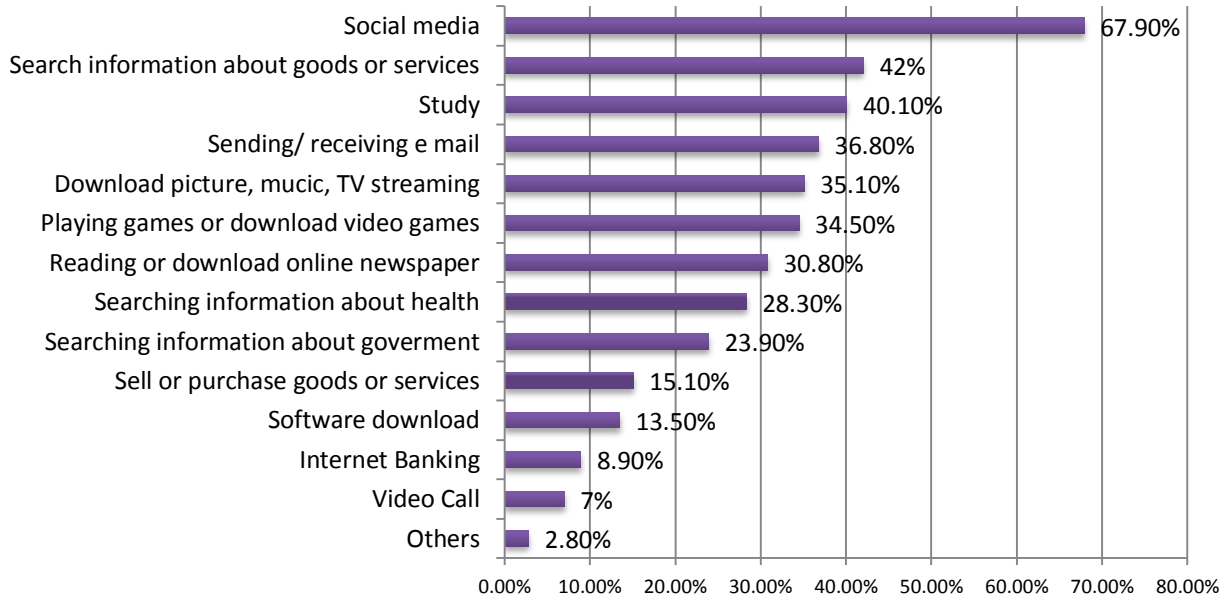


Figure 11. Internet activities undertaken by individuals in the past month

HH10 Proportion of Individuals Who Used a Mobile Cellular Telephone In The Past 3 Months

The survey ICT access and usage also measures mobile phone (Hand Phone) user as an individual. The individual must use a mobile phone / mobile at least 1 time in the last 3 months, even though the respondent do not own HP by himself/herself. they can borrowed from family, friends and others. Based on the survey results it appears that 80% of respondents admitted to using a mobile phone / mobile in the last 3 months, while 20% are not.

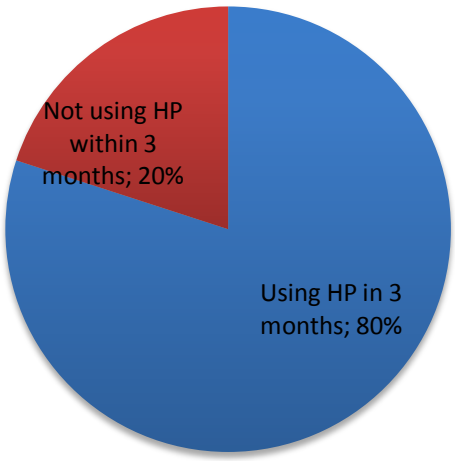


Figure 12. Proportion of individuals who used a mobile cellular telephone in the past 3 months

HH11 Proportion of Households With Access to the Internet by Type of Access

There are three types of Internet access that can be used in the Indonesia households, Narrowband (such as Telkomnet Instant), Fixed Broadband (Telkom speedy), and Mobile Broadband with internet service from operators like Indosat,Telkomsel, Smartfren, XL and another internet product with 3G networks. Based on the results of the survey can be seen that 65% of households choose Mobile broadband as a used Internet access at home, while 17% choose Fixed Narrowband and 18% of households have Fixed Broadband access.

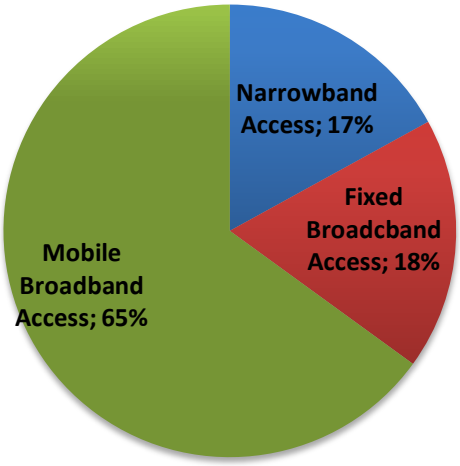


Figure 13. Proportion of households with access to the Internet by type of access

HH12 Frequency of Individual Use of the Internet In the Past 3 Months

Frequency of internet access for individuals in Indonesia based on survey datas, 51% of respondents stated that they access the internet at least once a week but not every day. Then 37% of respondents access the internet every day and only 12% of respondents who access the Internet less than once a week.

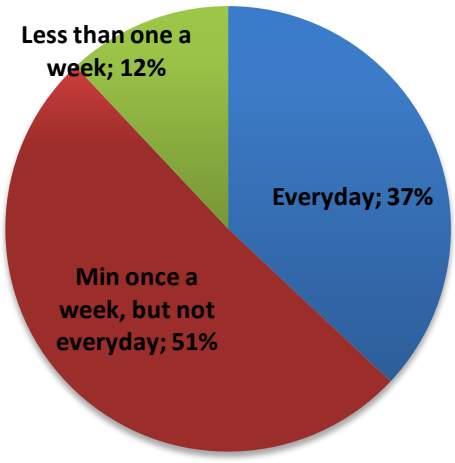


Figure 14. Frequency of individual use of the Internet in the past 3 months

Strategic Prioritization on ICT Development

GOI have working indefatigably to ensure that ICT is being transformed to best serve the interests of the national goals. In Indonesia, the rate of growth of the transport and communications sector in 2013 is still the highest among other sub-sectors, which reached 11.1%. Compared to, for example, the trade, hotels and restaurants which are only contribute 8.3% and construction of 7.3%.[1] The growth rate of the communications subsector was driven by data and internet sales factor, in addition to the increasing number of purchase of new communication tools such as mobile smartphones, tablet PCs, and other telecommunications equipment. How strategic ICT role in national level? A study of the National Development Planning Agency of GOI indicates that: 1) Every 10% increase in broadband penetration would increase economic growth by 0.8% -0.9%; and 2) Each 1% increase in broadband penetration of households, the growth of unemployment will be reduced by 8.6%; 3) The development of mobile broadband in the 700 MHz band are expected to increase productivity by 0.4% in the service industry and 0.2% in manufacturing with a total addition of the work-field creation for as many as 327,000 jobs. Supported by broadband, it is expected in 2025, Indonesia's GDP reached: US$ 4.0 to 4.5 billion.

As issued by National Development Planning Agency of GOI, there are still some critical problems in regard of ICT development. Based on the continous surveys, including ICT Household Indicators, it is important to prioritize our resources constraint to see what area of ICT development that need get more improvement. : 1) the unequal diffusion of information access (Yates, et al., 2010; Joseph & Nath, 2012). Due to a number of factors in Indonesia, including some of its islands which geographically remote and uneven distribution of population, affordable access and connectivity is the primary problem complicating basic effort to access information; 2) the shortage of ICT infrastructure allowing broadband access and the slowness of the establishment of the network infrastructure, the training of people to use it and to exploit commercially the information and knowledge that it makes available. Until mid-2014, only 361 districts/cities, or about 72% of the total 497 Indonesia’s districts/ cities has been reached by the national broadband fiber optic network. Most of the areas are in western Indonesia; 3) the unfavorably high price for broadband connection discouraging national broadband ecosystem. Based on data from the Indonesia’s Central Bureau of Statistics, the national income per capita in 2012 was around IDR 2,570,000 per month. As for the connection price of 1 Mbps is IDR 700,000, or equivalent to 27% of income per month. This figure is much higher than the amount of the average expenditure for the household's basic needs such as for education and health. The high price impedes the potential market arisen from broadband

[1] The 2014 Financial Report of Ministry of Finance of GOI

ecosystem; 4) The on-going coordination problems intra and extra-governmental bodies, not to mention the lack of an integrative policy framework from central and local government; 5) the issue of interoperability has become more important as countries deploy e-government systems. In order to provide high-quality services to citizens it is important that services can be accessed from the widest possible range of technology from all over Indonesia.

A major feature of the knowledge-based economy is the impact that ICTs have had on industrial structure, with a rapid growth of services and a relative decline of manufacturing (Houghton, 2009). GOI built an orderly systematic thinking that the ICT sector is a multistakeholder platform portfolio regulated by Government to reach national goals. As it implies numerous dimensions, we are dealing with multifaceted trends of technology, business, and industry (Tardiff, 2006; Shahid & Shoulian, 2007). The comprehension of trends has helped the identification of opportunity seizing and threat avoidance.

Conclussion and Implications

Until now, there has been little research on the practice of policy instruments of Government improving ICT enabled/enabling industry. Most literature about ICT is focused on business level and the impact of ICT expenditure to firm performance. Government intervention research implies contradictory findings. Some research proves that less government intervention is found to be beneficial for firms by allowing them to improve their performance (Wang & Cu, 2011), another research concludes that policy influence is negative and has no impact on particular area of improvement which is more easily affected by external variables (Tu, et al., 2014). For decades, a controversy has been raging about the case for Government interventions to accelerate economic growth in developing countries, as opposed to conducting international standard policies. This paper does not stand on any side of this debate. However, especially, in emerging economy such as Indonesia, there is a public pressure in favor of government interventions to promote a more healthy and mature ICT industry by protecting national goals including our SMEs growth.

From government's perspective, ICT development is the result of a specific form of dominant design which integrates national and global knowledge and trend of technology, business, and industry for the achievement of national goals. Until the beginning of the year 1990s, ICT plays a role as an enabler in which consumers have the option to utilize ICT to gain their own benefits. However, ICT has now moved to a more instrumental role as a driver of competitive advantage for both profit and nonprofit organizations. Regrettably, the phase of policy making do not just often overlap, they are also very late results in a loss of momentum.

This can lead to poorly conceived policies and government doesn't do enough programs to address the arisen problems in time. Moreover, such process are impeded by bureacratic inertia, a lack of appropriate institutional arrangements, and problems in public-civil service relationships. Government need better ways of ensuring that the policy problem has been fully considered and the instrument solution examined properly.

References

Acemoglu, D., Johnson, S. & Robinson, J., 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review,* Volume 91, pp. 13-69.

Al-Jaghoub, S. & Westrup, C., 2003. Jordan and ICT-led Development: toward a Competition State. *Information Technology & People,* 16(1), pp. 93-110.

Barney, J. & Hesterly, W., 2008. *Strategic Management and Competitive Advantage - Concept and Cases.* II penyunt. New Jersey: Pearson Prentice Hall..

Contini, F. & Cordella, A., 2007. Information System and Information Infrastructure Deployment: the Challenge of the Italian e-Justice Approach. *Electronic Journal of e-Government ,* 5(1), pp. 43-52.

Cortese, A., 2003. *The Critical Role of Higher Education in Creating a Sustainable Future in Planning for Higher Education.* [Online]Available at:http://www.aashe.org/resources/pdf/Cortese_PHE.pdf
[Diakses 14 September 2014].

Economist, T., 2012. *The Wealth of Nations*.:8 September 2012.

Fenton, F. & Sipes, J., 1996. Architectural and Technological Trends in Access: An Overview. *Bell Labs Technical Journal,* Issue Summer, pp. 1-20.

Finke, D. & König, T., 2009. Why Risk Popular Ratification Failure? A Comparative Analysis of the Choice of the Ratification Instrument in the 25 Member States of the EU. *Constitutional Political Economy,* 20(3-4), pp. 341-365.

Ghosh, R., 2006. *Economic impact of open source software on innovation and the competitiveness of the Information and Communication Technologies (ICT) sector in the EU,* The European Communities : MERIT.

Gruescu, R., Nanu, R. & Tanasie, A., 2009. Human Resources Development and ICT Contribution to the Tourist Destination Competitiveness. *European Research Studies,* 10(4), pp. 87-100.

Hafner-Burton, E. & Tsutsui, K., 2005. Human Rights in a Globalizing World: The Paradox of Empty Promises. *American Journal of Sociology,* 110(5), pp. 1372-1411.

Harris, I., Wang, Y. & Wang, H., 2014. ICT in Multimodal Transport and Technological Trends: Unleashing Potential for the Future. *Journal of Production Economics,* Volume September.

Heeks, R., 2006. Using Competitive Advantage Theory to Analyza IT Sectors in Developing Countries: A Software Industry Case Analysis. *MIT Information Technologies and International Development,* 3(3), pp. 5-34.

Houghton, J., 2009. *ICT and the Environment in Developing Countries: Opportunities and Developments.* Paris, OECD .

Joseph, I. & Nath, A., 2012. *Digital Divide: A comprehensive approach to Integrate State and community experiences into ICT Policy Framework from Indian perspective.* Trivandrum, IEEE.

Lou, H., Lu, F. & Lu, Y., 2010. *Regression model in Selecting Network Information Technology Companies and Analyzing Their Business Development Trends.* s.l., IEEE - First International Conference on Networking and Distributed Computing.

McMasters, M., Pitts, S., Prien, K. & Kamery, R., 2003. *Positive and Negative Aspects of International Trade on US Economy and Labor.* Las Vegas, Allied Academies International Conference.

Mihalic, T. & Buhalis, D., 2013. ICT as a New Competitive Advantage Factor - Case of Small Transitional Hotel Sector. *Economic and Business Review,* 15(1), pp. 33-56.

Narayan, V., 2001. *Managing Technology and Innovation for Competitive Advantage.* New Jersey: Prentice Hall.

Podobnik, B., Horvatic, D., Kenett, D. & Stanley, H., 2012. *The Competitiveness versus the Wealth of a Country,* s.l.: Scientific Reports.

Porter, A. & Detampel, M., 1995. Technology Opportunities Analysis. *Technological Forecasting & Social Change,* 49(3), pp. 237-255.

Porter, M. & Rivkin, J., 2012. *Harvard Business School Survey on Competitiveness,* s.l.: Harvard Business School.

Potočnik, J. & Verheugen, G., 2007. *Improving Knowledge Transfer between Research Institution and Industry accross Europe,* European Comission: Belgium.

Saha, S., 2013. *Emerging Business Trends in the Semiconductor Industry.* CA-USA, IEEE-Proceeding of PICMET: Technology Management for Emerging Technologies.

Shahid, M. & Shoulian, T., 2007. Convergence and Technological Innovations in ICT Industry: Trends and Strategies for Operators. *IEEE,* pp. 6456-6559.

Tardiff, T., 2006. *Changes in Industry Structure and Technological Convergence: Implications for Competition Policy and Regulation in Telecommunications.* The 34th Telecommunications Policy Research Conference , IEEE.

Tu, F., Yu, X. & Ruan, J., 2014. Industrial Land Use Efficiency under Government Intervention: Evidence from Hangzhou China. *Habitat International,* Volume 43, pp. 1-10.

UNESCO, 2010. *Measuring R&D: Challenges Faced by Developing Countries.* Montreal: UNESCO Institute for Statistics.

Wang, C. & Cu, H., 2011. Government intervention in investment by Chinese listed companies that have diversified into tourisme. *Tourism Management,* Volume 32, pp. 1371-1380.

Yates, D., Gulati, G. & Tawileh, A., 2010. *Explaining the Global Digital Divide: The Impact of Public Policy Initiatives on Digital Opportunity and ICT Development.* Honoluu, HICSS.

Yoon, B., Park, I. & Coh, B., 2014. Exploring technological opportunities by linking technology and products: Application of morphology analysis and text Mining. *Technological Forecasting and Social Change,* Volume 86, pp. 287-303.

Yunis, M. et al., 2012. ICT Maturity as a Driver to Global Competitiveness: a National Level Analysis. *International Journal of Accounting & Information Management,* 20(3), pp. 255-281.

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/2-E
18 November 2014

English

SOURCE: Statistics Netherlands

TITLE: Innovation of tourism statistics through the use of new big data sources

Innovation of tourism statistics through the use of new big data sources

Nico Heerschap, Shirley Ortega, Alex Priem en May Offermans, *Statistics Netherlands,* Netherlands

Email: nhrp@cbs.nl

The paper was written as a contribution to the 12th Global Forum on Tourism Statistics in Prague in May 2014. The Global Forum on Tourism Statistics has been providing since 1994 a unique platform for the regular exchange of views and experiences on developments in tourism statistics. The aim is to discuss major technical issues concerning the establishment of harmonised tourism statistics in an environment that strengthens co-operation among governments, the private sector, researchers, academics, OECD and EU member and non-member countries and international organisations.

1. Introduction

In the last decade, more and more data have been becoming available for research and development. There seems to be a move from data scarcity to data abundance ("data deluge"). This is primarily the result of increasing possibilities to digitize growing volumes of data at lower costs. Not only large administrative data sources are becoming available; data are also being generated by sensors (e.g. cameras, electronic traffic loops and public transport cards), mobile telephony and radio frequency identification (RFID chips). Most of these data are collected to control operational processes and for management. However, the data are increasingly being used for purposes other than those for which they are collected, such as the analysis of consumer behaviour for marketing strategies, predicting trends, detecting fraud, and also to produce new, faster and more detailed statistics. The development of the internet is a second important factor in this respect, especially the transition from web 1.0 to 2.0 at the beginning of the century. Internet users are no longer passive recipients of content; they absorb information from the internet and in return produce their own new content which in turn is spread via the internet. This change from unilateral to multi-way communication has also been the breeding ground for the emergence of social media. This is, finally, enforced by the introduction of mobile internet supported by smartphones and other portable devices. All these activities leave behind some form of digital footprints that can be detected, measured and analysed. This kind of analysis falls under the term "big data". Although not clearly defined, big data relates to datasets which are so large and complex in structure that they cannot be processed and analysed with traditional technology, software and methodology [1, 2].

This paper presents three examples of innovative (big) data research for tourism statistics. They were carried out by Statistics Netherlands (SN) in the period 2012–2013. The first example is the use of the internet for the compilation of the population of units for the tourism accommodation statistics, which produce indicators on inbound and domestic tourism. The second example focuses on the use of log data registered by an app installed on the mobile phones of a group of respondents (smartphone measurements). And the third example consists of an analysis of mobile metadata, which are generated when a mobile phone communicates with a telecom provider, so-called Call Detail Records. In the latter case Vodafone and Mezuro supplied anonymous and aggregated data. The paper concludes with some general remarks about the effects of these new and innovative data sources for the production of tourism statistics.

2. Internet as a source for tourism statistics

The internet has become part of society's basic infrastructure. More and more activities are taking place through the internet, leaving behind digital footprints which can be detected and measured in real time. This also applies for tourism accommodations and their characteristics, such as hotels, campsites, youth hostels, bed and breakfasts etc. This information can be collected from the internet by so-called robots or crawlers. Internet robots are small software programs which trawl the internet gathering the desired information (web page content). This can be done in one of two ways. Either with a *dedicated robot,* which gathers information from a specific website, such as Booking.com, Tripadvisor.com or other specific tourism websites where tourism accommodation providers post their information. Or with *general robots* that search or scrape the internet, or part of the internet, looking for web pages which contain designated keywords, such as "hotel", "accommodation" or "B&B"; these robots are not built for or limited to specific websites.

SN experimented with both methods to compile the population of units for the tourism accommodation statistics. The internet data collection focused on the smaller accommodations. The method was also applied to find the tourism accommodation units in the Caribbean Netherlands, where tourism accommodation data were incomplete. The Chamber of Commerce was not able to provide a usable set of population units.

The main conclusions of these pilots were:

- Technically it is not very difficult to do. The robots required, especially the dedicated robots, can be built and tested within a couple of days. This also applies if robots have to be updated if websites change their layout or technical structure.

- In addition to names and addresses of tourism accommodations, these methods can also collect various other characteristics: number of rooms, prices, tourist tax, available facilities and guest review scores, but also job vacancies and Chamber of Commerce registration numbers (see Table 1). The latter can be used to link the collected internet data to other relevant statistics like turnover, investment and innovation. The internet information can also help to improve the classification of accommodations.

Table 1. Example of data collected on tourism accommodations by a dedicated internet robot (fictitious data).

Name	TypeAcc	#Rooms	#Beds	Price p/n	City	Street	Num	Tel	Star rating	Swim. pool	Free wifi	Job vacancy	Website
Hotel Amsterdam	Hotel	174	400	$125-200	Kralendijk	J. W. Boulevard	18	(+599) 720 123	4	y	y	2	www. amsterdam.com
Henk's Beach Resort	Hotel Resort	102	210	$201-324	Hato	Kaya Apeldoorn	11	(+599) 720 567	4	y	y	8	www.henk.com
Dive Hotel	Hotel	38	76	$104-117	Kralendijk	Kaya Nice	40	(+599) 722 901	2	n	y	0	www.dive.com
Casino Resort	Hotel Resort	56	112	$120-200	Kralendijk	J. W. Boulevard	25	(+599) 723 098	3	y	y	3	www.club.com
Club Apartments	Appartment	40	80	$37 - 115	Kralendijk	Kaya Gold	104	(+599) 798 654	5	y	n	0	www.villas.com
Crown Villas	Villa	24	48	$231-403	Sabadeco	Lagoen road	80	(+599) 745 210	3	y	y	0	www.crown.com
Belair Apartments	Apartement	12	24	$235	Sorobon	Lighthouse	11	(+599) 745 123	2	n	y	1	www.belair.com
Hotel Rest	Hotel	10	22	$104-117	Kralendijk	Hill side	2	(+599) 723 456	2	y	n	3	www.rest.com
Eco Villa	Villa	7	16	$85-215	Belnem	Kaya Mooi	17	(+599) 745 123	4	y	n	0	www.eco.nl
Prince guesthouse	B&B	5	10	$90-125	Kralendijk	Kaya Prince	5	(+599) 723 456	1	n	y	n.a.	www.prince.com
Casa White	Studio	3	8	$200-250	Kralendijk	Kaya Red	8	(+599) 723 789	3	y	y	1	www.white.com
Villa Rose	Villa	2	4	$650-700	Hato	Flower hill	1	(+599) 745 012	5	y	n	0	www.rose.nl
Residence Yun	Appartement	2	4	$65-115	Sabadeco	J. W. Boulevard	19	(+599) 745 098	4	y	y	0	www.yun.com
Cottage High	Appartement	1	3	$100-150	Kralendijk	Kaya Linda	3	(+599) 745 876	1	n	n	n.a.	www.hill.nl

- Dedicated robots which collect data from specific websites result in better information than general robots which trawl the “entire” Internet. The advantage of dedicated robots is that we know beforehand what kind of information is available on the selected website and how the website has structured this information. The crucial step in this respect is then to find the group of websites which covers the whole domain as best as possible. Generally speaking, these include aggregation sites, such as Booking.com or Expedia.com, or specific tourism sites for, for example, campsites, holiday homes, hotel chains or bed and breakfasts. Although general robots cover the “entire” Internet, in this case the challenge is to determine whether a collected web page contains information on an actual tourism accommodation, and not, for example a web page of a hotel construction company, investment company or a general tourism site, which also contains words like “hotel”, “accommodation” etc. The collected information is often incomplete and much more unstructured.

Table 2. Example of information problems of data collected on tourism accommodations by an internet robot (fictitious data).

Name	Typacc	#Rooms	#Beds	Price p/n	City	Street	Num	Tel
2 bedroom house in Hato	Cottage	2	4	$50-100	Hato	Kaya J. Willem Boulevard		
3 bedroom oceanfront penthouse	Apartment	3		$210-260	Kralendijk			(+599) 723 345
3 bedroom villa - ocean views	Villa	3	7		Sorobon	J. Willem Boulevard		
Affordable charming house	Villa			$87	Kralendijk	Near the Lighthouse		
House in Sabadeco	Apartment	2	4	$95-135	Sabadeco			
Spacious condo in Nawati	Studio	1	2		Nawati	J. W. Boulevard	3	Contact Richard
Beautiful renovated apartments	Apartment			$85	Kralendijk			(+599) 722 123

- The major challenge is not collecting the data from the internet, but the next stage: processing them, i.e. extracting the desired information, cleaning them up and converting them into useful information for the production of statistics. For example, a tourist accommodation often will post its information on more than one website. It often has its own website, but will also post on several aggregation sites. This means that after the collection of the data, a de-duplication process has to be implemented, e.g. based on names or addresses of the accommodations. This is not an easy task in practice, as building a good de-duplication program requires some effort. This step becomes more complex when names and addresses of accommodations are not available, are spelled differently or when the information is ambiguous or out-dated (see Table 2). Other challenges are standardisation

of names and addresses; comparability and interpretation of prices, as it is not always clear which elements (e.g. cleaning, tourism tax and activities) are included in the price; calculating the number of bed places, as usually only the number of rooms or units (e.g. buildings or cottages) is available; determining how up-to-date and valid the information is (e.g. the website may still exist while the accommodation has been closed); the reliability of guest review scores (these may be biased or even fake).

- It is important to ensure continuously that the whole population is observed, and that the robots do not miss any accommodation groups. In the case of tourism accommodations, this is probably only an issue in theory: if a tourism accommodation can no longer be found on the internet, it is highly debatable whether it is still operational. However, in the case of dedicated robots and thus selected websites, websites in this group may no longer be updated as a consequence of a fall in popularity compared to other and new websites. Lastly, statisticians should also be aware that a tourism accommodation may change the way it posts its information on the internet, for example by using social media or advertising on eBay.
- Internet data can be collected in real time and are more comprehensive than data collected by traditional surveys and from other sources. However, to assess the quality of internet data it is always wise to compare them with data from other available sources. Also, the fact that information often overlaps between websites is helpful (part of the de-duplication process).

So the general conclusion is that the collection of information on tourism accommodations or in fact any other tourism supply - restaurants and museums, for example - from the internet by robots has some potential, as more information than name and address only can be collected at much lower costs. Therefore, using internet information can reduce the survey burden, especially for small businesses. Data are available faster, new statistics can be produced, and operational efficiency can be improved. However, practice shows that methods for the processing, interpretation and visualization of the data are not easy to implement and take time to develop. Moreover, the quality of the internet data and the question of whether the whole population is observed also remain issues. Therefore, at present, SN sees data collection by internet robots more as a supporting data source alongside already available sources, i.e. the Chamber of Commerce register and the Business Register. Internet data can be used for information on smaller accommodations and businesses whose main activity lies outside the tourism domain, such as farms with campsites. They can also be used to check the quality of the current population of units of the tourism accommodation statistics. However, information from the internet is currently not seen as the sole source of information to compile the population of units for the tourism accommodation statistics.

3. Using mobile phone data for tourism statistics

Two other innovative projects at SN in the domain of tourism used mobile phone data, i.e.: 1) log data from smartphone measurements and 2) anonymous aggregate mobile phone metadata based on Call Detail Records. The former project was carried out in collaboration with the Delft University of Technology and the latter in collaboration with Vodafone (telecom provider) and Mezuro.

3.1. Log data from smartphone measurements software

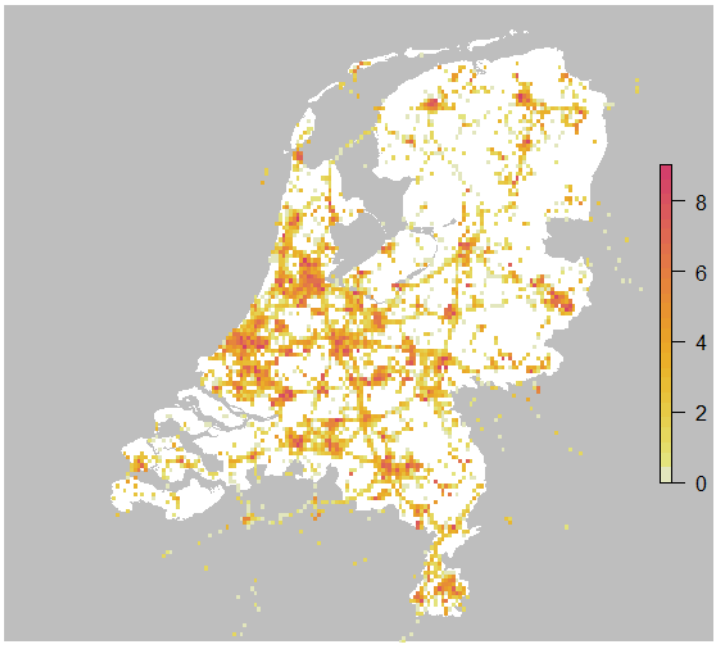
Log data from smartphone measurements are data collected directly from mobile phones of respondents through an *app installed on their phone or mobile device*. This app registers either everything or certain specific actions carried out on a mobile phone, including a time and location registration every five minutes. The latter two variables in particular are interesting for tourism statistics [3].

Besides the fact that movements of a person or family can be tracked quite accurately through time by GPS measurement, another advantage of this method is that specific questions - so-called pop-up questions - can be triggered on the basis of, for example, (the change in) time or location. This includes, questions about the purpose of the journey, mode of transport, price paid, type of accommodation, restaurant visits, satisfaction, activities etc. In more sophisticated apps, some of these variables can be derived on the basis of the location or registration of other data, such as mode of transport, journey duration and distance, mo-

bile payment and overnight stays. Another advantage of this method of data collection is that respondent background data are available: age, sex, income, region, composition of the family/group, making it possible to control the sample and weighting.

SN tested this method in two successive pilots in 2011 and in 2012. These pilots were not specifically conducted for tourism statistics but for mobility statistics and ICT use statistics (what do people do with their mobile phones?). The smartphone measurements were combined with two flanking surveys, one before the actual measurements and one at the end of the study. This was done to collect information about variables that cannot be registered on the basis of an installed app, to evaluate the project, and to compare this method of data collection with traditional surveys.

Figure 1. Visualisation (heat map) of the travel behaviour of a group of respondents based on log data from smartphone measurements, 10 October - 7 November 2012 [Bouwman et al].



Source: Statistics Netherlands and TU Delft.

Practice showed that the combination of information from traditional surveys and log data from smartphone measurements in particular produced a rich and valuable set of data with a lot of potential for research and statistics (see Figure 1, for example). Not only did this produce valuable data, the quality of the data was also improved considerably. A comparison between data registered by the app and the same data collected through the survey (self-report) at the end of the project showed a substantial number of wrong estimates in the self-reports (overestimation). A major challenge in this kind of research is the willingness of people to participate. Not only privacy concerns, but also technological issues and people's lack of knowledge about how to download applications played a role in this respect. The project dealt with privacy concerns by separating the data collection and analysis. The researchers did not have access to names and addresses of participants. In addition, participants could opt out anytime during the data collection.

In addition to the pilot described above, SN has also looked specifically at smartphone measurements software and tracking devices to support or replace questionnaires for the Mobility survey. This study was of a more technical nature, i.e.: to explore how accurate the mobility software was in the registration and determination of location, journeys and modes of transport [4].

3.2. Anonymous aggregated mobile phone metadata based on Call Detail Records[1]

In 2013 a pilot was conducted to see whether aggregated mobile phone metadata based on so-called Call Detail Records (CDR) could be used for tourism statistics. This study was carried out in collaboration with Vodafone and Mezuro. CDRs are generated by telecom providers through phone masts when mobile devices (e.g. mobile phones) make contact with the telecom provider or vice versa. For each contact (call, SMS or data session) these CDRs contain among other things an assigned identification number of the mobile device (not available for SN), the date and time of the call/SMS/data session, which kind of communication was used (voice, data or SMS), and location. The location is a reference to a radio cell in the operator's network, that is mapped to a geographical area. This is less accurate than a location measured with GPS. In addition, similar data were received from the telecom provider on the aggregated number of unique foreign mobile phone callers per month. In handling this kind of information, researchers must pay special attention to protecting and managing privacy by ensuring the anonymity of individuals. Therefore a special system was built and used to create aggregated, anonymised datasets.

The scope of the project was limited to inbound tourism specifically, because this group can be relatively well defined as a separate group. Specific data are also available from foreign mobile devices, i.e. so-called "roaming data" including the country code of the SIM card used. This makes it possible to gain an insight into times and places visited by groups of foreign tourists by country of residence, and to a lesser extent in the number of foreign visitors. To keep things simple, we assumed that every foreign device or unique mobile caller represented one inbound tourist, and that the country code of the SIM card was equal to the country of residence. It was also assumed that foreign mobile callers were randomly assigned to the nearest telephone mast and telecom provider[2].

To get a first indication of the quality of the data, the trend in the number of unique foreign mobile users of Vodafone NL Telecom services in the period April 2012 – March 2013 was compared to the trend in the number of foreign tourists visiting the Netherlands according to SN's survey-based tourism accommodation statistics for the same period. Similarly, a comparison was made between the trend in the aggregated counts obtained from the CDRs and the trend in the number of foreign visitors from the tourism accommodation statistics for the period January – March 2013 (see Figure 2). Subsequently, research was done for a selected set of specific cases concerning small tourist areas (e.g. the Dutch coast and specific events) and short time frames (e.g. calling patterns for a week and for national holidays).

The main conclusions of the pilot were:

- The *trends* in the data on the number of unique foreign mobile callers (blue line in Figure 2) and the data from the CDRs (green line in Figure 2) are comparable to the trends in the number of inbound tourists based on the tourism accommodation statistics (red line in Figure 2). In addition, predictions were made for the then still to be measured number of foreign tourists for the months April, May and June 2013 from the tourism accommodation statistics based on the mobile phone metadata. This was done by extrapolating the number of foreign tourists in March from the tourism accommodation statistics with the trends in the number of unique foreign mobile callers and in the CDR data. When the results (dotted lines in Figure 2) were later compared with the actual results from the tourism accommodation statistics, they only deviated by about 5 percent.
- The deviations between the *number* of foreign tourists according to the tourism accommodation statistics and the number based on the unique number of foreign mobile callers were quite large, for example 2.5 million and 3.5 million in the first quarter of 2013, respectively. However, this is not really surprising because, among other things:
    - the tourism accommodation statistics do not include foreign visitors in accommodations with fewer than 10 bed places and those staying with friends and relatives;
    - the number of unique mobile callers also includes day trips of residents from neighbouring countries, which account for many one-day visitors.

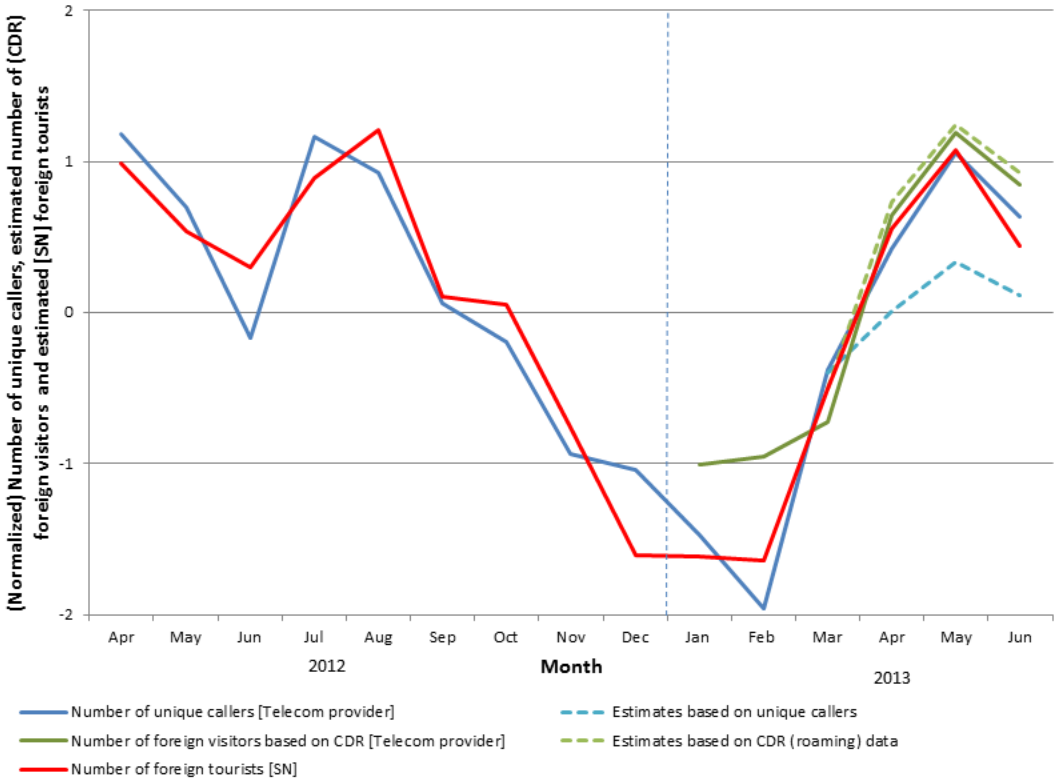
  Furthermore, it should also be taken into account that, besides Vodafone, there are other telecom providers in the Netherlands.

[1] Later this year a more elaborate paper will be published on this issue. See for other studies in this field for example Tiru and Ahas, 2013 [5].

[2] In a later stage of the study this proved not entirely to be true. This raises the question of representativeness.

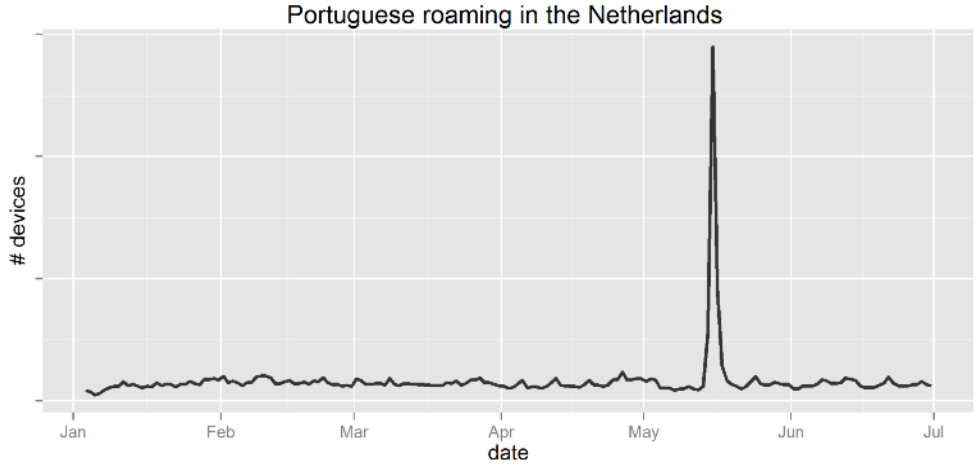
- The differences between the numbers of foreign tourists according to the tourism accommodation statistics and the numbers based on the CDR data are much larger. However, these numbers are not really comparable, as in the pilot a mobile device could only be followed for one day. Therefore, the next day the same mobile device was counted as a new device, resulting in double counting.

Figure 2. Comparisons between trends in inbound tourism based on SN Tourism accommodation statistics, unique foreign mobile callers, and CDR data[3].



Source: Vodafone, Mezuro and Statistics Netherlands

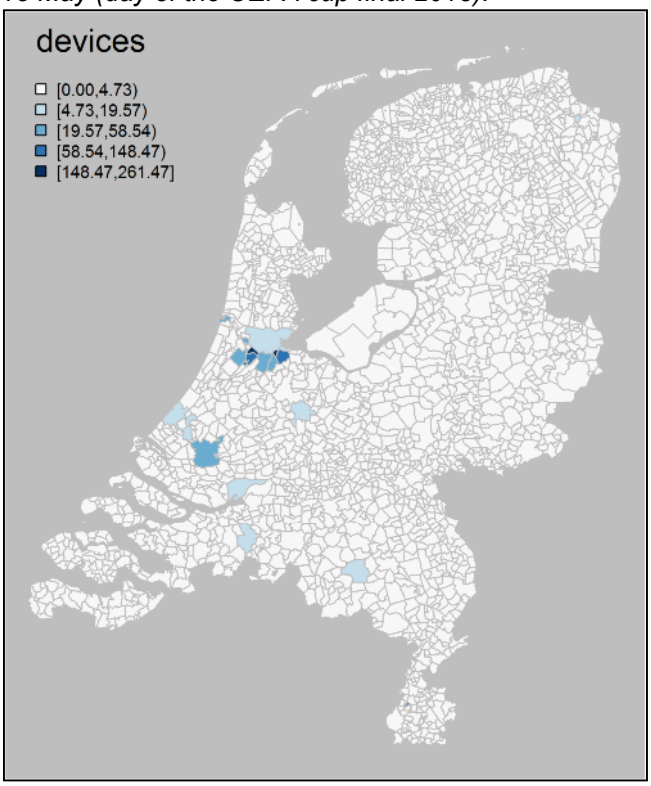
Figure 3. Call behaviour of Portuguese tourists (devices) in the Netherlands, January – July 2013.



Source: Vodafone and Mezuro, compiled by Statistics Netherlands

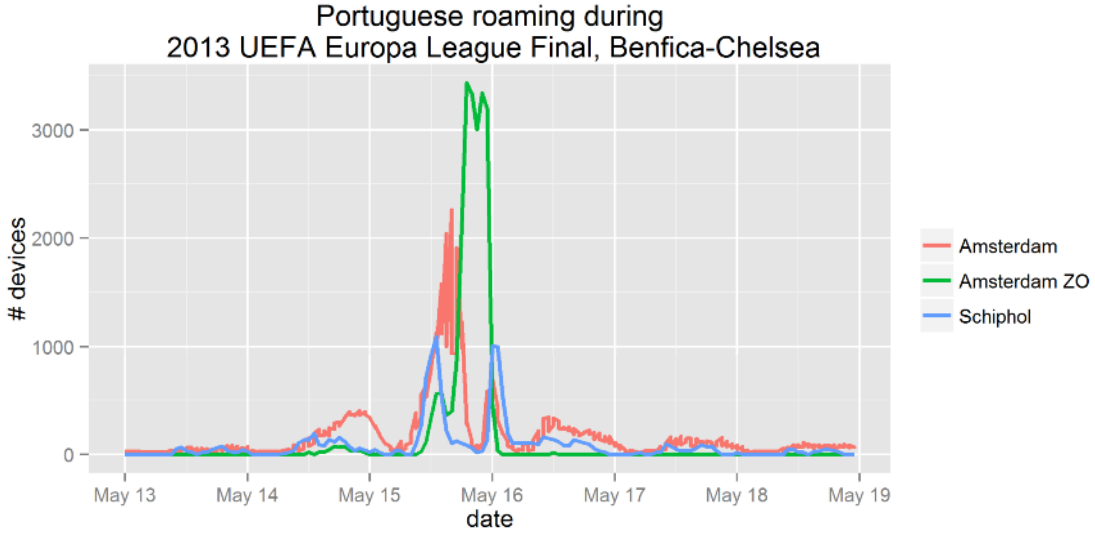
[3] To keep things simple and for the comparability of the trends in the data, all the time series have been standardised using their means and standard deviations. This ensures that all the time series have the same scale (same minimum and maximum and average zero).

*Figure 4. Call intensity of Portuguese tourists (devices) based on their **location** in the Netherlands on 15 May (day of the UEFA cup final 2013).*



Source: Vodafone and Mezuro, compiled by Statistics Netherlands

Figure 5. Calls by Portuguese tourists (devices) before, during and after the UEFA cup football final on 15 May 2013 (green line), Schiphol Amsterdam Airport (blue line) and Amsterdam in general (red line).



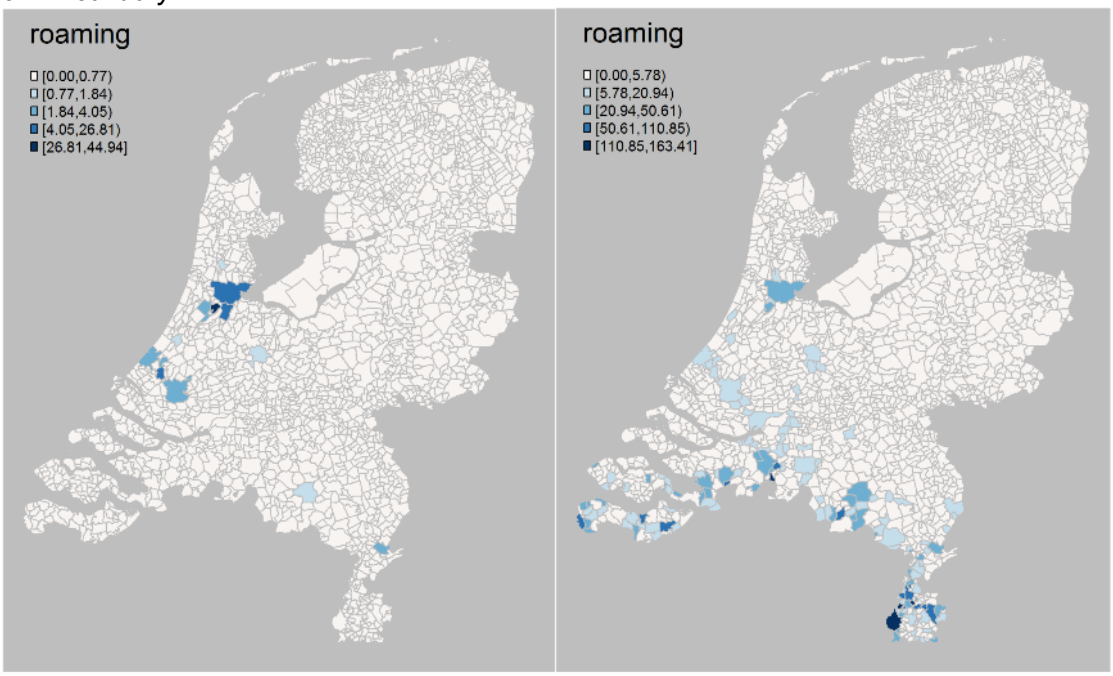
Source: Vodafone and Mezuro, compiled by Statistics Netherlands

- It proved quite easy to zoom in and present results on the patterns of call behaviour of foreign tourists on specific days (e.g. holidays like Easter) or in small areas (e.g. certain events or the Dutch coast). One specific example is given here: the UEFA cup football final between Benfica and Chelsea in Amsterdam on 15 May 2013 in the Arena Stadium in Amsterdam. Figure 3 clearly shows the peak in calls by Portuguese tourists (devices) on 15 May. This is also reflected in Figure 4, which shows

the intensity of calls by Portuguese tourists (devices) in the Netherlands on 15 May. The highest intensities are indeed in south Amsterdam, which is the location of the Arena stadium (Amsterdam ZO). Lastly, plotting calls by Portuguese tourists in Amsterdam and at Schiphol Amsterdam Airport, shows that apparently most Benfica fans arrived to watch the final and left the same day (see Figure 5).

- Figure 6 illustrates another example of the usability of mobile phone metadata. It shows the number of roaming events of Asian and Belgian tourists (= devices) in the Netherlands on 11 January from 12.00 – 16.00 hrs. The results show clearly that Asian tourists are concentrated in the big cities: The Hague, Rotterdam and Amsterdam, and Schiphol Amsterdam airport. As Belgium borders on the Netherlands to the south, roaming data of Belgian devices clearly show the expected border traffic. [6].

Figure 6. Roaming data for Asian (left) and Belgian (right) tourists (devices) from 12:00 – 16:00 hrs. on 11 January



Source: Vodafone and Mezuro, compiled by Statistics Netherlands

- Although mobile phone metadata are a rich source of information, it is clear that this kind of research is still in its infancy and requires a different approach than that of traditional data collection and analysis (e.g. surveys and administrative data sources). Many challenges still remain, not only in the field of tourism and mobility in particular, but also more generally in terms of processing and analysing big data sources. In the case of tourism, for example, one question is how many tourists does one mobile device represents? Foreign tourists may switch off their mobile phone, a family may carry only one mobile phone, while business travellers may have two mobile phones. Tourists from outside Europe may buy a SIM card in the Netherlands, which would result in underestimation of the number of foreign tourists for specific groups. And also the cost of calling abroad can play a different role for different groups of tourists. The representativeness of the data is therefore an issue. One obvious option would be to compare the number of devices at different events or areas with each other, to show the relative importance of these events or areas (benchmarking).
- Protecting privacy is extremely important when using this kind of data, and indeed their very use is sensitive to public opinion. Several measures were put in place in this pilot to guarantee confidentiality. First, all micro data remained with the telecom provider (Vodafone). Requests for data from SN where dealt with by an intermediary organization (Mezuro). Second, SN only received aggregated data and each area in the aggregated dataset contained at least 15 devices in a certain

timeframe. If there were fewer than 15 devices in an area, no data were supplied to SN. Third, mobility patterns of anonymised roaming visitors could only be analysed for one day. The following day the same phone would be assigned a new ID to prevent any risk of indirect identification if source data were to be disclosed unintentionally. And lastly, data were processed in a secure environment, only accessible for authorised researchers, and audited for this purpose by independent research organisation TNO.

- Results of measurements based on big data sources also require new methods of data visualisation, such as geo-location maps with time indicators. The challenge here is to produce visualisations that can be clearly understood and interpreted by users, such as policy-makers.

The overall conclusion of this pilot is, therefore, that data based on Call Detail Records and unique foreign mobile callers seem to have wide and feasible potential for tourism statistics, especially when they provide more details in terms of in place and time. It is possible to, for example, zoom in on small specific tourist areas and very short timeframes, such as certain events, coastal areas, or national holidays; to distinguish data per country of residence; or determine the number of transit passengers.

The potential for tourism statistics could be enlarged considerably if a mobile device could be followed for longer than one day. In a longer term development perspective, this would make it possible to distinguish between numbers, distances and frequency of daytrips, short and long holidays, and possibly derive modes of transport. Furthermore, it would also make it possible to distinguish flows of tourists, i.e. groups of tourists could be followed from the time they enter the Netherlands right up until they depart. We could then also determine which tourism locations are inter-related, even distinguished by the country of residence. This would generate new and important information for the tourism industry and policymakers that cannot be produced with traditional data collection and analysis. Similar methods could also be applied for outbound and domestic tourism, although for the latter case, the subjective definition of tourism – people spending time outside their usual environment – would be mean more of an obstacle; tourism would have to be objectified by variables like duration, distance and frequency. Future profiling of 'the usual environment' may be able to solve this issue: calls could be analysed to determine a person's usual environment by taking into account where they live, work, sport or go to school. Lastly, mobile phone metadata can also be combined with other data such as the effects of weather, the supply and quality of the accommodations, number of events, and the economic situation of a region or country of residence.

4. Final conclusions

The existence of new "big" data sources, like the internet and social media, mobile phone data, but also other sources like public transport data, data from electronic traffic loops, cameras and credit card data, may change the production landscape of tourism statistics significantly in the future. Not only will it affect the way existing tourism statistics are produced, it will also present tourism with new statistics which cannot be produced with traditional data collection and analysis. It won't happen tomorrow: there are still many challenges ahead, but the first results are promising. Data processing, methodology (e.g. representativeness) and, of course, privacy will remain important issues.

In essence big data sources have the potential to improve existing tourism statistics (better quality), to distinguish smaller areas and timeframes (more detail), to increase frequency and timeliness, to reduce the survey burden and produce new statistics and indicators (e.g. tourist flows) for policy-makers and the tourism industry. It is not technology and data availability that are putting the brakes on, but the development of methodology, the lack of big data knowledge and resources, privacy concerns, and the prevailing culture of statistics production.

The emergence of these new big data sources is comparable with the emergence of administrative data sources at the end of the last century. This also took a long time to sink in as an accepted and standard way of using data and making statistics. However, with big data there are differences. The internet is available for everybody, not only for statistical offices (democratisation of data). On the other side, companies who own big data sources are becoming more and more aware of the value of their data (commercialisation of data).

This affects the current role of national statistical institutes (NSIs). In this changing data landscape with new players and fields of force, they are slowly losing their relative monopoly on data and statistics. They will

find themselves in a more competitive environment, which is further enforced by the trend for open or public data. Therefore, NSIs will probably need to reposition themselves as the objective and unbiased party for high-quality reliable statistics for society as a whole.

5. References

1. Laney, D., 2001, 3D Management: controlling data volume, velocity and variety, Meta Group/Gartner;
2. Mayer-Schonberger, V. and K. Cukier, 2013, Big data: a revolution that will transform how we live, work and think, Eamon Dolan-Houghton Mifflin Harcourt;
3. Bouwman, H., N. Heerschap and M. de Reuver, 2013, Mobile handset study 2012: evaluation of mobile measurement software to monitor the use of smartphones and tablets in the Netherlands, TU Delft – CBS, Delft – The Hague, Netherlands;
4. Arends, J., M. Morren, F.Y. Wong and M. Roos, 2013, Smarter 2012, CBS, Statistics Netherlands, The Hague, Netherlands;
5. Tiru, M. and R. Ahas, 2013, Passive anonymous positioning data for tourism statistics, University of Tartu, Estonia;
6. Offermans, M., A. Priem and M. Tennekes, 2013, Rapportage project Impact ICT; mobiele telefonie, 16 April 2013, Statistics Netherlands, The Hague, Netherlands.

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/3-E
19 November 2014

English

SOURCE: Telefonica

TITLE: Big Data for Social Good: Opportunities and Challenges

Big Data for Social Good: Opportunities and Challenges

Nuria Oliver, PhD, nuria.oliver@telefonica.com

We live in a world of data, of big data. Our digital and physical interactions are increasingly leaving digital traces behind, which leads to a *big data revolution*.

Mobile phones are a powerful source of large-scale human behavioral data as it is the most widely adopted piece of technology in human history. In addition, mobile phones are personal devices that we carry with us all the time and are always connected, leaving a digital trace behind. Therefore, they can be seen as *sensors of large-scale human activity*.

Interestingly and very importantly, the analysis of mobile phone data can be done anonymously (*i.e.* without looking at any personal information) and in an aggregated fashion (*i.e.* always reporting aggregated results, never individually). For example, by only counting the total number of phone calls that each cell tower is handling or the number of phones that move between towers.

In my research team at Telefonica, we have been carrying our research in what we call Big Data for Social Good since 2008. The purpose of this research is to understand the value of large-scale human behavioral data (as it is captured passively by the mobile network infrastructure) for positive social impact.

Since then, we have been able to shed light on important questions such the impact that decisions made by governments have on the propagation of a pandemic [1,2,3,4,13,14], the automatic inference of socio-economic status in a developing nation [15] or the detection of floods from significant changes in the pattern of activity in cell towers [11]. These projects have arrived to important findings that are of great interest to governments and international health, development or emergency support organizations [1,2]

However, these projects are just a seed. For this seed to grow into something larger, a tree, there is a need to collaborate with the institutions that know the realities of the countries of study and have the decision making power to implement some of the findings. In addition, there should be a combination of mobile data with other types of data –such as transportation and traffic reports, weather and natural disaster information, emergency services locations….—to fully reach the potential of Big Data for Social Good.

And the tree is starting to grow. We have already ongoing collaborations with United Nations Global Pulse, the World Food Project, the Mexican Government, the

DataPopAlliance, the GSMA and with MIT, among others. We are learning that some of the world's most important challenges can be understood and tackled through the analysis of big data. As Robert Kirkpatrick from UNGP states, "we should be able to work together, private and public sector, to find ways to harness big data for massive positive social impact, both safely and responsibly".

In addition to these focused collaborations, which are the branches of our tree, the ambitious goal is to plant a forest; a forest of big data for social good. Hence, some of this anonymized data is also opening up to a variety of organizations and research teams so that there is collaboration regarding this topic. For example, in September 2013, Telefónica carried out our first Datathon for Social Good at Campus Party (London) in collaboration with the Open Data Institute. Twelve teams participated in this datathon and analyzed both aggregated mobile data and open data and proposed a number of interesting projects, including a crime prediction model for London [10,5,6,8].

While the promise of using big data for social good is large, so are the challenges encountered to be able to realize this promise, as has become clear in light of the recent Ebola outbreak –the worst in our history, with almost 5,000 deaths and over 13,000 infections [17]. Unfortunately and despite the existence of large-scale human behavioral data that could be tremendously useful to help combat Ebola, we are only now – a few months after the outbreak -- starting to put into place coordinated efforts towards the analysis of mobile phone network data and what this tells about the spread of the disease.

Understandably, an important concern is the impact on privacy. The good news is that extensive research conducted by a range of academic teams demonstrates that it is possible to both analyze large-scale human behavioral data (*e.g.* levels of activity at the cell tower level, mobility patterns…) and preserve privacy. All data is typically anonymized using state-of-the-art encryption algorithms. In addition, data is usually analyzed in a highly secure and protected environment (*e.g.* the mobile operator premises) by authorized personnel. No analysis should be undertaken that would ever identify individuals. In addition, only the resulting aggregated, non-sensitive analyses (*e.g.* population mobility estimates, aggregate statistics…) would be made available to relevant aid or government agencies. Of course, an extremely clear code of conduct and ethical protocol would need to be followed by the authorized personnel handling anonymized individual data. A framework to share best practices among mobile phone operators and world recognized research teams would be extremely helpful in this regard.

A second challenge arises from the lack of expertise to be able to carry out such analyses in a technically sound way. However, technical difficulties should not be a barrier either, as there is a body of work illustrating how to carry out this type of analysis.

Moreover, there are highly skilled data scientists worldwide — including ourselves — and strong support from organisations, such as the ITU, ISOC, GSMA and United Nations Global Pulse — who are ready and willing to assist in the process, particularly to ensure that all data handling is carried out in an ethical and anonymous manner, always respecting local data privacy laws.

Finally, there are regulatory barriers and legal uncertainties that would need to be addressed. In particular, standards, practices and regulation would need to incorporate trust mechanisms for giving access to data for humanitarian purposes and in specific conditions (*e.g.* handling of the data by trusted partners, clear code of conduct and ethics when handling personal data, definition of where and how the data would be stored, guarantees that the data will be used only for humanitarian purposes, etc…) [16].

The potential to have positive impact and help save lives is immense. I truly hope that we can find a way to realize this the full potential of big data for social good. It's an opportunity that we cannot afford to miss.

References:

[1]Using Big Data to Fight Pandemics by Nuria Oliver. TechCrunch. Nov 2014

[2]"Ebola: Can big data analytics help contain its spread?" by Matthew Wall. BBC News. Oct 2014

[3]"Nuria Oliver: What Big Data and the Mexican Pandemic Taught Us". WIRED. October 2013.

[4]"Combating Global Epidemics with Big Data" Nuria Oliver. The Guardian. September 2013.

[5]"Mobile Phone Data can help predict Crime one Month in Advance" Motherboard. Sept 2014

[6]"Mobile Data could be Used to Help Predict Where Crime's About to Happen" Gizmodo. Sept 2014

[8]"Researchers use cell-phone data, not precogs, to predict crime in London" Engadget. Sept 2014

[10]"Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data"
A. Bogomolov, B. Lepri, J. Staiano, **N. Oliver**, F. Pianesi, A. Pentland
16th ACM International Conference on Multimodal Interaction (ICMI 2014)

[11]"Flooding through the Lens of Mobile Phone Activity"
Pastor-Escuredo, D., Torres Fernandez, Y., Bauer, J.M., Wadhwa, A., Castro-Correa, C., Romanoff, L., Lee, J.G., Rutherford, A., Frias-Martinez, V., **Oliver, N.,** Frias-Martinez, E. and Luengo-Oroz, M.

[12]Younes Moumny, Vanessa Frias-Martinez and Enrique Frias-Martinez, "Characterizing Social Response to Urban Earthquakes using Cell-Phone Network Data: The 2012 Oaxaca Earthquake", Third Workshop on Pervasive Urban Applications @Pervasive'13, Zurich,Switzerland 2013.

[13]Enrique Frias-Martinez, Graham Williamson and Vanessa Frias-Martinez, "Simulation of Epidemic Spread Using Cell Phone Call Data: H1N1 Case Study", NetMob 2013 (Poster).

[14]Vanessa Frias-Martinez, Alberto Rubio and Enrique Frias-Martinez, "Measuring the Impact of Epidemic Alerts on Human Mobility using Cell-Phone Network Data", Second Workshop on Pervasive Urban Applications @Pervasive'12, Newcastle, UK, 2012.

[15] Vanessa Frias-Martinez, Cristina Soguero, Malvina Josephidou and Enrique Frias-Martinez, "Forecasting Socioeconomic Trends With Cell Phone Records", 3rd ACM Symposium on Computing for Development, DEV'13, Bangalore, India, 2013.

[16] Enabling Humanitarian Use of Mobile Phone Data Montjoye, Y.A., Kendall, J. and Kerry, C. Issues in Technology Innovation Nov 2014 http://www.brookings.edu/~/media/research/files/papers/2014/11/12%20enabling%20humanitarian%20mobile%20phone%20data/brookingstechmobilephonedataweb

[17] Waiting on Hold. Ebola and Big Data. The Economist Nov 2014 http://www.economist.com/news/science-and-technology/21627557-mobile-phone-records-would-help-combat-ebola-epidemic-getting-look

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/4-E
19 November 2014

English

SOURCE: Bangladesh Bureau of Statistics

TITLE: Global ICT indicators' status in Bangladesh

Global ICT indicators' status in Bangladesh

Sami Kabir
Assistant Programmer
Bangladesh Bureau of Statistics (BBS)
Statistics & Informatics Division (SID)
Ministry of Planning
Government of Bangladesh.

INTRODUCTION

Bangladesh, a developing country in the South Asia is focusing on Information and Communication Technology (ICT) with a view to realizing the vision of *Digital Bangladesh 2021.* Progress has been made over the past few years in terms of developing the ICT infrastructure in the country and delivering e-services at the doorsteps of the citizens. Government, in collaboration with different national and international agencies, is also training up people to make them capable of utilizing ICT facilities. This paper's contributions are as follows.

First, it highlights big data in the ICT sector. Second, it focuses on the existing online service providers. Third, it sheds light on trade opportunities in the ICT services. Fourth, it discusses about the telecommunication status of the country. Fifth, it finds country's household indicators to determine ICT progress. Finally, we provide an analysis of the future ICT goals and identify different merits and demerits.

2. ICT indicators

Accurate ICT indicators are critical to tracking the progress against the set ICT goals. In Bangladesh, several state-run agencies including Bangladesh Bureau of Statistics (BBS), Bangladesh Telecommunication Company Ltd. (BTCL) etc execute the task of formulating key ICT indicators to facilitate growth in this sector. BBS has an important strategic paper, National Strategy for the Development of Statistics (NSDS) to boost up statistical activities in a coordinated, participatory and scientific way. Statistics Act, 2013 has empowered BBS with a structured legal frame. BBS has recently deployed Computer Assisted Personal Interview (CAPI) technology to digitize the data collection method from field level.

2.1 Big data in the ICT sector

Some important and large-scale data in the field of ICT with respect to Bangladesh are furnished below.

(a) 115 Million Services provided to citizens from Union Information and Service Centres (UISC).
(b) BDT 1.38 Billion earnings from UISC Entrepreneurs.
(c) 30,000 local youths got ICT training from UISCs.
(d) More than 9000 leaders and entrepreneurs have been trained through leadership.
(e) 4 Million Students Learning from Multimedia Content Developed by 20,000 Teachers.
(f) 70 Million Results of Public Exams over Internet, 55 Million over SMS, 2.9 Million Admission Applications through SMS.

(g) 300+ electronic versions of text books (e-books) available in primary and secondary levels including technical, vocational and Madrasha education. Millions Land Records Delivered Electronically in DC Offices.
(h) 100 Million Birth Registered Electronically.
(i) 10 Million Electronic Money Orders Sent through Post Offices.
(j) 100 Million Birth Registered Electronically.
(k) 2.6 Million Purchase Orders Sent for Sugarcanes over SMS.
(l) 25 Million Utility Bills paid over Mobile Phones.
(m) 1.2 Million Railway Tickets bought over Mobile Phones.
(n) 300 Thousand Online tax calculations by citizens.
(o) In Bangladesh, 112 million wireless subscribers now receive early warning info on natural disasters.
(p) 1800 KM long optic fiber network under Bangladesh Railway is being utilized by the private Cellular Phone Operator.

2.2 Online Service Providers

To turn Bangladesh into a digital nation, initiatives have been taken to deliver e-services at people's doorsteps. To serve this purpose, a program titled "*Access to Information (a2i)*" has been commenced which is an UNDP and USAID supported project (program) having its office at the Prime Ministers' Office of Bangladesh. The overall objective of the project is to provide support in building a digital nation through delivering services at the citizen's doorsteps. The programme aims to improve quality, widen access, and decentralize delivery of public services to ensure responsiveness and transparency. Various initiatives taken by this a2i program are as follow.

(a) Union Information and Service Centres (UISC)

Union Information and Services Centres (UISC) are newly established one-stop service outlets operating at all 4,547 Union Parishads (UP, lowest tier of local government) of the country. Through use of ICT, UISC is able to bring various types of information related to government, livelihood and private services to the doorstep of citizens in rural areas. It ensures services providers and users to save time, cost and has made operations hassle free. Operating under the Public-Private-Peoples' Partnership (PPPP) modality, these centres are run by local entrepreneurs, hosted by UPs and supported by central administration.

KEY SERVICES:

- Public exam results
- Government forms download
- Birth and death registration
- Online university admission
- Data Entry
- VGD/VGF card database

- Livelihood information
- Employment information
- Visa processing /visa form printing
- E-mail and Internet browsing
- Computer Training
- Video conferencing
- Mobile Banking
- British Council's English Learning
- Photocopying/Scanning/Photo/Mobile phone services

(b) National Portal Framework (NPF)

The National web Portal Framework (NPF) is the single platform for accessing all public information from any government organization to ensure easy accessibility for citizens, easy updating of data by non-technical personnel and sharing of data amongst various organizations.

On June 23, 2014, Bangladesh formally launched the world's largest public web portal, Bangladesh Jatiya Tathya Batayon (National Portal) consisting of 25,043 websites of administrative levels including government agencies in one website titled *www.bangladesh.gov.bd* . A total Websites of 4550 unions, 14,640 upazila offices , 4,032 district offices, 455 divisional offices, 64 zila parishads, 488 upazila parishads, 55 ministries and divisions, 345 directorates as well as 414 pourasavas and city corporations are tied in one place through this National Portal (NP).

Counting all ministries, all directorates, all semi-government and all autonomous organizations, and all government offices at the division, district, upazila and union levels, there are about 27,000 government offices. Only a few hundred of these offices – mostly in Dhaka – have their own specific websites.

NPF ensures that all government organizations have online presence, and the existing officers and staffs of the organizations are able to update the sites without depending on technical personnel. In addition, NPF will allow the government to implement the Right to Information Act through proactive information disclosure.

The NPF is designed and developed by the Access to Information (a2i) Program of the Prime Minister's Office in consultation with government stakeholders in different tiers. The templates in NPF can be used as a guideline for website development of other affiliated organizations.

(c) District e-Service Centres

The District e-Service Centre (DESC) is an ICT facilitated one-stop service centre which provides an efficient electronic version of the century-old manual and heavily bureaucratic service

delivery system at every DC office. It is located in the Deputy Commissioner's (DC) office. DESC has been designed to improve the accessibility and transparency of public service delivery system at the district level to achieve the following.

Objectives:

• Ensure service delivery at the at the door steps of the people at the least possible time

• Uphold citizens' Rights to Information through extensive information flow

• Save time and labor in the processing period

• Increase the number of clients served everyday through the use of ICT

• Reduce corruption and increase accountability by ensuring enhanced flow of information and more transparent processes.

Services available at the DC office can be requested and received through the one stop service counters, online, by phone, by post, or by fax. The DESC was first piloted in Deputy Commissioner's Office of Jessore district and is currently operational in all the 64 districts of the country.

(d) Multimedia Class room

The education initiatives by Access to Information Project aim to make teaching and learning more effective and enjoyable for both students and teachers using ICT. a2i has followed a 3-pronged approach in its efforts to remodel education: establishing Multimedia Classrooms in secondary schools, training teachers on making ICT aided educational content on hard-to-grasp topics and make electronic versions of text books available in primary and secondary levels including technical, vocational and Madrasa institutions. As part of the education reform driven by the Ministry of Education, a2i through public private partnership has so far established 500 multimedia classrooms in secondary schools and trained about 4500 teachers through 400 trainers of public training institutes.

(e) Jatiyo e-Tathyakosh

In rural Bangladesh, information is vital for the betterment of lives and livelihood of the millions. However, availability of even the very basic information still remains a challenge for most citizens. Within this context and in the light of upholding the right to information, Access to Information (a2i) programme has taken the initiative to provide a one-stop solution for sharing livelihood-related information, through an online knowledge bank, the Jatiyo e-Tathyakosh (www.infokosh.gov.bd). This is a national e-Content repository having the largest pool of livelihood contents in Bangla delivered through audio-visual, text and animation formats.

(f) e-Purjee DIGITAL PURCHASE ORDERS FOR SUGARCANE FARMERS

An e-Purjee is a SMS-based purchase order that sugar mills all over the country send to sugarcane growers during crushing season. Formally known as Digital Sugarcane Procurement System, this new system promises timely delivery of purchase orders to farmers. This reduces

uncertainty of the previous paper-based system by enabling farmers with timely delivery of sugarcane to the mills benefiting farmers and sugar mills alike. With advisory and technical support from the Access to Information (a2i) Programme, Sugar and Food Industries Corporation is running the e-Purjee system in all 15 state owned sugar mills of Bangladesh since 2011-12 crushing season.

- ✓ Average 18,000 e-Purjees are being sent through SMS and published at the websites to cane farmers from 15 state-owned sugar mills.
- ✓ Approximately 2,00,000 sugar cane farmers around the country are benefiting from this service.

(g) NESS (National E-Service System)

The National e-Service System (NESS) is an initiative to consolidate all government e-Services within one framework. The NESS framework will accommodate all services offered by the government of Bangladesh to its citizens. An e-Service under NESS has capability to receive online or mobile-based applications, provide receipt numbers immediately for future tracking of status, and in most cases, completion of the service online. Any financial transaction associated with service application may be accommodated through a special e-Financial Inclusion platform. A forms engine to develop a digital version of any paper form is an integral part of NESS.

The major components of NESS are:

• e-Service delivery – a web service platform featuring initially 400+ e-Services and capability to integrate new e-Services from any service delivery organization of the government;

• e-Forms – an e-Forms engine that enables conversion of paper-based forms to paperless forms. Since many government services are initiated through a citizen's application using a form, e-Forms will allow integration of many services into the NESS platform;

• e-Filing – electronic management and archival of all files within and across government organizations;

• e-Communication – sending of SMS, email and other forms of electronic alerts for meetings, events, etc. to government officers. Also, sending of acknowledgment, progress and any other kind of notification to citizens waiting for service delivery.

• Government Directory Access Service (GDAS) – e-Directory service providing authentication and authorization of government users of the system. This also serves as a directory of government users with relevant organizational and contact information.

• Citizens' profile – usage profile of citizens in the NESS system so that citizens can access his or her own usage history;

• Dashboard – monitoring dashboard tracking all requests from citizens and all files generated from citizens' requests or internally within an office. A dashboard for a particular office will have personalized views containing all requests/files pertaining to that officer and all his/her subordinates.

Thus, an officer is able to view activities or lack thereof of all officers/staff under him/her. For example, a DC's dashboard will be able to monitor all e-service and e-file processing within the DC office and by all big UNOs under him/her, Cabinet Division's dashboard all Divisional Commissioner, DC and UNO offices, a DG's dashboard all divisional, district and upazila offices of the respective directorate.

(h) GIS Application

The GIS based interactive mapping tool, developed by BBS, has the capability of using the digital map of Bangladesh as well as census and survey results in order to prepare thematic maps and graphs for presenting different census and socio-demographic indicators.

(i) BDInfo

It is a national, social and economic database system to store data generated from censuses, surveys, records of management information systems. BBS has developed this database as a step towards 'Digital Bangladesh' with technical and financial support from UNICEF. It can be used for evidence-based planning, budgeting and decision-making; and to monitor the achievements with equity towards the MDGs and other national goals in Bangladesh.

2.3 Trade in ICT services

IT and software businesses have been exempted from paying income taxes and VAT has also been exempted on imports and at the production level of software. A 'Research and Development Center' has been established in Bangladesh by South Korean tech giant company Samsung Electronics Co., and offices have been set up here by organizations like Google, Microsoft and Dell. Currently, the country earns around US $250 million by exporting IT and IT enable services through firms or personal initiatives, which was only US $5 million five years ago. "We aim to develop it to $1 billion by 2018 and contribute 1 percent to the gross domestic product." Business opportunity exists in the field of e-governance, outsourcing, e-commerce, business software, mobile content and application, IT education, IT job fair, career counseling, multimedia and animation etc.

With the advent of ICT, conservative outlook about women in the country is diminishing gradually. As a result we find more women are employed in various knowledge based industries such as computer-aided designing, graphic designing, composing etc. Village Phones have increased income and savings accumulation among phone owners, mostly women. The women phone operators are generally poorer than the average villager is. However, the income that they earn is significant, generally accounting for 30-40% of household income and averaging US $300 per year in a country where average per capita income is US $286. The operators are likely to be married (90%), and half of them have no formal education. Another quarter has primary education and the remaining quarter, some secondary education. 36 % identify themselves as housewives, and only 6% have some kind of formal employment (in government or business). The women operate their phone businesses while doing household chores or operating another business. The phones are used primarily for calls relating to financial matters, particularly relating to remittances, which are a significant source of village income. Strikingly among poor villagers, 38% of phone users had one or more family member living abroad. Women phone operators have achieved economic and social empowerment within their households and communities.

2.4 Telecommunication status

2.4.1 Mobile Phone Subscribers in Bangladesh

The total number of Mobile Phone subscribers has reached 116.239 million at the end of May, 2014.

The Mobile Phone subscribers are shown below:

Operator	Subscriber (In Millions)
Grameen Phone Ltd. (GP)	49.088
Banglalink Digital Communications Limited	29.626
Robi Axiata Limited (Robi)	24.077
Airtel Bangladesh Limited (Airtel)	8.491
Pacific Bangladesh Telecom Limited (Citycell)	1.433
Teletalk Bangladesh Ltd. (Teletalk)	3.525
Total	116.239

* Subscribers in Millions

**The above subscribers' numbers are declared by the mobile operators

Source: Bangladesh Telecommunication Regulatory Commission (BTRC)

2.4.2 Internet Subscribers in Bangladesh May, 2014

The total number of Internet Subscribers has reached 38939.036 thousand at the end of May, 2014.

The Internet subscribers are shown below:

OPERATOR	SUBSCRIBER
Mobile Internet	37417.841
WiMAX	291.385
ISP + PSTN	1229.81

OPERATOR	SUBSCRIBER
Total	38939.036

* Subscribers in Thousands
**The above mentioned figure represents the number of Active subscribers only. A subscriber/ connection using the internet during the last Ninety (90) days is considered to be an Active subscriber.
Source: Bangladesh Telecommunication Regulatory Commission (BTRC)

2.4.3 PSTN Subscribers in Bangladesh May, 2014

The total number of Glossary Link PSTN Subscribers has reached 1,072.08 thousand at the end of May 2014.

The PSTN Phone subscribers are shown below:

OPERATOR	SUBSCRIBER
BTCL	845.00
Ranks Telecom Ltd	217.93
BanglaPhone Ltd	5.45
World Tel	3.70
Total	**1,072.08**

*Subscribers in Thousands

**The above subscribers' numbers are declared by the PSTN operators

Source: Bangladesh Telecommunication Regulatory Commission (BTRC)

Subscriber increase rate of Mobile & Fixed Phone and Tele-density till January-2014 are shown below.

SUBSCRIBER	Quantity
Mobile Subscribers (in Millions)	114.80
Fixed Phone Subscribers (in Millions)	1.10
Total Subscribers (in Millions)	115.90
Internet Users (in Millions)	35.50
Yearly Tele-density (%)	76.44

2.5 Household Indicators

Household Income and Expenditure Survey 2005 and 2010

Bangladesh Bureau of Statistics has conducted another survey Household Income and Expenditure Survey (HIES) 2005 & 2010. The questions focused on the use of fixed telephone, mobile phone, computer and email facilities. As is known HIES is a large survey with a long questionnaire. The inclusion of only four questions marks the beginning of a full effort in the future to conduct a survey especially on ICT.

As mentioned earlier, "Household Income and Expenditure Survey 2010" sought information on the use of ICT (use of computer, email, internet, telephone, mobile phone etc.) at the household and individual levels. The enumerators were provided with Laptop computers to collect and process field level data. The preliminary result of the survey has been published. However preliminary results show the following:

Table 01: Percentage of households having ICT facilities

Type of facilities	National		Rural		Urban	
Year	2010	2005	2010	2005	2010	2005
Telephone	2.07	2.87	0.70	0.33	5.79	10.36
Mobile Phone	63.74	11.29	56.77	6.05	82.74	26.73
Computer	3.01	1.36	0.97	0.17	8.58	4.88
email	1.39	0.20	0.39	-	4.10	0.81

Source: BBS, Household Income and Expenditure Survey 2005 & 2010

The table shows increasing tendency of use of computer facilities in Bangladesh especially in rural areas. The most notable information from the survey is that though mobile phone has been introduced in the country in

early nineties, there has been substantial growth of mobile phone use. The use of Internet has considerably increased both in urban and rural areas as focus by the recent survey. Understandably, the use of the ICT facilities is much higher in urban areas compared to rural areas.

There has been rapid growth of use of Mobile phones between 2005 and 2010. While only 11.29% of households used mobile phones in 2005, more than 63% of households were found to use them in 2010.Graph-1: Percentage of households having ICT facilities.

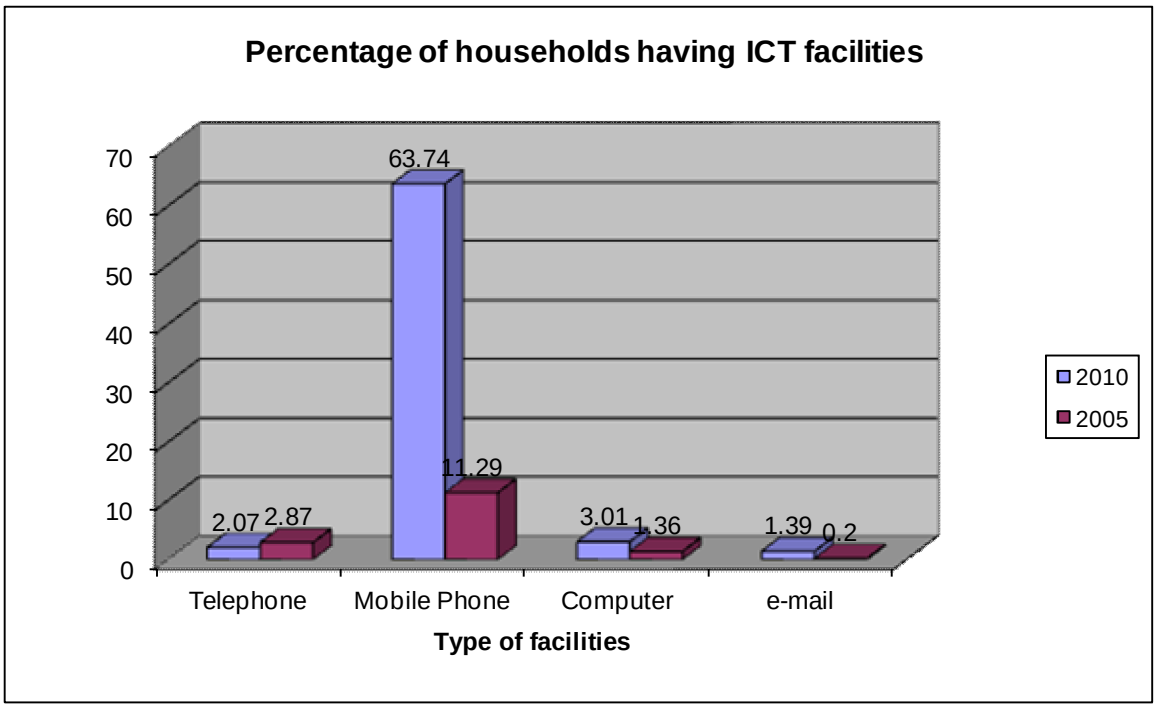


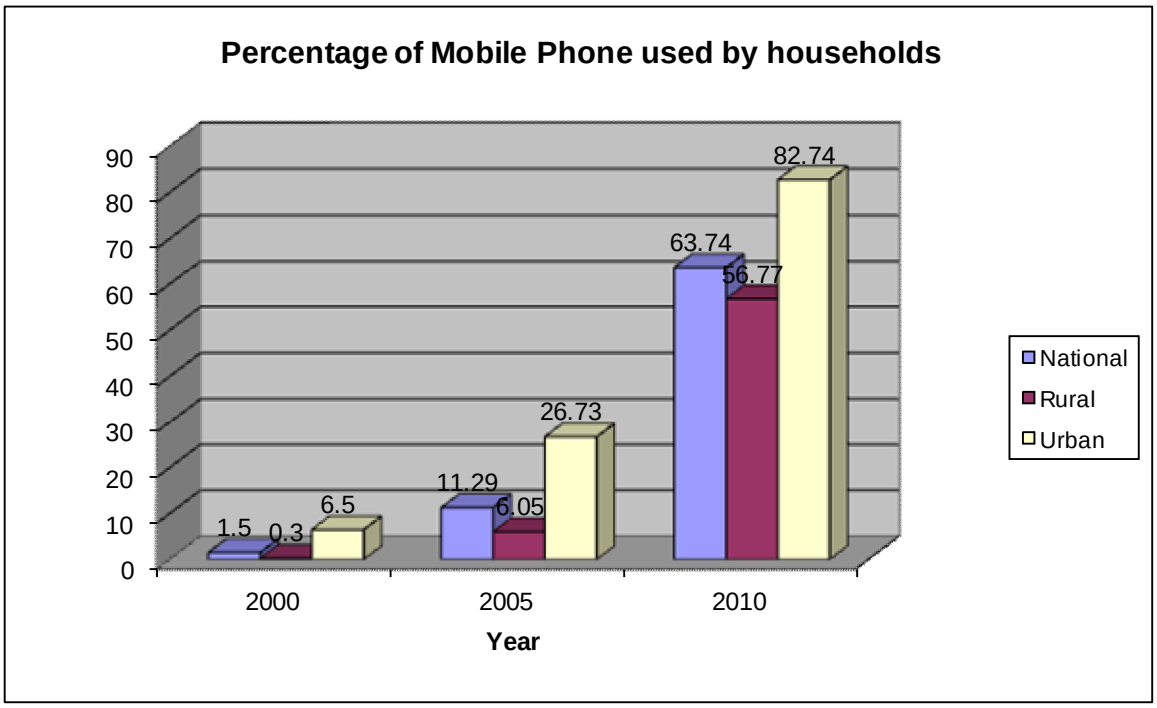
Table 02: Mobile phone used by households

Locality	**2000**	**2005**	**2010**
National	1.50	11.29	63.74
Rural	0.30	6.05	56.77
Urban	6.50	26.73	82.74

Source: BBS, HIES 2000 and HIES 2005 & 2010

It may be mentioned here that the government has been trying to popularize ICT in government offices to improve productivity and promote e-governance. The progress in this front has been captured in a recent survey which is discussed below.

Graph: Percentage of Mobile Phone used by households



In the more than 20000.00TK monthly income group the use of computer 51.27% and use of internet/e-mail is 82.5% in the year 2005. Whereas the percentage use regarding computer in the year 2010 is 67.17% and use of internet/e-mail is 72.78% in the year 2010. In this observation, it is shown that there has been an increment of 15.89% in the year 2010 compare to the year 2005. In case of the use regarding and use of internet/e-mail facilities there has been observed decreasing tendency of 9.77% in the year 2010 compare to the year 2005.

Table 03: **Core indicators on access to, and use of, ICT by households and individuals, latest available data (Year 2013).**

Country	**Percentage of households with** (*Year 2013*)	
Bangladesh	**Radio**	**13.7**
	TV	**46.2**
	Fixed Line Telephone	**2.5**
	Mobile-cellular telephone	**86.6**
	Computer	**5.8**
	Internet access	**4.6**

Country	**Percentage of individuals using a** (*Year 2013*)	
Bangladesh	**Ever used Computer**	**4.7**
	Ever used Mobile-cellular telephone	**78.5**

Country	**Percentage of individuals using** (*Year 2013*)	
Bangladesh	**Ever use Internet?**	**6.5**
	Where use Internet?	
	Home	**49.49**
	Work place	**16.99**
	Education Centre	**13.37**
	Another Person's home	**6.86**
	Cyber Cafe	**6.42**
	Other Places	**6.86**
	Frequency of internet use	
	At least once a day	**25.08**
	At least once a week but not everyday	**45.69**
	Once or more in 2 weeks	**29.23**

2.6 Future ICT Goals

The government will set up high-tech parks in every district in phases to develop the IT sector in every corner of the country in the areas of industry, health, employment generation, poverty alleviation, economic development and people's empowerment.

Digital Bangladesh Vision 2021 will need from both Public and Private sector an infusion of resources, leadership and ICT centered development if it is to be made meaningful. Over the last few decades, the world has been shifting from industrial to knowledge-based societies; the ability of a nation to use and create knowledge capital determines its capacity to empower and enable its citizens by increasing human capabilities. Easy access to knowledge, creation and preservation of knowledge systems, dissemination of knowledge and better knowledge services should be core concerns of the Digital Bangladesh Vision 2021.Bangladesh should be part of a well-crafted national strategy and "Digital Bangladesh", needs to be the cornerstone strategy for Bangladesh. We have to build a people-centered, development-oriented Information Society, where everyone would be able to access, utilize and share information and knowledge easily and efficiently. The concept of Digital Bangladesh should be centered on the creation of what is popularly termed as a "knowledge- based society," Information and communication technologies (ICTs) are a critical component for building this knowledge-society. Our ability in creating and disseminating knowledge will eventually drive the nation's growth in the coming days. A digital society ensures an ICT-driven knowledge-based society where information will be readily available online and where all possible tasks of the government, semi-government and also private spheres will be processed using state of the art technology. The first and foremost challenge to materialize the Digital Bangladesh Vision 2021 would be to ensure overall connectivity at an affordable cost. With the intent to enhance connectivity emphasis should be provided on the establishment of infrastructures to "Connect the Unconnected" and importance must be given on laying more optical fiber to reach the marginal people of the country. Digital Bangladesh Vision 2021 should establish technology-driven e-governance which includes e-administration, e-education, e-health, e-commerce, e- production, e-agriculture, etc. in the five focus areas of the knowledge paradigm:

(a) Access to Knowledge:

Providing access to knowledge is the most fundamental way of increasing the opportunities and reach of individuals and groups. Therefore, means must exist for individuals who have the ability to receive and comprehend knowledge to readily obtain it. This also includes making accurate knowledge of the state and its

activities available to the general public. Project, should be immediately initiated with an objective to facilitate the establishment of affirm presence of Bangladesh Government entities on the Web with two way communication capability or Web 2.0. The Programme requires provision of an entire spectrum of web services to the Government sector as well as running specialized Portals for the benefit of citizens and other stakeholders.

(b) Knowledge Concepts:

Knowledge concepts are organized, distributed and transmitted through the education system and that's why we need a *National Research & Education Network (NREN)* in Bangladesh. It is through education that an individual can make better informed decisions, keep abreast of important issues and trends around him or her and most importantly, question the socio-economic arrangements in a manner that can lead to change and development. In fact, a successful "Digital Bangladesh" would need amore literate population. A mass computer-literacy programme or even a government-sponsored computer course, offered perhaps as an incentive for every student who completes his or her secondary-school education, would benefit everyone. If there is will - backed by investment - there is a way.

(c) Creation of Knowledge:

A nation can develop in two ways – either it learns to use existing resources better, or it discovers new resources. Both activities involve creation of knowledge. This makes it important to consider all activities that lead to the creation of knowledge directly or help in protecting the knowledge that is created. To realize the aspirations of the 2021 vision, the country must be able to produce its own engineers, scientists and technological know-how.

(d) Knowledge Applications:

Knowledge can be productively applied to promote technological change and facilitate reliable and regular flow of information. This requires significant investment in goal-oriented research and development along with access models that can simplify market transactions and other processes within an industry. Initiatives in the areas of agriculture, small and medium enterprises (SMEs) and traditional knowledge can demonstrate that knowledge can be very effectively applied for the betterment of the rural poor.

(e) Delivery of Services:

Knowledge services have the potential to simplify many different points at which citizens interact with the State. Traditionally, these points of interaction have been vulnerable to unscrupulous activities and rent-seeking. We need to set the bureaucracy under an e-governance initiative, with a transparent file tracking system that the public can access. This will, right away, reduce corruption, because everyone involved in the process can be tracked down. Technology provides us with an opportunity to ensure accountability, transparency and efficiency in government services. E-governance is one of the ways in which citizens can be empowered to increase transparency of government functioning, leading to greater efficiency and productivity. E-Governance aims to place the government within the reach of all citizens increasing transparency and citizen's participation. Thus, the development of e-Governance should promote universal access to government's services, integrate administrative systems, networks, and databases, and make such information available to the citizen via Internet. In a nutshell, such e-Governance should transform the government into a citizen centric technological driven one.

3. Conclusion

Bangladesh has been advancing fast overcoming all the obstacles in the ICT sector under the auspicious leadership of the pro-ICT present government. If this pace is maintained, Digital Bangladesh 2021 vision can definitely be realized. The young generation of the country is highly tech-savvy and the members of the earlier generations are also adjusting themselves with the emerging ICT environment. Various public and private universities of the country have been producing ICT graduates who are capable of contributing to this sector. Adequate physical infrastructure is also being developed to track down the progress and hurdles in this arena.

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/5-E
19 November 2014

English

SOURCE: ITU

TITLE: Data Quality Assessment Framework for ITU

International Telecommunication Union

Data Quality Assessment Framework for ITU

Data Quality Assessment Framework for ITU

1. Introduction

Several other international organisations, including OECD, ECB, and Eurostat have defined a *quality assurance framework* for their internal use in collecting, processing and disseminating statistical data. Although ITU has a very much smaller statistical program than these organisations there is every reason to suppose that such a framework, scaled appropriately to the scale of ITU's statistical activities, would be equally useful.

As the main mission of ITU is "to enable and foster the growth and sustained development of telecommunication networks and services…", the framework is referred to as a *data quality assessment framework (DQAF)* to emphasise that it refers to activities related to the collection processing and dissemination of statistical data, not to ITU as a whole.

The main benefits expected from the DQAF are that it will:

- provide a systematic mechanism for facilitating the ongoing identification of quality problems and possible actions for their resolution;
- provide a basis for creating and maintaining a data quality culture within ITU;
- stimulate and maximize the interaction among ITU staff involved in production or use of statistics;
- give greater transparency to the processes by which statistics are produced and their quality is assured and thereby reinforce ITU's image as a trustworthy provider of good quality statistics;
- provide reference material that can be helpful for training;
- provide a mechanism for the exchange of ideas on quality assurance with other producers and users of statistics, at international and national levels.

Whilst the DQAF was developed with the statistical operations of ICT Data and Statistics Division in mind, it is equally applicable to operations of other statistics producing units within the ITU, specifically the Regulatory and Market Environment Division (RME).

The intended readership/users of the DQAF are:

- IDS staff – the DQAF provides a framework for assessment of statistical activities;
- ITU and BDT senior management – the DQAF provides an indication how quality may be assessed;
- NSOs and NRAs – the DQAF provides quality guidelines and an indication of roles NSOs and NRAs as data providers can play in quality assurance;
- Data users – the DQAF provides users of ICT statistics with evidence of quality assurance by ITU.

Development of the DQAF

The DQAF was prepared by an expert in statistical data quality, Michael Colledge under the supervision and with the support of the ICT Data and Statistics Division. A preliminary outline

was discussed at a combined meeting of the Expert Group on Telecommunication/ICT Indicators (EGTI) and the Expert Group on ICT Household Indicators (EGH). Comments were obtained from members of both working groups and were taken into account in preparing this version.

In developing the DQAF there was no need to reinvent the wheel as several quality assurance frameworks for statistical organisations have been developed in recent years and advantage could be taken of their contents and experience in their use. Particularly influential international documents include the following.

- *Fundamental Principles of Official Statistics* – indicating how national statistical systems should be organized in order to produce appropriate and reliable data that adhere to appropriate professional and scientific standards.
- *Principles Governing International Statistical Activities* – comprising principles and practices that were developed and publicized by the CCSA and that should underpin the production of statistics by an international organisation;
- *UN Statistical Commission National Quality Assurance Framework (NQAF)* – a template that provides the general structure within which a country can formulate and operationalize a national quality assurance framework.
- *European Statistics Code of Practice (ESCoP)* – developed by Eurostat, comprising principles and indicators relating to statistical environment, processes and outputs of European NSOs and agencies.
- *European Statistical System Quality Assurance Framework (ESS QAF)* – developed by Eurostat, the focus of the framework is to assist in implementation of the ESCoP by European NSOs and Eurostat.
- *Data Quality Assessment Framework (IMF DQAF)* – developed by the IMF Statistics Division for use by NSOs and other national government agencies collecting and disseminating statistics.
- *Quality Framework and Guidelines for OECD Statistical Activities (OECD QFG)* – developed by the OECD for managing quality within its own organization.
- *European Central Bank Statistics Quality Framework (ECB SQF)* - developed by the ECB for managing quality within its own organisation.

As the organisations authoring these documents have much bigger statistical programmes than the ITU, although the quality principles remain much the same whatever the organisation, the framework itself has been appropriately simplified for ITU.

DQAF Components

The DQAF has four components.

- The first component is a set of *underlying principles*, providing the basis for formulating the DQAF.
- The second component is a set of *quality dimensions*, highlighting the various aspects of data and process quality.
- The third component is a set of *quality guidelines*, comprising good practices for assuring quality.

- The fourth component is *quality assessment and improvement program*, comprising a set of procedures for ensuring that quality is regularly assessed and appropriate quality improvement actions are implemented.

2. Underlying Principles

The underlying *Principles Governing International Statistical Activities* upon which the DQAF is based were formulated by the Committee for the Coordination of Statistical Activities and endorsed by the chief statisticians/coordinators of statistical activities of UN agencies in 2005. They are as follows.

1. ***High quality international statistics, accessible for all, are a fundamental element of global information systems***

Good practices include:

- Having regular consultations with key users both inside and outside the relevant organisation to ascertain that their needs are met
- Periodic review of statistical programmes to ensure their relevance
- Compiling and disseminating international statistics based on impartiality
- Providing equal access to statistics for all users
- Ensuring free public accessibility of key statistics

2. ***To maintain the trust in international statistics, their production is to be impartial and strictly based on the highest professional standards***

Good practices include:

- Using strictly professional considerations for decisions on methodology, terminology and data presentation
- Developing and using professional codes of conduct
- Making a clear distinction, in statistical publications, between statistical and analytical comments on the one hand and policy prescriptive and advocacy comments on the other

3. ***The public has a right to be informed about the mandates for the statistical work of the organisations***

Good practices include:

- Making decisions about statistical work programmes publicly available
- Making documents for and reports of statistical meetings publicly available

4. ***Concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics are chosen to meet professional scientific standards and are made transparent for the users***

Good practices include:

- Aiming continuously to introduce methodological improvements and systems to manage and improve the quality and transparency of statistics

- Enhancing the professional level of staff by encouraging them to attend training courses, to do analytical work, to publish scientific papers and to participate in seminars and conferences.
- Documenting the concepts, definitions and classifications, as well as data collection and processing procedures used and the quality assessments carried out and making this information publicly accessible
- Documenting how data are collected, processed and disseminated, including information about editing mechanisms applied to country data
- Giving credit, in the dissemination of international statistics, to the original source and using agreed quotation standards when reusing statistics originally collected by others
- Making officially agreed standards publicly available

5. Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimise the reporting burden for data providers

Good practices include:

- Facilitating the provision of data by countries
- Working systematically on the improvement of the timeliness of international statistics
- Periodic review of statistical programmes to minimise the burden on data providers
- Sharing collected data with other organisations and collecting data jointly where appropriate
- Contributing to an integrated presentation of statistical programmes, including data collection plans, thereby making gaps or overlaps clearly visible
- Ensuring that national statistical offices and other national organisations for official statistics are duly involved and advocating that the Fundamental Principles of Official Statistics are applied when data are collected in countries

6. Individual data collected about natural persons and legal entities, or about small aggregates that are subject to national confidentiality rules, are to be kept strictly confidential and are to be used exclusively for statistical purposes or for purposes mandated by legislation

Good practices include:

- Putting measures in place to prevent the direct or indirect disclosure of data on persons, households, businesses and other individual respondents
- Developing a framework describing methods and procedures to provide sets of anonymous micro-data for further analysis by bona fide researchers, maintaining the requirements of confidentiality

7. Erroneous interpretation and misuse of statistics are to be immediately appropriately addressed

Good practices include:

- Responding to perceived erroneous interpretation and misuse of statistics
- Enhancing the use of statistics by developing educational material for important user groups

8. Standards for national and international statistics are to be developed on the basis of sound professional criteria, while also meeting the test of practical utility and feasibility

Good practices include:

- Systematically involving national statistical offices and other national organisations for official statistics in the development of
- international statistical programmes, including the development and promulgation of methods, standards and good practices
- Ensuring that decisions on such standards are free from conflicts of interest, and are perceived to be so
- Advising countries on implementation issues concerning international standards
- Monitoring the implementation of agreed standards

9. Coordination of international statistical programmes is essential to strengthen the quality, coherence and governance of international statistics, and avoiding duplication of work

- Good practices include:
- Designating one or more statistical units to implement statistical programmes, including one unit that coordinates the statistical work of the organisation and represents the organisation in international statistical meetings
- Participating in international statistical meetings and bilateral and multilateral consultations whenever necessary
- Working systematically towards agreements about common concepts, classifications, standards and methods
- Working systematically towards agreement on which series to consider as authoritative for each important set of statistics
- Coordinating technical cooperation activities with countries between donors and between different organisations in the national statistical system to avoid duplication of effort and to encourage complementarities and synergy
- to the improvement of statistics in the organisations and in countries

10. Bilateral and multilateral cooperation in statistics contribute to the professional growth of the statisticians involved and to the improvement of statistics in the organizations and in countries

Good practices include:

- Cooperating and sharing knowledge among international organisations and with countries and regions to further develop national and regional statistical systems
- Basing cooperation projects on user requirements, promoting full participation of the main stakeholders, taking account of local circumstances and stage of statistical development
- Empowering recipient national statistical systems and governments to take the lead
- Advocating the implementation of the Fundamental Principles of Official Statistics in countries
- Setting cooperation projects within a balanced overall strategic framework for national development of official statistics

3. Quality Dimensions

Introductory Remarks

It is generally agreed that, whilst *statistical product quality* can be summarized in line with the definition in the ISO 9000 Series for any product as *fitness for use*, there is a need to elaborate this definition in terms of the various quality aspects or *dimensions*. Many versions of quality dimensions have been proposed over the last 20 years, most of which contain essentially the same ideas and all of which include a significant expansion of the original narrow interpretation of quality as simply *accuracy*.

In developing the following set of dimensions particularly influential documents are:

- the *OECD Quality Framework and Guidelines* (the primary source);
- the generic *National Quality Assurance Framework (NQAF) Template* developed by UN Statistical Division and endorsed by the UN Statistical Commission; and
- the European Statistical System quality dimensions that were subsequently incorporated in the *European Statistics Code of Practice*.

The ITU quality dimensions are in two groups: those relating to *data quality* and those relating to *process quality*. The latter group is important as well designed and executed processes provide the basis for data quality.

Data Quality Dimensions

1. Relevance

The *relevance* of a data product is the degree to which the data serve to address the purposes for which they are sought by users. Relevance has three aspects: coverage of the required population (completeness); inclusion of the appropriate content: and use of appropriate concepts. Relevance is further characterised by the merit of the data uses in terms of the ITU mandate.

Typically a data product has multiple users and uses. Thus, measuring relevance requires the identification of user groups and their needs, whilst recognizing that these may change over time.

Relevance may be indirectly assessed by ascertaining whether there are processes in place to determine the views of users and the uses they make of the data.

Users of ITU data may be divided into two main groups;

- *internal users* - primarily analysts within ITU divisions; and
- *external users* - including other UN organisations, other international organisations, national governments, national statistical offices and other national organisations, academic institutions and the media.

Whilst internal users are the very important, it is essential that the content and format of published outputs be adapted to the full range of audiences.

2. Accuracy

The *accuracy* of a data product is the degree to which the data correctly estimate or describe the quantities or characteristics they are designed to measure. Accuracy refers to the closeness between the values provided in the product and the (unknown) true values. Accuracy has many attributes, and in practical terms there is no single overall measure of it. Typically, accuracy is

described in terms of the errors, or the potential significance of errors, introduced at various stages in the production process from initial acquisition of the data to dissemination of aggregates.

In the case of data from sample surveys, the major sources of error are coverage, sampling, non-response, response, processing, and seasonal adjustment. For data from censuses there are no sampling errors. For data from administrative sources, there are also no sampling errors, but there are additional problems due to mismatching of administrative concepts or classifications to statistical requirements.

An aspect of accuracy, commonly referred to as *reliability*, is the closeness of the initially released values to the subsequent values of data releases. In this context it useful to consider the sources of revision, which include (1) replacement of preliminary source data with later data, (2) replacement of projections with source data, (3) changes in definitions or estimating procedures, and (4) updating of the base year for constant-price estimates.

The accuracy of the data produced by ITU is largely determined by the accuracy of the data received from the contributing organisations. ITU activities can improve accuracy; for example, quality checks may detect errors in data provided by contributing organisations and lead to improvements in these data. Alternatively ITU activities can have an adverse effect, for example by introducing errors during the processing stages.

3. Credibility

The *credibility* of a data output refers to the confidence that users place in that product based primarily on their image of the data producer and the product, i.e., the *brand image*. It is based on the users' *perceptions* of accuracy as well as the actual accuracy.

Credibility is built over time. An important aspect is trust in the objectivity of the data. This implies that the data are perceived to be produced professionally in accordance with appropriate statistical standards, and that policies and practices are transparent. In particular, data are not manipulated, nor their release timed in response to political pressure.

Another aspect of credibility is trust in the integrity of the production process. To obtain complete coverage ITU may impute data for missing countries; to improve accuracy it may adjust data received. The extent to which this is well done and well understood affects credibility. Also, once agreement between ITU and an organisation has been reached on how data will be provided or imputed, the agreement should not be subsequently withdrawn in response to political pressure.

4. Coherence

The *coherence* of a data product reflects the degree to which it is logically connected and mutually consistent with other data products. Coherence implies that the same term should not be used without explanation for different concepts or data items; that different terms should not be used without explanation for the same concept or data item; and that variations in methodology that might affect data values should not be made without explanation.

Coherence in its loosest sense implies the data are "at least reconcilable." For example, if two data series purporting to cover the same phenomena differ, the differences in time of recording, valuation, and coverage should be identified so that the series can be reconciled.

Coherence has four important sub-dimensions.

- *Coherence within a dataset* implies that the elementary indicators are based on compatible concepts, definitions, and classifications and can be meaningfully combined. Incoherency within a dataset occurs, for example, when indicator values that should add up to a total do not.
- *Coherence across datasets* implies that the data are based on common concepts, definitions and classifications, or that any differences are explained and can be allowed for. An example of incoherency across datasets would be household ICT usage could not be reconciled with ICT supply. Unexplained inconsistencies across datasets can seriously reduce the interpretability and credibility of ITU statistics.
- *Coherence over time* implies that the data are based on common concepts, definitions, and methodology over time, or that any differences are explained and can be allowed for. Incoherence over time refers to breaks in a series resulting from changes in concepts, definitions, or methodology.
- *Coherence across countries* implies that, from country to country, the data are based on common concepts, definitions, classifications and methodology, or that any differences are explained and can be allowed for. Ensuring coherence across countries, commonly referred to as *harmonization*, is one of the major sources of value added by ITU.

Metadata plays a fundamental role in explaining possible changes in concepts or methodologies over time and across countries.

5. Timeliness and Punctuality

The *timeliness* of a data product is the length of time between its availability and the event or phenomenon it describes. Timeliness is assessed in terms of a time scale that depends upon the period for which the data are of value, i.e., are sufficiently timely to be acted upon. The concept applies equally to short-term or structural indicators, the only difference is the time scale.

Although ITU processes themselves can have an adverse effect, for the most part the timeliness of ITU data products is determined by the timeliness of the data it receives from the contributing organisations.

Punctuality implies the existence of and adherence to a *data product dissemination schedule.* A data product is punctual if it is disseminated in accordance with the schedule. In the case of data published externally the schedule may comprise a set of target release dates or may involve a commitment to release data within prescribed time period. (Here "release date" refers to the date on which the data are first made publicly available, by whatever medium, typically, but not inevitably the web site).

A dissemination schedule assists:

- internal users, by enhancing their capacity to plan their work based on target internal dissemination dates for data they require;
- external users, by improving their capacity to make timely use of ITU statistics;

There may be occasions when ITU simply cannot adhere to the dissemination schedule due to the late acquisition of data from input sources. In such circumstances advance warning regarding the delay in dissemination should be communicated to users.

Although timeliness and punctuality are different concepts they are grouped together in a single quality dimension for two reasons, first because their separate achievements are heavily inter-related in practice, and second to be in line with international standards and practices.

6. Accessibility

The *accessibility* of a data product reflects how readily the data can be discovered, located and accessed from within ITU data holdings. It includes the suitability of the formats in which the data are available, the media of dissemination, the availability of metadata and user support services, and, in the event that there is a charge, the affordability of the data to users.

From the perspective of data availability, ITU users are divided into two very distinct groups: internal users; and external users. Typically, because of the differences in access methods, internal users can access data earlier and in more detail than external users. Thus these two groups may have quite different perceptions of accessibility.

The range of different external users leads to the need for multiple dissemination formats and selective presentation of metadata. A publication policy should be articulated and made publicly known.

7. Interpretability

The *interpretability* (sometimes called *clarity*) of a data product reflects the ease with which users can understand and properly use the data. The degree of interpretability is largely determined by the adequacy of the definitions of concepts, target populations, indicators and other terminology describing the data, and its limitations.

If there are several dissemination mechanisms they should be harmonised in order to avoid confusing users.

Coping with the needs of the broad range of external users leads to the use of metadata presentation in layers of increasing detail. The content and format of published products should be adapted to the different target groups.

Balancing Data Quality Dimensions

The data quality dimensions are not mutually exclusive in the sense that there are relationships between the factors that contribute to them. Factors leading to improvements with respect to one dimension may result in deterioration with respect to another. Thus, in designing a data collection and products, it is often necessary to trade-off quality in one dimension with quality in another. The most significant trade-offs to consider are as follows.

- *Accuracy and timeliness.* This is probably the most frequently occurring and important quality trade-off. Improvement in timeliness can be obtained by terminating data acquisition earlier and compiling products based on a smaller number of countries and/or reduced editing. However, as this reduces accuracy, there needs to be a trade-off. For major products a compromise is to disseminate a preliminary version of the data product based on partial acquisition and then one or two revised products based on successively more acquisition and editing. The size of the revisions between preliminary and revised products is an indicator of degree of accuracy that is being sacrificed in order to produce the increased timeliness.
- *Relevance and accuracy*. Relevance can be increased by acquiring more data items, but accuracy may be diminished because the additional data are less reliable. Conversely elimination of inaccurate data items will increase accuracy but reduce relevance.

- *Relevance and timeliness*. Timeliness may be improved by reducing the number of data items collected or by replacing those that are difficult to collect by ones that are easier. This will have a negative effect on relevance.
- *Relevance and coherence*. Improvements in relevance, for example by redefining the indicators for which data are collected, or moving to a later version of a classification, will reduce comparability over time, perhaps to the point of requiring a series break. Conversely, the desire to retain comparability over time may inhibit changes in content required to improve relevance.
- *Accuracy and coherence*. Improved methods may increase accuracy but reduce coherence by introducing changes in data that are attributable to changes in methods not in what is being measured. Conversely, the desire to retain coherence may inhibit the changes required to improve accuracy.

Process Quality Dimensions

8. Sound Methodology

Sound methodology refers to the use of international standards and best practices through all stages of a data collection from identification of requirements, through design, data collection, processing, analysis, dissemination, archiving and evaluation. Application of standards and best practices not only engenders ITU process and product quality, it fosters comparability across organisations and countries.

Sound methodology includes both theory and its application in the sense of ensuring that, not only are procedures well designed, but also they are well implemented and documented, and that staff are well informed and trained.

9. Sound Systems

Sound systems refers to the use of international standards and best practices in systems development, including liaising with systems developers in other statistical organisations and making optimum use of off-the-shelf or shared statistical products where available.

Sound systems also includes both theory and its application in the sense of ensuring that systems are well designed, developed, implemented and documented, and that staff are well trained in their use.

10. Cost-Efficiency

The *cost-efficiency* with which data products are produced is a measure of the costs incurred and resources expended relative to the benefits of the products. The aim is to produce a specified set of products of specified quality at minimum cost.

Efficiency can affect all dimensions of product quality in the sense that, if a product can be produced more efficiently with the same quality, then the resources released can be used to improve any dimension of the quality of that product or other products, or to create new products.

Two types of costs are incurred:

- the costs to ITU of acquiring, processing and disseminating the data;

- the costs incurred by the NSOs and NRAs from which the data are acquired. These costs depend significantly on whether or not these national organisations collect the data for their own purposes. If they do, then the costs are essentially those of repackaging and transmitting data already collected. Otherwise they are the full costs of data collection.

ITU never collects data directly from basic units (operators, households, and individuals) to which the data refer and that provide the original elementary data. Thus their costs in responding are of only indirect concern.

ITU acquires ICT data for EU member countries from Eurostat. Eurostat is collecting this information in pursuance of an EU regulation and thus little or no additional burden is imposed by ITU.

4. Quality Guidelines

Introductory Remarks

The guidelines are presented in two broad groups: those applying to any *individual statistical production process* within ITU; and those applying to the ITU *statistical infrastructure.*

For the purposes of presenting guidelines for a statistical production process, the various activities associated with the collection are subdivided into eight subgroups corresponding to the *phases* defined in the *Generic Statistical Business Process Model (GSBPM),* published by UNECE, which is the international standard. The groups have been slightly expanded to include certain closely related statistical infrastructure activities, namely managing *provider relations, managing user relations and assuring quality*

The remaining activities associated with the statistical infrastructure are presented in three subgroups, namely, *assure methodology impartiality, objectivity and transparency; manage metadata,* and *manage human, financial and technological resources.*

Even though not explicitly referenced, staff training is part of every subgroup.

In principle, for each subgroup of statistical activities thus defined, the guidelines should include the following:

1. *Scope* – a short description of the statistical activities to which the guidelines refer;
2. *General guidelines* – statements of best practice reflecting the aims of the guidelines in general terms;
3. *Detailed guidelines* – covering all aspects of quality and performance to be addressed;
4. *Monitoring mechanisms* – the methods by which adherence to the guidelines might be monitored, including quality and performance indicators and quality assessments; and
5. *Reference documentation* – documents that elaborate the guidelines and/or that were instrumental in their formulation.

This first version of the DQAF contains only Scope and General Guidelines.

The General Guidelines may also be used as a self-assessment checklist. The degree to which current activities correspond to each statement of best practice can be assessed on a five point scale:

- +2: completely

- +1: for the most part
- 0: not applicable
- -1: to a limited extent
- -2: not at all

The scores can be aggregated across each activity group, and over all groups, to give summary quality scores that can be monitored over time, or compared with other statistical production processes and products assessed using the same checklist.

1. Specify Needs and Manage Users

Scope

This group includes all activities associated with engaging users to identify their statistical needs, proposing strategic options for meeting these needs, and, if required, preparing a business case for changes to the production process and products. It includes: maintaining a knowledge of, and good relationships, with users, and examining their statistical needs in detail; checking the extent to which current statistical production process and other data sources meet these needs; defining the general content of new or changed statistical products to meet the needs; and preparing a business case where needed to secure approval and resources to re-engineer the production process and/or produce new or changed statistics

General Guidelines

ITU has:

- recently determined the needs, or changes in needs, for the statistical data products it produces;
- categorised the users and potential users by type of use and data needs;
- reviewed user requests, queries and comments with a view to identifying new or changing user needs.
- identified the most important, say 20, individual users and discussed their needs and specified in detail the corresponding data requirements;
- established memoranda of understanding, service level agreements or equivalent for provision of data to key users;
- conducted user satisfaction surveys on a regular basis;
- identified the relevant populations and indicators for which data are required and the appropriate definitions and classifications;
- evaluated data currently available from other sources and determined the extent to which the data requirements can be met with data from these other sources;
- prepared the business case for development, or substantial revision, of the statistical process and products;
- consulted with key users on proposed changes to statistical products;
- shared information about new or revised statistical products with ITU divisions thereby maximizing the possibility of data coherence and minimizing the risk of duplication of effort and waste of resources;

- provided opportunities to interested staff in different divisions to contribute to the development or redevelopment of the data production process;
- maximised use of existing data sources before embarking on data additional collection;
- facilitated sharing of ICT related data between divisions within ITU;
- ensured the coherence of all ICT related data acquired and produced by ITU;

2. Design Statistical Production Process and Statistical Infrastructure

Scope

The statistical activities in this group follows from, and build on, the results of specifying user needs. They include the research and design work needed to define or redefine, as needed, the statistical concepts and indicators, and the data collection, processing, storage and dissemination procedures required to produce the envisaged statistical products. They include specification of the metadata that are inputs to and outputs from to the production process, and general design of the systems and workflows that support and enable the efficient conduct of the procedures.

General Guidelines

ITU has:

- defined and justified the scope and content, i.e., countries to be included, indicators and classifications;
- identified and used the most appropriate concepts, definitions, and classifications, where possible taking advantage of those already developed and used by the international statistical community;
- identified and reviewed all possible available data sources and selected the most appropriate;
- designed efficient and effective data collection systems and procedures, including making formal agreements with providers where required;
- designed efficient, effective and integrated data processing procedures and systems for coding, verifying, imputing, estimating, integrating, and disseminating data products, using internationally accepted methods to the fullest extent possible;
- designed statistical products, including tables, datasets, databases and analyses in accordance with the specified data requirements;
- identified and included all the metadata required to support data collection and production and to inform users;
- involved experts in data collection and processing, and in ICT statistics, wherever available;

3. Build Statistical Procedures and Systems

Scope

The statistical activities in this group cover building procedures and systems to support the production process and infrastructure and testing them prior to production. They include building the data collection tools, data repositories, processing tools, management control functions, and metadata management tools. They include configuring workflows to handle data

transformations from acquisition to archiving. They include producing documentation and training production staff, providers, and users in use of the systems and procedures.

General Guidelines

ITU has:

- built efficient data collection tools;
- built appropriate and efficient data storage mechanisms, for example using databases rather than worksheets for data storage;
- built efficient and effective processing tools and metadata management tools;
- configured smooth workflows in which all activities within the entire production process fit together efficiently without gaps or redundancies;
- documented all systems and procedures;
- trained production staff, providers and users in use of the systems;
- recently specified and met deadlines for introduction of new statistical products and/or reengineering of statistical processes;
- and if so, tested all new procedures and systems before putting them into production.

4. Collect Data and Manage Provider Relations

Scope

The statistical activities in this group refer to the actual acquisition of data, using the various sources and collection modes specified in the design phase, and storing the data acquired securely in an appropriate repository. They include implementing procedures and systems for data collection from NSOs, NRAs and databases of other international or commercial organisations, including follow-up in the event of non-response or dubious data values. They include liaising with providers, making them aware of the reasons for and specifics of the data required, and responding to their comments, queries and complaints. They include ensuring that the data and associated metadata are loaded into a suitable repository.

General Guidelines

ITU has:

- ensured that NRAs and NSOs are totally familiar with the reasons for data collection and the precise meanings of the indicators to be provided?
- implemented efficient and effective procedures for collecting data and associated metadata from these providers?
- made provision for responding to comments, queries and complaints from these providers?
- monitored and followed up non-responses and partial responses and thus ensured that as much of the required data and metadata are collected?
- where appropriate, implemented procedures to obtain data and corresponding metadata from the databases of other international or commercial organisations?
- ensured the risk of errors in data and metadata acquisition is minimised?

- minimised the reporting burden on organisations providing data and metadata?
- minimised ITU resources spent in data and metadata acquisition.
- ensured that all information concerning data and metadata flows between any organization within a country and ITU is reported to the country NSO (in view of its role as statistical coordinator for the country);
- implemented procedures that ensure data and metadata are loaded into a suitable repository;

5. Process Data

Scope

The statistical activities in this group refer to the verification, editing, harmonisation and imputation of incoming data and their preparation for analysis. They include integrating data from various sources and classifying and coding them where needed; applying checks that identify missing, invalid or inconsistent data or metadata; imputing missing values for which no data have been received or for which data are inadequate or in error; applying adjustments to harmonise data across countries; deriving values for indicators for which data are directly acquired, compiling totals, averages and ratios for regions, and measures of dispersion; and storing data and metadata in databases from which data can be readily extracted for analysis purposes, statistical products can be compiled, and users internal to ITU can make extracts.

General Guidelines

ITU has:

- implemented efficient and effective procedures and systems for data collection;
- implemented efficient and effective procedures and systems for data, verification, editing, harmonisation, including a full range of validity checks and edits;
- implemented efficient and effective procedures and systems for derivation and imputation;
- implemented efficient and effective procedures and systems for aggregation, including production of totals, averages and ratios for regions, and measures of dispersion;
- implemented efficient and effective procedures and systems for storage of data and metadata in an internal database from which statistical products can readily be compiled and internal users can make extracts;

ITU does:

- minimize the human resources required for data and metadata processing by good workflow management;
- collect, review and analyse all operational metadata, including non-response rates, indicator non-response rates, indicator imputation rates;
- facilitate internal user access to the data;
- facilitate internal user understanding of the data, and their limitations, through provision of metadata describing procedures and operations;
- facilitate integration of the data with data from other production processes and checked their coherence;

6. Analyse Data

Scope

The statistical activities in this group refer to the analysis required for verification of the statistical products and their preparation for dissemination. They include preparing draft statistical products, including associated metadata and quality indicators and checking that they will support the analyses for which they were designed. They include undertaking the analyses required to explain the data. They include scrutinizing, analysing and explaining the data in relation to expectations and identifying divergences from expectations. They include finalising the statistical products including interpretation notes, briefings, measures of uncertainty and other relevant metadata. (In the event that statistical products include data for individual operators or very small groups of operators, the activities also include checking that the data did not breach any relevant confidentiality rules.)

General Guidelines

ITU does:

- prepare and review draft data products and scrutinize, analyse and explain the data in relation to expectations;
- compare the data with data for previous reference periods, and confront the data with any related data from other sources;
- view the data from all other perspectives and ensured there is an in-depth understanding of the data content before dissemination;
- produce and analyse quality indicators;
- discuss the results with internal experts and otherwise check that statistical products and associated metadata are fit for purpose before external dissemination;

7. Disseminate Statistical Products

Scope

The statistical activities in this group refer to the dissemination of the statistical products to users within ITU and to external users. They include formulating and applying a dissemination strategy, including a release schedule and pricing policy; reformatting the data and metadata as required and loading them into publicly accessible databases, and preparing and disseminating printed publications in accordance with ITU publishing and presentation guidelines; notifying users of the release of the statistical products, giving briefings for key users and user groups; promoting the products to ensure that they reach the widest possible range of users; and managing communications with internal and external users, including ensuring that user queries are recorded and that responses are provided

General Guidelines

ITU has:

- implemented efficient and effective procedures and systems for data dissemination;

ITU does:

- notify internal users and enable their access to statistical products at the earliest possible opportunity;
- prepare statistical products in accordance with relevant ITU publishing and presentation guidelines;
- maximize internal and external interpretability of the data products by accompanying them with appropriate metadata;
- ensure that products are timely and punctual by maintaining and adhering to detailed activity and release schedules;
- disseminate statistical products externally, including preparing printed publications and making databases accessible via the web site;
- give briefings to key users, including senior officials of international and national organisations, large ICT service providers, members of think tanks, academia and the media;
- promote externally products via wikis and blogs and social media to ensure that they reach the widest possible audience, for example;
- manage communications with internal and external users by ensuring that user queries are recorded, that responses are provided;

8. Evaluate and Assure Quality

Scope

The statistical activities in this phase focus on evaluation (assessment) of the statistical process products and infrastructure within the context of data quality assurance framework. They include developing and implementing the framework, including quality principles, dimensions, guidelines and evaluation procedures.

Depending upon the scale of the evaluation, activities may include: defining the evaluation objectives and procedures; establishing the evaluation team; assembling the relevant documentation; analysing the documentation; discussing the statistical process and products with the staff responsible for them and with key users and providers; identifying quality and performance issues; and making recommendations for improving quality and performance.

The evaluation may refer to one specific repetition (cycle, instance) of the process and product, or to a particular set of repetitions, for example the process and all the products during the course of the previous five years. The evaluation may be lightweight or detailed, ranging from self-assessment of the values of key quality and performance indicators for a single repetition to a comprehensive external assessment. In its most detailed form, evaluation involves a complete review of the products relative to the original data requirements and of the production process relative to the original design.

General Guidelines

ITU has:

- reviewed quality concepts and prepared a data quality assurance framework, including quality principles, dimensions, guidelines and assessment checklist;
- defined evaluation procedures and implemented an evaluation programme;
- recently conducted a self-assessment of its production process and/or products;

- recently commissioned an external assessment of its production process and/or products;

In the event that an evaluation has been recently conducted, ITU did:

- set the objectives and scale of the evaluation;
- provide the evaluation team with comprehensive documentation;
- review and discuss the evaluation results;
- based on the results, implement quality and performance improvements;
- based on the results, prepare a business case for quality and performance improvements and submit to senior management for support and resources;

9. Assuring Sound Methodology, Impartiality, Objectivity and Transparency

Scope

ITU needs to be seen to collect, produce and disseminate statistics in a manner that is based on sound methods and that is professional, transparent, impartial and objective.

General Guidelines

ITU has:

- made known the commitment to follow professional standards in collecting, producing and disseminating statistics that are impartial and objective, including by publicizing its data quality assurance framework;
- developed and made public a data dissemination policy;
- developed and made public a release calendar in which dissemination dates and times are pre-announced and ensure that any deviations from the calendar are announced and justified to the users;

ITU does:

- base the recruitment and promotion of statistical staff on aptitude and expertise in statistics;
- ensure that statistics are produced on an objective basis, determined only by statistical considerations;
- select sources, concepts, methods, processes and dissemination practices on the basis of statistical considerations, and national and international principles and best practices;
- explain major changes in the methodologies and data revisions to users;
- ensure that statistical releases are clearly distinguished from political/policy statements and issued separately from them;
- ensure that statistical releases and statements made at press conferences are objective and non-partisan;
- correct errors in data outputs as soon as possible after they are detected and inform users;

10. Manage Metadata

Scope

In broad terms metadata are *data about data*. In the specific context of quality assessment, metadata are data about every aspect of the IDS statistical production process and its products. They are divided into four broad groups:

- *Data-related metadata* are those metadata that directly describe the data as they are input and output during the various phases constituting the production process. Examples are descriptions of rows and columns in output tables, and descriptions of the contents of a database
- *Definitional metadata* are the metadata that describe the concepts and definitions used in the production process. Examples are definitions of indicators, description of target and actual populations, question wordings in questionnaires.
- *Procedural metadata* are those metadata that describe the particular procedures that make up the phases of the production process. An example of procedural metadata is the specification of the automated verification rules applied during data entry. Procedural metadata are further divided into two types: *active* procedural metadata drive a procedure in the sense that the procedure cannot commence without them; and *passive* procedural metadata simply document the procedure
- *Operational metadata* are metadata that describe the inputs and outputs of a procedure, other than the actual data. Operational metadata are also further divided into two types. *Input operational metadata* comprise metadata that *enable and control* the execution of each particular instance of a procedure. *Output operational metadata* comprise metadata that *result from the execution* of a particular instance of a particular process. They include information to be passed to a later activity. They also include *process metrics*, also called *paradata*, that are generated during the various production phases. Such metadata are the source of quality and performance and quality indicators.

Activities associated with managing metadata include identifying and classifying the various types of metadata of interest; determining how they are to be managed, i.e., obtained, recorded, accessed and used; building appropriate procedures and systems to do so; and analysing them in order to better understand and improve the production process and products.

General Guidelines

ITU has:

- identified and fully documented the various types of metadata, and the needs for and uses of them;
- designed, built and operated a comprehensive metadata management infrastructure;
- appointed a single, authoritative registration authority for each metadata type;
- registered metadata using a registration process that is well documented so there is clear identification of ownership;
- ensured that metadata are active to the greatest extent possible, thereby ensuring they are accurate and up-to-date, and paving the way for automation;

- ensured that passive metadata are recorded at the time they are created, preferably automatically as a by-product of the processes that generate them;
- ensured that there is a single copy of each metadata value, which is entered once and can be accessed or superseded, but not overwritten, earlier values being retained to allow historical access;
- constructed different views of metadata corresponding to the differing needs of the various users;
- reused metadata wherever possible rather than recreating them.

11. Manage Human, Financial and Technological Resources

Scope

Managing resources involves ensuring that the financial, human, and technological resources available for statistical work are adequate both in magnitude and quality, and are sufficient to meet the needs associated with the development, collection, production and dissemination of statistics.

General Guidelines

ITU does:

- make maximum use of ICT in supporting the statistical process;
- measure and analyse the work effort involved in each phase of the production process;
- design and conduct statistical activities in each phase such a way as to promote efficiency;
- ensure user feedback is taken into account in determining priorities when resources are limited;
- ensure that the financial and human resources available match the statistical activities envisaged;
- review resource allocations on a regular basis.

5 Quality Assessment and Improvement Programme

Introductory Remarks

Quality has to be managed through a comprehensive quality assessment and improvement programme that draws on and puts into action the quality principles, dimensions and guidelines described above. The programme comprises three types of quality assessment:

- monitoring quality and performance indicators for each repetition of each statistical production process;
- annual or biennial quality self-assessment of each production process or component of the statistical infrastructure;
- external quality review of a particular process or infrastructure component on an occasional basis.

These types of assessment differ from one another in their aims, the amount of detail and effort involved, the frequency with which they are conducted and the sort of recommendations they

may generate. In all cases the goal is to improve quality. So there has also to be a process for ensuring that all recommendations arising from an assessment are implemented where this can be done with existing resources or are advanced to senior management for consideration if addition resources are required.

Monitoring Quality and Performance Indicators

The objectives of identifying and monitoring quality and performance indicators (QPIs) are to quickly check ongoing operations, to monitor performance with respect to target objectives, and to identify sources of operational errors and correct them.

QPIs monitor statistical operations in terms of quality (i.e., effectiveness) and performance (i.e., efficiency). They may be divided into two groups:

- *product QPIs* - monitoring output indicators and analyses;
- *process QPIs* - monitoring all phases of statistical processes and infrastructure;

QPIs for the DQAF have to be developed by the staff with intimate knowledge of the statistical activities to be monitored. They must be very carefully chosen. Too few QPIs, or the absence of QPIs for key procedures or outputs, result in ineffective monitoring. Too many QPIs, or ill-chosen ones, overload the production procedures and are a waste of resources.

The procedures involved in development and use of QPIs are:

- define a preliminary set of QPIs;
- designate selected QPIs as being *key* and set targets for each of these;
- analyse the values of process and product QPIs for each repetition of each statistical process;
- take immediate action to address the *operational* problems thereby identified; and
- document *structural* problems, i.e., problems that cannot be solved at operational level, and provide them as input to the next quality self-assessment, and, if serious, let them trigger a quality self-assessment.

Quality Self-Assessment of a Statistical Production Process and Its Products

A self-assessment of quality is conducted by the staff responsible for the statistical process on an annual basis. Its objectives are to help the staff responsible to develop an impression of the quality of the process and products, and hence to identify structural weaknesses and to propose quality improvements. It involves:

- Assembling documentation about the process and its products;
- Convening one or more meetings with the staff responsible for all aspects of processing, and at these meetings reviewing the documentation, completing the relevant sections of the checklist, and identifying process and product weaknesses and potential improvements;
- Convening one or more meetings with the principal users and at these meetings reviewing the products, completing the relevant sections of the checklist and identifying product weaknesses and potential improvements; and
- Taking action on any improvements that can be implemented with existing resources and documenting improvements that would require additional resources and/or support from other areas.
- Presenting a summary of the results to senior management.

Self-assessment is facilitated by use of a self-assessment checklist. The checklist should cover all aspects of the production process from identification to dissemination each viewed in terms of the quality dimensions. As previously noted, the quality guidelines

External Quality Assessment

External quality assessment is appropriate on a regular but infrequent basis or if concerns about quality of products or processes reach a high level. The assessment objectives are to provide Division Head and ITU senior management and with an objective view of the quality of the statistical production process, and hence to identify any structural weaknesses, to propose quality improvements to address them, and to indicate the resource implications.

An external quality assessment involves the following steps:

- Defining the terms of reference for the assessment;
- Identification of the assessment team, involving an external expert, with the manager of the data collection as a resource person;
- Obtaining documentation and the results of recently completed self-assessments;
- Convening meetings with relevant ITU staff to further elaborate the problem areas and improvements required to address them;
- Convening meetings with the principal users and further investigating the problem areas as reflected in product weaknesses;
- Reporting the results of the assessment to ITU management in accordance with the terms of reference.

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/6-E
19 November 2014

English

SOURCE: ITU

TITLE: Final Report - 2nd Meeting of the ITU Expert Group on ICT Household Indicators (EGH), Geneva, 15-16 September 2014

2nd Meeting of the ITU Expert Group on ICT Household Indicators (EGH)

Geneva, 15-16 September 2014

Final Report

1. The 2nd Meeting of the ITU Expert Group on ICT Household Indicators (EGH) took place in Geneva, Switzerland on 15-16 September 2014. The meeting was held back-to-back with the Expert Group on Telecommunication/ICT Indicators (EGTI) meeting that took place from 17-18 September. A joint session of the two expert groups was held in the morning of 17 September. In addition, a half-day session on information sharing was held on 19 September where a number of international organizations and a representative from Google presented their work related to ICT measurements.

2. The EGH meeting was attended by 85 participants representing national statistical offices, ministries, regulators and international and regional organizations, as well as the private sector. ITU staff was also present during the meeting and acted as the secretariat of EGH. The meeting was chaired by Mr Alexandre Barbosa, head of the NIC.br's Center of Studies on Information and Communication Technologies (CETIC.br), Brazil.

3. The topics discussed in the 2nd EGH meeting were those identified by the 11th WTIS (held in Mexico City in December 2013). In particular, the agenda covered indicators on individuals using the Internet by type of portable device and network used, measuring Internet security, measuring child and youth online protection, mobile phone ownership, mobile phone activities, barriers to Internet use by individuals, ICT employment and future work of the EGH.

4. The meeting discussed the contributions received in the EGH online forum between July 2013 and 10 September 2014. The discussions focused on the suggested indicators, their relevance and definition and whether to include them in the future (2015) ITU data collection. The meeting further discussed the future work of EGH, identifying topics that required further discussion as well as new areas related to ICT measurement.

5. The EGH meeting agreed on the definition and response categories of the indicator on **individuals using the Internet by type of portable device and network used to access the Internet.** The meeting decided to include all Internet use irrespective of location of use and purpose of Internet use. It was suggested that the terminologies used in the response categories of the indicator should be adapted at the country level to better reflect the local context and facilitate the responses. In addition, it was agreed that the discussion in the forum related to this item will be closed and that the indicator will be included in the ITU ICT household questionnaire 2015.

6. The EGH meeting recognized that **Internet security** is relevant for policy making but difficult to measure. The meeting agreed keeping the discussion open in the EGH forum for experience sharing and reviewing the indicator and its definition in the future, based on the experiences shared. It was suggested that if countries want to collect the data, they can use the proposed questions available in the EGH forum or those used by other countries (such as the EUROSTAT model questions). It was further suggested that national Computer Incident Response Team (CIRTs), which exist in many countries, could be used as an alternative data source for compiling information on Internet security. The meeting decided that the indicator will not be included in the ITU ICT household questionnaire.

7. The EGH meeting further recognized that **child online protection (COP)** is relevant for policy making but difficult to measure. The meeting highlighted that collection of data for this topic is different from other ICT surveys since the sample is designed for a different in-scope population (for children using the Internet). In addition, collection of data for this topic requires a stand-alone survey, which will allow collecting more questions related to the topic. The meeting agreed keeping the discussion open in the EGH forum for experience sharing and to review the indicator and its definition in the future, based on the experiences shared. It was suggested that if countries want to collect the data, they can use the proposed questions available in the EGH forum or those used by other countries (such as the EU Kids Online survey and the ITU COP Statistical Framework document). The meeting decided that the indicator will not be included in ITU ICT Household questionnaire.

8. The EGH meeting recognized that the indicator on **proportion of individuals who own a mobile phone** is relevant for policy making and for the analysis related to gender. The meeting agreed on the following definition of ownership: *An individual owns a mobile cellular phone if he/she has a mobile cellular phone device with at least one active SIM card for personal use*. Mobile cellular phones supplied by employers that can be used for personal reasons (to make personal calls, access the Internet, etc.) are included. Individuals who have only active SIM card(s) and not a mobile phone device are excluded. Individuals who have a mobile phone for personal use that is not registered under his/her name is included.

9. The meeting suggested to provide a clarification on the reference period for “active SIM card”. It was decided to use three months as the reference period to be consistent with other core ICT indicators on ICT usage. The meeting suggested replacing “own” with “have” in the question to clarify that mobile phones not paid by the person who has the device are counted. It was emphasized that this indicator is different from mobile phone users and mobile cellular subscribers. The definition of mobile ownership will be refined in the EGH forum and the discussion will be closed by end of October, with the final definition to be presented in the WTIS to be held in Georgia in November 2014. The meeting decided that the indicator **proportion of individuals who own a mobile phone** will be included in the ITU ICT household questionnaire 2015.

10. The EGH meeting discussed the indicator **proportion of individuals using a mobile phone, by type of activity**. Several countries are collecting this indicator for their national policy making needs. Several countries also expressed concern on the usefulness of collecting this indicator at the international level due to overlap with HH9 (Proportion of individuals using the Internet, by type of activity), HH10 (Proportion of individuals using a mobile cellular telephone) and the new indicator on portable devices. The meeting decided that the discussion related to this indicator in the forum will be closed and that the indicator will not be included in the ITU ICT household questionnaire. Countries can collect the data for the indicator for their national policy making needs.

11. The EGH meeting discussed the indicator **proportion of individuals not using the Internet, by type of barriers**. Many countries highlighted that this indicator is important for policy making. However, the meeting highlighted that the response categories suggested on "barriers for individuals not using the Internet" should take into consideration HH14 (barriers to Internet access by households). The meeting suggested that the discussion on the indicator and the response categories should continue in the EGH online forum. It was further suggested to review in the future the response categories of HH14 to ensure that they refer only to household barriers.

12. ILO presented the progress of the work related to measuring **ICT employment** and highlighted that there is an increasing demand for data from both the industry and policymakers. However, there is lack of comparable data at the international level. The meeting highlighted the need to assess the availability of disaggregated data by sex at more detailed level for both employment in the ICT Sector and ICT occupations. It was further highlighted that there is a need to continue the methodological work on ICT occupations and employment using ICT skills and tools, under the leadership of ILO. The meeting agreed that the discussion related ICT employment will be opened in the EGH forum. Countries are encouraged to share their experience in collecting the data on ICT occupations at detailed level of ISCO and to participate in the discussion related to the methodology, definition and classifications.

13. During the joint EGH/ETI session, the consultant to the ITU presented the ongoing work to assess the data quality of ITU statistics and to develop a data quality assurance framework (DQAF), which was welcomed by participants. The meeting highlighted the importance of assessing data quality of ICT statistics both at international and national levels. The meeting agreed that a discussion item on data quality will be opened in the EGH discussion forum. The draft report on data quality assessment and DQAF will be posted in the EGH forum before presenting to the WTIS2014.

14. The meeting discussed the future work of the EGH. A large number of suggestions were made by experts. They include 1) continuation of discussions on indicators that were discussed during the meeting (such as ICT employment, proportion of individuals not using the Internet, by type of barriers, and review of HH14); 2) experience sharing related to measuring Internet security; and children and youth online protection; 3) data quality; 4) methodological issues related to data

collection (tools, processes, standards, vehicles, challenges), data verification/processing, and data dissemination (visualization, open data); 5) other data sources (big data); 6) national coordination (between NSOs, ministries, regulators); 7) barriers to mobile phone ownership; and 8) other areas (SDGs/post2015, digital inclusion, e-waste, e-government, e-health, ICT in education).

15. The meeting agreed that the outcomes of the EGH meeting will be presented by the Chair of the EGH to the WTIS 2014, to be held from 24 to 26 November 2014 in Georgia.

12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/7-E
19 November 2014

English

SOURCE: ITU

TITLE: Final Report - 5th Meeting of the ITU Expert Group on Telecommunication/ICT Indicators (EGTI), Geneva, Switzerland, 17-18 September 2014

5th Meeting of the ITU Expert Group on Telecommunication/ICT Indicators (EGTI)

Geneva, Switzerland, 17-18 September 2014

SUMMARY

1. The 5th Meeting of the ITU Expert Group on Telecommunication / ICT Indicators (EGTI) took place in Geneva, Switzerland, on 17-18 September 2014, back-to-back with the 2nd Meeting of the Expert Group on ICT Household Indicators (EGH), which was held on 15-16 September 2014.

2. There were 92 participants attending the meeting, including experts from regulators, ministries and national statistical offices from 48 countries, as well as AHCIET, América Móvil, ECTEL, the European Commission, Eurostat, Google, GSMA Intelligence, Intel, OECD, UIS, UNCTAD and Telefónica. ITU staff was also present during the meeting and acted as the secretariat of EGTI. The meeting was chaired by Mr Iñigo Herguera, National Markets and Competition Commission, Spain.

3. The topics discussed in the 5th EGTI meeting were those identified by the 11th WTIS (held in Mexico City in December 2013) as requiring further work, as well as long standing issues (such as the ICT Development Index). In particular, the agenda covered the separation of subscription data on individuals and public and private organizations, indicators on international Internet bandwidth, the revision of the classification of broadband services, M2M indicators, LTE indicators, data on bundled telecommunication services and the revision of the list of indicators included in the ITU World Telecommunication/ICT Indicators Long Questionnaire.

4. The meeting discussed the inputs received in the EGTI online forum in 2014. The discussions focused on the suggested indicators, their definitions and whether to include them in the future (2015) ITU data collection. The meeting further discussed the future work of EGTI, identifying topics that required further discussion as well as new additions to the list of indicators.

5. EGTI agreed that the indicator "Percentage of fixed-telephone subscriptions that are residential" would be kept as a measure of the split between residential and non-residential fixed-telephone subscriptions. For mobile services, EGTI concluded that there were fewer differences between organization and residential subscriptions and no specific indicator was envisaged at the international level. EGTI acknowledged that fixed-broadband plans for organizations had different features than most residential plans, although there is a grey area when it comes to micro and small organizations. Even if these differences are not fully captured, EGTI highlighted that data on fixed-broadband subscriptions for organizations were relevant and feasible to collect. To that end, EGTI agreed to define an indicator on subscription and/or bandwidth data for organizations through the online forum by the deadline of 15 November 2014, with a view to presenting a proposal at the 12th WTIS. Finally, EGTI noted that the separation between public

and private organizations was difficult to make based on a common international definition, and agreed to leave the topic open for discussion in the EGTI for the next working period.

6. EGTI decided that ITU would discontinue the collection of the two indicators on leased lines – "Revenue from leased lines" and "Leased-line subscriptions" – because they were not relevant anymore at the international level.

7. EGTI noted that measurements on international Internet bandwidth remained relevant in many countries and data for different bandwidth metrics were collected. EGTI agreed to refine the definition of lit/equipped capacity and used capacity through the online forum by the deadline of 15 November 2014, with a view to presenting a proposal at the 12th WTIS. Based on the refined definitions, data will be collected on a trial basis in the ITU World Telecommunication/ICT Indicators Questionnaires in 2015, with the understanding that each country will report the international bandwidth indicator(s) that are feasible to collect nationally.

8. EGTI agreed that the classification of broadband services would be based on a fixed/mobile categorization, rather than the previous wired/wireless classification. EGTI noted that data reported for the sub-categories of active mobile-broadband subscriptions (i.e. standard and dedicated mobile-broadband subscriptions) were not fully harmonized at the international level, and decided that the discussion on the most relevant sub-categories of active mobile-broadband subscriptions would be carried over into the next working period of EGTI.

9. EGTI noted that data on M2M subscriptions from cellular operators were already being collected in several countries. Although EGTI acknowledged that the Internet of things (IoT) spanned beyond M2M services offered by mobile operators, EGTI decided to start the data collection by proposing a harmonized definition of M2M subscriptions from mobile operators. The definition will build upon the OECD definition of M2M and the input received during the 5th EGTI meeting, and it will be finalized through the online forum by the deadline of 15 November 2014, with a view to presenting a proposal at the 12th WTIS.

10. EGTI agreed that data on the coverage of LTE and other advanced mobile technologies would be collected starting from the ITU World Telecommunication/ICT Indicators Long Questionnaire 2015. Based on the ITU standards, the concept of "LTE and other advanced mobile technologies" will be clarified in the EGTI forum in order to complete the definition of the indicator. In view of the methodological difficulties of collecting subscription data on advanced mobile-broadband technologies, EGTI decided to continue this discussion for the next working period.

11. EGTI agreed to collect subscription data for bundled services starting from the ITU World Telecommunication/ICT Indicators Long Questionnaire 2015 for the following bundled services: (i) Mobile voice and mobile broadband; (ii) Fixed broadband and fixed telephony; and (iii) Fixed telephony, fixed broadband and pay TV. It was clarified that bundled services would be counted under the respective bundle indicator (as one) AND separately under each category for the

services included in the bundle. The definition of "bundled" will build upon the following principles: (i) includes two or more services (fixed voice, mobile voice, fixed broadband, mobile broadband, pay TV); (ii) is marketed as a single offer, with a single price for the set of services included in the bundle, with a single invoice; (iii) is subscribed under conditions that cannot be obtained by adding single play offers together. The definition of bundle will be finalized through the EGTI online forum by the deadline of 15 November 2014, with a view to presenting a proposal at the 12th WTIS.

12. EGTI reviewed the list of indicators included in the ITU World Telecommunication/ICT Indicators Long Questionnaire and decided to discontinue the collection of the following indicators: Percentage of fixed-telephone subscriptions in urban areas; Percentage of localities with telephone service; Fixed (wired) Internet subscriptions; Domestic Internet bandwidth; Faults per 100 fixed-telephone lines per year; Percentage of fixed-telephone faults cleared by next working day; Complaints per 100 mobile-cellular subscriptions; Complaints per 100 fixed (wired)-broadband subscriptions; Complaints per 100 mobile-broadband subscriptions; SMS international; VoIP traffic; Revenue from fixed-telephone services; Revenue from leased lines; Fixed telephone service investment; Annual investment in fixed (wired)-broadband services; Annual investment in mobile communication services; and Leased-line subscriptions.

13. EGTI agreed on proposing the following topics for discussion in 2015: the revision of the 2011 ITU Handbook for the Collection of Administrative Data on Telecommunications/ICT; additional sources for administrative data (big data, OTT); indicators on fixed-broadband services in public and private organizations (carry over); sub-categories of mobile broadband subscriptions (carry over); and subscription data on advanced mobile-broadband technologies (LTE and plus, carry over).

14. The ICT Development Index (IDI) was discussed in a joint session with the Expert Group on ICT Household Indicators (EGH). EGH and EGTI discussed the feasibility of producing an indicator on unique mobile phone subscribers, and welcomed the efforts undertaken by the Unesco Institute for Statistics (UIS) to improve the IDI Skills sub-index. UIS presented two indicators: (i) Learner-to-computer ratio; and (ii) Gross enrolment ratio in programmes with computer-assisted instruction.

15. EGH and EGTI agreed that the discussion on the IDI would continue in the EGTI forum. New proposals regarding the IDI should address the following points: (i) how would they improve the measurement of ICT developments? (ii) How would they fit in the conceptual framework of the IDI? (iii) Would data be available for the suggested new indicators and, if so, from which source?

16. The EGTI Chair will present the conclusions of the 5th EGTI meeting to the 12th World Telecommunication/ICT Indicators Symposium (WTIS), which will take place from 24-26 November 2014, in Tbilisi, Georgia.