International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Series Y

Supplement 65

(07/2020)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

ITU-T Y.3600-series – Big data adoption in developing countries

ITU-T Y-series Recommendations – Supplement 65

International Telecommunication Union

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899
IPTV over NGN	Y.1900–Y.1999
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Enhancements to NGN	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Network control architectures and protocols	Y.2500–Y.2599
Packet-based Networks	Y.2600–Y.2699
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899
Carrier grade open environment	Y.2900–Y.2999
FUTURE NETWORKS	Y.3000–Y.3499
CLOUD COMPUTING	Y.3500–Y.3599
BIG DATA	Y.3600–Y.3799
QUANTUM KEY DISTRIBUTION NETWORKS	Y.3800–Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	
General	Y.4000–Y.4049
Definitions and terminologies	Y.4050–Y.4099
Requirements and use cases	Y.4100–Y.4249
Infrastructure, connectivity and networks	Y.4250–Y.4399
Frameworks, architectures and protocols	Y.4400–Y.4549
Services, applications, computation and data processing	Y.4550–Y.4699
Management, control and performance	Y.4700–Y.4799
Identification and security	Y.4800–Y.4899
Evaluation and assessment	Y.4900–Y.4999

For further details, please refer to the list of ITU-T Recommendations.

Supplement 65 to ITU-T Y-series Recommendations

ITU-T Y.3600-series – Big data adoption in developing countries

Summary

Supplement 65 to ITU-T Y-3600 series of Recommendations seeks to assess the situation of the use of big data technologies in developing countries based on data collected mainly via a survey that targeted ITU-T members from developing countries.

This Supplement studies the main use cases for big data technologies within organizations from developing countries as well as their requirements, mainly in terms of standards. Challenges for big data deployment as well as the opportunities it can offer, mainly to achieve the United Nations Sustainable Development Goals, were also identified.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y Suppl. 65	2020-07-31	13	11.1002/1000/14384

Keywords

Analytics, applications, barriers, big data, developing countries, opportunities, privacy, requirements, security, skills, standards, sustainable development goals, use cases.

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this publication may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the publication development process.

As of the date of approval of this publication, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this publication. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Introduction

Big data technologies could offer diverse opportunities for developing countries, which are becoming more and more interesting particularly with the rapid growth in data creation.

Big data technologies can derive value from large datasets in ways that were previously impossible. The business community is not alone in approaching the opportunities of big data, but development actors are also realizing the benefits.

In addition to the advantages offered to companies, helping them to improve customer experience, enhance operational efficiency and develop new business models, big data technologies are considered as important tools for development in many sectors (healthcare, agriculture, banking, education, etc.) that could help smartly manage data to offer better quality of life and sustainability for citizens and provide improved services and enhanced convenience in daily life.

However, many challenges should be addressed to achieve the full potential of big data, some of them are technical while the others are specific to developing countries (lack of skills, inadequate technological infrastructure, data protection concerns, etc.).

In this context, this study aims to assess several aspects related to the use of big data technologies in developing countries, including standardization requirements, and to propose some guidelines to help these countries benefit from the opportunities offered through big data adoption while addressing the encountered challenges.

Supplement 65 to ITU-T Y-series Recommendations

ITU-T Y.3600-series – Big data adoption in developing countries

1 Scope

This Supplement specifies the main opportunities, requirements, use cases and challenges related to dig data adoption in developing countries with the objective to identify issues that could be addressed by standards and propose some guidelines that could accelerate the adoption of big data technologies in these countries.

2 References

[ITU-T Y.3600] Recommendation ITU-T Y.3600 (2015), *Big data – Cloud computing based requirements and capabilities*.

3 Definitions

3.1 Terms defined elsewhere

This Supplement uses the following term defined elsewhere:

3.1.1 big data [ITU-T Y.3600]: A paradigm for enabling the collection, storage, management, analysis and visualization, potentially under real-time constraints, of extensive datasets with heterogeneous characteristics.

NOTE 1 – Paradigm is used here to express the necessity to find new approaches to address identified issues and phenomena.

NOTE 2 – Examples of datasets characteristics include high-volume, high-velocity, high-variety, etc.

3.2 Terms defined in this Supplement

None.

4 Abbreviations and acronyms

This Supplement uses the following abbreviations and acronyms:

API	Application Programming Interface
ASCII	American Standard Code for Information Interchange
ASN	Abstract Syntax Notation
ASN-1	Abstract Syntax Notation One
bDDN	Big Data Driven Networking
BDSP	Big Data Service Provider
CAPEX	Capital Expenditure
CoE	Center of Excellence
CIO	Chief Information Officer
CSV	Comma-Separated Values
DAaaS	Data Analysis as a Service
DCAT	Data Catalog Vocabulary

DCIM	Data Centre Infrastructure Management System
DMG	Data Mining Group
DPI	Deep Packet Inspection
GDP	Gross Domestic Product
GPS	Global Positioning System
HPC	High Performance Computing
HTML	Hypertext Markup Language
ICT	Information and Communication Technology
IoT	Internet of Things
IT	Information Technology
M2M	Machine To Machine
MOOC	Massive Open Online Course
NAT	Network Address Translation
NIST	National Institute of Standards and Technology
NoSQL	Not only Structured Query Language
OPEX	Operational Expenditure
PFA	Portable Format for Analytics
PMML	Predictive Model Markup Language
QoE	Quality of Experience
QoS	Quality of Service
SC&C	Smart Cities and Communities
SDG	Sustainable Development Goals
SDO	Standards Developing Organization
SQL	Structured Query Language
VAS	Value-Added Service
XML	Extensible Markup Language

5 Conventions

None.

6 Overview on big data

With the rapid development of information and communication technology (ICT), Internet technologies and services, huge amount of data are generated, transmitted, and stored with explosive growth [ITU-T Y.3600]. Data are generated by many sources and not only sensors, cameras, network devices, web pages, email systems, but also social networks and many other sources. Datasets are becoming so large and complex or are arriving so fast that traditional data processing methods and tools are inadequate. Efficient analytics of data within tolerable elapsed times becomes very challenging. The paradigm being developed to resolve the above issues are called big data [ITU-T Y.3600]. Within big data ecosystem, data types include structured, semi-structured and unstructured data. Structured data are often stored in databases which may be organized in different

models, such as relational model, document model, key-value model, graph model, etc. Semi-structured data does not conform to the formal structure of data models but contain tags or markers to identify data. Unstructured data do not have a pre-defined data model and are not organized in any defined manner. Within all data types data can exist in formats, such as text, spreadsheet, video, audio, image, map, etc. [ITU-T Y.3600].

6.1 Characteristics of big data

Big data is used in many fields, where data processing is characterized by scale (volume), diversity (variety), speed (velocity) and possibly others like credibility (veracity) or business value, if traditional methods and tools are not efficient. These characteristics, usually called "v's", can be explained as follows [ITU-T Y.3600]:

- **Volume**: refers to the amount of data collected, stored, analysed and visualized, which big data technologies need to resolve;
- **Variety**: refers to different data types and data formats that are processed by big data technologies;
- **Velocity**: refers to both how fast the data is collected and how fast the data is processed by big data technologies to deliver expected results.

NOTE – Additionally, veracity refers to the uncertainty of data, and value refers to the business results from gaining new information using big data technologies. Other "v's" can be considered as well.

6.2 Challenges addressed by big data

Taking into account the characteristics of the "v's" described in clause 6.1, big data technologies and services can resolve many new challenges, and can also create new opportunities more than ever before [ITU-T Y.3600]:

- **Heterogeneity and incompleteness**: data processed using big data can miss some attributes or introduce noise into data transmission. Even after data cleaning and error correction, some incompleteness and some errors in data are likely to remain. These challenges can be managed during data analysis.
- **Scale**: processing of large and rapidly increasing volumes of data is a challenging task. Using data processing technologies, the data scale challenge is mitigated by evolution of processing and storage resources. However, nowadays data volumes are scaling faster than resources are evolving. Technologies such as parallel databases, in-memory databases, non-SQL databases and analytical algorithms resolve this challenge.
- **Timeliness**: the acquisition rate and timeliness, to effectively find elements in a limited-time period that meet a specified criterion in a large dataset, are new challenges faced by data processing. Other new challenges are related to the types of criteria specified and the need to devise new index structures and responses to the queries having tight response-time limits.
- **Privacy**: data about human individuals, such as: demographic information, Internet activities, commutation patterns, social interactions, energy, or water consumption, are being collected and analysed for different purposes. Big data technologies and services are challenged to protect personal identities and sensitive attributes of data throughout the entire data processing process, while respecting applicable data retention policies. Positive resolutions of the above challenges open new opportunities to discover new data relationships, hidden patterns or unknown dependencies.

7 Introduction to the survey and to the respondents

In order to collect data for the Supplement, SG13 disseminated a survey on "Big Data Adoption in Developing Countries" in September 2018 [b-TSB Circular 120].

This survey includes four sections, given as follows:

- **Section I** (questions 1.1 – 1.11) on "**Big data usage**", including questions on:
  - The percentage of enterprises having already adopted big data technologies in the respondent's country,
  - The most sector using big data in the respondent's country,
  - The need for big data and its importance for the respondent's organization,
  - The main domains for big data usage,
  - The motivations for the engagement in the big data area and the replacement of traditional solutions for the storage, the processing and the exploitation of data,
  - The types of analysed data/records,
  - The needed skills to deploy big data solutions,
  - The most important data management functions/features,
  - The use of big data technologies in government national projects.
- **Section II** (questions 2.1 – 2.8) is on "**Big data requirements**". The following questions were asked in this section and were essentially about big data requirements in relation to:
  - Data sources,
  - Data transformation,
  - Infrastructure,
  - Data consumption,
  - Security & Privacy,
  - Lifecycle management,
  - Big data skills,
  - Other requirements.
- **Section III** (questions 3.1 – 3.3) focuses on "**Big data standardization**" through questions related to:
  - Awareness regarding the standardization activities of ITU-T on big data,
  - Areas of standardization which are important for developing countries,
  - Urgent standardization areas to work on among those identified in Supplement 40 to ITU-T Y.3600 – Big data standardization roadmap.
- **Section IV** (questions 4.1 – 4.5) aims to collect data on "**Big data opportunities and challenges in developing countries**". The questions included in this section are mainly about the following:
  - The benefits that organizations would like to achieve through big data,
  - The role of big data in achieving the United Nations' Sustainable Development Goals
  - Data challenges addressed with big data,
  - The main barriers that could be encountered in the deployment of big data solutions in developing countries,
  - The scenarios that can spur the deployment of big data in developing countries.

The initial deadline for responses to the questionnaire was 30 November 2018, but since the number of received responses was insufficient, this deadline was extended until 30 August 2019 in order to collect more responses. The questionnaire was available in dynamic PDF format and accessible online via the ITU-T Study Group 13's website.

The total number of received responses was 18; these responses came from different regions (14 responses from Africa, 2 from Asia, one from Latin America and one from Europe) and representing 13 different countries as follows (in alphabetic order):

Table 7-1 – Number of the survey's respondents by country

	Country	Number of respondents
1	Bosnia and Herzegovina	1
2	China	1
3	Iran	1
4	Ghana	2
5	Malawi	1
6	Mexico	1
7	São Tomé e Príncipe	1
8	Nigeria	4
9	Rwanda	1
10	Tchad	1
11	Tunisia	1
12	Uganda	1
13	Zimbabwe	2
	Total	18

39% of respondent organizations have already started using big data technologies; around two thirds of them started using big data technologies since 2015, while the others started using it quite earlier.

Around 39% of organizations revealed that they will start using big data technologies in the upcoming 5 years, while 11% of the organizations do not plan to use big data technologies , because of various reasons, mainly the lack of competences and appropriate infrastructures and the lack of awareness on the benefits of big data at the management level.

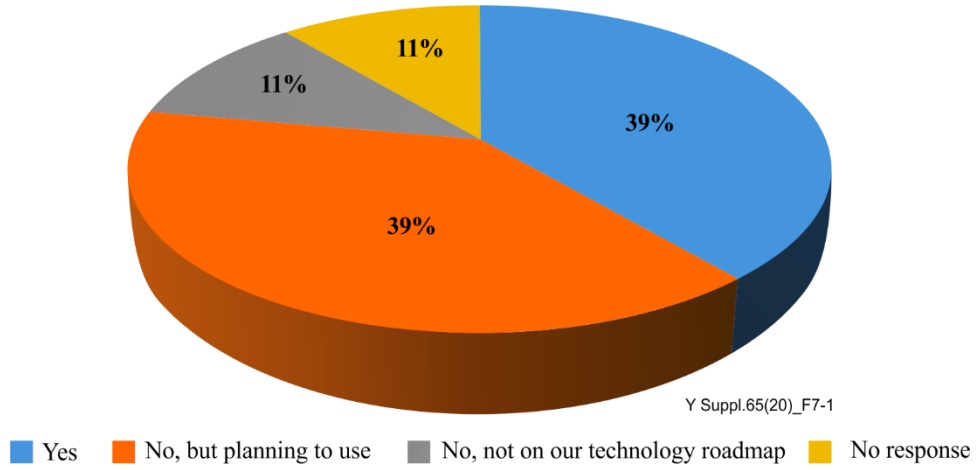

Figure 7-1 – Usage of big data by respondent organizations

8 Big data usage in developing countries

This section covers several aspects related to big data usage in developing countries, mainly the motivation for big data adoption, big data use cases and launched projects at national level.

8.1 Awareness regarding the need for big data usage

The survey's results show that there is a strong need for the use of big data in the developing countries (100% of respondents). 6% of respondent organizations think that the use of big data is mission-critical for their work while 89% of respondents think that it is important or quite important for their work.

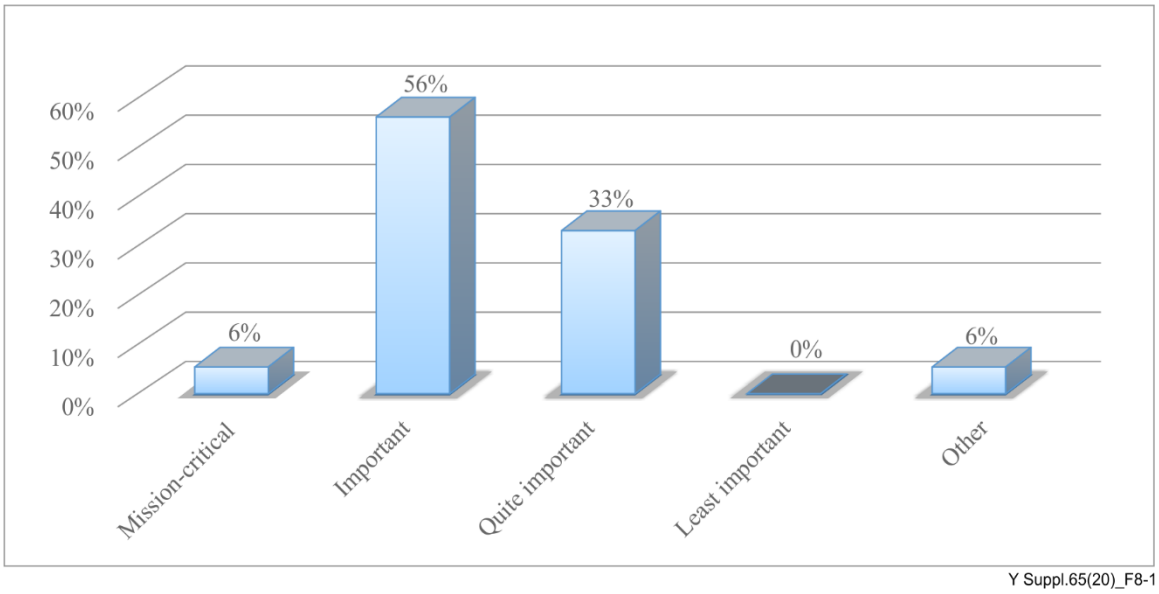

Figure 8-1 – Importance of the use of big data for respondent organizations

Different reasons where behind the decision of organizations to get engaged in the big data area and replace their traditional solutions for the storage, the processing, and the exploitation of data.

The first reason according to respondents is the need for analysis-driven requirements instead of requirements-driven analysis (61% of responses) as they are conscious that they need to change the way that they manage their work and rely on data analysis to better decision making and more accurate definition of strategic orientations.

The second reason is cost optimization (56% of responses) which is a constant goal for most companies whatever their activity sector is. In fact, big data and data analytics can assist with keeping costs down in several ways such as reducing logistic expenses (such as fleet costs and energy usage), avoiding employees' churn, reducing marketing costs (by the use of targeting campaigns for example), minimizing frauds, etc.

Respondents also clarify that they need new and solid approaches to manage the increasing volume of their organization's data (50% of responses) and to benefit from related opportunities.

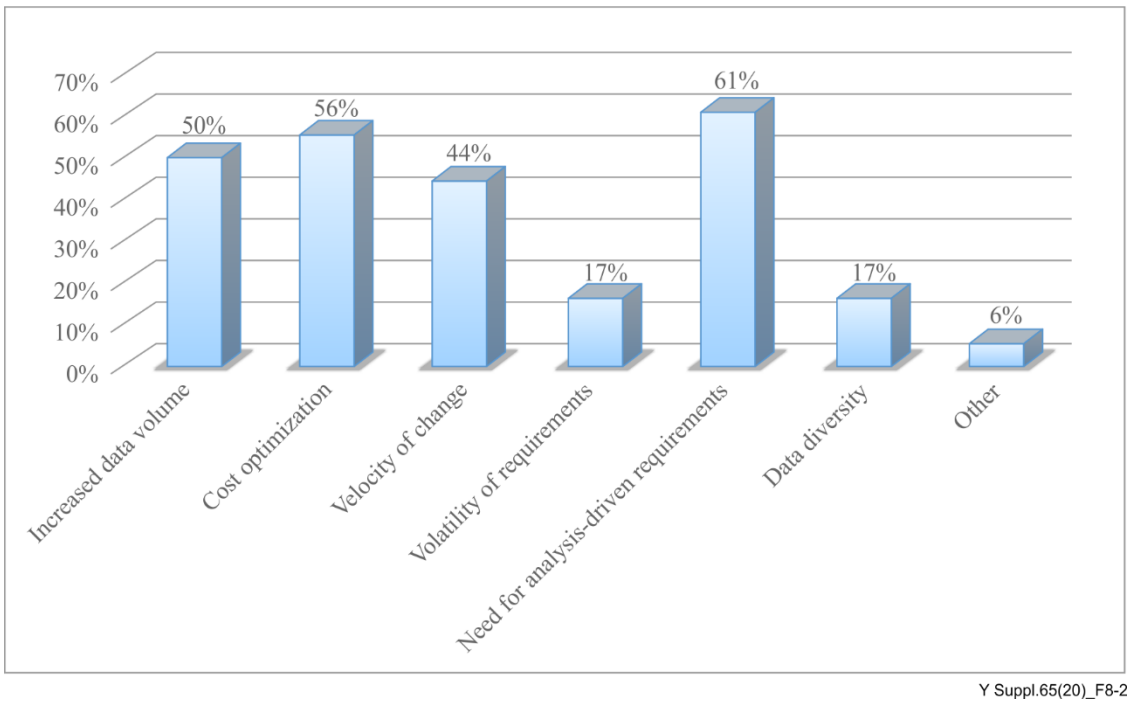

Figure 8-2 – Reasons behind the organizations' engagement in the big data area

8.2 Big data usage and management at organization level

The respondent organizations use big data technologies in many activity domains, mainly in information technology (IT) activities (61% of responses), marketing and customer service (39% of responses) and sales and fraud management activities (22% of responses).

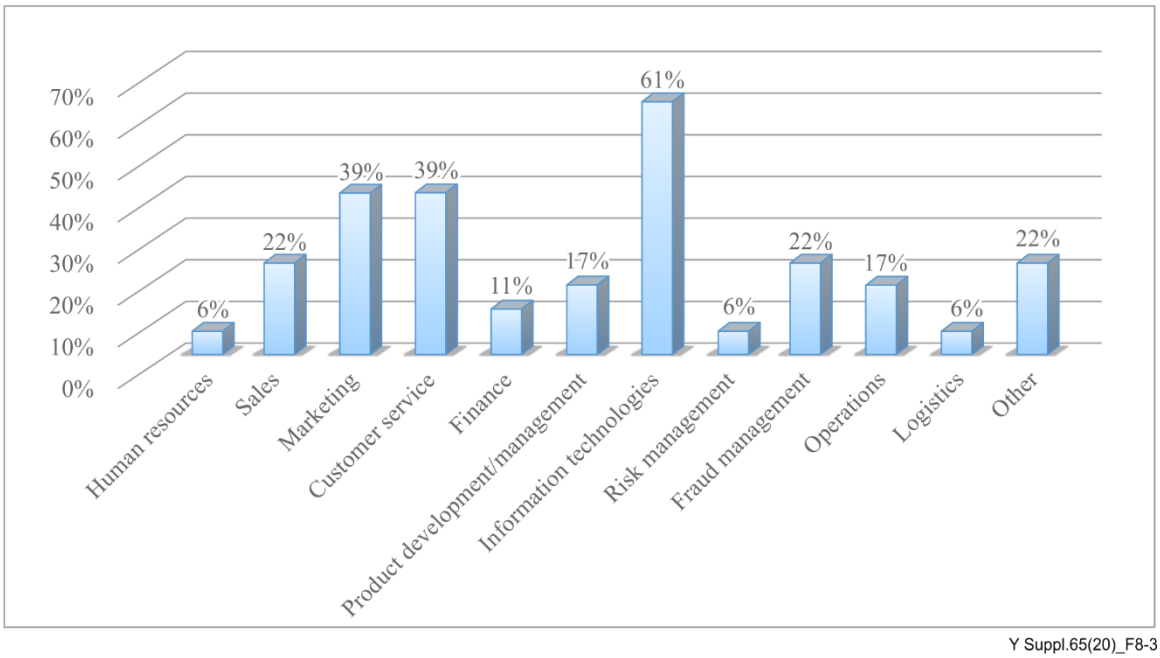

Figure 8-3 – Work areas where big data technologies are used

Other work areas where big data technologies are used**,** were also mentioned such as innovation, utility monitoring and regulation compliance.

Different types of data/records are being analysed by the use of big data technologies, mainly transactional data from enterprise applications (67% of responses), data from social media (56% of responses), and locational/geospatial data (44% of responses).

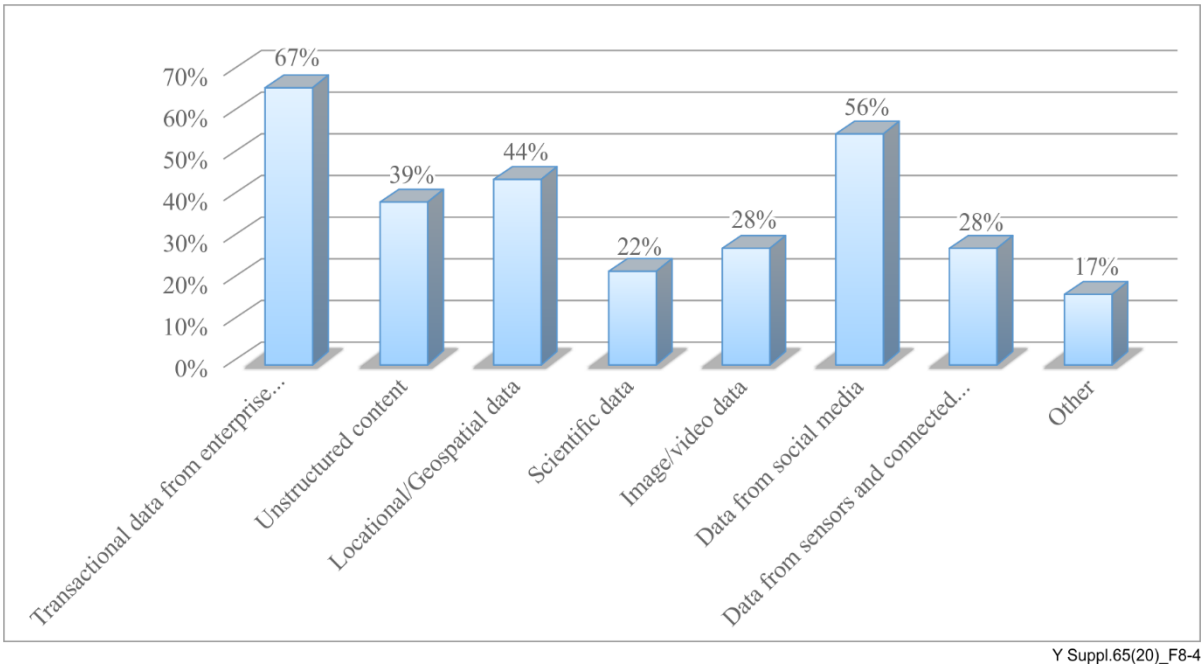

Figure 8-4 – Data records analysed by the use of big data technologies

The other types of data records, which were mentioned in the responses to the survey, were related to economic and financial data such as the industry market share, and data collected from technical platforms and systems such as security logs and event detail records.

The type of data records to be analysed depends on the use case to be implemented. For example, data from social media are mainly used for marketing and customer loyalty purposes.

In order to make different types of data from various data sources available for advanced analytics, Big Data management is needed as it covers the policies, procedures and technologies used for the collection, storage, governance, organization, administration and delivery of huge volumes of Data [b-BD Management]. Developing countries are interested in some data management functions, in particular, the high-performance and inexpensive processing power (67% of responses), the high-capacity and inexpensive storage and the data integration and quality capabilities (56% of responses each).

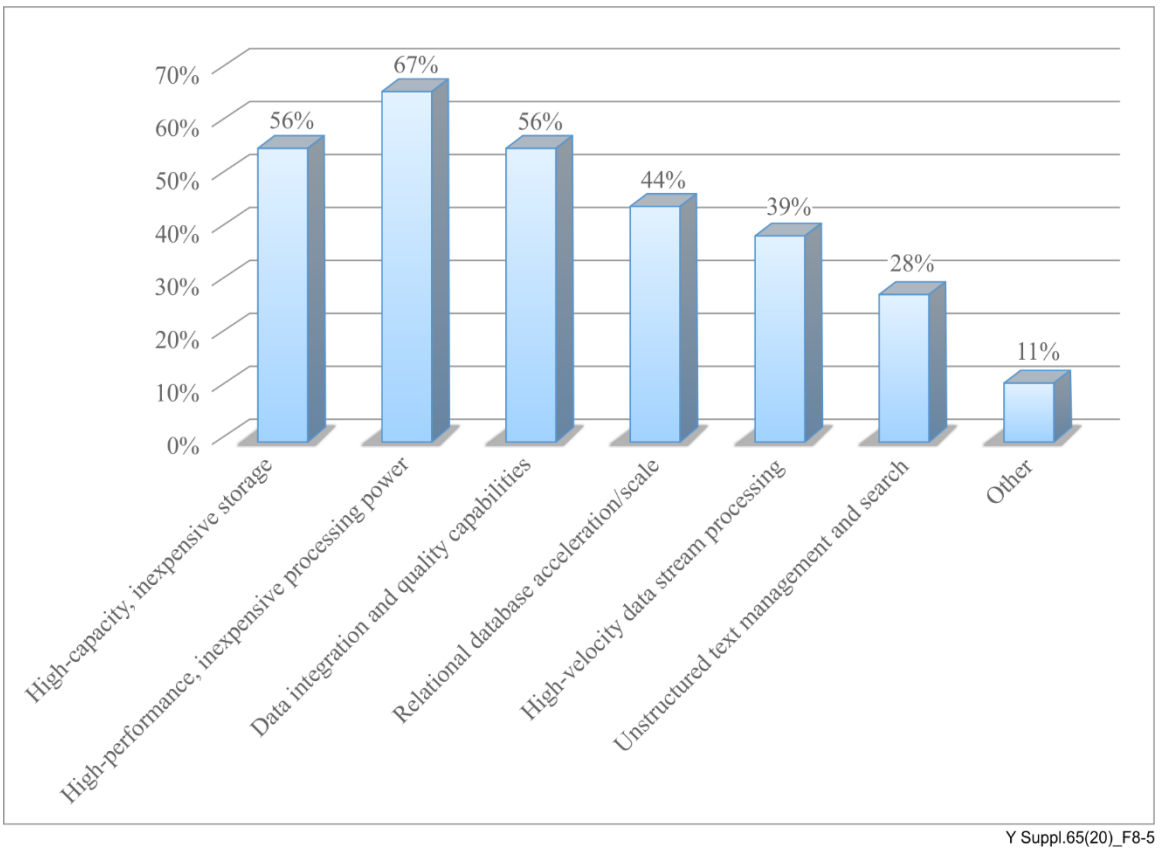

Figure 8-5 – Data management functions

8.3 Big data usage in developing countries: A sectoral perspective

According to the survey's results, respondents estimate that big data adoption in sectors such as utilities, retail and healthcare is still low (11% of responses) and needs to be improved in the upcoming years.

However, the sectors which are using big data technologies the most in the developing countries were the communications sector (56% of responses), followed by the banking sector (39% of responses) then the education and research sector (28% of responses).

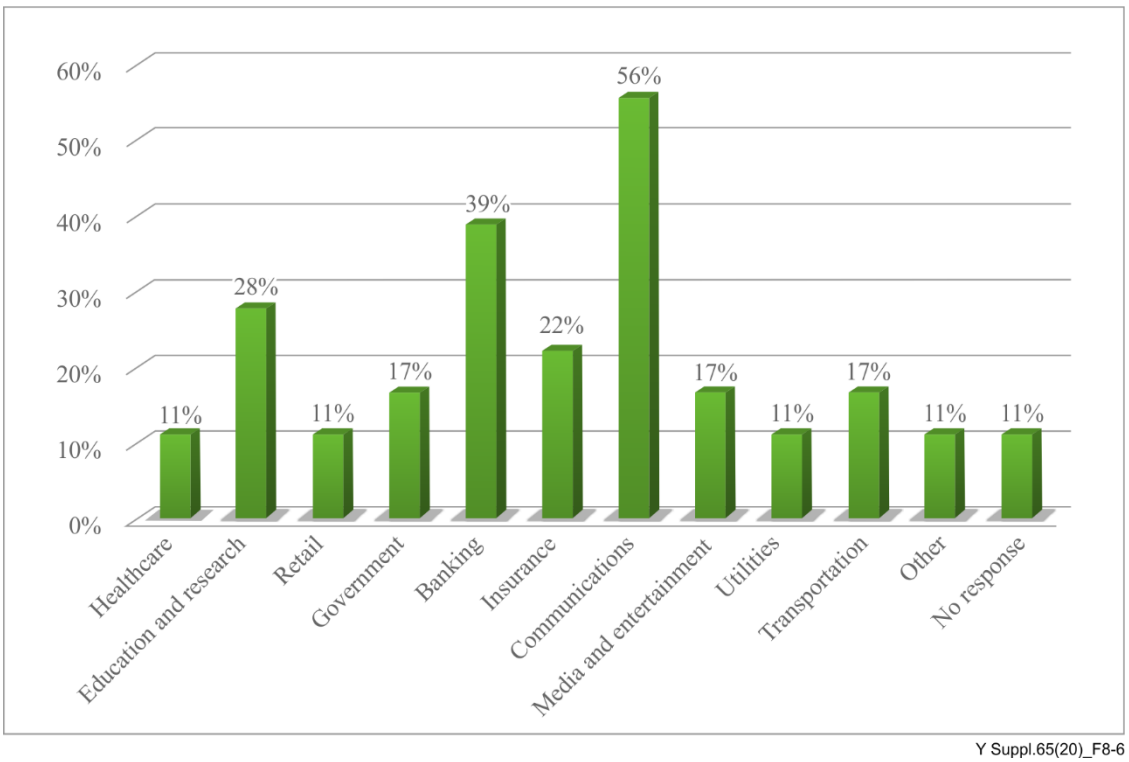

Figure 8-6 – Big data usage by sector

8.3.1 Big data usage in the communications sector

Nowadays, the massive use of mobile broadband services in the developing countries and the growing number of used smartphones is enabling telecom operators to have access to huge volumes of data sources related to customer profiles, customer usage patterns, devices, networks, value-added service (VAS) platforms, location, etc.

Telecom operators in developing countries should be aware of the opportunities provided by an appropriate exploitation and analysis of these exceptional volumes of data, which would help them improve customer experience, optimize networks and create new revenue streams.

In fact, telecom operators could use big data analytics to have a true 360-degree view of their customers, taking into account their demographic details, usage patterns, purchase history, location, etc., and consequently, perform targeted micro-segmentation of their consumer base and provide personalized campaigns, offers and products. Big data analytics is also an efficient tool to provide a proactive customer support and predict and prevent churn which is a painful issue for telecom operators in developing countries with the availability of multiple telecom operators at the national market proposing quite similar offering and service tariffs.

Regarding networks, big data analytics have the potential to assist telecom operators in dimensioning their networks, managing and monitoring their networks' capacities and effectively planning their investments in networks extension and evolution which will strongly assist developing countries in 5G rollouts. Analytics could also help telecom operators improve their networks' performance by managing quality of service (QoS) and identify and fix network issues on a real time basis.

Additionally, big data analytics allow telecom operators to detect frauds in real time and take appropriate actions instantly, as well as to prevent unlawful access and cyber security threats.

Furthermore, the analysis of big data contributes to decrease the capital expenditure (CAPEX) and operational expenditure (OPEX) related to the business operations and it should be noted that telecom operators have found new ways to monetize the available data on their customers and they are starting to provide data analytics as a service (DAaaS) to other businesses and verticals, mainly retail and advertising.

8.3.2 Big data usage in the banking sector

Financial institutions in developing countries are storing large amounts of data on their customers including personal data and account data and they are starting to use them to improve their operations and to become more efficient, more customer-centric and, consequently, more gainful.

Use cases for big data in the banking sectors could include the improvement of fraud detection, as big data could enable financial institutions to analyse customer's data (such as financial transactions, account balances, credit history, payment patterns, customer location, etc.) in order to keep track of any suspicious transactional behaviours and take timely preventive actions.

Besides, big data, coupled with business intelligence tools, automated analysis features and predictive capabilities, could help financial institutions enhance their risk assessment capacities through a better measurement and detection of customer profiles that present higher risk than others [b-BD Banking].

Moreover, the access of financial institutions to in-depth customer profiles make it easier for them to build more robust and longer-lasting customer relationships and thus guarantee better customer retention [b-BD Banking] which is becoming quite challenging for banks taking into account the tough competition from actors from outside the banking sector in developing countries, which provide mobile wallets and mobile payment services. Customer retention could be also ensured by the developing and providing of products, services, and offerings tailored to the specific requirements of customers, based on customer profiles as well as on the better insights and understanding financial institutions are having on each individual customer [b-BD Banking].

Big data could also improve customer service and enable financial institutions, through the collection of customers' questions, suggestions, comments, claims and concerns, to provide appropriate and timely responses to new customer requests [b-BD Banking].

Another use case for big data, which is not only limited to the banking sector is the improvement of the workplace. In fact, using big data could help create a better work environment for employees that foster productivity and innovation and allow monitoring performance metrics, evaluating employees' feedback and company culture, and measuring overall employees' satisfaction [b-BD Banking].

8.3.3 Big data usage in the education sector

In the education sector, big data and analytics technologies are being of great interest as large volumes of data are created on a daily basis by students and teachers.

Big data and analytics technologies have the potential to enhance teaching and learning experiences, improve educational research and progress education governance.

In fact, customized programs for each student could now be created with the help of big data technologies, regardless of the number of students. Big data helps to design intelligent and interactive teaching systems, adapted to learners' personal needs and weaknesses, and creating enjoyable learning experiences. Students are offered the opportunity to attend the classes they are interested in and work at their own rhythm, while still having the possibility of offline guidance by teachers. This is already happening in the case of massive open online courses (MOOCs) that are attracting millions of students worldwide [b-BD Education].

Besides, with the use of big data technologies, student results could be improved and answers to assignments and exams will no longer be the only measurement of students' performance. The analysis and tracking of data generated by each student in real-time will help have a better understanding of the individual behaviour of students and create an optimal learning environment for them. Moreover, the use of big data in the education sector would help reduce dropout rates at schools and colleges in different ways, mainly by improving student results and using predictive analysis to get insights on their future outcomes [b-BD Education].

Predictive analysis could also be used to test the efficiency of a course program before introducing it in the curriculum, and could be used also to help the future students to select the right university and

course based on the monitoring of the way graduated students are performing in the job market [b-BD Education].

Regarding recruiting in universities, institutions can use Big Data technologies to more precisely predict applicants and also analyse the possible factors that affect the application process, which will allow them to adapt their recruitment strategies and allocate resources accordingly. In the other hand, students will be able to analyse information about universities at a global scale, which facilitates the search and application process for international students too [b-BD Education].

So, the use of Big Data in the education sector will maximize the effectiveness of learning in the coming years and will have a positive impact on the society and the economy.

8.4 Big Data Projects in Developing Countries

Most governments in development countries are aware of the importance of Big Data to achieve sustainable socio-economic development and to provide citizens with services and facilities in total transparency and with high quality.

39% of respondents to the survey revealed that some projects based on Big Data have been already launched in their countries. Among these projects:

- Beijing Smart City 2015 in China,
- Projects for the power distribution company in Ghana,
- Voters registration cards and national citizens identity cards in Nigeria,
- E-health projects, E-education projects and elections in Tunisia,
- Biometric electronic voter registration system in 2015 to improve the voter's roll in Zimbabwe,
- High performance computing (HPC) for meteorological analysis and research in Zimbabwe.

22% of respondents mentioned that their countries haven't yet launched projects based on big data but they have projects in roadmap, such as the national registration and SIM card registration in Malawi, and the citizen registration, national identification and electoral commission in Uganda.

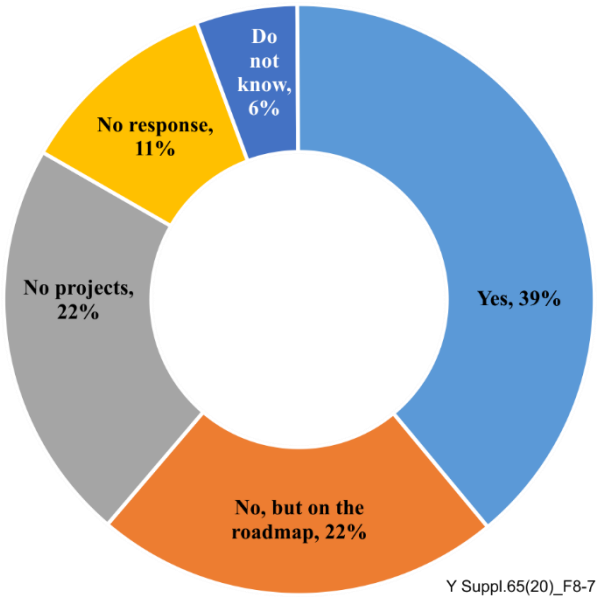

Figure 8-7 – The launch of National Projects based on Big Data

It should be noted that United Nations agencies and groups have launched many Big Data projects for the benefit of developing countries in the context of achieving Sustainable Development Goals.

9 Big data requirements in developing countries

Box 1: Big data for development: preventing the spread of epidemics

As part of the International Telecommunication Union's (ITU) efforts to support its Members States in the area of emergency telecommunications, a big data project was launched in 2015 which was based on a co-financed partnership between ITU and the Government of Japan.

The project showcased the potential of big data to facilitate the timely exchange of information to combat the Ebola epidemic, which had gripped West Africa in 2014, and future health crises. The project used call detail record (CRD) data, which includes information on the use of the mobile phone, including the location, from mobile network operators in Liberia, Guinea, and Sierra Leone. The project demonstrated how analysed CDR data can provide information on human mobility, including cross-border movement, and the spatiotemporal distribution of people, while safeguarding individual privacy. In the case of the outbreak of a disease this information is critical for governments as well as for humanitarian aid agencies, for effective intervention, and to tackle the disease. It can further be used to build models of population flow patterns over time, and at specific events, and to combine these data with other information [b-BD Development].

In the context of adopting big data technologies, respondents mentioned a list of requirements, some of them were general where others are related to big data standardization activities.

9.1 General requirements

According to the respondents, the main requirements of big data in developing countries are grouped as follows:

9.1.1 Requirements related to data sources

Interoperability/compatibility of formats, real-time data, data quality, meaningful sources, constant flow of data, secured digital data sources, granularity of data acquisition, data size, collections of call data records, ASN and plain text ASCII, Media, and web databases (Petabytes, Video, HTML, CSV).

9.1.2 Requirements related to data transformation

Ability to plan and anticipate changes and innovation, diverse sources correlation, security, technical support, data fusion, analytics, data conversion ASN-1 text to ASCII text, batch data analytic processes, and streaming data analytics.

9.1.3 Requirements related to Infrastructure

Server fields, shared infrastructure, advanced software tools based on advanced algorithms, vital infrastructure costs, trade off over on premise vs cloud storage/analysis/reporting, cloud storage, back-up storage, high processing power, high connection bitrate, secured connectivity, Hadoop, spark, Hive, SQL, Cloudera, computing and networking, storage data lake and storage access network, and software/platform tools.

9.1.4 Requirements related to data consumption

Table visualization, information interpretation, user friendly interfaces, self-service analytics and reports, support of different formats, simple presentation, support of structured and unstructured data, business needs, text to SQL tables and databases, as well as output template/formats.

9.1.5 Requirements related to security and privacy

Need to develop related strategy, data desensitization, personal data protection, mostly trust and protection of intellectual property and beyond, well defined data governance procedures, accountability, encryption of data, and access controls management.

9.1.6 Requirements related to lifecycle management

Collection of new knowledge for the added value of the data, pre-analytic conversion, dealing with cleaning of data and authentication, to have the right collection – processing – disposal procedures in place, compliance and governance, data protection, efficiency, data count and data verification, and quality check are necessary.

9.1.7 Requirements related to big data skills

The deployment of big data technologies in any organization needs the availability of some specific skills, which may not be found among the existent organization's staff. For this reason, all the respondent organizations confirmed that they could not totally rely on internal skills to deploy big data technologies, and 56% of them need external expertise from third parties while 44% rely mostly or totally on outsourcing (see Figure 9-1).

Organizations from developing countries will then need to develop big data related skills among their staff or to hire big data experts, but this could incur additional expenses.

According to the survey's results, the requirements related to skills are: advanced math and algorithms, high speed communications infrastructure, machine to machine (M2M), data combining, creation of opportunities and added advantages in the field, Apache Hadoop, SQL, NoSQL, data visualization, machine learning, Apache Spark, quantitative analysis, programming languages, data mining, data protection, data analytics, data management, AI skills, business expertise and computer science.

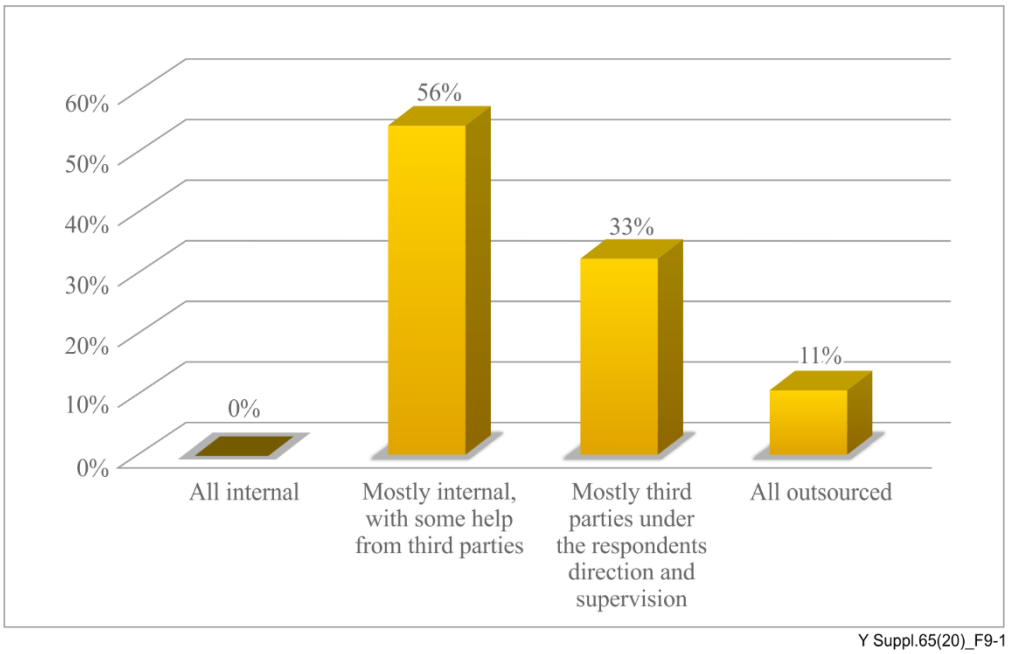

Figure 9-1 – Availability of big data skills

9.1.8 Other requirements

Respondents mentioned some specific requirements that do not fall in one of the previous categories. These requirements are as follows:

– The elaboration of a legal and regulatory framework for big data which ensures data security and privacy and foster big data competition and innovation,

– The acceleration of 5G deployments in developing countries in order to enable and support IoT-based services and generate huge volumes of data collected from sensors and connected devices which will be beneficial in many sectors such as agriculture,

– The necessity to set up a data driven culture to take decisions based on data-derived facts.

9.2 Big data standardization

Clauses 9.2.1 to 9.2.2 give an overview on the main standardization activities on big data inside ITU-T and other standardization bodies and identifies the requirements of developing countries in terms of big data standards.

9.2.1 Overview on big data standardization activities

ITU-T and other standardization bodies worldwide are making remarkable efforts to produce international standards dealing with many aspects of big data and related technologies.

Clause 9.2.1.1 will present the main activities of ITU-T conducted to date on this subject as well as the related work achieved by other standardization bodies.

9.2.1.1 ITU-T activities

The majority of studies on big data within ITU-T are conducted by its Study Group 13 *"Future networks, with focus on IMT-2020, cloud computing and trusted network infrastructures"*. These studies were initiated in the last study period 2013-2016 and continued during this current study period under the following four questions of Working Party 2/13 on *"Cloud Computing & Big Data"*:

- **Q7/13: Big data driven networking (bDDN) and deep packet inspection (DPI)**

  This Question is intended to develop new Recommendations on DPI mechanism for future networks in the emerging application context, DPI performance model, DPI functional requirements for future networks, DPI functional architecture for future networks, framework of data driven networking, requirements of data driven networking and mechanism of data driven networking application.

- **Q17/13: Requirements, ecosystem, and general capabilities for cloud computing and big data**

  This Question is intended to develop new Recommendations for cloud computing and big data definitions, overview, ecosystem, and use cases – data requirements and capabilities – requirements for interoperability, data portability, and exchange information in cloud computing and big data – relationship between cloud computing and big data.

- **Q18/13: Functional architecture for cloud computing and big data**

  This Question is intended to develop new Recommendations for big data functional architectures, including big data exchange functional architecture, and cloud computing based big data architecture.

- **Q19/13**: **End-to-end cloud computing management, cloud security and big data governance**

  This Question is intended to develop new Recommendations for big data governance including data management, data preservation as well as lifecycle management of big data.

Besides, other ITU-T study groups have developed recommendations related to big data or are having related work items under study, each in the context of its mandate as follows:

- **ITU-T Study Group 3 (SG3): Tariff and accounting principles and international telecommunication/ICT economic and policy issues**

SG3 is studying economic and policy aspects related to big data in international telecommunication services and networks including data protection.

- **ITU-T Study Group 5 (SG5): Environment, climate change and circular economy**

SG5 developed Recommendation ITU-T L.1305: Data centre infrastructure management system based on big data and artificial intelligence technology, which was approved in November 2019.

- **ITU-T Study Group 16 (SG16): Multimedia**

SG16 developed Recommendation ITU-T F.743.7: Requirements for big data enhanced visual surveillance services which was approved in May 2019 and its current studies are related to the application of big data in visual surveillance system and assessment framework for big data infrastructure.

- **ITU-T Study Group 17 (SG17): Security**

Through its Question 8/17: "Cloud computing and big data infrastructure security", SG17 is aiming to develop Recommendations for security solutions, best practices and guidelines to big data platform and infrastructure security. SG17 has already developed Recommendation ITU-T X.1147: Security requirements and framework for big data analytics in mobile internet services and is developing a Recommendation on "Security-related misbehaviour detection mechanism using big data for connected vehicles" under Question 13/17 (X.mdcv).

- **ITU-T Study Group 20 (SG20): Internet of things (IoT) and smart cities and communities (SC&C)**

Q1/20 on "End to end connectivity, networks, interoperability, infrastructures and big data aspects related to IoT and SC&C" is focusing on developing Recommendations related to IoT and SC&C big data overview, requirements and ecosystems, including developing standardized efficient systems for data analytics, distributed data computation, real time big data encryption. Moreover, Recommendation ITU-T Y.4114: Specific requirements and capabilities of the IoT for big data was approved in July 2017 under the scope of SG20.

Until July 2020, 12 Recommendations and 2 Supplements on big data were approved. More details on these Recommendations and Supplements are given in Table 9-1.

Table 9-1 – List of ITU-T approved Recommendations and Supplements on big data

Recommendation/ Supplement	ITU-T study group	Date of approval	Description
ITU-T Y.3600: Big data – Cloud computing based requirements and capabilities	SG13 (Q17/13)	2015-11-06	This Recommendation provides requirements, capabilities and use cases of cloud computing based big data as well as its system context. Cloud computing based big data provides the capabilities to collect, store, analyse, visualize and manage varieties of large volume datasets, which cannot be rapidly transferred and analysed using traditional technologies.
ITU-T Y.3600-series Supplement 40: Big data standardization roadmap	SG13 (Q17/13)	2016-07-08	This Supplement provides the standardization roadmap of big data area in the telecommunication sectors and describes landscape and conceptual ecosystem of big data from ITU-T perspective, related technical areas, SDO's activities, and gap analysis.

Table 9-1 – List of ITU-T approved Recommendations and Supplements on big data

Recommendation/ Supplement	ITU-T study group	Date of approval	Description
ITU-T Y.4114: Specific requirements and capabilities of the IoT for big data	SG20 (Q2/20)	2017-07-07	This Recommendation specifies requirements and capabilities of the IoT for big data. It complements the developments on common requirements of the IoT [ITU-T Y.2066] and functional framework of the IoT [ITU-T Y.2068] in terms of the specific requirements and capabilities that IoT is expected to support in order to address the challenges related to big data, and it also constitutes a basis for further standardization work (e.g., functional entities, application programming interfaces (APIs) and protocols) concerning big data in the IoT.
ITU-T Y.3650: Framework of big data driven networking	SG13 (Q7/13)	2018-01-13	This Recommendation specifies the overview and the model architecture of big data driven networking as well as the high-level capability of big data driven networking based on deep packet inspection (DPI).
ITU-T Y.3601: Big data – framework and requirements for data exchange	SG13 (Q17/13)	2018-05-07	This Recommendation provides a framework for data exchange in a big data ecosystem, introducing direct and intermediary exchange patterns. It provides a description of the big data activities for the support of big data exchange by extending the activities defined in [ITU-T Y.3600] and identifies functional requirements which are derived from relevant use cases.
Supplement 50 to ITU-T Y.3650-series: Use case and application scenario for big data driven networking	SG13 (Q7/13)	2018-11-02	This Supplement presents a set of use cases and several scenarios supported by big data driven networking including: 1) network management; 2) network active maintenance; 3)network optimization; 4) network operation; 5) network attack prevention; 6) root cause tracking of QoS; 7) QoE improvement; 8) resource management; 9) network planning and design; 10) traffic engineering; 11) cross layer design; 12) content delivery network. 13) NAT devices detection; 14) bDDN in future network; 15) bDDN in data centre network; and 16) bDDN in industrial Internet.
ITU-T X.1147: Security requirements and framework for big data analytics in mobile Internet services	SG17 (Q7/17)	2018-11-13	This Recommendation analyses the security requirements of big data analytics in mobile Internet services and provide security framework.
ITU-T Y.3602: Big data – Functional	SG13 (Q17/13)	2018-12-14	This Recommendation describes a model and operations for big data provenance and provides the functional requirements for big

Table 9-1 – List of ITU-T approved Recommendations and Supplements on big data

Recommendation/ Supplement	ITU-T study group	Date of approval	Description
requirements for data provenance			data service provider (BDSP) to manage big data provenance.
ITU-T Y.3519: Cloud computing – Functional architecture of big data as a service	SG13 (Q18/13)	2018-12-14	This Recommendation describes the functional architecture for big data as a service (BDaaS). The functional architecture is defined on the basis of the analysis of requirements and activities of cloud computing based big data described in ITU-T Y.3600. Following the methodology of ITU-T Y.3502, the BDaaS functional architecture is described from a set of functional components and cross-cutting aspects.
ITU-T F.743.7: Requirements for big dataenhanced visual surveillance services	SG16 (Q21/16)	2019-05-14	This Recommendation defines requirements for big data enhanced visual surveillance service. It promotes the value of visual surveillance service by using big data analytics method and tools. This Recommendation also provides application scenarios, service requirements, functional requirements, performance, and security requirements for big data enhanced visual surveillance service.
ITU-T L.1305: Data centre infrastructure management system based on big data and artificial intelligence technology	SG5 (Q6/5)	2019-11-13	This Recommendation contains technical specifications of data centre infrastructure management system (DCIM), covering many aspects such as management system scheme, data collection function requirements, operational function requirements, maintenance function requirements, etc.
ITU-T Y.3603: Big data – Requirements and conceptual model of metadata for data catalogue	SG13 (Q17/13)	2019-12-14	This Recommendation describes the general concept of metadata and its utilization in a big data ecosystem and provides requirements and a conceptual model of metadata for data catalogue as well as the XML schema of metadata as an example.
ITU-T Y.3604: Big data – Overview and requirements for data preservation	SG13 (Q19/13)	2020-02-06	This Recommendation provides the overview of big data preservation and its requirements which are derived from the corresponding use cases. It addresses the subjects of overview of big data preservation, functional requirements of big data preservation as well as use cases of big data preservation.

Table 9-1 – List of ITU-T approved Recommendations and Supplements on big data

Recommendation/ Supplement	ITU-T study group	Date of approval	Description
ITU-T Y.3652: Big data driven networking – requirements	SG13 (Q7/13)	2020-06-22	This Recommendation specifies requirements of big data driven networking (bDDN) and it studies general requirements for bDDN, requirements of big data plane for bDDN, requirements of network plane for bDDN, requirements of management plane for bDDN, and interface requirements for bDDN.

9.2.1.2 Activities of other standardization bodies

In addition to the ITU-T activities, some other leading Standards Developing Organizations (SDOs) and industry consortia works on big data-related standards, among them are those listed in Table 9-2.

Table 9-2 – Other standardization bodies working on big data

SDO/Consortium	Areas of big data standardization activities
Joint Technical Committee of the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC JTC 1) https://www.iso.org	Overview and vocabulary, framework and application process reference architecture, standards roadmap, security and privacy, use cases and derived technical considerations, cloud services and devices: data flow, data categories and data use, etc.
World Wide Web Consortium (W3C) http://www.w3.org	Metadata Vocabulary for Tabular Data, Model for Tabular Data and Metadata on the web, Data Catalog Vocabulary (DCAT), The Organization Ontology, Linked Data Platform 1.0, etc.
Organization for the Advancement of Structured Information Standards (OASIS) https://www.oasis-open.org/,	Advanced Message Queuing Protocol, Message Queuing Telemetry Transport
Data Mining Group (DMG) http://dmg.org/	Predictive Model Markup Language (PMML), Portable Format for Analytics (PFA);
TeleManagement Forum (TM Forum) https://www.tmforum.org/	Guidebook for big data analytics describing best practices on big data
Institute of Electrical and Electronics Engineers (IEEE) https://www.ieee.org/	Medical data, Big Data Governance, Metadata Management, etc.
National Institute of Standards and Technology (NIST) https://www.nist.gov/	Big Data Definitions and Taxonomies, Use Case and Requirements, Security and Privacy, interoperability, Reference Architecture Subgroup, Technology Roadmap

9.2.2 Standardization requirements of developing countries

In spite of the consistency on ITU-T work on big data and related technologies since the study period 2013-2016, it was noted that 50% of respondent organizations have no idea on the activities of ITU-T on big data standardization.

39% of respondents mentioned that they are following ITU-T activities on big data standardization including the activities of Q17/13, Q18/13 and Q19/13, ITU-T Y.3600 and ITU-T Y-Sup.40, forums and workshops discussing big data, work on big data quality of service and security, etc.

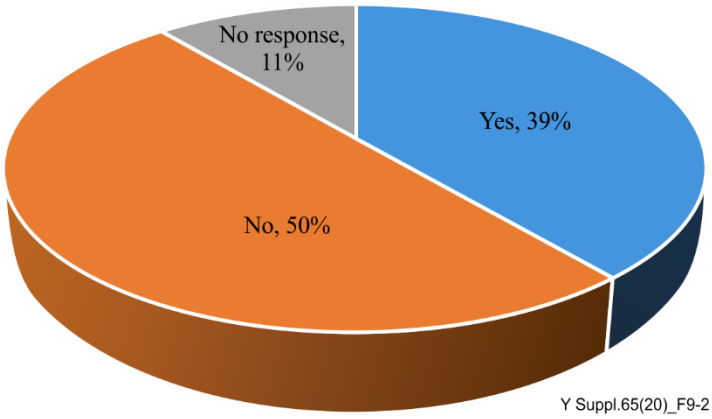

Figure 9-2 – Awareness about ITU-T work on big data

ITU-T would be invited to further promote its activities on big data while developing countries should be encouraged to get involved in these activities. ITU-T is also invited to focus on areas of standardization, which are considered important for developing countries, mainly infrastructure, security, privacy, architecture, and networking, according to the survey's results.

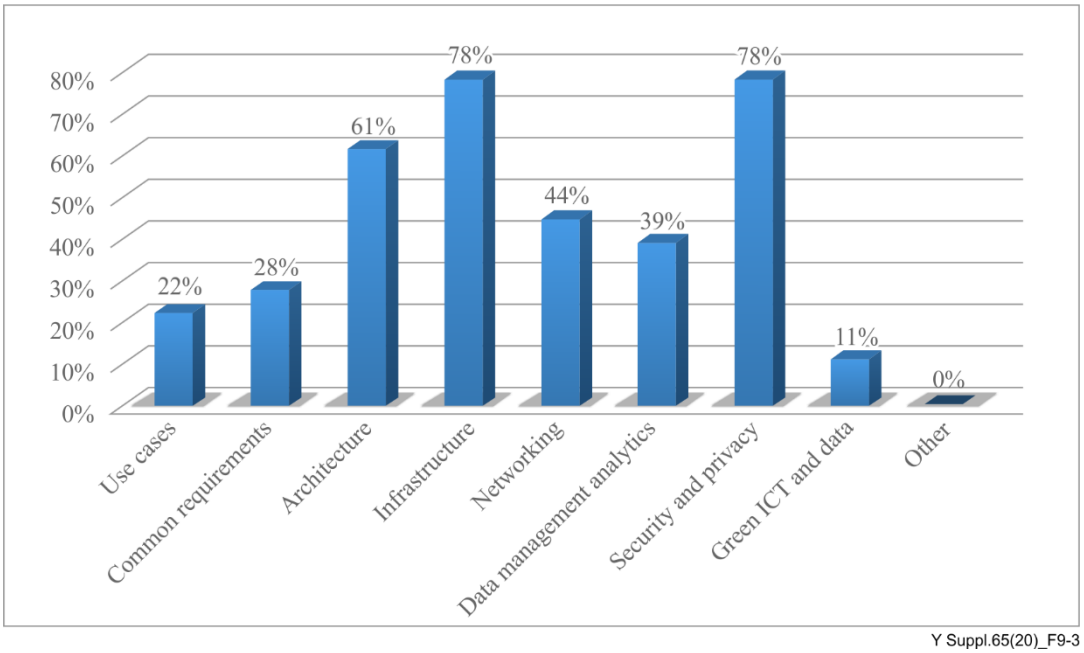

Figure 9-3 – Important standardization areas according to developing countries

Besides, respondent organizations from developing countries consider that working of security and data protection standards should be among the priorities of ITU-T (78% of respondents), as well as legal implications of big data (67% of respondents) and big data exchange (50% of respondents).

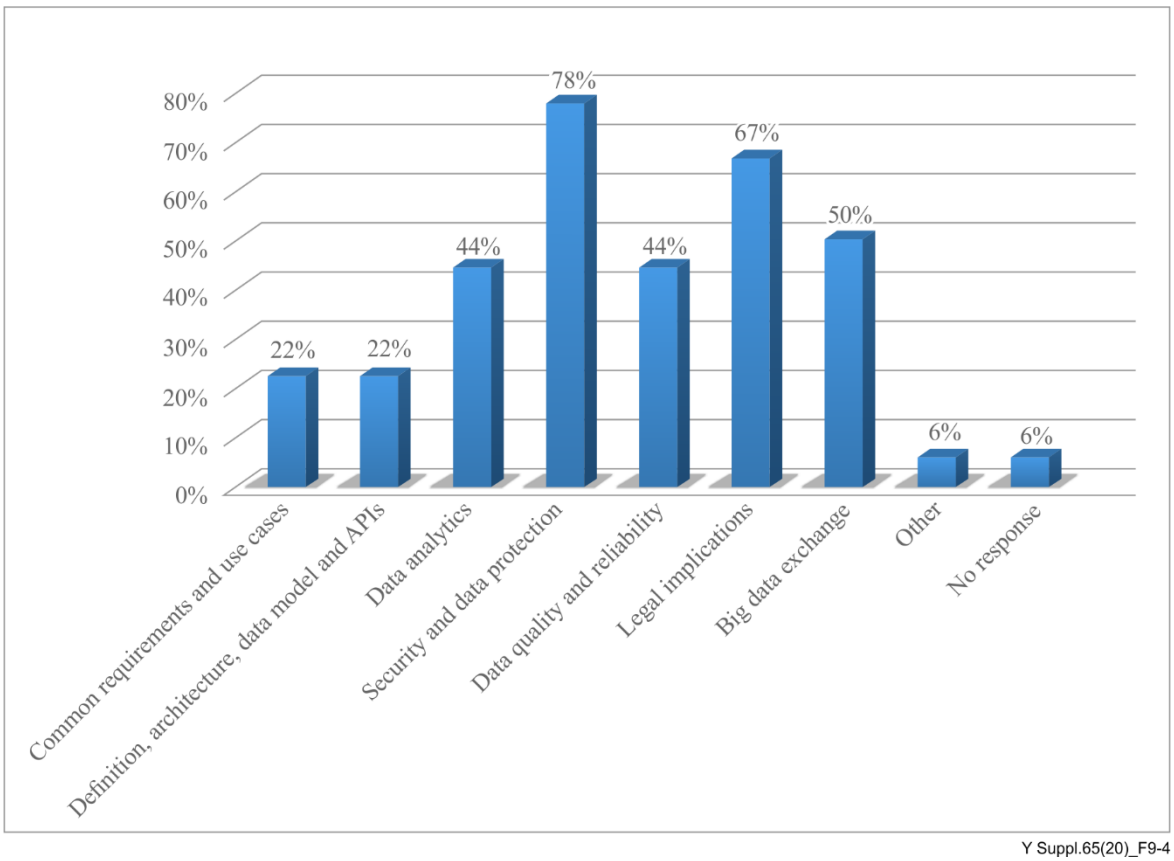

Figure 9-4 – Big data standardization requirements of developing countries

10 Opportunities of big data adoption

This clause highlights the main opportunities of big data adoption from organizations' perspective and showcases how big data and analytics could help achieve the United Nation's Sustainable Development Goals.

10.1 Opportunities of big data adoption at the organization level

The survey's results show that, through the adoption of big data and analytics technologies, organizations from developing countries could achieve many benefits, mainly (from the most to the least ranked).

10.1.1 More efficient operations

With the use of data analytics, companies will be able to streamline their processes, reduce expenses and improve their bottom lines. As they will have a better understanding of the needs of their customers, companies will avoid wasting time in developing products and services or providing contents that will not match the interest of their clients, resulting in monetary savings and improve business efficiency.

10.1.2 Better decision making

As market environments are evolving rapidly, the decision-making process within companies should also evolve to be more flexible and effective. Companies have then to make decisions faster, react quicker, be more predictive and decide collaboratively.

This could be ensured thanks to business analytics, which allows decision makers to better understand the dynamics of their business, anticipate market changes and mitigate risks. Rather than following intuition or instinct when maintaining inventory, pricing solutions, or hiring talents, companies are

embracing analytics and systematic statistical reasoning to make decisions that enhance efficiency, risk management and revenues [b-Decision Making].

10.1.3 Improved customer experience

Companies are able nowadays to collect huge amounts of data on their customers, from purchase history to social media comments through multiple touchpoints. A thoughtful and customer-centric approach to interpret the collected data may lead to improved service practices and robust customer experiences. For instance, contact center metrics such as average handling time and first contact resolution provide data on how customer experience is affected by service practices. The analysis of both quantitative and qualitative feedback from customers helps companies to make improvements to their service, such as using softer language, more emotionally charged marketing campaigns, and more personalized service. Companies are also using big data to streamline processes, save customers time and enhance targeted marketing practices, and they use customer feedback and satisfaction scores to improve communications with customers [b-Customer Experience].

10.1.4 Higher quality products and services

New forms of data and analytics are revolutionizing the area of product quality and safety, like many others business areas.

Businesses are shifting from descriptions of past issue frequencies to predictions of needed asset servicing. Post-manufacturing inspection work is being extended to include product design, sourcing of components, and logistics and distribution. New sources of data – from sensors to call center conversations – are supplementing traditional product inspections. All these innovations are transforming together the management of quality and safety and improving the quality of provided products and services [b-BD Quality].

10.1.5 New product innovations

Through the use of business intelligence tools, data mining, predictive analytics, and other big data tools with traditional market research techniques, companies will have the capacity to collect data on customers' needs and similar or related brands/products, and then develop a proactive approach to new product development. This proactive approach allows them to gain a competitive advantage and guarantee product quality, brand reliability, and marketing efficiency. The same approach could be applicable for the development of entirely new products, the introduction of new product features, the extension of new products or the improvement of existing product lines. Moreover, companies will be able to mitigate the risks related to the development of new products and associated costs [b-New Product].

10.1.6 Increased sales

Sales is one of the areas which is strongly changing by the use of big data, at practically all stages of the sales process, from enhancing the accuracy of prospecting lists to improving sales lead data, increasing sales lead quality, streamlining sales territory planning, discovering sales lead engagement and others [b-Sales].

The way prices are determined, handled, and disseminated through sales networks is also being changed by big data in order to generate the most profits. By using big data algorithms and machine learning techniques, businesses have gained the ability to set the optimal market price for their products and services [b-BD Sales].

10.1.7 Reduced risk

The contribution of big-data and data analytics in improving the risk management practices of companies is becoming more and more significant at many levels. In fact, big data and predictive analytics can help identify potential churn and reduce customer defection based on historical data. Big data makes it possible also to identify potential fraud by analyzing risk factors and locating

unusual behaviour and divergences in a highly streamlined and filtered process. Besides, big data could help companies to adapt to change and adjust their plans according to market conditions thereby mitigating risks, and could also reduce risk for new business as it help predict whether setting up a business at a particular location or for a particular target group will be viable or not. For financial institutions that require risk management more than other companies, big data provides the statistics these organizations need to assess and mitigate financial risks such as credit card fraud, market risk, and asset liability. By using predictive modeling and creating risk-free services based on analytics, financial organizations can maintain business continuity and improve customer satisfaction [b-BD Risk].

10.1.8 Increase customer loyalty and reduce churn

Having a 360-degree view of their customers thanks to data analytics tools, companies can build a personal level of communication with them and enhance their loyalty through addressing issues before they arise, reducing inappropriate or poorly targeted advertising and motivational campaigns and proposing the products and services that meet exactly the customers' expectations [b-BD Experience].

As customers loyalty increase, the risk of churn is reduced. This risk could also be reduced by the use of predictive models of retention, which detects any change in user's behavior and help companies intervene before churn occurs. These analytics can also identify factors that contribute to customer defection, thus helping to reduce and prevent high churn rates [b-BD Risk].

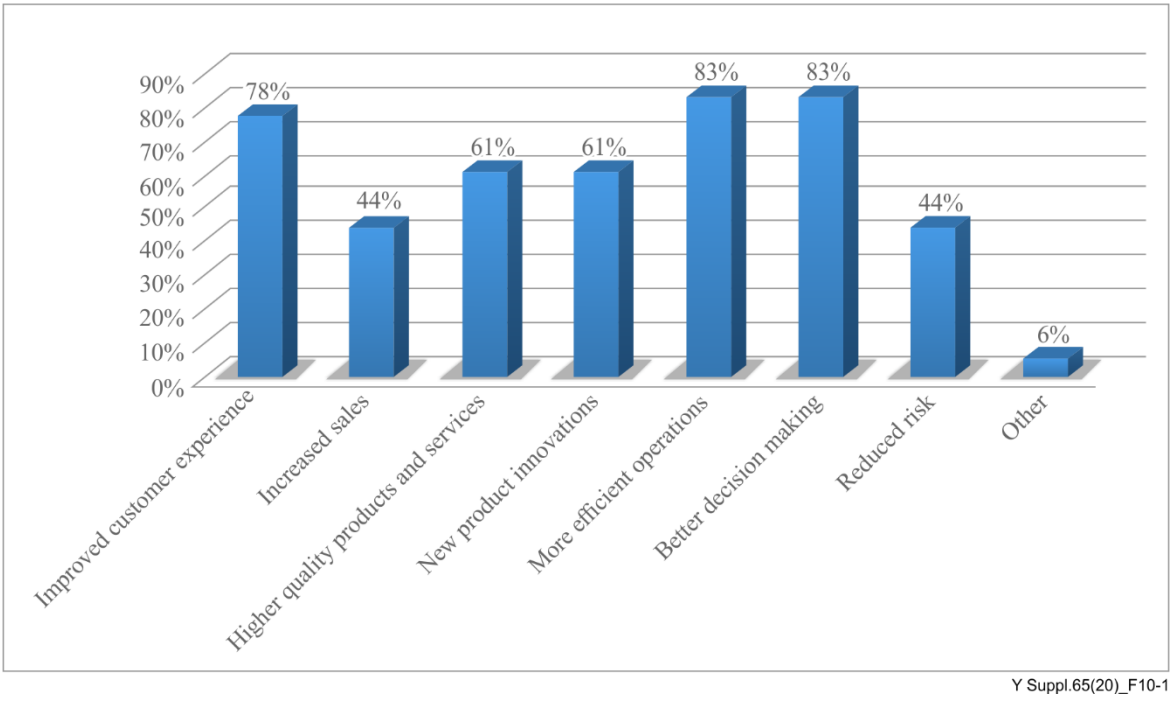

Figure 10-1 – Benefits of big data at the organization's level

10.2 Big data and the United Nations Sustainable Development Goals

The Sustainable Development Goals are the blueprint to achieve a more enhanced and more sustainable future for the world's population. They address critical global challenges, including those related to poverty, inequality, well-being, environmental deterioration, peace, and justice [b-UN SDG]. They also offer definite, time-limited, and measurable targets in synchronization with national development plans and priorities.

83% of the respondents to the survey agreed on the important contribution of big data and big data analytics in achieving sustainable development worldwide.

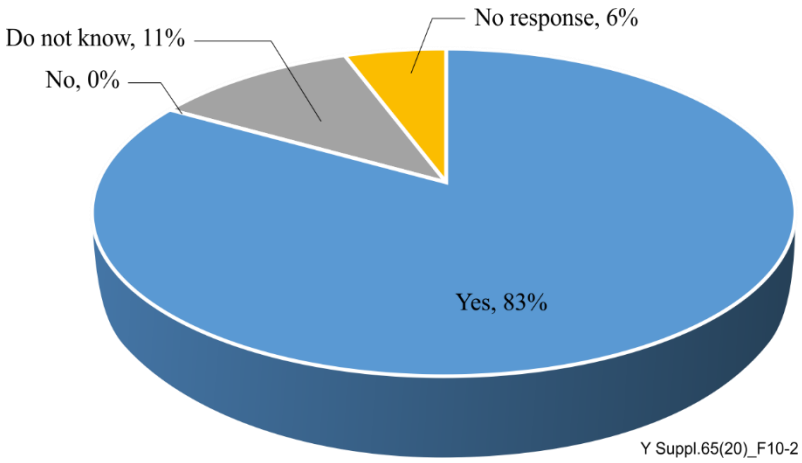

Figure 10-2 – Importance of big data to achieve SDGs

Respondents believe that the new available technologies, the new sources of data (such as call logs, mobile-banking transactions, online user-generated content, satellite images, data collected via IoT sensors, etc.) and the new analytical approaches, if applied responsibly, could enable more agile, efficient and evidence-based decision-making [b-UN BD SDG]. They could also help increase the effectiveness of the implementation of national programs and policies focused on the achievement of the SDGs.

Moreover, the use of big data and analytics could provide a solution to governments that still lack access to adequate data on their entire populations, in order to leave no one behind, and could offer policymaking support for global development in three main ways. First, policymakers will be able to detect anomalies earlier and respond faster to population in times of crisis; second, big data can paint a fine-grained and up-to-date representation of reality that can inform the design and targeting of programs, enabling thus a real-time awareness; and finally, monitoring population in real time makes it possible to understand where policies and programs are failing and to make adequate adjustments [b-SDG development].

The contribution of data science and analytics to sustainable development is becoming more and more obvious. Table 10-1 provides some examples of this contribution in relation to each of the 17 SDGs according to a study of the UN Global Pulse in 2017 [b-UN BD SDG]:

Table 10-1 – Contribution of big data in SDGs [b-UN BD SDG]

SDGs		Examples of big data contribution
1	No poverty	Spending patterns on mobile phone services can provide proxy indicators of income levels
2	Zero hunger	Crowdsourcing or tracking of food prices listed online can help monitor food security in near real-time
3	Good health and well-being	Mapping the movement of mobile phone users can help predict the spread of infectious diseases
4	Quality education	Citizen reporting can reveal reasons for student drop-out rates
5	Gender equality	Analysis of financial transactions can reveal the spending patterns and different impacts of economic shocks on men and women
6	Clean water and sanitation	Sensors connected to water pumps can track access to clean water

Table 10-1 – Contribution of big data in SDGs [b-UN BD SDG]

SDGs		Examples of big data contribution
7	Affordable and clean energy	Smart metering allows utility companies to increase or restrict the flow of electricity, gas, or water to reduce waste and ensure adequate supply at peak periods
8	Decent work and economic growth	Patterns in global postal traffic can provide indicators such as economic growth, remittances, trade, and gross domestic product (GDP)
9	Industry innovation and infrastructure	Data from GPS devices can be used for traffic control and to improve public transport
10	Reduced inequality	Speech-to-text analytics on local radio content can reveal discrimination concerns and support policy response
11	Sustainable cities and communities	Satellite remote sensing can track encroachment on public land or spaces such as parks and forests
12	Responsible consumption and production	Online search patterns or e-commerce transactions can reveal the pace of transition to energy efficient products
13	Climate action	Combining satellite imagery, crowd-sourced witness accounts and open data can help track deforestation
14	Life below water	Maritime vessel tracking data can reveal illegal, unregulated, and unreported fishing activities
15	Life on land	Social media monitoring can support disaster management with real-time information on victim location, effects and strength of forest fires or haze
16	Peace, justice and strong institutions	Sentiment analysis of social media can reveal public opinion on effective governance, public service delivery or human rights
17	Partnerships for the goals	Partnerships to enable the combining of statistics, mobile and internet data can provide a better and real time understanding of today's hyper-connected world

Besides, many indicators of SDGs require classification by location, gender, age, salary, and other relevant dimensions and the collection of the required granular data to monitor all SDGs and targets is not an easy task to accomplish [b-SDG Performance].

In this regard, big data from many sources, such as mobile phones, IoT sensors, social media, online search engines, and satellite imagery, could provide an additional data source to complement statistics from traditional sources such as official statistics and surveys. The new data could add depth and nuance to information on human behaviors and experiences, and when integrated with traditional data, it could produce high-quality information that is more detailed, opportune and pertinent [b-UN BD SDG] and could better measure progress on the SDGs at national and global scale.

11 Challenges of big data adoption

Organizations from developing countries use big data technologies to address some internal issues. These issues are mainly the integration of a wide variety of data (72% of responses), using real time data (67% of responses) and understanding unstructured data (61% of responses).

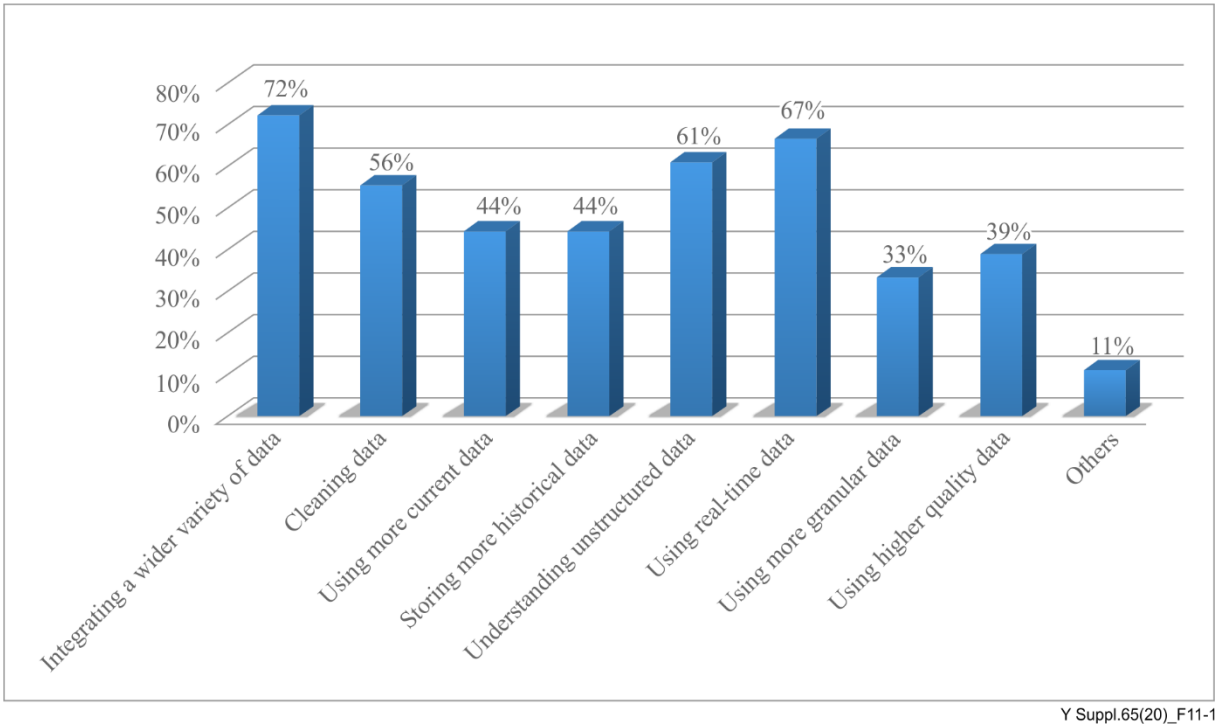

Figure 11-1 – Issues addressed with the use of big data

However, the use of big data technologies is not an easy task for organizations in developing countries that are facing many barriers to adopt big data. These barriers, according to the survey's results, are given in the following clauses.

11.1 Lack of big data experts/skills

The lack of skills is one of the persistent pain points of organizations, mainly when they seek to adopt new technologies such as big data, analytics, and artificial intelligence.

Since 2014 and for four consecutive years, the Harvey Nash/KPMG CIO Survey has found that big data and analytics are top of the skills shortage critical list [b-BD Skills].

The newness of these technologies justifies the shortage of skilled data scientists. Currently, only few universities provide pure data science degrees and it will take some years before there will be a sufficient number of skilled data scientists in the workforce [b-BD Skills].

Organizations, which are planning to adopt big data technologies, should make additional investments to hire big data experts, and/or provide additional training to their IT and development teams or even opt for outsourcing.

11.2 Security and privacy

When an organization started a big data adoption project, its existing data security methods become insufficient to meet the new big data requirements. Additional measures should then be taken to secure data generation and access in order to avoid possible compliance issues, accidental data loss, access to data by non-authorized users, etc.

NIST created a list of the following eight major characteristics that set big data projects apart, making these projects a security and privacy challenge [b- BD Interoperability]:

1) Big data projects often encompass heterogeneous components in which a single security scheme has not been designed from the outset,

2) Most security and privacy methods have been designed for batch or online transaction processing systems,

3) The use of multiple big data sources not originally intended to be used together can compromise privacy, security, or both,

4) A huge increase in the number of sensor streams for the Internet of things create vulnerabilities in the Internet connectivity of the devices, in the transport, and in the eventual aggregation,

5) Certain types of data thought to be too big for analysis, such as geospatial and video imaging, will become commodity big data sources,

6) Issues of veracity, context, provenance, and jurisdiction are greatly magnified in big data,

7) Volatility is significant because big data scenarios envision that data is permanent by default,

8) Data and code can more readily be shared across organizations, but many standards presume management practices that are managed inside a single organizational framework.

Moreover, it is crucial to guarantee the privacy and confidentiality of personal data, especially for big data projects related to some critical sectors, mainly healthcare. Privacy and confidentiality of personal data should be considered as part of the structure of each big data project from the initial phase.

11.3 Investment costs

Big data adoption projects require high investment costs, whether the organization opt for an on-premises big data solution or for a cloud-based solution.

For the first option, investment costs could include the costs of new hardware, new employees′ hiring and/or staff training, energy consumption, as well as costs related to the new software (development, setup, configuration, and maintenance).

In case of the adoption of a cloud-based big data solution, organizations will always have to invest in skills development (by hiring new employees or training current staff) and also pay for cloud services, big data solution development as well as setup and maintenance of required frameworks [b-BD Challenges].

In both cases, costs related to future expansions should be also considered.

11.4 Lack of capacity building

Capacity building is a key action item in the data revolution policy. Unfortunately, there is a lack of capacity building related to big data analytics in developing countries, but various institutions are starting developing and implementing programs that aim to increase skills for big data analytics [b-Capacity Building].

For example, Rwanda has established four African Centre of Excellence in Data Science under the University of Rwanda:

- African Centre of Excellence for Data Science (ACEDS), based in College of Business and Economics and in partnership with Carnegie Mellon University (CMU)-Rwanda;
- African Centre of Excellence in Internet of Things (ACEIoT) based at the College of Science and Technology;
- African Centre of Excellence in Energy for Sustainable Development (ACEESD), also based in the College of Science and Technology;
- African Centre of Excellence in Innovative Teaching and Learning Mathematics and Science (ACEITLMS), based in College of Education.

Collectively, these centres are expected to enroll over 3 500 students in regional development priority areas, out of which over 700 will be PhD students and more than 1 000 will be female. The centres

are expected to publish 1 500 journal articles and 500 studies in collaboration with the private sector and other academic institutions within and beyond the region [b-Capacity Building].

11.5 Resistance to change

Resistance to change is a key factor that may cause the failure of the innovative system in organizations and may influence the intention of an organization to migrate from legacy systems to big data technologies.

The origin of the resistance to change could be decision-makers, mainly because of the lack of information about the reasons of the change. Resistance could also originate from employees due to fear of not being able to develop the new required skills.

11.6 Lack of adequate standards

Standardization has the potential to make any new technology secure, stable, trustworthy, interoperable, safe for human health, energy efficient, and operating without technical constraints and at lower cost.

Many efforts to standardize big data technologies have been started during the last decade, but many standards are still required in some specific fields, as shown in clause 9.2.2.

Besides, in order to achieve big data objectives of both business and consumers, the interworking of multiple systems and technologies, both legacy and new, will be required as well as standards to facilitate interoperability among all the components of the big data value chain.

11.7 Inadequate technological infrastructure

Many organizations in developing countries may suffer from the lack of the storage and communications infrastructure required to organize and integrate huge volumes of data.

They may also lack the computing capacity needed to analyse data. Shortage in electrical grids and telecommunication networks also represents important barriers to some organizations to adopt big data technologies.

11.8 Difficulties of migration

These difficulties depend on the state and nature of the legacy system that needs to be updated. Among these difficulties, there are the evaluation of the state and condition of data in the legacy system and the selection of which data sets will have to be migrated in order to decrease migration costs [b-BD Migration].

Another issue with data migration occurs when it is not possible for organizations to automate the migration, and then data needs to be entered manually [b-BD Migration].

Other challenges involve dealing with the "containers" or systems that handle the data, problems with proprietary systems and nonstandard tooling or coding as being central to many legacy migration projects [b-BD Migration].

Another inherent challenge is the need to have a detailed legacy migration plan ready before initiating such project [b-BD Migration].

Migration is a very critical task as legacy systems often hold valuable data and many organizations fear losing data in the process of an upgrade or transition to a new platform.

11.9 The lack of an innovation culture

The presence of an innovation culture inside an organization is important to ensure its sustainability, mainly in the digital era.

With the adoption of big data, organizations should follow a different approach in the investment in technologies and data collection than the one in traditional business intelligence. In fact, organizations that are used to adopt a top-down approach, starting with a strategy or an idea then elaborating a business case, will have to invest first, then find the value later with big data, which is a totally different way of doing things.

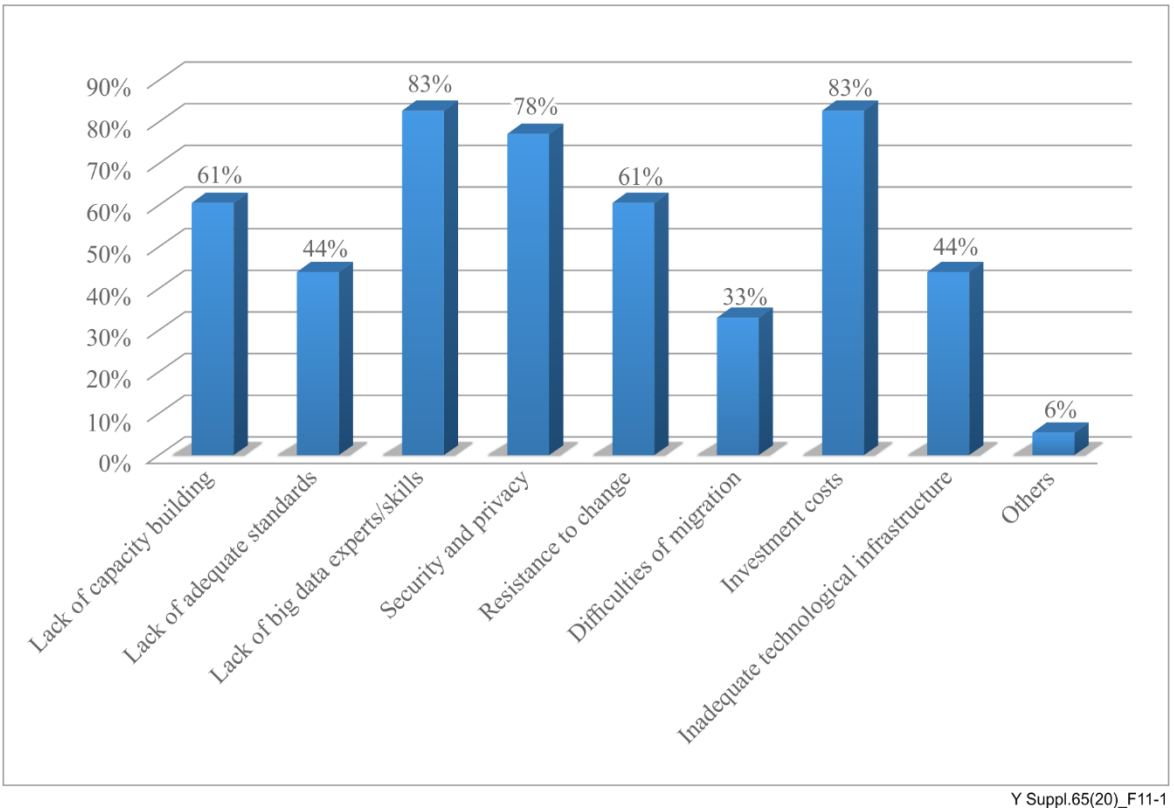

Figure 11-2 – Barriers for the use of big data in developing countries

12 Recommendations for big data adoption in developing countries

This clause suggests some recommendations that may help accelerate the adoption pace of big data in developing countries and overcome the existent challenges taking into account the responses to the survey as well as the scenarios recommended by the respondents to spur the deployment of big data in developing countries in response to question 4.5 of the survey.

12.1 Raising awareness on big data benefits

Although organizations in developing countries are becoming more and more aware of the positive impact that big data usage can have on their business, operations, performance and profit, this awareness is more obvious in some sectors (such as communications, banking and education) compared to others (such as utility and retail).

Therefore, national, and regional initiatives should be launched to create awareness and sensitisation on the use of big data and related common and specific opportunities per sector. These initiatives may include awareness campaigns through workshops, bootcamps, and conferences as well as the elaboration of specific studies on big data use cases and impacts and the launch of blogs and online platforms to share experience and lessons learnt. This could reduce resistance to change effect at enterprises level and help set innovation culture.

12.2 Skills development and capacity building

Universities and academic institutions in developing countries could play an important role in developing skills for the 21st century digital economy.

They should adapt their academic programs and syllabuses to embrace emerging trends and technologies related to big data and help students develop the necessary technical, business and entrepreneurial skills to enable them to manage or contribute to big data projects and meet the local market demand.

In this context, both academic institutions and industry should collaborate to establish a roadmap for the required big data skills at the local market which will help academic institutions to adapt their programs as well as to establish specific training programs tailored for the needs of companies.

Universities in developing countries may also benefit from a close collaboration with ITU Centers of Excellence (CoEs). This collaboration includes delivery of specialized ICT courses for university students, the access to equipment for practical exercises required by the university, joint research, and consultation for public market.

Moreover, the establishment of centers of excellence related to data science in developing countries could help develop national expertise and drive research and development in this field.

12.3 Financial assistance

One of the greatest challenges related to big data adoption projects is the high investment costs which makes medium and small businesses in developing countries reluctant to use of big data technologies even if they are aware that the benefits of leveraging big data will outweigh IT investment.

In this context, more research, development, and innovation efforts are needed to establish more cost-effective big data and analytics tools and techniques.

Besides, governments in developing countries still need financial and technical assistance to launch big data initiatives. They often encounter difficulties to allocate the necessary budget to such initiatives as they have many expenses related to urgent issues such as health, education, clean water, food, etc.

The implementation of strategic partnerships and seeking external support from international development community and major donors in project co-funding, may help developing countries reduce expenses and minimize the complexity of big data projects' implementation.

12.4 Infrastructure considerations

Developing countries need to develop their ICT infrastructure, including telecommunication networks, electrical grids, computing and storage capacities, in order to gain benefits from big data projects.

The building of data Centres at the national level, with high storage and computing capacities, high availability, high connectivity and compliance to international standards especially in terms of security, would be essential to store and analyse the increasing amounts of data generated in developing countries and avoid storing data abroad which could be prohibited by regulations in some countries.

12.5 Big data national program

The establishment of a big data national program or a national plan for big data, with clear definition of objectives, milestones and expected outcomes, could help organizations have a better understanding of the big data potential and could help achieve SDGs.

However, big data is tightly related to other technologies and concepts, mainly cloud computing, IoT and artificial intelligence. Therefore, the type of national program to be developed should also take

into account these technologies and concepts or should be harmonized with existent related national programs, if any.

Moreover, developing countries should encourage digitalization, and boost data generation by improving their mobile broadband services and facilitating mobile and Internet access at affordable prices. They may also encourage the development of online contents, the development of mobile and IoT applications, the use of electronic payment facilities and further access to e-commerce and e-learning platforms.

12.6 Building a regulatory framework for big data

Building a regulatory framework for big data is essential to support the uptake of big data in developing countries, protect personal data and increase trust in big data technologies. It allows the different stakeholders to understand their legal status and the applicable jurisdiction in case of any issues related to their data.

This framework should be in harmonization with current rules relating to privacy and data protection, while being adapted to foster innovation. It should also deal with issues such as data ownership, data location, roles, and responsibilities of involved stakeholders, etc.

12.7 Standards development

Awareness regarding the importance of standards, especially for innovation, should be raised more and more in developing countries in order to enhance their participation in the standardization activities of international SDOs while setting mechanisms to encourage the participation of SME, start-ups and academia from these countries.

Big data standards should follow the accelerating rhythm of big data technologies adoption driven by the uptake of IoT and mobile applications and the increasing market demand for data analytics.

Standardization should focus on the most important challenges of big data which are slowing down the adoption of big data around the world, and take into account the priorities of developing countries which are mainly infrastructure, security, privacy, architecture and networking.

Big data standardization stakeholders should further coordinate and work together in order to avoid standards fragmentation.

Appendix

Survey on big data adoption in developing countries

This appendix provides a questionnaire for a survey on bid data adoption in developing countries.

Responder's information

Country:	……
Sector:	……
Organization:	……..……
Name:	……
Title:	……
Address:	……
Telephone:	……
Fax:	……
E-Mail:	……

1 Big data usage

1.1 What percentage of enterprises in your country has already adopted big data technologies?

- ❑ Approximately: …..……..% (please specify)
- ❑ Do not know

1.2 What sector in your country uses big data the most?

- ❑ Healthcare
- ❑ Education and research
- ❑ Retail
- ❑ Government
- ❑ Banking
- ❑ Insurance
- ❑ Communications
- ❑ Media and entertainment
- ❑ Utilities
- ❑ Transportation
- ❑ Other (please specify): …….

1.3 Do you need to use big data in your organization?

- ❑ Yes
- ❑ No

1.4 How important is the use of big data for your organization?

- ❑ Mission-critical
- ❑ Important
- ❑ Quite important
- ❑ Least important
- ❑ Other (please specify): …….

1.5 Have your organization already used big data technologies?

❑ Yes	Please mention when has your organization started using big data technologies: …….
❑ No, but planning to use	Please mention when will your organization start using big data technologies: ……..
❑ No, not on our technology roadmap	Please mention the reasons behind not using big data technologies in your organization: …….....

1.6 In what domain(s) in your organization do you use big data:

- ❑ Human resources
- ❑ Sales
- ❑ Marketing
- ❑ Customer service
- ❑ Finance
- ❑ Product development/management

- ❑ Information technologies
- ❑ Risk management
- ❑ Fraud management
- ❑ Operations
- ❑ Logistics
- ❑ Other (please specify): …….

1.7 What are the reasons behind your engagement in the big data area and the replacement of traditional solutions for the storage, the processing, and the exploitation of data?

- ❑ Increased data volume
- ❑ Cost optimization
- ❑ Velocity of change
- ❑ Volatility of requirements
- ❑ Need for analysis-driven requirements instead of requirements-driven analysis
- ❑ Data diversity
- ❑ Other (please specify): …….

1.8 What types of data/records are you analysing/are you planning to analyse by the use of big data technologies?

- ❑ Transactional data from enterprise applications
- ❑ Unstructured content from email, office documents, etc.
- ❑ Locational/Geospatial data
- ❑ Scientific data
- ❑ Image/video data
- ❑ Data from social media
- ❑ Data from sensors and connected device
- ❑ Other (please specify): …….

1.9 Have you relied on internal skills of your organization to deploy big data solutions or have you requested help from third parties (consultants, SMEs, etc.)?

- ❑ All internal
- ❑ Mostly internal, with some help from third parties
- ❑ Mostly third parties under our direction and supervision
- ❑ All outsourced

1.10 What data management functions/features are most important to you?

- ❑ High-capacity, inexpensive storage
- ❑ High-performance, inexpensive processing power
- ❑ Data integration and quality capabilities
- ❑ Relational database acceleration/scale
- ❑ High-velocity data stream processing
- ❑ Unstructured text management and search
- ❑ Other (please specify): …….

1.11 Did your government use big data technologies in a project/some projects at a national scale?

- ❑ Yes

❑ No, but on the roadmap

❑ No projects

If you select one of the two first choices, please give some details on this (theses) project(s) (such as objective, date of launch….):

……

……

……

2 Big data requirements

In your opinion, what are the main requirements related to big data:

2.1 Data sources (data size, file formats, rate of growth, etc.):

……

2.2 Data transformation (data fusion, analytics, etc.)

……

2.3 Infrastructure (software tools, platform tools, hardware resources such as storage and networking)

……

2.4 Data consumption (format of the processed results (table, visuals, ...))

……

2.5 Security and privacy

……

2.6 Lifecycle management (conversion, quality check, pre-analytic processing, etc.)

……

2.7 Big data skills

……

2.8 Other requirements (please specify):

……

3 Big data standardization

3.1 Do you have any idea about the standardization activities of ITU-T on big data?

❑ Yes

❑ No

If yes, what activities do you follow or participate to?

.....…….....

3.2 What areas of standardization do you consider important for big data?

❑ Use cases

❑ Common requirements

❑ Architecture

❑ Infrastructure

❑ Networking

❑ Data management analytics

❑ Security and privacy

❑ Green ICT and data

❑ Other (please specify): ……

3.3 Among the potential areas of standardization for big data that may be of interest to ITU-T, as identified in Supplement 40 to ITU-T Y.3600 – Big data standardization roadmap, what areas of standardization do you consider most urgent to work on?

❑ Common requirements and use cases

❑ Definition, architecture, data model and APIs

❑ Data analytics

❑ Security and data protection

❑ Data quality and reliability

❑ Legal implications of big data in the telecommunications sector (e.g., data ownership)

❑ Big data exchange

❑ Other (please specify):…….

4 Big Data Opportunities & Challenges

4.1 What benefits would your organization like to achieve through big data:

❑ Improved customer experience

❑ Increased sales

❑ Higher quality products and services

❑ New product innovations

❑ More efficient operations

❑ Better decision making

❑ Reduced risk

❑ Other (please specify): …….

4.2 Do you think big data could help reach the United Nations Millennium Development Goals?

❑ Yes

❑ No

❑ Do not know

If you select "Yes", please explain how this could be achieved:

……

……

4.3 What data challenges is your organization addressing with big data?

❑ Integrating a wider variety of data

❑ Cleansing data

❑ Using more current data

❑ Storing more historical data

❑ Understanding unstructured data

❑ Using real-time data

❑ Using more granular data

❑ Using higher quality data

❑ Other (please specify):……..

4.4 In your opinion, what are the main barriers that could be encountered in the deployment of big data solutions in developing countries?

- ❑ Lack of capacity building
- ❑ Lack of adequate standards
- ❑ Lack of big data experts/skills
- ❑ Security and privacy
- ❑ Resistance to change
- ❑ Difficulties of migration
- ❑ Investment costs
- ❑ Inadequate technological infrastructure
- ❑ Other (please specify): …….

4.5 What are the scenarios that can spur the deployment of big data in your country and in developing countries in general?

……

Bibliography

[b-TSB Circular 120] ITU-T TSB (2018) *Questionnaire on Big Data Adoption in Developing* Countries.
https://www.itu.int/md/T17-TSB-CIR-0120

[b-BD Banking] *Your Go-to Guide to Big Data Analytics in Banking*, Hitachi
https://us.hitachi-solutions.com/blog/big-data-banking/

[b-BD Challenges] *The 'Scary' Seven: big data challenges and ways to solve them*, ScienceSoft
https://www.scnsoft.com/blog/big-data-challenges-and-their-solution

[b-BD Development] *Big Data for development: preventing the spread of epidemics*, ITU
https://www.itu.int/en/ITU-D/Emergency-Telecommunications/Pages/BigData/default.aspx

[b-BD Education] *4 ways big data is transforming the education sector*, Allerin, August 2017
https://www.allerin.com/blog/4-ways-big-data-is-transforming-the-education-sector

[b-BD Experience] *Using Big Data to Improve Customer Experience*, JIGSAW Academy
https://www.jigsawacademy.com/using-big-data-to-improve-customer-experience/

[b-BD Interoperability] *NIST Big Data Interoperability Framework: Volume 4, Security and Privacy*, National Institute of Standards and Technology, October 2018
https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-4r2.pdf

[b-BD Management] *What is Big Data Management*, Datamation, June 2017
https://www.datamation.com/big-data/big-data-management.html

[b-BD Migration] *What are some of the biggest challenges with legacy migration?*, Technopedia
https://www.techopedia.com/7/32194/technology-trends/big-data/what-are-some-of-the-biggest-challenges-with-legacy-migration

[b-BD Quality] *A revolution in data-driven quality improvement*, Deloitte, April 2018
https://www2.deloitte.com/us/en/insights/topics/analytics/a-revolution-in-data-driven-quality-improvement.html

[b-BD Risk] *4 Ways Big Data is Evolving Risk Management*, Techwave, July 2019
https://techwave.net/4-ways-big-data-is-evolving-risk-management/

[b-BD Sales] *How to Harness the Power Of Big Data and Increase Sales, OSI Affiliate*
https://www.osiaffiliate.com/blog/harness-power-big-data-increase-sales/

[b-BD Skills] *Big data skills shortages – and how to work around them*, ComputerWeekly.com, June 2018
https://www.computerweekly.com/opinion/Big-data-skills-shortages-and-how-to-work-around-them

[b-Capacity Building] *Ongoing Capacity building initiatives*, National Institute of Statistics of Rwanda
http://statistics.gov.rw/content/data-revolution/ongoing-capacity-building-initiatives

[b-Customer Experience] *5 Ways Big Data Can Improve the Customer Experience*, VOCALCOM, August 2017
https://www.vocalcom.com/en/blog/customer-experience/5-ways-big-data-can-improve-the-customer-experience/

[b-Decision Making] *Using Analytics for Better Decision-Making*, Towards Data Science, December 2018
https://towardsdatascience.com/using-analytics-for-better-decision-making-ce4f92c4a025

[b-New Product] *Big Data and New Product Development*, CLEVERISM, September 2019
https://www.cleverism.com/big-data-new-product-development/

[b-SDG development] *Big data and global development*, SAS
https://www.sas.com/en_us/insights/articles/big-data/big-data-global-development.html

[b-SDG Performance] *Big Data can transform SDG performance. Here's how*, Asian Development Blog, June 2018
https://blogs.adb.org/blog/big-data-can-transform-sdg-performance-here-s-how

[b-UN BD SDG] *Big Data for Sustainable Development*, United Nations
https://www.un.org/en/sections/issues-depth/big-data-sustainable-development/index.html

[b-UN SDG] *About the Sustainable Development Goals*, United Nations
https://www.un.org/sustainabledevelopment/sustainable-development-goals/

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	**Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2020