International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.4114

(07/2017)

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

Internet of things and smart cities and communities – Requirements and use cases

---

Specific requirements and capabilities of the Internet of things for big data

Recommendation ITU-T Y.4114

ITU-T

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.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.

Recommendation ITU-T Y.4114

Specific requirements and capabilities of the Internet of things for big data

Summary

Recommendation ITU-T Y.4114 specifies requirements and capabilities of the Internet of things (IoT) for big data. This Recommendation complements the developments on common requirements of the IoT described in Recommendation ITU-T Y.4100/Y.2066 and the functional framework and capabilities of the IoT described in Recommendation ITU-T Y.2068 in terms of the specific requirements and capabilities that the IoT is expected to support in order to address the challenges related to big data. This Recommendation also constitutes a basis for further standardization work such as functional entities, application programming interfaces (APIs) and protocols concerning big data in the IoT.

History

Edition	Recommendation	Approval	Study Group	Unique ID[*]
1.0	ITU-T Y.4114	2017-07-07	20	11.1002/1000/13265

Keywords

Big data, Internet of things, IoT big data characteristics, IoT data operations, IoT data roles

[*] 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

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation 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 Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. 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

Recommendation ITU-T Y.4114

Specific requirements and capabilities of the Internet of things for big data

1 Scope

This Recommendation provides specific requirements and capabilities of the IoT for big data. This Recommendation complements the developments on common requirements of the IoT [ITU-T Y.4100] and functional framework of the Internet of things (IoT) [ITU-T Y.2068] in terms of the specific requirements and capabilities that the IoT is expected to support in order to address the challenges related to big data.

In order to enhance the IoT system for dealing with the big data challenges in the IoT environment, the big data characteristics of IoT data and the IoT ecosystem from the IoT data point of view are investigated. From the perspective of the IoT data roles, big data challenges in the IoT are described.

The requirements of the IoT for big data are specified addressing different IoT data operations in the IoT which are impacted by big data. IoT data operations include data collection, data pre-processing, data analysis, data transfer, data storage, data query and data visualization.

Building on the identified specific requirements of the IoT for big data, the capabilities of the IoT for big data are then specified.

The scope of this Recommendation includes:

– Overview of big data in the IoT

– Requirements of the IoT for big data

– Capabilities of the IoT for big data

An IoT use case with big data characteristics is provided in Appendix I.

Details on relationships among IoT data roles, IoT business roles, IoT data operations and IoT components are provided in Appendix II.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T Y.2068] Recommendation ITU-T Y.2068 (2015), *Functional framework and capabilities of the Internet of things.*

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

[ITU-T Y.4000] Recommendation ITU-T Y.4000/Y.2060 (2012), *Overview of the Internet of things.*

[ITU-T Y.4100] Recommendation ITU-T Y.4100/Y.2066 (2014), *Common requirements of the Internet of things.*

[ITU-T Y.4113] Recommendation ITU-T Y.4113 (2016), *Requirements of the network for the Internet of things.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 Internet of things (IoT) [ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

NOTE 1 – Through the exploitation of identification, data capture, processing and communication capabilities, the IoT makes full use of things to offer services to all kinds of applications, whilst ensuring that security and privacy requirements are fulfilled.

NOTE 2 – From a broader perspective, the IoT can be perceived as a vision with technological and societal implications.

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

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 IoT big data characteristics: IoT data set characteristics of high-volume, high-velocity and/or high-variety related to the challenges of IoT data set operations.

NOTE 1 – The term "volume" refers to the size of the IoT data sets, the term "velocity" refers to the speed of IoT data streams in and out of the IoT data sets and the term "variety" refers to the diversity of IoT data types of the IoT data sets.

NOTE 2 – Additional dimensions of data, such as veracity, variability, etc., may also be associated with the IoT big data characteristics.

NOTE 3 – Examples of operations on IoT data sets include collection, pre-processing, transfer, storage, query, analysis and visualization.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

API	Application Programming Interface
EHM	e-Health Monitoring
IoT	Internet of Things
QoE	Quality of Experience
QoS	Quality of Service
SLA	Service Level Agreement

5 Conventions

In this Recommendation:

The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

6 Overview of big data in the IoT

6.1 Data characteristics in the IoT

It is expected that in the future the quantity of connected things will be so huge that the IoT data will constitute a predominant part of the data carried by networks. These IoT data include not only data collected from things via devices and/or gateways within the IoT system and data processed within the IoT system itself, but also data injected from sources which are external to the IoT system (e.g., Web, social networks and industry domain specific information systems).

In order to enhance the IoT system enabling it to deal with big data challenges in the IoT environment, it is essential to investigate application independent characteristics (i.e., common characteristics) of IoT data. Consequently, according to the common characteristics of IoT data, the IoT system is enabled to handle the IoT data by the means of common capabilities facilitating data operations in the IoT.

With respect to traditional data operations, operations on IoT data require additional study due to some unique characteristics of IoT data, named here as "IoT big data characteristics", essentially due to the diversity of collected IoT data types and the broad variety of application scenarios. The prominent IoT big data characteristics are IoT data's high volume (the size of the data sets), IoT data's high variety (heterogeneity of data types and sources) and IoT data's high velocity (the speed of data streams in and out of the data sets). These prominent IoT big data characteristics are the main driver of the requirements study in this Recommendation, helping to identify related common capabilities of the IoT system with respect to the variety of concrete IoT deployments.

6.2 IoT data roles

Based on consideration of the IoT system and the IoT big data characteristics, five key IoT data roles, i.e., the key roles which are relevant in an IoT deployment from a data operation perspective, are identified for the IoT ecosystem as shown in Figure 1.

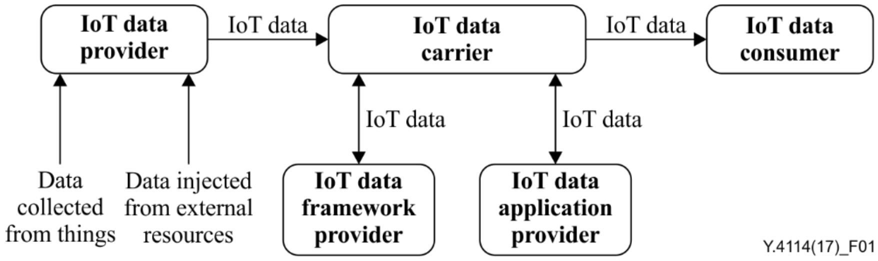
IoT data provider
IoT data
IoT data carrier
IoT data
IoT data consumer
IoT data
IoT data
Data collected from things
Data injected from external resources
IoT data framework provider
IoT data application provider
Y.4114(17)_F01

Figure 1 – IoT data roles

– **IoT data provider**: The IoT data provider collects data from things, injects data processed within the IoT system as well as data from external sources and provides them via the IoT data carrier to the IoT data consumer (optionally, the applications provided by the IoT data application provider may execute relevant data operations with the support of the IoT data framework provider).

– **IoT data application provider**: The IoT data application provider provides applications related to the execution of IoT data operations (e.g., applications for data analysis, data pre-processing, data visualization and data query).

NOTE 1 – The applications provided by the IoT data application provider can interact with the infrastructure (e.g., storage cloud) provided by the IoT data framework provider through the IoT data carrier or run on the infrastructure (e.g., scalable distributed computing platforms) provided by the IoT data framework provider.

- **IoT data framework provider**: The IoT data framework provider provides general IoT data processing capabilities and related infrastructure (e.g., storage and computing resources, data processing run time environment) as required by IoT data provider, IoT data carrier, IoT data application provider and IoT data consumer for the support of data operations execution.
- **IoT data consumer**: The IoT data consumer consumes IoT data. Usage of the consumed data depends on the application purposes.
- **IoT data carrier**: The IoT data carrier carries data among the IoT data provider, the IoT data framework provider, the IoT data application provider and the IoT data consumer.

  NOTE 2 – An actor of a concrete IoT deployment can play multiple roles. As an example, an actor executing data analysis plays the role of IoT data application provider, but it also plays the role of IoT data provider when it sends the results of this data analysis to other actors.

Figure 2 shows key possible mappings from IoT business roles [ITU-T Y.4000] to the above described IoT data roles.

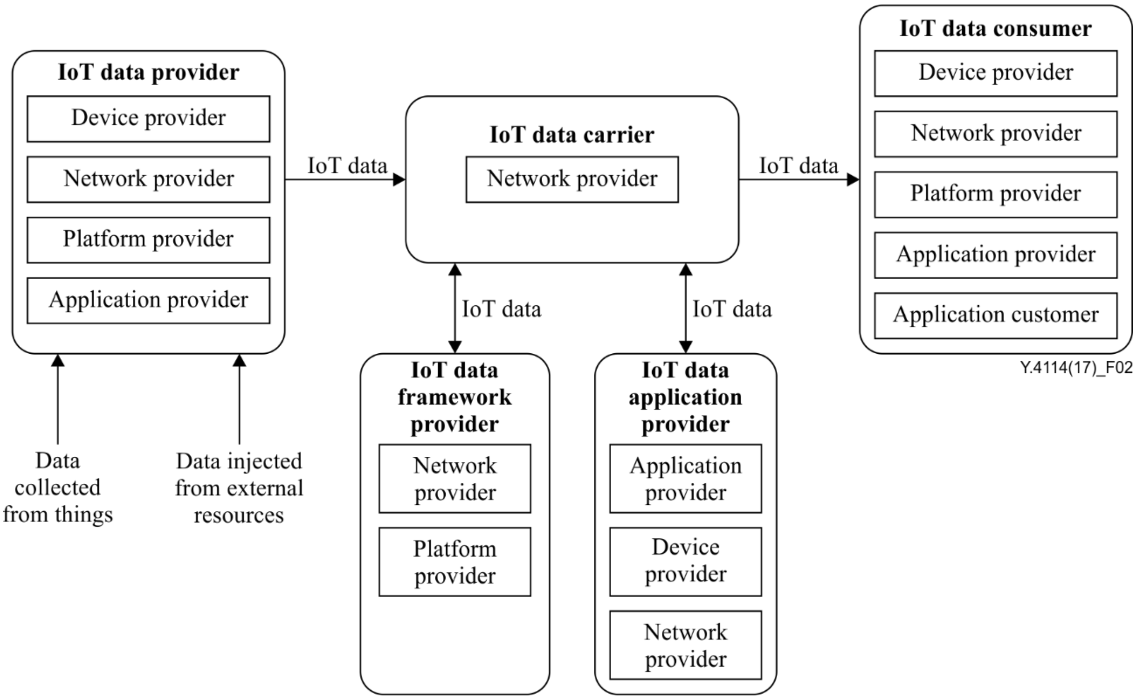

Figure 2 – Key possible mappings from IoT business roles to IoT data roles

NOTE 3 – Figure 2 is an abstract representation of multiple possible IoT deployments. In fact, as shown in Figure 2, from an IoT data perspective, the same IoT business role may be mapped to multiple IoT data roles.

As an example, concerning the device provider, when it provides its collected data to other IoT business roles (including platform provider, application provider and application customer), it acts as IoT data provider; on the other hand, when it consumes IoT data provided by other IoT business roles (including network provider, platform provider and application provider), it acts as IoT data consumer; finally, when it provides applications related to IoT data processing capabilities to other IoT business roles (including network provider, platform provider and application provider), it acts as IoT data application provider.

An example of deployment scenarios with respect to IoT data roles is described in Appendix II.1.

6.3 Challenges of the IoT from the IoT data roles perspective

The IoT is a system that can connect things, capture, aggregate, transfer, store, analyse and query IoT data and provide things-related services. The IoT faces the challenges of IoT data from the perspective of the various roles related to IoT data as described in clause 6.2.

From the IoT data provider point of view, the IoT needs to cope with a rapidly increasing number and types of devices and needs to provide IoT data in an efficient way to facilitate data operation execution and data consumption.

From the IoT data carrier point of view, the IoT needs to deal with a large number of devices accessing the network and transferring large amounts of data. These IoT data, related to different IoT applications, have different service level agreement (SLA) and quality of service (QoS) requirements: e.g., some IoT data have real time QoS requirements (for example, real-time monitoring IoT data and emergency alarm IoT data), others have no strict QoS requirements (for example, IoT data from electricity meters). These aspects raise challenges in terms of IoT data transfer handling, including the avoidance of possible network congestions from concurrent access of devices and gateways and in terms of active scheduling data transmission (e.g., among IoT application servers and IoT platforms).

From the IoT data framework provider point of view, the IoT needs to leverage general IoT data processing capabilities and related infrastructure, in order to be accessible via open and standard interfaces, provide APIs to the IoT data application provider, interoperate with other IoT data processing capabilities and with data processing related IoT applications, while providing integrity, privacy and security protection for IoT data.

From the IoT data application provider point of view, the IoT needs to be scalable and flexible for support of applications based on IoT data processing capabilities, enabling the processing of variable sizes of data sets, differentiated speed of data streams in and out of the IoT application, diversity of data formats to be processed by the IoT application while providing integrity, privacy and security protection for IoT data.

From the IoT data consumer point of view, the IoT needs to provide consumers with the desired data according to their requirements, including quality of experience (QoE).

Besides, collection, transfer, processing and consumption in a secure way and assurance of appropriate privacy of IoT data, including personal data, are challenges which concern all IoT data roles.

7 Requirements of the IoT for big data

This clause complements the common requirements of the IoT [ITU-T Y.4100] in terms of the specific requirements that the IoT is expected to support in order to address the challenges related to big data.

7.1 IoT data operations in the IoT for big data

Big data in the IoT are IoT data which have characteristics of high volume, high variety, high velocity and high volatility as described in clause 6.1 and are hard to operate using existing tools and methods. To handle challenges faced by big data in the IoT as described in clause 6.3, requirements of the IoT for big data are specified for the IoT components, i.e., device, gateway, network, IoT platform and IoT application server as described in [ITU-T Y.4113], from the view point of IoT data operations as shown in Figure 3.

Considering that the diverse set of concrete IoT deployments do not imply a unique logical sequencing of the various IoT data operations, Figure 3 provides an abstract representation of the various IoT data operations and related data flows.

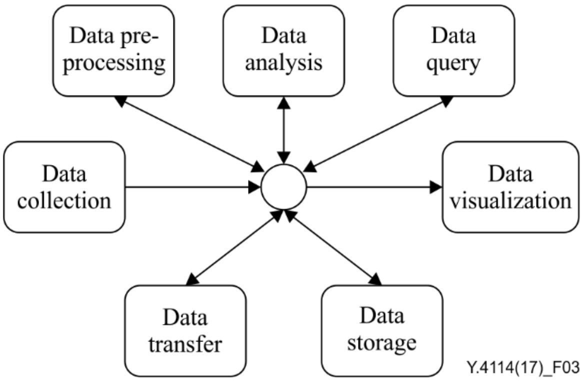

Figure 3 – IoT data operations

In Figure 3, the boxes represent the IoT data operations and the arrows represent the data flows with associated directionality. The circle, cross point of the data flows, is merely intended to represent the fact that there is no unique logical sequencing of the various IoT data operations.

In particular, for the operation of data collection, a unidirectional data flow is shown since this data flow can only be from data collection to other operations; similarly, for the operation of data visualization, the data flow can only be from other operations to data visualization, so it is also shown as unidirectional data flow; for the other IoT data operations, bi-directional data flows are used to describe their possible data flows.

NOTE 1 – The sequencing of IoT data operations in IoT depends greatly on the service and deployment scenarios. Cloud computing and edge computing are two technologies that may be implemented in the IoT for support of different IoT data operation sequences: e.g., cloud computing can be used to perform data analysis after the data are transferred to the remote IoT platform, while edge computing can be used to perform data analysis locally such as at the gateway.

NOTE 2 – Detailed information concerning relationships between IoT data roles and IoT data operations is provided in Appendix II.2.

NOTE 3 – Detailed information concerning relationships between IoT data operations and IoT components are described in Appendix II.3.

7.1.1 Data collection

Data collection is the data operation related to measurement processes and/or data information gathering from components of IoT.

NOTE – Data collected by an IoT platform directly from sensors and data collected by an IoT platform via a gateway can be seen as two typical examples of data collection scenarios.

7.1.2 Data transfer

Data transfer is the IoT data operation related to the physical transfer of IoT data between components of IoT as well as between the IoT system and external systems.

7.1.3 Data storage

Data storage is the IoT data operation related to the physical storage of IoT data.

NOTE – The data storage local to an edge IoT component such as a gateway and/or device and the remote data storage in a core IoT component such as an IoT platform and IoT application server (e.g., cloud based storage) can be seen as two typical examples of the data storage scenarios.

7.1.4 Data pre-processing

Data pre-processing is the IoT data operation on the IoT data aiming to ensure and/or enhance the quality or efficiency of other IoT data operations.

NOTE – Data format transformation after data collection and data de-noising before the data analysis can be seen as two typical examples of data pre-processing scenarios.

7.1.5 Data analysis

Data analysis is the IoT data operation related to the computing process on IoT data aiming to derive additional information according to application requirements, e.g., for making decisions, predicting future results, etc.

7.1.6 Data visualization

Data visualization is the IoT data operation related to the visual representation of IoT data.

7.1.7 Data query

Data query is the IoT data operation related to querying IoT data, including raw IoT data and processed IoT data.

NOTE – Data query on historical IoT data and real-time IoT data can be seen as two typical examples of data query scenarios.

7.2 Requirements of the device for big data

7.2.1 Data collection

The device is required to have enough precision to sense IoT data from monitored things in order to measure and record the characteristics of the monitored things.

The device is recommended to mark the sensed IoT data with labels to indicate the types (related to the monitored things' characteristics) of the sensed IoT data, so that other IoT components can understand the types of the sensed IoT data (e.g., physiological data for healthcare).

7.2.2 Data pre-processing

The device is recommended to organize sensed IoT data according to pre-defined information models and represent them according to the pre-defined formats.

The device is recommended to associate time stamps to the collected IoT data, in order to keep track of chronology of sensed data.

The device is recommended to associate semantic description to the collected IoT data, so that other technical components of the IoT, i.e., gateway, IoT platform and IoT application server, can understand the semantics of the sensed IoT data.

NOTE 1 – Examples of associated semantic description may include provenance, location, etc.

The device is recommended to have the ability to clean IoT data, e.g., removing redundant IoT sensed data in order, for example, to minimize the amount of transferred IoT data.

NOTE 2 – The specific data cleaning capabilities of a device depend on its computing resources.

7.2.3 Data storage

The device is recommended to store the sensed IoT data according to rules (including e.g., temporally or permanently).

7.2.4 Data analysis

The device is recommended to make appropriate data analysis for sensed IoT data and operational status of the device itself (e.g., battery status, transmission bandwidth) to help its decision making process, including e.g., when to send out what signaling data to other IoT components.

7.2.5 Data transfer

The device is recommended to provide other IoT components with transfer configuration parameters for sensed IoT data in order to apply required network resources (e.g., data buffer and network QoS) for data transfer quality support.

7.2.6 Data visualization

The device is recommended to show the sensed IoT data according to application requirements, e.g., in real time.

7.3 Requirements of the gateway for big data

7.3.1 Data collection

The gateway is required to efficiently collect the sensing IoT data or monitoring IoT data from connected devices. For example, the gateway can arrange the collection period for each connected device according to application or network requirements to avoid large amount of useless IoT data being frequently collected.

7.3.2 Data pre-processing

The gateway is recommended to pre-process the collected IoT data to verify and improve their quality. For example, some inaccurate sensing IoT data collected from one connected device can be detected by the gateway via comparing with the sensing IoT data from other similar connected devices.

The gateway is recommended to associate a time stamp to collected IoT data if there is no associated time stamp from connected devices.

The gateway is recommended to associate a semantic description to collected IoT data if there is no associated semantic description from connected devices.

The gateway is recommended to have the ability to perform local operations, e.g., in case the collected IoT data is not required by other remote components of the IoT system.

7.3.3 Data storage

The gateway is recommended to support data storage for the collected IoT data and operation log files.

NOTE – The size of data storage depends on the application requirements. The data storage of a gateway can be based on a local disk or based on cloud storage.

7.3.4 Data transfer

The gateway is recommended to aggregate the collected IoT data for transfer in order to reduce the transfer overhead.

NOTE 1 – Aggregation is particularly beneficial in cases where the destination of the IoT data is the same.

The gateway is recommended to provide buffer mechanisms for transfer in order to ease transfer pressure on the network.

NOTE 2 – If the network is busy, the gateway can buffer IoT data within a tolerable time before transmission.

The gateway is recommended to compress IoT data before transfer in order to reduce the transfer load.

7.3.5 Data analysis

The gateway is recommended to have the ability to perform data analysis on its collected data. For example, these results can be then sent to local applications according to their requests so that the latency can be reduced compared with performing the data analysis on the remote IoT platform.

NOTE – The gateway data analysis capability can be used to support edge computing.

7.4 Requirements of the network for big data

General network requirements of the IoT can be found in [ITU-T Y.4100].

7.4.1 Data transfer

In addition to the general network requirements of the IoT, three main challenges, as described in clause 6.3, for a network handling data with "IoT big data characteristics" are identified:

- Access of a large number of devices to the network;
- Transfer of data in the network with different SLA and QoS requirements;
- Transfer in the network of high volumes of IoT data.

The access of a large number of devices to the network may cause network congestion. Network requirements related to congestion issues caused by massive device access (e.g., in application scenarios using a large number of smart meters and sensors) are identified in [ITU-T Y.4113].

Concerning SLA and QoS, according to the specific service requirements, the network is required to adopt appropriated network technologies in order to support diverse services (e.g., time-critical services and non-time-critical services).

NOTE – Examples of network technologies include IoT dedicated cellular technologies and IoT dedicated non-cellular technologies.

In the IoT, in particular for data migration and data backup purposes, the volume of IoT data transferred among different IoT platforms and among IoT platforms and IoT application servers may be huge. The network is recommended to provide appropriate bandwidth according to requirements of IoT platforms and IoT application servers in order to support the huge volume of these data flows. Figure 4 describes typically relevant IoT data flows among different IoT components. The thickness of the lines in Figure 4 symbolizes the relative volume of exchanged data.

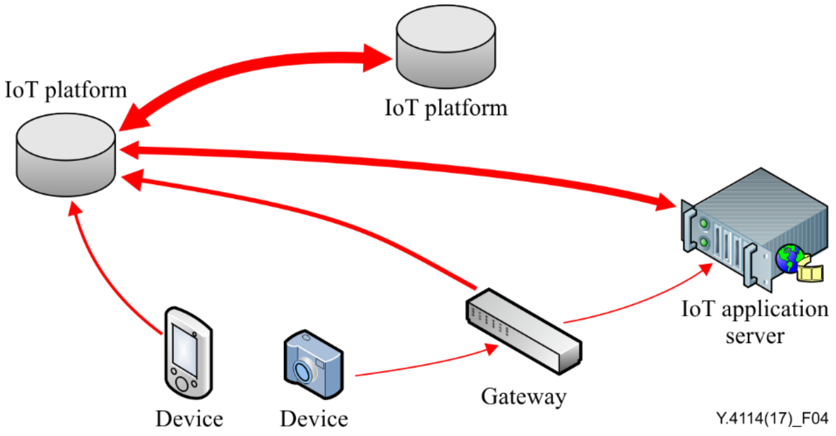

Figure 4 – Typically relevant data flows among different IoT components

7.5 Requirements of the IoT platform for big data

7.5.1 Data collection

The IoT platform is required to have the ability to efficiently collect the IoT data received from a gateway or devices. For example, data collection schedulers can be used in the IoT platform to arrange the collection period and collection sequence according to application or network requirements in order to relieve the network and the IoT platform load when dealing with a large amount of IoT data.

7.5.2 Data pre-processing

The IoT platform is required to have the ability to pre-process the collected IoT data in order to verify and improve the accuracy of the IoT data, e.g., de-noising operation.

The IoT platform is required to associate time stamps to collected IoT data if data have no associated time stamps.

The IoT platform is recommended to have the ability to associate semantic description to collected IoT data if there is no associated semantic description for the IoT data, assuming that the IoT platform can understand the meaning of the IoT data (e.g., according to the associated data model).

7.5.3 Data storage

The IoT platform is required to support secure and high-volume data storage for the IoT data.

The IoT platform is also recommended to support secure and high-volume data storage for the IoT data with semantic annotation.

7.5.4 Data transfer

The IoT platform is required to support buffer mechanisms for data transfer in order to reduce network load.

NOTE – If the network is busy, the IoT platform can buffer IoT data within a tolerable time before transmission.

The IoT platform is recommended to have the ability to compress IoT data before their transfer in order to reduce the transfer load.

7.5.5 Data analysis

The IoT platform is recommended to have the ability to perform some pre-defined universal data analysis (e.g., universal data statistics and data mining algorithms) according to application requests and send the data analysis results to the requesting applications.

The IoT platform, for its internal use, is also recommended to have the ability to perform data analysis for classifying which service type the IoT data is associated with, so that the IoT platform can efficiently deal with data collection, data access and data transfer according to the service type.

7.5.6 Data query

The IoT platform is required to support data query on the stored historical IoT data and raw IoT data.

The IoT platform is recommended to support data query on real-time processed results of collected IoT data.

The IoT platform is recommended to support semantic data query.

The IoT platform is required to support data query with access control enabling query scope restriction.

7.6 Requirements of the IoT application server for big data

7.6.1 Data collection

The IoT application server is required to have the ability to collect IoT data from devices, gateways, IoT platforms and other IoT application servers.

7.6.2 Data pre-processing

The IoT application server is required to have the ability to pre-process IoT data. Examples of data pre-processing include validation, cleaning of invalid and irrelevant IoT data, insertion of default values for incomplete IoT data, compression of IoT data and format transformation of IoT data.

Data received by the IoT application server from devices, gateways, IoT platforms or other IoT application servers may be invalid or irrelevant. The IoT application server is required to identify invalid and irrelevant IoT data by certain methods (e.g., data correlation) and then remove them.

Required data fields of IoT data received by the IoT application server from devices, gateways, IoT platforms or other IoT application servers may be incomplete. The IoT application server is required to supplement missing data fields with default values.

Not all data fields of IoT data received by the IoT application server from devices, gateways, IoT platforms or other IoT application servers may be needed. The IoT application server is required to compress received IoT data (e.g., lossy compression by removing data fields which are not needed by the IoT application server).

Data received by the IoT application server from devices, gateways, IoT platforms or other IoT application servers may have an inappropriate data format for data analysis. The IoT application server is required to transform a received data format for efficient data analysis.

7.6.3 Data storage

The IoT application server is required to have the ability to store IoT data temporarily and/or permanently.

Before pre-processing, IoT data received by the IoT application server from devices, gateways, IoT platforms or other IoT application servers are recommended to be stored temporarily and locally.

Data concerning analysis results are required to be stored permanently. IoT data stored permanently can be stored locally or remotely.

7.6.4 Data transfer

The IoT application server is required to support application specific data transfer requirements including for, but not limited to, latency, reliability, bandwidth and connectivity.

7.6.5 Data analysis

The IoT application server is required to have the ability to analyse IoT data according to dedicated algorithms and customized profiles. For example, e-health monitoring (EHM) application servers should be able to process EHM IoT data according to specific medical algorithms and specific profiles configured by doctors.

The IoT application server is recommended to have the ability to execute distributed data analysis. In case of complex data analysis, the IoT application server should be able to make usage of resources from an IoT platform and other IoT application servers.

7.6.6 Data visualization

The IoT application server is required to have the ability to show results of data analysis according to a consumer's customized needs.

Personal identity and other privacy sensitive information are required to be protected.

7.6.7 Data query

The IoT application server is required to support data query on the IoT data related to its provided applications.

The IoT application server is recommended to support semantic data query on the IoT data related to its applications.

The IoT application server is required to support data query with access control enabling query scope restriction.

8 Capabilities of the IoT for big data

This clause complements the functional framework of the IoT [ITU-T Y.2068] in terms of the specific capabilities that the IoT is expected to support in order to address the challenges related to big data.

8.1 Overview

Nine capabilities are identified based on the requirements provided in clause 7, some capabilities are deduced from corresponding requirements and others from analysis of multiple requirements:

- Big data collection
- Big data pre-processing
- Big data storage
- Big data transfer
- Big data time synchronization
- Big data analysis
- Big data query
- Big data visualization
- Big data security and privacy protection

All IoT components are concerned by capabilities of the IoT for big data.

NOTE – Specific deployments may vary in terms of the location of one or more of these capabilities in the different parts of the IoT system.

8.2 Big data collection

Big data collection capability implements the data collection operation as described in clause 7.1.1.

According to the requirements in clauses 7.2.1, 7.3.1, 7.5.1 and 7.6.1, the following abilities are provided by the big data collection capability:

- The ability to mark the collected IoT data to indicate sufficient information related to the usage of the collected IoT data.

  NOTE 1 – Data marking with tags and semantic annotation can be seen as two examples of the IoT data marking.
- The ability to ensure sufficient precision of the collected IoT data according to the quality required by the usage of these collected IoT data.
- The ability to support efficient data collection in order to relieve the load of the IoT system.

  NOTE 2 – As an example, when collecting IoT data from multiple devices, the usage of a data collection scheduler, which schedules the sequence and frequency of the IoT data collection, can relieve the load of the IoT system.

8.3 Big data pre-processing

Big data pre-processing capability implements the data pre-processing operation as described in clause 7.1.4.

According to the requirements in clauses 7.2.2, 7.3.2, 7.5.2 and 7.6.2, the following abilities are provided by the big data pre-processing capability:

- The ability to support the data format transformation of IoT data according to pre-defined data models.

  NOTE – Missing data fields are supplemented with default values during transformation.
- The ability to support the association of time stamp and collected IoT data.
- The ability to support the semantic annotation for IoT data.
- The ability to support data quality enhancement operations, e.g., de-noising and de-duplication.

8.4 Big data analysis

Big data analysis capability implements the data analysis operation as described in clause 7.1.5.

According to the requirements in clauses 7.3.5, 7.5.5 and 7.6.5, the following abilities are provided by the big data analysis capability:

- The ability to analyse IoT data according to user algorithms and profiles.
- The ability to provide pre-defined data analysis functions (e.g., data statistics and data mining algorithms).
- The ability to support data analysis for classifying the type of IoT services, the operational states of devices and the network status.
- The ability to support distributed data analysis.

8.5 Big data transfer

Big data transfer capability implements the data transfer operation as described in clause 7.1.2.

According to the requirements in clauses 7.2.5, 7.3.4, 7.4.1, 7.5.4 and 7.6.4, the following abilities are provided by the big data transfer capability:

- The ability to support IoT data transfer configuration concerning required network resources (e.g., data buffer and network QoS) to enable transfer quality.
- The ability to support aggregation of IoT data to enable reduction of transfer overhead.
- The ability to support transfer buffer management to mitigate network congestion issues.
- The ability to compress IoT data before transfer in order to reduce transfer load.

8.6 Big data storage

Big data storage capability implements the data storage operation as described in clause 7.1.3.

According to the requirements in clauses 7.2.3, 7.3.3, 7.5.3 and 7.6.3, the following abilities are provided by the big data storage capability:

- The ability to support secure and high-volume data storage for the IoT data with various formats. The data storage can be either locally or remotely.

  NOTE – For high-volume data storage, cloud based data storage infrastructure can be used. For time critical services, edge computing based data storage can be used.
- The ability to support secure and high-volume data storage for the IoT data with semantic annotation. The semantic data and non-semantic data can be separately stored in different databases to facilitate the IoT data processing.

- The ability to support high-volume data read and write operations with fast speed. Data caching, synchronization and indexing mechanisms can be used to achieve fast speed IoT data read and write.
- The ability to support data integrity checking and life cycle management of stored data (e.g., classification of data activation level and storage in different storage spaces as appropriate).

8.7 Big data time synchronization management

Big data time synchronization management capability provides support to IoT data operations and implements time synchronization and time calibration among different components of the IoT.

According to the requirements in clauses 7.2.2, 7.3.2 and 7.5.2, the following abilities are provided by the big data time synchronization management capability:

- The ability to provide time synchronization and time calibration among different components in the IoT to ensure correct time stamping of IoT data.
- The ability to support the following features, but not limited to:
  - high precision and low power consumption

  NOTE – Different components of the IoT may require different time precision and power consumption levels.

  - Time conflict management and negotiation among different components in the IoT.
  - Support of centralized synchronization mode or distributed synchronization mode.

8.8 Big data visualization

Big data visualization capability implements the data visualization operation as described in clause 7.1.6.

According to the requirements in clauses 7.2.6 and 7.6.6, the following abilities are provided by the big data visualization capability:

- The ability to show IoT data according to a consumer's customized needs (e.g., via charts, tables, figures).
- The ability to support data visualization for the various types of IoT data (according to their variety, velocity, volume, etc.).

8.9 Big data query

Big data query capability implements the data query operation as described in clause 7.1.7.

According to the requirements in clauses 7.5.6 and 7.6.7, the following abilities are provided by the big data query capability:

- The ability to support query on stored historical IoT data and raw IoT data.
- The ability to support query on the real-time processed results of the collected IoT data.
- The ability to support semantic query.

8.10 Big data security and privacy protection

Big data security and privacy protection capability, providing support to IoT data operations, implements security and privacy protection of data.

According to the requirements in clauses 7.5.3 and 7.6.6, the following abilities are provided by the big data security and privacy protection capability:

- The ability to support the security of IoT data access control, including access control on data query scope, so that the stored IoT data can be only accessed by authorized users.

- The ability to support the security of IoT data storage during the storage process and once data are stored.
- The ability to support the security of data transfer.
- The ability to support the protection of the privacy sensitive information during IoT data visualization.

Appendix I

Use case with IoT big data characteristics

(This appendix does not form an integral part of this Recommendation.)

I.1 City environment monitoring

Figure I.1 shows a city environment monitoring use case as an example showing how the IoT can handle data with IoT big data characteristics. In this use case, a large number of environment monitoring sensors and cameras will be deployed in the city to monitor the real-time conditions of the air, water, traffic, power and waste in the city.

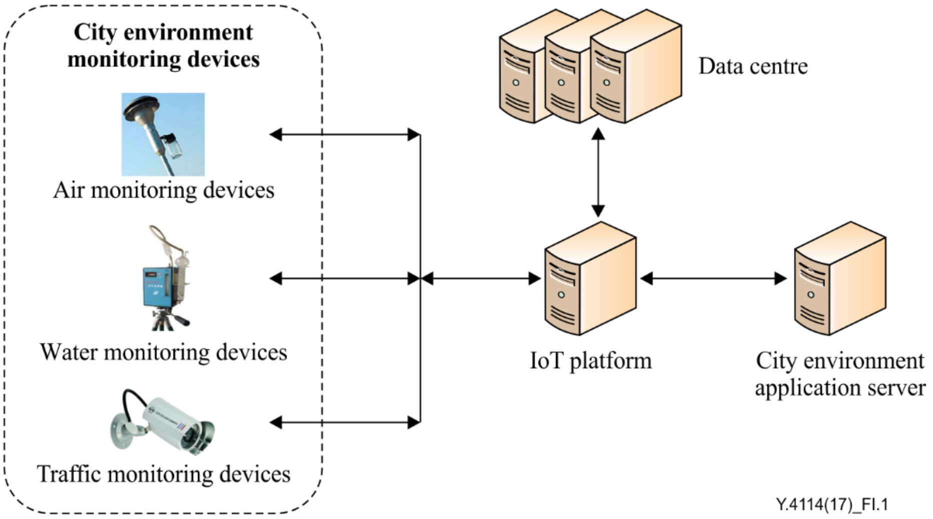
City environment monitoring devices
Air monitoring devices
Water monitoring devices
Traffic monitoring devices
Data centre
IoT platform
City environment application server
Y.4114(17)_FI.1

Figure I.1 – City environment monitoring use case

In each second, a large amount of data will be transferred from those sensors and cameras to the IoT platform. All the data will experience data collection, data transfer, data pre-processing, data storage and data analysis procedures.

– During the data collection procedure, the data collection schedulers in the IoT platform will arrange the collection period and collection sequence to relieve the network and the IoT platform load. The collected data from the city environment sensors will be partially annotated with time stamp and semantic information according to the capability of sensors.

– During the data transfer procedure, small data packets will converge into bigger packets to improve the transmission efficiency and some redundant data (e.g., video data) will be compressed before the data transmission.

– During the data pre-processing procedure, the collected environment data without time stamps will be annotated with time stamps by the IoT platform and the collected environment data in different formats will be transformed into a unified standard data format. For time-series environmental data, the missing or incorrect data detection and interpolation operation may be adopted.

– During the data storage procedure, the collected data will be stored in the data centre after a data pre-processing procedure.

– During the data analysis procedure, the city environment analysis applications will analyse the collected environmental data. Some city environment indexes (e.g., air quality index) will be calculated according to the collected environmental data and some city environment events (e.g., traffic jams) will be detected.

Appendix II

Relationships among IoT data roles, IoT business roles, IoT data operations and IoT components

(This appendix does not form an integral part of this Recommendation.)

II.1 Relationships between IoT data roles and IoT business roles

This appendix describes some typical relationships between IoT data roles (as described in clause 6) and IoT business roles (as described in [ITU-T Y.4000]) via an example of deployment scenarios.

Figure II.1 shows three IoT components, i.e., device, IoT platform and IoT application server, connected to each other via the network, another IoT component.

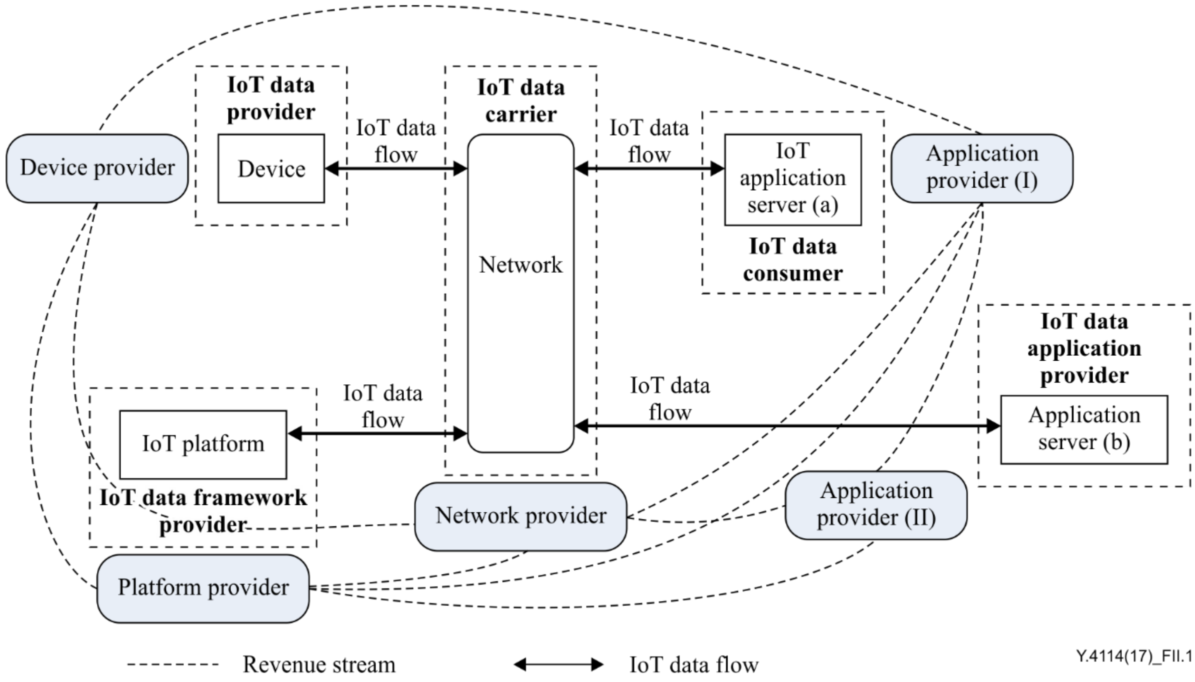

Figure II.1 – Example of deployment scenarios with respect to the relationship between IoT data roles and IoT business roles

In Figure II.1, the IoT business roles [ITU-T Y.4000] associated to the above four IoT components, i.e., device provider, platform provider, network provider and application provider, describe the relationships in terms of revenue streams. The two application providers shown in Figure II.1 aim to describe two different sub-roles in a given deployment scenario from the IoT data perspective:

– Application provider (I) provides IoT applications to the application customer (described in [ITU-T Y.4000]);

– Application provider (II) provides IoT data processing capabilities to application provider (I) via handling IoT data and providing the analysis results to application provider (I).

In Figure II.1, the interactions between the IoT data roles associated with the five IoT components, i.e., IoT data provider, IoT data carrier, IoT data framework provider, IoT data application provider and IoT data consumer, describe the relationships in terms of IoT data operations in this specific scenario:

– The IoT data provider role is played by the device provider collecting the IoT data;

– The IoT data framework provider role is played by the platform provider providing general IoT data processing capabilities and related infrastructure;

– The IoT data application provider role is played by the application provider (II) providing specific (e.g., application-domain specific) data processing related applications;

– The IoT data carrier role is played by the network provider supporting data transfer;

– The IoT data consumer role is played by the application provider (I) getting the desired data from the specific data processing related applications provided by the application provider (II) and providing customized applications (e.g., data visualization) to the application customer.

II.2 Relationships between IoT data roles and IoT data operations

In the IoT ecosystem, different IoT data roles are in charge of different IoT data operations. The Table II.1 gives a non-exhaustive association of the key IoT data operations with the IoT data roles described in clause 6.

NOTE – Table II.1 is not exhaustive. According to the specific deployment scenarios, the roles below can perform other IoT data operations not identified in Table II.1. For example, a device provider, when acting as IoT data consumer, can also perform data storage on devices.

Table II.1 – Key IoT data operations associated to the IoT data roles

IoT data roles\ IoT data operations	Data collection	Data transfer	Data pre-processing	Data analysis	Data query	Data storage	Data visualization
IoT data provider	√						
IoT data framework provider						√	
IoT data application provider			√	√	√		√
IoT data carrier		√					
IoT data consumer							

II.3 Relationships between IoT data operations and IoT components

The IoT components perform different operations on the IoT data. Table II.2 summarizes the key IoT data operations associated with the different IoT components.

NOTE – Table II.2 is not exhaustive. According to the specific deployment scenarios, the different IoT components can perform other IoT data operations not identified in Table II.2.

Table II.2 – Key IoT data operations performed by the different IoT components

IoT data operations\IoT components	Device	Gateway	Network	IoT platform	IoT application server
Data collection	√	√		√	√
Data pre-processing	√	√		√	√
Data storage	√	√		√	√
Data analysis	√			√	√
Data transfer	√	√	√	√	√
Data query		√		√	√
Data visualization/application	√				√

Bibliography

[b-ITU-T Y.4111] Recommendation ITU-T Y.4111/Y.2076 (2016), *Semantics based requirements framework of the Internet of things.*

[b-FG M2M D3.1] Focus Group on Machine-To-Machine Service Layer, *Deliverable 3.1, M2M service layer: APIs and protocols overview.* <http://www.itu.int/en/ITU-T/focusgroups/m2m/Pages/default.aspx>

[b-NIST Special Publication 1500-7] NIST Big Data Interoperability Framework: Volume 7, *Standards Roadmap.* <https://www.nist.gov/publications/nist-big-data-interoperability-framework-volume-7-standards-roadmap>

[b-JTC 1 Big data report] ISO/IEC JTC I Big data *Preliminary Report 2014.* <https://www.iso.org/files/live/sites/isoorg/files/developing_standards/docs/en/big_data_report-jtc1.pdf >

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, 2017