

**Question(s):** 17/13

Virtual, 16 July 2021

**TD****Source:** Editor**Title:** Draft revised Supplement 40 to ITU-T Y-3600 series (Y.sup.bdsr2): “Big data Standardization roadmap”**Purpose:** Proposal**Contact:** Suwook Ha  
ETRI  
Korea (Republic of)  
Tel: +82 42 860 5256  
Fax: +82 42 861 5404  
E-mail: sw.ha@etri.re.kr**Keywords:** big data, big data standardization, roadmap**Abstract:** This document is the draft Supplement of Y.sup.bdsr2, “Big data standardization roadmap”. This document includes the results of discussion on the Q17/13 meeting which was held in Virtual, 5–16 July 2021.

The following table shows discussion results for contributions.

Contribution No.	Source	Contribution title	Result and action
C63	ETRI	Y.sup.bdsr2: A proposal of modification of the standardization matrix in clause 10	Accepted with modification
C64	ETRI	Y.sup.bdsr2: A proposal of modification of the scope and structure	Accepted with modification

During this meeting, it is agreed as follows.

- Change title from ‘Big data standardization roadmap’ to ‘Big data and data handling standardization roadmap’;
- Extend the scope of document to cover data handling;
- Modify Landscape and Technical areas in clause 6 and clause 7 to cover big data and data handling.

It is recommended that future contributions cover following topics;

- Describe the related technical area in clause 7;
- Update/Add SDOs activities in clause 9 and Appendix I;

- Update/Add SDOs activities in the standardization matrix in clause 10;
- Modify/Split the standardization matrix in clause 10;
- But not limited to.

<Attachment> Revised draft Supplement Y.Suppl.40\_revision

<Attachment>

## **Draft Revised Supplement 40 to ITU-T Y-3600 series (formerly Y.sup.bdsr2)**

### **Big data and data handling standardization roadmap**

#### **Summary**

Supplement 40 to ITU-T Y-series Recommendations provides the standardization roadmap for big data in the telecommunication sector. It describes the landscape and conceptual ecosystem of big data from an ITU-T perspective, related technical areas, activities in standards development organizations (SDOs) and gap analysis.

#### **Table of Contents**

	<b>Page</b>
1	Scope..... 5
2	References..... 5
3	Definitions ..... 5
3.1	Terms defined elsewhere ..... 5
3.2	Terms defined in this Supplement ..... 5
4	Abbreviations and acronyms ..... 5
5	Conventions ..... 7
6	Landscape of big data and data handling from an ITU-T perspective ..... 7
7	Technical areas of big data and data handling ..... 8
7.1	Big data..... 8
7.2	Data quality and management ..... 9
7.3	Machine learning ..... 9
7.4	Cloud computing ..... 9
7.5	Internet of things..... 10
7.6	Security and privacy ..... 10
7.7	Deep packet inspection..... 10
7.8	Big data-driven networking ..... 11
7.9	Open data..... 11
7.10	Smart city..... 11
8	Conceptual model of bDig data ecosystem ..... 11
9	SDO activities ..... 12
9.1	ITU-T..... 12
9.2	ISO/IEC JTC 1 ..... 16

9.3	ISO/TC 69 .....	18
9.4	W3C.....	18
9.5	OASIS.....	20
9.6	Data Mining Group.....	21
9.7	TM Forum .....	21
10	Gap analysis in big data standardization .....	21
Appendix I Summaries of referenced standardization work items .....		25
I.1	ITU-T references and associated summaries.....	25
I.2	ISO/IEC JTC 1 References and associated summaries .....	33
I.3	ISO/TC 69 References and associated summaries .....	42
I.4	W3C references and associated summaries.....	42
I.5	OASIS references and associated summaries.....	44
I.6	Data Mining Group references and associated summaries.....	44
I.7	TM Forum references and associated summaries .....	45
Bibliography.....		46

## **Draft Revised Supplement 40 to ITU-T Y-3600 series (formerly Y.sup.bdsr2)**

[Editor's note in 22 July 2020] *Based on the discussion result of Q17 meeting, these were recommended to update SDO's activities, in particular the change of the roles or responsibilities of each ITU-T SG and sub-committee for the next study period. Contributions are invited.*

### **Big data and data handling standardization roadmap**

#### **1 Scope**

Data is of high value to building applications and services based on future computing. This Supplement provides the standardization roadmap for data in the telecommunication sector including not only big data but also data usage, processing, analysing, exchanging, sharing, and data quality assessment in terms of data handling. It addresses the following subjects:

- landscape of data handling from an ITU-T perspective;
- conceptual model of big data ecosystems;
- related technical areas of data handling;
- data activities in standards development organizations (SDOs);
- standardization gap analysis.

#### **2 References**

- [ITU-T Y.2060] Recommendation ITU-T Y.2060 (2013), *Overview of the Internet of things*.
- [ITU-T Y.3300] Recommendation ITU-T Y.3300 (2014), *Framework of software-defined networking*.
- [ITU-T [Y.3500](#)] Recommendation ITU-T Y.3500 (2014), *Information technology – Cloud computing – Overview and vocabulary*.
- [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 – 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:

- AI                    Artificial Intelligence
- AMQP                Advanced Message Queuing Protocol

API	Application Program Interface
BDaaS	Big Data as a Service
BDC	Big Data service Customer
bDDN	big Data-Driven Networking
BDIP	Big Data Infrastructure Provider
BDSP	Big Data Service Provider
CSC	Cloud Service Customer
CSP	Cloud Service Provider
CSV	Comma-Separated Values
DCAT	Data Catalogue Vocabulary
DCIM	Data Centre Infrastructure Management
DMG	Data Ming Group
DP	Data Provider
DPI	Deep Packet Inspection
DPVCG	Data Privacy Vocabularies and controls CG
DSA	Data Sharing Agreement
ERCIM	the European Research Consortium for Informatics and Mathematics
GDPR	General Data Protection Regulation
HTTP	Hypertext Transfer Protocol
ICT	Information and Communications Technology
IEC	International Engineering Consortium
IoT	Internet of Things
IRIs	Internationalized Resource Identifiers
ISMS	Information Security Management System
ISO	International Organization for Standardization
JSON	Java Script Object Notation
JSON-LD	JSON for Linked Data
JTC 1	Joint Technical Committee 1
KVDB	Key-Value Database Application Interface
LDP	Linked Data Platform
M2M	Machine to Machine
ML	Machine Learning
MQTT	Message Queuing Telemetry Transport
NP	New work item Proposal
OASIS	Organization for the Advancement of Structured Information Standards
ODRL	Open Digital Rights Language
OEM	Original Equipment Manufacturers

OGC	Open Geospatial Consortium
OWL	Web Ontology Language
PFA	Portable Format for Analytics
PIA	Privacy Impact Assessment
PII	Personally Identifiable Information
PIMS	Privacy Information Management System
PMML	Predictive Model Markup Language
RDF	Resource Description Framework
RIF	Rule Interchange Format
SC	Subcommittee
SC&C	Smart Cities and Communities
SDN	Software-defined Networking
SDO	Standards Development Organization
SG	Study Group
SHACL	Shapes Constraint Language
SMB	Standard Management Board
SPARQL	SPARQL Protocol and RDF Query Language
TC	Technical Committee
TMB	Technical Management Board
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WG	Working Group
WTSA	World Telecommunications Standardization Assembly
XMILE	XML Interchange Language
XML	Extensible Markup Language

## **5 Conventions**

None.

## **6 Landscape of big data and data handling from an ITU-T perspective**

Our future relies on the wise use of data. With ICTs everywhere, data is generated everywhere. The huge amount of data accumulation has created a resource of enormous productive value. In the knowledge-driven modern economy, data will fuel innovation in all industrial sectors and public sector institutions. In this manner, standardization of big data and data handlings, such as ecosystems, data lifecycles, data security and management, and ICT applications supported by data will help make data more valuable in the following areas:

- Data brings artificial intelligence and machine learning to life;
- Data analytics brings more automation, precision and dynamics to network management and control;

- Data generated by smart cities informs public service improvement and environmental sustainability;
- Advances in multimedia depend on advances in data compression and analysis;
- Cloud computing systems record event data from aircraft, automobiles and other connected machines to pinpoint the cause of performance anomalies;
- In addition, security measures to protect data are very important for building trust in ICT.

## 7 Technical areas of big data and data handling

### 7.1 Big data

[ITU-T [Y.3600](#)] describes the characteristics and general concepts of the big data ecosystem. 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](#)].

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 following [ITU-T [Y.3600](#)]:

- **Volume:** refers to the amount of data collected, stored, analyzed 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.

Taking into account the described above v's characteristics, big data technologies and services can resolve many new challenges, and can also create more new opportunities 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 [b-CRA].
- **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 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 analyzed 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 resolution of the above challenges opens new opportunities to discover new data relationships, hidden patterns or unknown dependencies [ITU-T [Y.3600](#)].

## 7.2 Data quality and management

TBD

## 7.3 Machine learning

Machine learning (ML) is a significant trend in the industry. There is a broad agreement about the remarkable potential of ML to lead innovation, boost commerce and drive progress among leaders in the industry, academia and governments. IMT-2020 networks, featured by diverse services, e.g., mobile Internet, Internet of things, cloud computing and other types of communication, will lead to the growth of data traffic and to the need of handling large amounts of data in the network [ITU-T [Y.3174](#)].

## 7.4 Cloud computing

Cloud computing is a paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand. Key characteristics of cloud computing are [ITU-T [Y.3500](#)]:

- **Broad network access:** a feature where physical and virtual resources are available over a network and accessed through standard mechanisms that promote use by heterogeneous client platforms;
- **Measured service:** a feature where the metered delivery of cloud services is such that usage can be monitored, controlled, reported, and billed. This is an important feature needed to optimize and validate the delivered cloud service;
- **Multi-tenancy:** a feature where physical or virtual resources are allocated in such a way that multiple tenants and their computations and data are isolated from, and inaccessible to, one another;
- **On-demand self-service:** a feature where a cloud service customer can provision computing capabilities, as needed, automatically or with minimal interaction with the cloud service provider;
- **Rapid elasticity and scalability:** a feature where physical or virtual resources can be rapidly and elastically adjusted, in some cases automatically, to quickly increase or decrease resources;
- **Resource pooling:** a feature where a cloud service provider's physical or virtual resources can be aggregated in order to serve one or more cloud service customers.

Big data needs on-demand high-performance data processing and distributed storage as well as a variety of tools required to accomplish activities of the big data ecosystem. The burst nature of workloads makes cloud computing more appropriate for big data challenges such as scalability and timeliness [ITU-T [Y.3600](#)].

The relationship of cloud computing and big data mainly concerns two aspects:

- 1) Cloud computing can support big data using cloud infrastructure and services;
- 2) Big data services can provide public cloud analysis services, such as big data as a service (BDaaS).

## 7.5 Internet of things

The Internet of things (IoT) is 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 [ITU-T Y.2060].

The IoT can be perceived as a far-reaching vision with technological and societal implications. From the perspective of technical standardization, the IoT can be viewed as a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable ICT. 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, while ensuring that security and privacy requirements are fulfilled.

Big data in the context of IoT has some specific characteristics which do not necessarily pertain to big data in other technical areas. The prominent characteristics of big data in the context of IoT are: high variety (heterogeneity of data types and sources), high velocity (high frequency of data generation) and high volatility (data generated in a non-persistent stateless manner).

Some identified challenges concerning big data in the context of IoT are the following [b-Chen]:

- An increasing number of connected things generates huge amounts of data;
- The generated data are mainly semi-structured or even unstructured;
- The generated data may have different confidence and precision levels;
- The generated data are generally not useful until they are adequately "processed" (including pre-processing, analysis, etc.).

## 7.6 Security and privacy

Data security and privacy comprise the people, process and technology required to prevent destructive forces and unwanted actions [b-IBM]. From a big data perspective, security and privacy requirements are magnified by the characteristics of big data. Some identified challenges concerning big data in the context of security and privacy are the following [b-CSA]:

- Secure computations in distributed programming frameworks;
- Secure data storage and transactions logs;
- End-point input validation/filtering and data provenance;
- Real-time security/compliance monitoring;
- Scalable and composable privacy-preserving data mining and analytics;
- Anonymization and de-identification.

NOTE – The technical area of big data in clause 7.4 and clause 7.5 also helps to solve a number of security and privacy challenges by being able to integrate vast quantities of heterogeneous log data (from firewalls, hosts, routers, etc.) arriving at network speed and apply advanced analytics to identify threat and intrusion patterns.

## 7.7 Deep packet inspection

Deep packet inspection (DPI) is a form of filtering used to inspect data packets sent from one computer to another over a network. Software-based DPI, provides advanced traffic analysis and multidimensional reporting, showing the possibility of making off-the-shelf hardware work at actual line rates. Software-based DPI can be pervasively deployed in the network, providing much better analysis capabilities, as well as simpler mechanisms for deployment, update, testing and scaling to changing workloads [b-ITU-T DPI].

## 7.8 Big data-driven networking

Big data-driven networking (bDDN) is a group of technologies and methods to facilitate network operation, administration, maintenance and optimization, etc., based on the big data generated by the network and a series of methods and tools. That is to say, big data generated by the network is used to serve the network and make the network better. bDDN solves this problem by introducing and applying the big data technology to the framework of future networks [b-ITU-T DDN].

## 7.9 Open data

Open data is accessible public data that people, companies, and organizations can use to launch new ventures, analyze patterns and trends, make data-driven decisions, and solve complex problems. Open data includes two basic features: the data must be publicly available for anyone to use, and it must be licensed in a way that allows for its reuse [b-theguardian]. Open data is more focused on a horizontal scaling of big data sources.

The main technical issues for open data are as follows:

- Data publication: metadata supporting machine readability, data format, and licenses;
- Data finding: data identification, data semantics, and data access;
- Data provenance: data quality, data lineage tracking, and data versioning.

## 7.10 Smart city

The accelerating population density in urban areas is increasing the pressure on the existing infrastructures to meet the needs of inhabitants. Accordingly, there is an increasing demand for connected cities with pervasive embedded devices, to improve quality of IoT and SC&C services.

While traditional urban information databases and analytics architectures and infrastructures remain essential, with the growing data management demands, specific capabilities and capacities are required to be able to handle diverse and complex data streams from different sources. This data needs to be processed and managed properly to maximize its value in a secure manner, while complementing it with other information sources.

Urban areas increasingly require defined and comprehensive data processing and management frameworks and guidelines which incorporate reasonable measures to achieve a layered, data-centric paradigm to support SC&C. [b-FG-DPM]

## 8 Conceptual model of bDig data ecosystem

[ITU-T [Y.3600](#)] describes the roles and sub-roles of the big data ecosystem as shown in Figure 8-1.

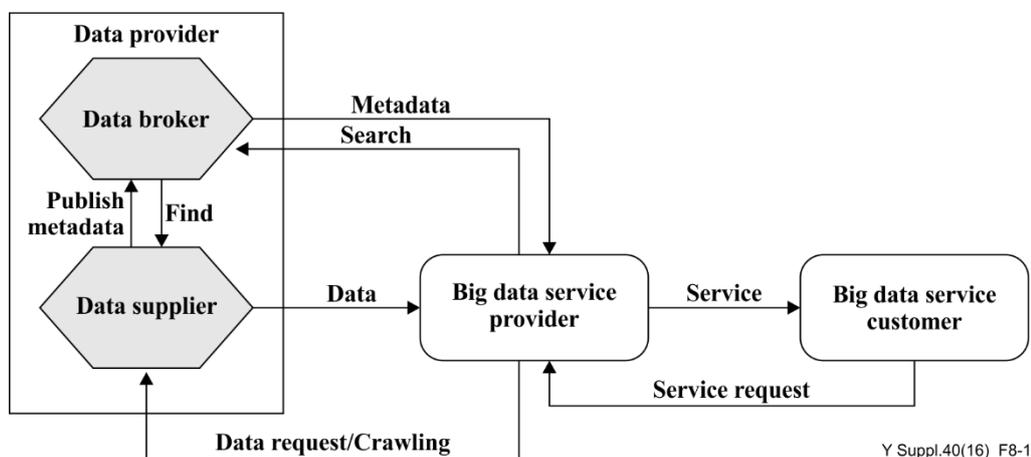


Figure 8-1 – Big data ecosystem (from [ITU-T [Y.3600](#)])

Data provider (DP) roles consists of two sub-roles:

- data supplier;
- data broker.

The data supplier provides data from different sources to the data broker, which can be accessed by the big data service provider (BDSP). The data supplier's activities include:

- generate data;
- create metadata information describing the data source(s) and relevant attributes;
- publish metadata information to access it.

The data broker serves as the connection between the data supplier and the BDSP. The data broker can act as a clearinghouse, open data mart, etc., and its activities include:

- providing a meta-information registry to data suppliers for publishing their data sources;
- finding on-line open-data sources and registering corresponding meta-information;
- providing a service catalogue to the BDSP for searching usable data.

The BDSP supports capabilities for big data analytics and infrastructure. The BDSP can act as a form of big data platform, extension of existing data analytics platform, etc. BDSP activities include:

- searching data sources (from data broker) and collecting data by requesting and crawling;
- storing data to a data repository;
- integrating data;
- providing tools for data analysis and visualization;
- supporting data management such as: data provenance, data privacy, data security, data retention policy, data ownership.

The big data service customer (BDC) is the end-user or a system, that uses the results or services from a BDSP. The BDC may produce new services or knowledge on consumer activities and furnish them outside of the big data ecosystem. BDC activities include:

- requesting big data services to the BDSP;
- using the outputs of big data services.

## 9 SDO activities

This clause describes SDO's activities with big data and data handling in order to identify the current status of standardization.

NOTE – A summary of each standard item is described in Appendix I.

### 9.1 ITU-T

*[Editor's note in 2018-10-22] The alignment of the description depth of each sub-clause is needed.*

#### 9.1.1 SG 13

ITU-T Study Group 13 (SG13) has been studying requirements, capabilities and mechanisms of future networks.

- **Q17/13** deals **requirements, ecosystem, and general capabilities** for big data. The primary focus of this question is to provide the necessary overall frameworks, definitions, and ecosystems including requirements, capabilities related to the integration or support of the cloud computing, big data model and technologies in telecommunication ecosystem [[ToR](#)] [[Work Program](#)]

- **Q18/13** deals with **functional architecture** for big data based on preceding study in Q17/13. Main focus of this question on big data is to provide big data functional architectures [[ToR](#)] [[Work Program](#)]
- **Q19/13** is deals with big **data governance** including data management, data preservation as well as lifecycle management of big data to provide the necessary overall **frameworks, definitions**, and ecosystems including **requirements**, capabilities [[ToR](#)] [[Work Program](#)]
- **Q7/13** has been studying big data driven networking (bDDN) and deep packet inspection (DPI). In terms of **big data application**, bDDN is one of good example meanwhile DPI has been playing an important role in collecting data from network and pre-processing the data. [[ToR](#)] [[Work Program](#)]

Table 9-1 lists the ITU-T SG13 deliverables and work items related to big data.

**Table 9-1 – ITU-T SG 13 deliverables and work items related to big data**

Study group	Reference	Title	Status
SG 13	[ITU-T <a href="#">Y.3600</a> ]	Big data – Cloud computing based requirements and capabilities	Published 2015
SG 13	[ITU-T Y.3601]	Big data – Framework and requirements for data exchange	Published 2018
SG 13	[ITU-T Y.bdi-reqts]	Big data – Overview and functional requirements for data integration	3Q 2021
SG 13	[ITU-T Y.3602]	Big data – Functional requirements for data provenance	Published 2018
SG 13	[ITU-T Y.3603]	Big data – Requirements and conceptual model of metadata for data catalogue	Published 2019
SG 13	[ITU-T Y.3604]	Big data – Overview and requirements for data preservation	Published 2019
SG 13	[ITU-T Y.3505]	Cloud computing – Overview and functional requirements for data storage federation	Published 2018
SG 13	[ITU-T Y.3519]	Cloud computing – Functional architecture of big data as a Service	Published 2018
SG 13	[ITU-T Y.3605]	Big data – Reference architecture	Published 2020
SG 13	[ITU-T Y.3652]	Big data driven networking – requirements	Published 2020
SG 13	[ITU-T Y.bDDN-FunArch]	Functional architecture of big data driven networking	Mar. 2021
SG 13	[ITU-T Y.bDDN-MCMec]	Management and control mechanisms of big data driven networking	Mar. 2021
SG 13	[ITU-T Y.3650]	Framework of big data driven networking	Published 2017
SG 13	[ITU-T Y.3651]	Big data driven mobile network traffic management and planning	Published 2018
SG 13	[ITU-T Y.Sup.50]	Use cases and application scenarios of big data driven networking	Published 2018
SG 13	[ITU-T Y.bDPI-Mec]	Mechanism of deep packet inspection applied in network big data context	Mar. 2021

Study group	Reference	Title	Status
SG 13	[ITU-T Y.3302]	Functional architecture of software-defined networking	Published 2017
SG 13	[ITU-T Y.bDDN-MLMec]	Mechanisms of machine learning for big data driven networking	Oct. 2021

### 9.1.2 SG 17

ITU-T SG 17 is responsible for building confidence and security in the use of ICTs, and deals with the security and privacy issues of cloud computing. These activities on cloud computing can be applied to the area of big data as well. Table 9-2 lists the ITU-T SG17 deliverables and work items related to big data.

**Table 9-2 – ITU-T SG 17 deliverables and work items related to big data**

Study group	Reference	Title	Status
SG 17	[ITU-T X.1601]	Security framework for cloud computing	Published 2014
SG 17	[ITU-T X.1641]	Guidelines for cloud service customer data security	Published 2016
SG 17	[ITU-T X.1147]	Security requirements and framework for big data analytics in mobile internet services	Published 2018
SG 17	[ITU-T X.1750]	Guidelines on security of big data as a service for Big Data Service Providers	Published 2020
SG 17	[ITU-T X.sgBDIP]	Security guidelines for big data infrastructure and platform	Mar. 2021
SG 17	[ITU-T X.1751]	Security guidelines of big data lifecycle management for telecommunication operators	Published 2020
SG 17	[ITU-T X.1376]	Security-related misbehaviour detection mechanism using big data for connected vehicles	Published 2020

### 9.1.3 SG 20

ITU-T SG 20 is responsible for Internet of Things (IoT) and its applications, and smart cities and communities (SC&C) including Big Data aspects of IoT and SC&C, e-services and smart devices for SC&C.

- **Q1/20** deals end to end connectivity, networks, interoperability, infrastructures and Big Data aspects related to IoT and SC&C. In terms of Big Data, Q1/20 is responsible for 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, and more. [\[ToR\]](#) [\[Work Program\]](#)

Table 9-3 lists the ITU-T SG 20 deliverables and work items related to big data.

**Table 9-3 – ITU-T SG 20 deliverables and work items related to big data**

<b>Study group</b>	<b>Reference</b>	<b>Title</b>	<b>Status</b>
SG 20	[ITU-T Y.4114]	Specific requirements and capabilities of the Internet of Things for Big data	Published 2017
SG 20	[ITU-T Y.4461]	Framework of Open Data in Smart Cities	Published 2020
NOTE – Clause I.1 contains a description of each cited reference.			

#### 9.1.4 SG 3

ITU-T SG 3 is responsible for tariff and accounting principles including related telecommunication economic and policy issues. In November 2016, Study Group 3 was entrusted by the World Telecommunications Standardization Assembly (WTSA-16) with the study of 11 questions, including a new question on the economic and policy aspects of big data and digital identity in international telecommunications services and networks. Table 9-4 lists the ITU-T SG 3 deliverables and work items related to big data.

**Table 9-4 – ITU-T SG 3 deliverables and work items related to big data**

<b>Study group</b>	<b>Reference</b>	<b>Title</b>	<b>Status</b>
SG 3	[ITU-T D.princip_bigdata]	ITU-T D. policy framework and principles for data protection in the context of big data relating to international telecommunication services	Apr. 2020
SG 3	[ITU-T Study_bigdata]	Technical Paper on economic and policy aspects of Big Data in international telecommunication services and networks	2020
NOTE – Clause I.1 contains a description of each cited reference.			

#### 9.1.5 SG 16

ITU-T SG 16 is responsible for multimedia coding, systems and applications, including the coordination of related studies across the various ITU-T SGs. Table 9-5 lists the ITU-T SG 16 deliverables and work items related to big data.

**Table 9-5 – ITU-T SG 16 deliverables and work items related to big data**

<b>Study group</b>	<b>Reference</b>	<b>Title</b>	<b>Status</b>
SG 16	[ITU-T F.743.7]	Requirements for big data enhanced visual surveillance service	Published 2019
SG 16	[ITU-T F.743.20]	Assessment framework for big data infrastructure	Published 2020
SG 16	[ITU-T H.VSBD]	Architecture for Big Data Application in Visual Surveillance System	2020
NOTE – Clause I.1 contains a description of each cited reference.			

### 9.1.6 SG 5

ITU-T Study Group 5 (SG5) is responsible for studies on methodologies for evaluating ICT effects on climate change and publishing guidelines for using ICTs in an eco-friendly way. Under its environmental mandate SG5 is also responsible for studying design methodologies to reduce ICTs and e-waste's adverse environmental effects, for example, through recycling of ICT facilities and equipment. Table 9-6 lists the ITU-T SG 5 deliverable and work item related to big data.

**Table 9-6 – ITU-T SG 5 deliverable and work item related to big data**

Study group	Reference	Title	Status
SG 5	[ITU-T L.1305]	Data centre infrastructure management system based on big data and artificial intelligence technology	Published 2019
NOTE – Clause I.1 contains a description of each cited reference.			

### 9.2 ISO/IEC JTC 1

In November 2014, the ISO/IEC joint technical committee 1 (JTC 1) established the working group (WG) 9 on big data to begin big data standardization. With the establishment of the SC 42 (Artificial Intelligence) in November 2017 (confirmed in 2018 by ISO TMB and IEC SMB), WG 9's Program of the work was transferred to SC 42 and became SC42/WG 2 with the Terms of Reference on "Standardization in the relation to data in the context of artificial intelligence, big data, and data analytics".

JTC 1 subcommittee S) 27 has been developing standards for ICT security, which include generic methods, techniques and guidelines to address aspects of both security and privacy. SC 27 security and privacy are one of the cross-cutting aspect for ICT, and their methodologies are also applicable to big data security and privacy.

JTC 1/SC 38 focuses on the area of "Cloud Computing and Distributed Platforms". JTC1/SC 38 is developing ISO/IEC 19944 which describes data and their flow across devices and cloud services. Table 9-7 lists the JTC 1 deliverables and work items related to big data.

**Table 9-7 – JTC 1 deliverables and work items related to big data**

Sub group	Reference (Note)	Name/Title	Status
SC 42	[ISO/IEC 20546]	Information technology – Big data – Overview and vocabulary	Published 2019
SC 42	[ISO/IEC 20547-1]	Information technology – Big data – Reference architecture – Part 1: Framework and application process	Published 2020
SC 42	[ISO/IEC20547-2]	Information technology – Big data – Reference architecture – Part 2: Use cases and derived requirements	Published 2018
SC 42	[ISO/IEC 20547-3]	Information technology – Big data – Reference architecture – Part 3: Reference architecture	Published 2020
SC 42	[ISO/IEC 20547-5]	Information technology – Big data – Reference architecture – Part 5: Standards roadmap	Published 2018
SC 42	[ISO/IEC 24668]	Information technology - Artificial Intelligence – Process management framework for big data analytics	WD
SC 42	[ISO/IEC 5259-1]	Artificial intelligence – Data analytics and ML – Part 1: Overview, terminology and examples	WD

**Table 9-7 – JTC 1 deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
SC 42	[ISO/IEC 5259-2]	Artificial intelligence – Data analytics and ML – Part 2: Data quality measures	AWI
SC 42	[ISO/IEC 5259-3]	Artificial intelligence – Data analytics and ML – Part 3: Data quality manage requirements and guidelines	WD
SC 42	[ISO/IEC 5259-4]	Artificial intelligence – Data analytics and ML – Part 4: Data quality process framework	WD
SC 27	[ISO/IEC 20547-4]	Information technology – Big data – Reference architecture – Part 4: Security and privacy	Published 2020
SC 27	[ISO/IEC 20889]	Privacy enhancing data de-identification terminology and classification of techniques	Published 2018
SC 27	[ISO/IEC 27000]	Information technology – Security techniques – Information security manage systems – Overview and vocabulary	Published 2014
SC 27	[ISO/IEC 27001]	Information technology – Security techniques – Information security manage systems – Requirements	Published 2013
SC 27	[ISO/IEC 27002]	Information technology – Security techniques – Code of practice for information security controls	Published 2013
SC 27	[ISO/IEC 27018]	Information technology — Security techniques — Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors	Published 2019
SC 27	[ISO/IEC 27555]	Establishing a PII deletion concept in organizations	CD
SC 27	[ISO/IEC 27701]	Security techniques -- Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management -- Requirements and guidelines	Published 2019
SC 27	[ISO/IEC 29100]	Information technology – Security techniques – Privacy framework	Published 2011
SC 27	[ISO/IEC 29101]	Information technology -- Security techniques -- Privacy architecture framework	Published 2018
SC 27	[ISO/IEC 29134]	Information technology -- Security techniques -- Guidelines for privacy impact assessment	Published 2017
SC 27	[ISO/IEC 29151]	Information technology -- Security techniques -- Code of practice for personally identifiable information protection	Published 2017
SC 27	[ISO/IEC 29184]	Information technology -- Online privacy notices and consent	DIS
SC 27	[ISO/IEC 29190]	Information technology -- Security techniques -- Privacy capability assessment model	Published 2015
SC 38	[ISO/IEC 19944]	Information technology – Cloud computing – Cloud services and devices: Data flow, data categories and data use	Published 2017
SC 38	[ISO/IEC 19944:AMD1]	Information technology – Cloud computing – Cloud services and devices: Data flow, data categories and data use	PDAM <sup>1</sup> 1

<sup>1</sup> Proposed Draft Amendment

**Table 9-7 – JTC 1 deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
SC 38	[ISO/IEC TR 23186]	Information technology – Cloud computing – Framework of trust for processing of multi-sourced data	Published 2018
SC 38	[ISO/IEC 22624]	Information technology – Cloud computing – Taxonomy based data handling for cloud services	PRF <sup>2</sup>
SC 38	[ISO/IEC 23751]	Information technology – Cloud computing – Data sharing agreement (DSA) framework	WD
NOTE – Clause I.2 contains a description of each cited reference.			

### 9.3 ISO/TC 69

ISO/TC 69 is responsible for standardization in the application of statistical methods, including generation, collection (planning and design), analysis, presentation and interpretation of data [b-TC 69]. Under the TC 69, WG 12 was initiated to cover big data analytics. Table 9-8 lists ISO/TC 69 deliverables and work items related to big data.

**Table 9-8 – ISO/TC 69 deliverables and work items related to big data**

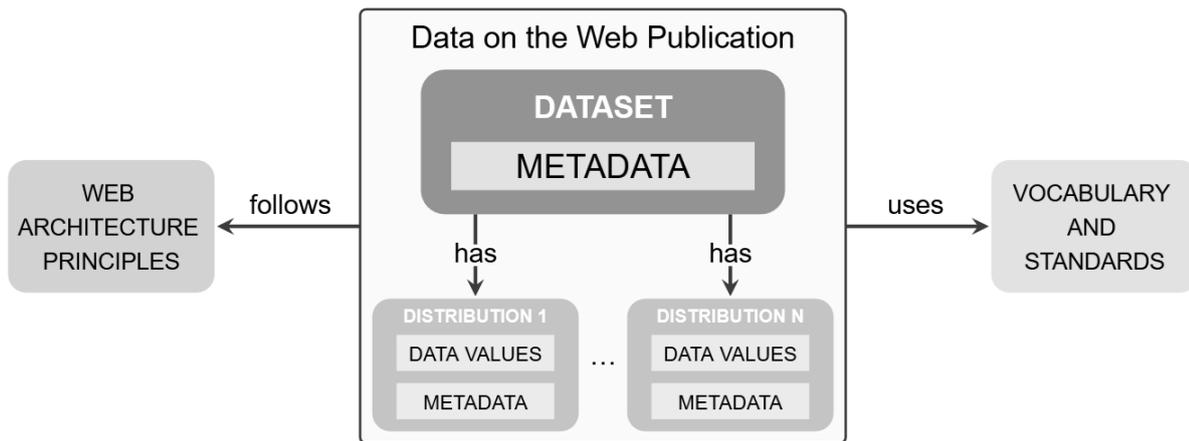
<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
-	[ISO 3534-5]	Statistics -- Vocabulary and symbols -- Part 5: Terms used in big data (predictive analytics)	NP
WG 12	[ISO 23347]	Statistics -- Big Data Analytics -- Data Science Life Cycle	NP
WG 12	[ISO 23348]	Statistics -- Big Data Analytics -- Model Validation	NP
NOTE – Clause I.3 contains a description of each cited reference.			

### 9.4 W3C

W3C doesn't have a Big Data chartered activity, but a more general Data Activity. The mission of W3C's Data Activity is to overcome the lack of data models and common vocabularies and facilitate potentially Web-scale data integration and processing, by providing standard data exchange formats, models, tools, and guidance. This builds upon W3C's previous work on RDF and Linked Data, and the corresponding suite of W3C Recommendations, e.g. for RDF, OWL, and SPARQL.

---

<sup>2</sup> Proof of a new International Standard



**Figure 9-1 – W3C Data activity**

There is a relationship to the Web of Things, where W3C is seeking to define standards for an object model as an abstraction layer over existing IoT standards, using programming language independent descriptions of things and their relationships.

The Dataset Exchange Working Group updated Data Catalog Vocabulary (DCAT) to version 2, and JSON for Linking Data W3C Community Group updates JSON-LD to version 1.1 as a Community Group Report.

W3C recently released the results of a study on tooling and practices for web data standardization, and we gratefully acknowledge funding for the study by the Open Data Institute. The next step is further discussions on how to accelerate W3C's role as a venue for developing community standards.

The Spatial Data on the Web Interest Group, a collaboration with the Open Geospatial Consortium (OGC) has started to develop additional best practices and to evaluate technology proposals that could improve the publication and consumption of spatial data on the Web (see on-going projects within the group)

W3C/ERCIM is participating in European projects "Boost 4.0" on big data in Industry 4.0, where we're responsible for standardization, data governance and certification. The project will field 10 lighthouse and 3 replication pilots across European manufacturers and accelerate the transition to Industry 4.0.

Data privacy vocabularies and controls CG (DPVCG) started at May 2018 to develop a taxonomy of privacy terms, which include in particular terms from the new European General Data Protection Regulation (GDPR), such as a taxonomy of personal data as well as a classification of purposes (i.e., purposes for data collection), and events of disclosures, consent, and processing such personal data [b-W3C-DPVCG].

The Permissions and Obligations Working Group recently closed after bringing to W3C Recommendation the Open Digital Rights Language (ODRL) Vocabulary & Expression and its Information Model.

Our current WG's with a connection to Big data also include: Verifiable Claims and there are plenty of existing relevant W3C Recommendations e.g. RDF core, RDF Schema, OWL, SPARQL, Turtle, JSON-LD, CSV, RIF, SHACL, Linked Data Platform (LDP), as well as those relating to XML, Privacy, or Web Services.

**Table 9-9 – W3C deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
CSV on the Web WG	[W3C MVTD]	Metadata Vocabulary for Tabular Data	Published 2015
CSV on the Web WG	[W3C MTDM]	Model for Tabular Data and Metadata on the web	Published 2015
Government Linked Data WG	[W3C DCAT2]	Data Catalog Vocabulary (DCAT) version 2	Published 2020
Government Linked Data WG	[W3C OO]	The Organization Ontology	Published 2014
Linked Data Platform WG	[W3C LDP 1.0]	Linked Data Platform 1.0	Published 2015
RDF WG	[JSON-LD 1.1]	A JSON-based Serialization for Linked Data	Published 2018
RDF WG	[RDF 1.1]	RDF 1.1 Concepts and Abstract Syntax	Published 2014
Permissions and Obligations WG	[ODRL V&E 2.2]	Open Digital Rights Language (ODRL) Vocabulary & Expression 2.2	Published 2018
NOTE – Clause I.4 contains a description of each cited reference.			

## 9.5 OASIS

The following Organization for the Advancement of Structured Information Standards (OASIS) technical committees (TCs) are relevant to big data [b-OASIS]:

- OASIS Advanced Message Queuing Protocol (AMQP) TC: defines a ubiquitous, secure, reliable and open Internet protocol for handling business messaging;
- OASIS Key-Value Database Application Interface (KVDB) TC: defines an open application programming interface for managing and accessing data from database systems based on a key-value model;
- OASIS Message Queuing Telemetry Transport (MQTT) TC: provides a lightweight publish/subscribe reliable messaging transport protocol suitable for communication in machine to machine (M2M) and IoT contexts where a small code footprint is required and/or network bandwidth is at a premium;
- OASIS XML Interchange Language (XMILE) for System Dynamics TC: defines an open XML protocol for sharing interoperable system dynamics models and simulations.

Table 9-10 lists the OASIS deliverables and work items related to big data.

**Table 9-10 – OASIS deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
AMQP TC	[OASIS AMQP 1.0]	Advanced Message Queuing Protocol Version 1.0	Published 2012

**Table 9-10 – OASIS deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
MQTT TC	[OASIS MQTT 5]	Message Queuing Telemetry Transport Version 5	Published 2019
NOTE – Clause I.4 contains a description of each cited reference.			

## 9.6 Data Mining Group

The Data Mining Group (DMG) is a vendor led consortium that develops data mining related standards. The DMG develops Predictive Model Markup Language (PMML); which provides a way for applications to describe and exchange models produced by data mining and machine learning algorithms, and Portable Format for Analytics (PFA); which provides a mini-language for mathematical calculations.

**Table 9-11 – DMG deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
–	[DMG PMML 4.4]	Predictive Model Markup Language 4.3	Published 2019
–	[DMG PFA]	Portable Format for Analytics	Published 2016
NOTE – Clause I.5 contains a description of each cited reference.			

## 9.7 TM Forum

The TM Forum (formerly TeleManagement Forum) is a global member association for digital business. The TM Forum published "Guide book for big data analytics" describing best practices on big data. Table 9-12 lists the TM Forum deliverables and work items related to big data.

**Table 9-12 TM Forum deliverables and work items related to big data**

<b>Sub group</b>	<b>Reference</b> (Note)	<b>Name/Title</b>	<b>Status</b>
–	[TMF BDAG]	The Big Data Analytics Guidebook	Published 2015
NOTE – Clause I.6 contains a description of each cited reference.			

## 10 Gap analysis in big data standardization

This clause provides a matrix for gap analysis and the related standardization activities with data handling to identify standardization gaps.

The matrix is composed of two axes. The horizontal axis describes document categories that cover the subject of applications as follows:

- **General, definition:** the standard which provides general descriptions or terms and definitions of the technology;

- **Common requirement:** the standard which provides use cases and derived general/functional requirements;
- **System architecture:** the standard which provides a reference architecture, and functional architecture for data application;
- **Common interface:** the standard which provides a common interface, API and/or its profile;
- **Data model and format:** the standard which provides data model or protocol including scheme and/or its encoding format;
- **Guideline:** the standard which provides a guideline for data handling;
- **Informative document:** the document which provides technical and/or not normative information about data handling (e.g. service scenarios and examples).

The vertical axis describes the related technical scope to support data handling as follows:

- **Overall concept:** the common concept of data handling and related applications;
- **Data sharing and data collection:** a technical scope of data publishing, sharing, collection, and transaction, etc.;
- **Data preparation:** a technical scope of manipulating raw data into a form that can be used for certain data processing;
- **Data service:** a technical scope of service that provides data analytics, machine learning, etc.;
- **Data quality and trustworthiness:** a technical scope of data management to ensure the data quality and trustworthiness of data;
- **Security and privacy:** a technical scope of security for big data system including handling of personal identification information;
- **Data governance, management and policy:** a technical scope of data handling procedure, rules and policy within the organization level;
- **Applications:** a technical scope of applications with an enhancement of its capabilities through data-driven technology or applying data-driven technology to applications;
- **Others:** data handling related technologies which are not described above.

NOTE 1 – The items on the horizontal axis are not subordinated to the different technologies.

NOTE 2 – The items on the vertical axis can be modified with technology change.

NOTE 3 – A standard has more than one location on the matrix. In the case that one standard is included in multiple document categories (horizontal axis) or related technologies (vertical axis), it can be mapped several times.

Table 10-1 shows the standardization matrix related to big data and data handling.

**Table 10-1 – Standardization matrix of big data and data handling**

	General/ Definition	Requirement	System architecture	Common interface including API	Data model and format	Guideline	Informative document
<b>Overall concept</b>	ITU-T Y.3600 ITU-T Y.3601 ISO/IEC 20546 ISO/IEC 19944 ISO/IEC 19944 AMD1	ITU-T Y.3600	ITU-T Y.3519 ITU-T Y.3605 ISO/IEC 20547-3				ISO/IEC 20547-1

**Table 10-1 – Standardization matrix of big data and data handling**

	General/ Definition	Requirement	System architecture	Common interface including API	Data model and format	Guideline	Informative document
<b>Data sharing and collection</b>	ITU-T Y.3601 ITU-T Y.4461 ISO/IEC 23751	ITU-T Y.3601 ITU-T Y.4461			ITU-T Y.3603 OASIS AMQP OASIS MQTT W3C DCAT W3C JSON-LD W3C LDP W3C RDF W3C OO W3C MVT W3C MTDW		
<b>Data preparation</b>	ITU-T Y.bdi-reqts	ITU-T Y.bdi- reqts					
<b>Data service</b>	ISO 3534-5			DMG PFA	DMG PMML DMG PFA		ISO 23347 ISO 23348 TMF BDAG ISO/IEC 20547-2
<b>Data quality and trustworthiness</b>	ITU-T Y.3602 ISO/IEC 23751 ISO/IEC 5259-1 ISO/IEC 5259-2 ISO/IEC 5259-4	ITU-T Y.3602 ISO/IEC 5259-3 ISO/IEC 5259-4			ITU-T Y.3602	ISO/IEC 5259-2 ISO/IEC 5259-3	ISO/IEC TR 23186
<b>Security and Privacy</b>	ITU-T X.1601 ITU-T D.princip_bigdata ITU-T X.1376 ISO/IEC 20889 ISO/IEC 27000 ISO/IEC 27018 ISO/IEC 27555 IEO/IEC 29100 ISO/IEC 29184	ITU-T X.1147 ITU-T Y.3604 ISO/IEC 20547-4 ISO/IEC 27001 ISO/IEC 27701	ISO/IEC 29101		W3C ODRL V&E 2.2	ITU-T X.1641 ITU-T X.1750 ITU-T X.sgBDIP ITU-T X.1751 ISO/IEC 20889 ISO/IEC 27002 ISO/IEC 27701 ISO/IEC 29134 ISO/IEC 29151 ISO/IEC 29190 ISO/IEC 27018	
<b>Data governance, management and policy</b>	ISO/IEC 24668	ITU-T F.743.20 ISO/IEC 5259-3				ISO/IEC 22624	
<b>Applications</b>	ITU-T Y.bDPI-Mec ITU-T Y.3650 ITU-T Y.3651 ITU-T Y.bDDN- MCMec ITU-T Y.bDDN- MLMec	ITU-T F.743.7 ITU-T Y.4114 ITU-T Y.3651 ITU-T Y.3652 ITU-T Y.3505 ITU-T L.1305	ITU-T H.VSBD ITU-T Y.bDDN- FunArch				ITU-T Y.Sup.50
<b>Others</b>			ITU-T Y.3302				ISO/IEC 20547-5 ITU-T Study_big data

NOTE 4 – The bold letter items in Table 10-1 are ITU-T work in progress activities.

According to the gap analysis in Table 10-1:

- ITU-T has been focusing on 'general/definition', 'requirement' and 'guideline' with each technical area described in vertical axis;
- It is expected that standardization efforts of ITU-T will be moved to 'system architecture' of each technical area;
- Consideration on standardizing 'data preparation, and data service' are needed;

- The entries under the column 'Common interface' of each of technical standardization areas are empty except 'DMG PFA'. These areas are being developed by open source projects, so ITU-T has to consider establishing relationships with them.

## Appendix I

### Summaries of referenced standardization work items

This appendix provides the summaries of the big data related SDO standardization items specified in clause 9.

NOTE – The summary text comes from the 'scope' or the corresponding part of each item such as 'overview', 'introduction', etc.

#### I.1 ITU-T references and associated summaries

[ITU-T [Y.3600](#)]

Big data – Cloud computing based requirements and capabilities

This proposed Recommendation provides an approach to use cloud computing to meet existing challenges in the use of big data.

The scope of this proposed Recommendation includes:

Overview of big data:

- Cloud computing based big data system context and benefits;
- Cloud computing based big data requirements;
- Cloud computing based big data capabilities.

Overview of cloud computing based big data:

- Big data system context and its activities;
- Cloud computing based big data requirements;
- Cloud computing based big data capabilities;
- Cloud computing based big data use cases and scenarios.

URL: <https://www.itu.int/rec/T-REC-Y.3600-201511-I/en>

[ITU-T Y.3601]

Big data - framework and requirements for data exchange

This Recommendation specifies the framework and requirements for data exchange in a big data ecosystem. This Recommendation identifies general concepts, patterns, activities, and functional requirements based on the big data ecosystem and capabilities defined in [ITU-T Y.3600]. The functional requirements provided in this Recommendation are derived from use cases.

URL: <https://www.itu.int/rec/T-REC-Y.3601/en>

[ITU-T Y.bdi-reqts]

Big Data - Overview and functional requirements for data integration

This Recommendation provides overview for big data integration. Compared with traditional data integration, big data needs to consider the characteristics for integration in terms of integration process. This Recommendation also provides functional requirement, which is derived from use cases.

The scope of this Recommendation is to specify overview and functional requirements for big data integration:

- Overview of big data integration
- Functional requirements of big data integration
- Use cases of big data

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14074](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14074)

[ITU-T Y.3602]

Big data – Functional requirements for data provenance

This Recommendation specifies the functional requirements for data provenance in a big data ecosystem defined in [ITU-T Y.3600]. This Recommendation introduces data provenance as well as data provenance in big data ecosystem, and also provides conceptual model, operations, logical components, and functional requirements for big data provenance. The functional requirements provided in this Recommendation are derived from use cases.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=13651](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=13651)

[ITU-T Y.3603]

Big data - Requirements and conceptual model of metadata for data catalogue

This Recommendation introduces metadata concept as well as its usages in a big data lifecycle. This Recommendation provides requirements and a conceptual model of metadata for data catalogue to support a big data ecosystem defined in [ITU-T Y.3600]. The XML profiles based on the conceptual model described by UML and examples of metadata are provided in the appendix.

URL: [https://www.itu.int/rec/dologin\\_pub.asp?lang=e&id=T-REC-Y.3603-201912-P!!PDF-E&type=items](https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-Y.3603-201912-P!!PDF-E&type=items)

[ITU-T Y.3604]

Big data – Overview and requirements for data preservation

This Recommendation provides overview and requirements of big data preservation. It addresses the following subjects:

- Overview of big data preservation;
- Functional requirements of big data preservation;
- Use cases of big data preservation.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14076](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14076)

[ITU-T Y.3505]

Cloud computing – Overview and functional requirements for data storage federation

This Recommendation provides overview and functional requirements of data storage federation including benefits, configuration for logical components, and ecosystem of data storage federation as well as cloud computing based data storage federation. The functional requirements provided in this Recommendation are derived from use cases.

URL: <https://www.itu.int/rec/T-REC-Y.3505>

[ITU-T Y.3519]

Cloud computing - Functional architecture of Big Data as a Service

This Recommendation provides an overview of the big data as a service (BDaaS) functional architecture and defines the BDaaS functional architecture and cross-cutting aspects by specifying the functional components for the support of BDaaS.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=13627](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=13627)

[ITU-T Y.3605]

Big data – Reference architecture

This Recommendation specifies the big data reference architecture (BDRA). The Recommendation provides a description of reference

architecture concepts, user view, functional view and cross cutting aspects..

URL: <https://www.itu.int/rec/T-REC-Y.3605-202009-1>

[ITU-T Y.3652]

Big data driven networking – requirements

This Recommendation specifies requirements of big data driven networking.

The scope of this Recommendation includes:

- general requirements for big data driven networking;
- requirements of big data plane for big data driven networking;
- requirements of network plane for big data driven networking;
- requirements of management plane for big data driven networking;
- interface requirements for big data driven networking;
- security aspect requirements for big data driven networking.

URL: <https://www.itu.int/rec/T-REC-Y.3652-202006-1>

[ITU-T Y.bDDN-FunArch]

Functional architecture of big data driven networking

This Recommendation specifies functional architecture of big data driven networking.

- Functional architecture of big data plane for big data driven networking;
- Functional architecture of network plane for big data driven networking.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14620](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14620)

[ITU-T Y.bDDN-MCMec]

Management and control mechanisms of big data driven networking

This Recommendation specifies the management and control mechanisms of big data driven networking. Its scope includes:

- general mechanisms related to management and control aspects of big data driven networking;
- management mechanisms of big data driven networking;
- control mechanisms of big data driven networking;
- coordination mechanisms between management and control for big data driven networking;
- other consideration related to management and control mechanisms of big data driven networking.

URL: <https://www.itu.int/md/T17-SG13-190628-TD-WP2-0439>

[ITU-T Y.3650]

Framework of big data driven networking

This Recommendation specifies a framework for big-data-driven networking. The scope of this Recommendation includes the model architecture of big-data-driven networking (bDDN), the high-level capabilities of bDDN and the interface capabilities among different planes and layers.

URL: <https://www.itu.int/rec/T-REC-Y.3650/en>

- [ITU-T Y.3651] Big-data-driven networking – mobile network traffic management and planning
- This Recommendation specifies some technology aspects related to big-data-driven networking – mobile network traffic management and planning. The scope of this Recommendation includes: requirements, framework, reference points, performance and security considerations of big-data-driven networking – mobile network traffic management and planning.
- URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=13647](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=13647)
- [ITU-T Y.Sup.50] Use case and application scenario of big data driven networking
- This supplement specifies the use case and application scenario of big data driven networking.
- URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=13638](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=13638)
- [ITU-T Y.bDPI-Mec] Mechanism of deep packet inspection applied in network big data context
- This proposed Recommendation specifies mechanism of DPI for network big data. The scope of this proposed Recommendation includes:
- overview of big data processing procedure;
  - analysing role of DPI in big data processing procedure;
  - data classification mechanism used for DPI for big data;
  - data pre-processing mechanism used for DPI for big data;
  - coordination processing mechanism of DPI in network big data context;
  - interfaces between DPI and the upper-layer big data related method.
- URL: [http://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=10966](http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=10966)
- [ITU-T Y.3302] Functional architecture of software-defined networking
- This Recommendation defines the functional architecture of software-defined networking (SDN) by describing a layered architecture, the functional components of the architecture and its reference points. Details of multi-layer management are provided in Annex A and the role of orchestration is addressed in Appendix I. The architecture is based on [ITU-T Y.3300] which defines the SDN framework and [ITU-T Y.3301] which describes the functional requirements of the SDN architecture which the present Recommendation has to fulfil.
- URL: <https://www.itu.int/rec/T-REC-Y.3302/en>
- [ITU-T Y.bDDN-MLMec] Mechanisms of machine learning for big data driven networking
- This Recommendation specifies the mechanisms of machine learning for big data driven networking, its scope includes the following aspects:
- studying the procedures of machine learning applied in bDDN;
  - studying the general machine learning approach for bDDN;
  - studying the interfaces related to machine learning for bDDN;

- studying the learning and control path based on machine learning for bDDN;
- studying other aspects related to machine learning for bDDN.

URL: [https://www.itu.int/itu-t/workprog/wp\\_item.aspx?isn=15183](https://www.itu.int/itu-t/workprog/wp_item.aspx?isn=15183)

[ITU-T X.1601]

#### Security framework for cloud computing

This Recommendation analyses security threats and challenges in the cloud computing environment, and describes security capabilities that could mitigate these threats and address security challenges. A framework methodology is provided for determining which of these security capabilities will require specification for mitigating security threats and addressing security challenges for cloud computing.

URL: <https://www.itu.int/rec/T-REC-X.1601-201510-I>

[ITU-T X.1641]

#### Guidelines for cloud service customer data security

This Recommendation provides guidelines for cloud service customer (CSC) data security in cloud computing, for those cases where the cloud service provider (CSP) is responsible for ensuring that the data is handled with proper security. This is not always the case, since for some cloud services the security of the data is the responsibility of CSCs themselves. In other cases, the responsibility may be mixed.

This Recommendation identifies security controls for CSC data that can be used in different stages of the full data lifecycle. These security controls can differ when the security level of the CSC data changes. Therefore, this Recommendation provides guidelines on when each control should be used for best security practice.

URL: <https://www.itu.int/rec/T-REC-X.1641/en>

[ITU-T X.1147]

#### Security requirements and framework for big data analytics in mobile internet services

This Recommendation provides security framework and requirements for big data analytics in mobile internet services. The intent of this Recommendation is to study the challenges brought forward by big data analytics, and the specific security requirements for the mobile internet services as well as the security framework. The scope of this Recommendation will focus on security threats analysis, security requirements, and a security framework.

URL: <https://www.itu.int/rec/T-REC-X.1147-201811-I>

[ITU-T X.1750]

#### Guidelines on security of big data as a service for Big Data Service Providers

This Recommendation analyses security challenges faced by big data as a service (BDaaS) and provides guidelines for big data service providers (BDSPs) to secure BDaaS. It identifies security roles and responsibilities of BDaaS components and specifies a security framework for a big data infrastructure, including platforms, applications, analytics, interfaces and the BDaaS ecosystem. This Recommendation also specifies security protection measures that should be taken for activities or components related to BDaaS. This Recommendation is a high-level description of security requirements

for BDaaS implementation that focuses on BDaaS. BDaaS involves big data infrastructure providers (BDIPs) and big data application providers (BDAPs). Guidelines for BDIPs and BDAPs, as well as detailed guidance on BDaaS implementation lie outside the scope of this Recommendation.

URL: <https://www.itu.int/rec/T-REC-X.1750-202009-I>

[ITU-T X.sgBDIP]

Security guidelines for big data infrastructure and platform

This work item will start by referring to the appropriate Big Data infrastructure and platform from existing standardization work in the relevant fora.

Then this work item will develop a threat landscape assessment and methodology, will map the threats, and will propose security guidelines to protect the Big Data infrastructure and platform.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14567](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14567)

[ITU-T X.1751]

Security guidelines on big data lifecycle management for telecommunication operators

This Recommendation describes security vulnerabilities and establishes lifecycle management guidelines for telecommunication big data services. This Recommendation:

- introduces characteristics of telecommunication big data services and data categories;
- analyses security vulnerabilities of lifecycle management for telecommunication big data services;
- specifies security guidelines for data lifecycle management for telecommunication big data services.

When telecommunication operators provide big data services, the basic prerequisite is that the explicit consent of subscribers has been obtained. In addition, for telecommunication operators, provision of necessary data protection measures is recommended throughout the entire big data service process. Protection mechanisms for various data categories lie outside the scope of this Recommendation.

URL: <https://www.itu.int/rec/T-REC-X.1751-202009-I>

[ITU-T X.1376]

Security-related misbehaviour detection mechanism using big data for connected vehicles

This Recommendation describes a security-related misbehaviour detection mechanism for connected vehicles. The mechanism includes the following steps:

- Data capture. Definition of the types of data and information that can be captured from different sources, including automotive, infrastructure, original equipment manufacturers (OEMs) and suppliers, for misbehaviour detection. Data capture methods and procedures lie outside the scope of this Recommendation.
- Detection. Analysis of the data captured to detect misbehaviour.

This Recommendation applies to connected vehicles to detect misbehaviour by designers and security solution providers.

Notification utilization methods lie outside the scope of this Recommendation. URL: <https://www.itu.int/rec/T-REC-X.1376-202101-P>

[ITU-T Y.4114]

Specific requirements and capabilities of the IoT for Big Data

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. 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

URL: <http://www.itu.int/rec/T-REC-Y.4114>

[ITU-T Y.4461]

Framework of Open Data in Smart Cities

This Recommendation defines a framework of Open Data in Smart Cities, in order to promote the sharing of data between different entities in smart city, fully exploit potentialities of data in smart cities, and ultimately build better and smarter cities. The scope of this Recommendation includes:

- The concept of Open Data in Smart Cities.
- The benefits of Open Data in Smart Cities.
- The key phases of Open Data in Smart Cities.
- The key roles and activities in Open Data in Smart Cities.
- The framework of Open Data in Smart Cities.
- The general requirements of Open Data in Smart Cities.

URL: <https://www.itu.int/rec/T-REC-Y.4461-202001-I/en>

[ITU-T D.princip\_bigdata]

Policy framework and principles for data protection in the context of big data relating to international telecommunication services

This Recommendation proposes a policy framework and a set of principles for data protection in the context of big data relating to international telecommunication services.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14604](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14604)

[ITU-T Study\_bigdata]

Technical Paper on economic and policy aspects of Big Data in international telecommunication services and networks

The aim of the technical paper is to study the economic and policy aspects of big data in international telecommunication services and networks.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14127](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14127)

[ITU-T F.743.7]

Requirements for big data enhanced visual surveillance services

This Recommendation specifies Requirements for big data enhanced visual surveillance service.

The scope of this recommendation includes:

- Application scenarios;
- Service requirements;
- Functional requirements;
- Performance requirements;
- Security requirements.

URL: <https://www.itu.int/md/T17-SG16-190319-TD-PLN-0282>

[ITU-T F.743.20]

Assessment framework for big data infrastructure

This Recommendation specifies the requirements in assessment methods for big data infrastructure system. The scope of this Recommendation includes:

- Operation and maintenance management requirement in big data infrastructure;
- High availability requirement in big data infrastructure;
- Functional requirement in big data infrastructure;
- Compatibility requirement in big data infrastructure;
- Performance requirement in big data infrastructure;
- Security requirement in big data infrastructure;
- Multi-tenant requirement in big data infrastructure;
- Extensibility requirement in big data infrastructure.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14704](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14704)

[ITU-T H.VSBD]

Architecture for big data application in visual surveillance system

This Recommendation specifies the reference architecture for big data application in visual surveillance system. The scope of this Recommendation includes:

- Reference architecture for big data application in visual surveillance system;
- Reference point for big data application in visual surveillance system;
- Service flow for big data application in visual surveillance system;
- Key Parameter for big data application in visual surveillance system;
- Interworking for legacy visual surveillance system and big data analytics system.

URL: [https://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=14702](https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=14702)

[ITU-T L.1305]

Data centre infrastructure management system based on big data and artificial intelligence technology

This Recommendation describes specifications of a data centre infrastructure management (DCIM) system based on big data and artificial intelligence (AI) technology. The system will manage all infrastructure in the data centre at the same time through a comprehensive platform.

The scope of this Recommendation includes:

- network infrastructure of management systems;
- standardization of data collection of the installed module;
- interconnection among various kinds of monitoring sub-systems;
- requirements for different functions in a DCIM.

URL: <https://www.itu.int/rec/T-REC-L.1305-201911-I/en>

## I.2 ISO/IEC JTC 1 References and associated summaries

- [ISO/IEC 20546] Information technology – Big Data – Overview and vocabulary
- This document provides an overview of big data along with a set of terms and definitions. It provides a terminological foundation for big data-related standards.
- URL: <https://www.iso.org/standard/68305.html>
- [ISO/IEC 20547-1] Information technology – Big data reference architecture – Part 1: Framework and application process
- This technical report (under development) describes the framework of the big data reference architecture and the process for how a user of the standard can apply it to their particular problem domain.
- URL: <https://www.iso.org/standard/71275.html>
- [ISO/IEC 20547-2] Information technology – Big data reference architecture – Part 2: Use cases and derived requirements
- This document provides examples of big data use cases with application domains and technical considerations derived from the contributed use cases.
- URL: <https://www.iso.org/standard/71276.html>
- [ISO/IEC 20547-3] Information technology – Big data reference architecture – Part 3: Reference architecture
- This International Standard (under development) specifies the big data reference architecture. The reference architecture includes the big data roles, activities, and functional components and their relationships.
- URL: <https://www.iso.org/standard/71277.html>
- [ISO/IEC 20547-4] Information technology – Big data reference architecture – Part 4: Security and privacy fabric
- This International Standard (under development) specifies the underlying Security and Privacy fabric that applies to all aspects of the big data reference architecture including the big data roles, activities, and functional components.

URL: <https://www.iso.org/standard/71278.html>

[ISO/IEC 20547-5]

Information technology – Big data reference architecture – Part 5: Standards roadmap

This document describes big data relevant standards, both in existence and under development, along with priorities for future big data standards development based on gap analysis.

URL: <https://www.iso.org/standard/72826.html>

[ISO/IEC 24668]

Information Technology – Artificial Intelligence – Process management framework for Big data analytics

This document provides a framework for developing processes to effectively leverage big data analytics across the organization irrespective of the industries/sectors.

URL: <https://www.iso.org/standard/78368.html>

[ISO/IEC 5259-1]

Artificial Intelligence – Data quality for analytics and machine learning – Part 1: Overview, terminology, and examples

This document provides an overview of data quality for analytics and machine learning. It also discusses associated technologies and examples (e.g. use cases and usage scenarios).

URL: <https://www.iso.org/standard/81088.html>

[ISO/IEC 5259-2]

Artificial intelligence – Data analytics and ML – Part 2: Data quality measures

This document provides a data quality model, data quality measures, and guidance on reporting data quality in the context of analytics and machine learning (ML). This document builds on ISO 8000 series, ISO/IEC 25012 and ISO/IEC 25024. The aim of this document is to enable organizations to achieve their data quality objectives and is applicable to all types of organizations.

URL: <https://www.iso.org/standard/81860.html>

[ISO/IEC 5259-3]

Artificial Intelligence – Data quality for analytics and machine learning – Part 3: Data quality management requirements and guidelines

This document specifies requirements and provides guidance for establishing, implementing, maintaining and continually improving the quality for data used in the areas of analytics and machine learning.

URL: <https://www.iso.org/standard/81092.html>

[ISO/IEC 5259-4]

Artificial Intelligence – Data quality for analytics and machine learning – Part 4: Data quality process framework

This document provides general common organizational approaches, regardless of type, size or nature of the applying organization, to ensure data quality for training and evaluation in analytics and machine learning.

URL: <https://www.iso.org/standard/81093.html>

[ISO/IEC 20889]

Privacy enhancing data de-identification terminology and classification of techniques

This document provides a description of privacy-enhancing data de-identification techniques, to be used to describe and design de-identification measures in accordance with the privacy principles in ISO/IEC 29100. In particular, this document specifies terminology, a classification of de-identification techniques according to their characteristics, and their applicability for reducing the risk of re-identification. This document is applicable to all types and sizes of organizations, including public and private companies, government entities, and not-for-profit organizations, that are PII controllers or PII processors acting on a controller's behalf, implementing data de-identification processes for privacy enhancing purposes.

URL: <https://www.iso.org/standard/69373.html>

[ISO/IEC 27018]

Information technology — Security techniques — Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors

This document establishes commonly accepted control objectives, controls and guidelines for implementing measures to protect Personally Identifiable Information (PII) in line with the privacy principles in ISO/IEC 29100 for the public cloud computing environment. In particular, this document specifies guidelines based on ISO/IEC 27002, taking into consideration the regulatory requirements for the protection of PII which can be applicable within the context of the information security risk environment(s) of a provider of public cloud services. This document is applicable to all types and sizes of organizations, including public and private companies, government entities and not-for-profit organizations, which provide information processing services as PII processors via cloud computing under contract to other organizations. The guidelines in this document can also be relevant to organizations acting as PII controllers. However, PII controllers can be subject to additional PII protection legislation, regulations and obligations, not applying to PII processors. This document is not intended to cover such additional obligations.

URL: <https://www.iso.org/standard/76559.html> [ISO/IEC 27555]

Information technology — Security techniques — Establishing a PII deletion concept in organizations

The standard is intended for organizations that store and process PII “and other personal data”.

It will not address:

- Specific provisions in laws and contracts;
- Specific deletion rules for particular types of PII;
- Deletion mechanisms including those for cloud storage;
- Security of the deletion mechanisms; nor
- Specific techniques for de-identification of data.

The standard will enable organizations to meet the increasing demands of privacy/data protection regulation, supporting them in fulfilling the requirements. Standardizing the approach may facilitate harmonized catalogues of PII deletion rules for industrial sectors, clarifying requirements for IT systems processing personal data.

URL: <https://www.iso.org/standard/71673.html>

[ISO/IEC 27701]

Security techniques -- Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management -- Requirements and guidelines

This document specifies requirements and provides guidance for establishing, implementing, maintaining and continually improving a Privacy Information Management System (PIMS) in the form of an extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy management within the context of the organization. This document specifies PIMS-related requirements and provides guidance for PII controllers and PII processors holding responsibility and accountability for PII processing. This document is applicable to all types and sizes of organizations, including public and private companies, government entities and not-for-profit organizations, which are PII controllers and/or PII processors processing PII within an ISMS.

URL: <https://www.iso.org/standard/71670.html>

[ISO/IEC 29101]

Information technology -- Security techniques -- Privacy architecture framework

This document defines a privacy architecture framework that:

- specifies concerns for ICT systems that process PII;
- lists components for the implementation of such systems; and
- provides architectural views contextualizing these components.

This document is applicable to entities involved in specifying, procuring, architecting, designing, testing, maintaining, administering and operating ICT systems that process PII.

URL: <https://www.iso.org/standard/75293.html>

[ISO/IEC 29134]

Information technology -- Security techniques -- Guidelines for privacy impact assessment

This document gives guidelines for

- a process on privacy impact assessments; and
- a structure and content of a PIA report.

It is applicable to all types and sizes of organizations, including public companies, private companies, government entities and not-for-profit organizations. ISO/IEC 29134:2017 is relevant to those involved in designing or implementing projects, including the parties operating data processing systems and services that process PII.

URL: <https://www.iso.org/standard/75293.html>

[ISO/IEC 29151]

Information technology -- Security techniques -- Code of practice for personally identifiable information protection

ISO/IEC 29151:2017 establishes control objectives, controls and guidelines for implementing controls, to meet the requirements identified by a risk and impact assessment related to the protection of personally identifiable information (PII). In particular, this Recommendation | International Standard specifies guidelines based on ISO/IEC 27002, taking into consideration the requirements for processing PII that may be applicable within the context of an organization's information security risk environment(s). ISO/IEC 29151:2017 is applicable to all types and sizes of organizations acting as PII controllers (as defined in ISO/IEC 29100), including public and private companies, government entities and not-for-profit organizations that process PII.

URL: <https://www.iso.org/standard/62726.html>

[ISO/IEC 29184]

Information technology -- Online privacy notices and consent

This document specifies controls which will shape the content and the structure of online privacy notices as well as the process of asking for consent to collect and process personally identifiable information (PII) from PII principals. This document is applicable in any online context where a PII controller or any other entity processing PII informs PII principals of processing.

URL: <https://www.iso.org/standard/70331.html>

[ISO/IEC 29190]

Information technology -- Security techniques -- Privacy capability assessment model

ISO 29190:2015 provides organizations with high-level guidance about how to assess their capability to manage privacy-related processes. In particular, it

- specifies steps in assessing processes to determine privacy capability;
- specifies a set of levels for privacy capability assessment;
- provides guidance on the key process areas against which privacy capability can be assessed;
- provides guidance for those implementing process assessment; and
- provides guidance on how to integrate the privacy capability assessment into organizations operations.

URL: <https://www.iso.org/standard/45269.html>

[ISO/IEC 19944]

Information technology – Cloud computing – Cloud services and devices: data flow, data categories and data use

Establish common and functional ways of understanding and describing the breadth of the cloud service ecosystem.

Enumerate and define the types of connections that can exist between cloud services and customers where their devices are mobile.

Provide foundational concepts necessary to enable others to provide guidance concerning data locality, mobile ecosystem issues, and identity issues.

Identify the types of data that flow across the customers and cloud services ecosystem and that can help cloud customers'

URL:

[http://www.iso.org/iso/home/store/catalogue\\_tc/catalogue\\_detail.htm?csnumber=66674](http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=66674)

[ISO/IEC 20889]

Privacy enhancing data de-identification terminology and classification of techniques

This document provides a description of privacy-enhancing data de-identification techniques, to be used to describe and design de-identification measures in accordance with the privacy principles in ISO/IEC 29100. In particular, this document specifies terminology, a classification of de-identification techniques according to their characteristics, and their applicability for reducing the risk of re-identification. This document is applicable to all types and sizes of organizations, including public and private companies, government entities, and not-for-profit organizations, that are PII controllers or PII processors acting on a controller's behalf, implementing data de-identification processes for privacy enhancing purposes.

URL: <https://www.iso.org/standard/69373.html>

[ISO/IEC 27000]

Information technology – Security techniques – Information security manage systems – Overview and vocabulary

This International Standard provides the overview of information security management systems, and terms and definitions commonly used in the ISMS family of standards. This International Standard is applicable to all types and sizes of organization (e.g., commercial enterprises, government agencies, not-for-profit organizations).

URL: [http://www.iso.org/iso/catalogue\\_detail?csnumber=63411](http://www.iso.org/iso/catalogue_detail?csnumber=63411)

[ISO/IEC 27001]

Information technology – Security techniques – Information security manage systems – Requirements

This International Standard specifies the requirements for establishing, implementing, maintaining and continually improving an information security management system within the context of the organization. This International Standard also includes requirements for the assessment and treatment of information security risks tailored to the needs of the organization. The requirements set out in this International Standard are generic and are intended to be applicable to all organizations, regardless of type, size or nature.

URL: <https://www.iso.org/obp/ui/#iso:std:iso-iec:27001:ed-2:v1:en>

[ISO/IEC 27002]

Information technology – Security techniques – Code of practice for information security controls

This International Standard gives guidelines for organizational information security standards and information security management practices including the selection, implementation and management of controls taking into consideration the organization's information security risk environment(s).

URL: [http://www.iso.org/iso/catalogue\\_detail?csnumber=54533](http://www.iso.org/iso/catalogue_detail?csnumber=54533)

[ISO/IEC 27018]

Information technology — Security techniques — Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors

This document establishes commonly accepted control objectives, controls and guidelines for implementing measures to protect Personally Identifiable Information (PII) in line with the privacy principles in ISO/IEC 29100 for the public cloud computing environment. In particular, this document specifies guidelines based on ISO/IEC 27002, taking into consideration the regulatory requirements for the protection of PII which can be applicable within the context of the information security risk environment(s) of a provider of public cloud services. This document is applicable to all types and sizes of organizations, including public and private companies, government entities and not-for-profit organizations, which provide information processing services as PII processors via cloud computing under contract to other organizations. The guidelines in this document can also be relevant to organizations acting as PII controllers. However, PII controllers can be subject to additional PII protection legislation, regulations and obligations, not applying to PII processors. This document is not intended to cover such additional obligations.

URL: <https://www.iso.org/standard/76559.html>

[ISO/IEC 27555]

Information technology — Security techniques — Establishing a PII deletion concept in organizations

The standard is intended for organizations that store and process PII “and other personal data”.

It will not address:

- Specific provisions in laws and contracts;
- Specific deletion rules for particular types of PII;
- Deletion mechanisms including those for cloud storage;
- Security of the deletion mechanisms; nor
- Specific techniques for de-identification of data.

The standard will enable organizations to meet the increasing demands of privacy/data protection regulation, supporting them in fulfilling the requirements. Standardizing the approach may facilitate harmonized catalogues of PII deletion rules for industrial sectors, clarifying requirements for IT systems processing personal data.

URL: <https://www.iso.org/standard/71673.html>

[ISO/IEC 27701]

Security techniques -- Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management -- Requirements and guidelines

This document specifies requirements and provides guidance for establishing, implementing, maintaining and continually improving a Privacy Information Management System (PIMS) in the form of an extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy management within the context of the organization. This document specifies PIMS-related requirements and provides guidance for PII controllers and PII processors holding responsibility and accountability for PII processing. This document is applicable to all types and sizes of organizations, including public and private

companies, government entities and not-for-profit organizations, which are PII controllers and/or PII processors processing PII within an ISMS.

URL: <https://www.iso.org/standard/71670.html>

[ISO/IEC 29100]

Information technology – Security techniques – Privacy framework

This International Standard provides a privacy framework which:

- specifies a common privacy terminology;
- defines the actors and their roles in processing personally identifiable information (PII);
- describes privacy safeguarding considerations; and
- provides references to known privacy principles for information technology.

URL:

[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=45123](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45123)

[ISO/IEC 29101]

Information technology -- Security techniques -- Privacy architecture framework

This document defines a privacy architecture framework that:

- specifies concerns for ICT systems that process PII;
- lists components for the implementation of such systems; and
- provides architectural views contextualizing these components.

This document is applicable to entities involved in specifying, procuring, architecting, designing, testing, maintaining, administering and operating ICT systems that process PII.

URL: <https://www.iso.org/standard/75293.html>

[ISO/IEC 29134]

Information technology -- Security techniques -- Guidelines for privacy impact assessment

This document gives guidelines for

- a process on privacy impact assessments; and
- a structure and content of a PIA report.

It is applicable to all types and sizes of organizations, including public companies, private companies, government entities and not-for-profit organizations. ISO/IEC 29134:2017 is relevant to those involved in designing or implementing projects, including the parties operating data processing systems and services that process PII.

URL: <https://www.iso.org/standard/75293.html>

[ISO/IEC 29151]

Information technology -- Security techniques -- Code of practice for personally identifiable information protection

ISO/IEC 29151:2017 establishes control objectives, controls and guidelines for implementing controls, to meet the requirements identified by a risk and impact assessment related to the protection of personally identifiable information (PII). In particular, this Recommendation | International Standard specifies guidelines based on ISO/IEC 27002, taking into consideration the requirements for processing PII that may be applicable within the context of an

organization's information security risk environment(s). ISO/IEC 29151:2017 is applicable to all types and sizes of organizations acting as PII controllers (as defined in ISO/IEC 29100), including public and private companies, government entities and not-for-profit organizations that process PII.

URL: <https://www.iso.org/standard/62726.html>

[ISO/IEC 29184]

Information technology -- Online privacy notices and consent

This document specifies controls which will shape the content and the structure of online privacy notices as well as the process of asking for consent to collect and process personally identifiable information (PII) from PII principals. This document is applicable in any online context where a PII controller or any other entity processing PII informs PII principals of processing.

URL: <https://www.iso.org/standard/70331.html>

[ISO/IEC 29190]

Information technology -- Security techniques -- Privacy capability assessment model

ISO 29190:2015 provides organizations with high-level guidance about how to assess their capability to manage privacy-related processes. In particular, it

- specifies steps in assessing processes to determine privacy capability;
- specifies a set of levels for privacy capability assessment;
- provides guidance on the key process areas against which privacy capability can be assessed;
- provides guidance for those implementing process assessment; and
- provides guidance on how to integrate the privacy capability assessment into organizations operations.

URL: <https://www.iso.org/standard/45269.html>

[ISO/IEC TR 23186]

Information technology – Cloud computing – Framework of trust for processing of multi-sourced data

This document describes a framework of trust for the processing of multi-sourced data that includes data use obligations and controls, data provenance, chain of custody, security and immutable proof of compliance as elements of the framework.

URL: <https://www.iso.org/standard/74844.html?browse=tc>

[ISO/IEC 22624]

Information technology – Cloud computing – Taxonomy based data handling for cloud services

URL: <https://www.iso.org/standard/73614.html?browse=tc>

[ISO/IEC 23751]

Information technology – Cloud computing – Data sharing agreement (DSA) framework

URL: <https://www.iso.org/standard/76834.html?browse=tc>

### I.3 ISO/TC 69 References and associated summaries

- [ISO 3534-5]                      Statistics – Vocabulary and symbols – Part 5: Terms used in big data (predictive analytics)
- Defines terms used in the statistical analysis of very large data sets and may be used in the drafting of other International Standards.
- More specifically, it defines terms used in the field of statistics dealing with data sets that occur in the realm of so-called Big Data applications. The data sets can be large owing to one or more of the characteristics of volume, variety, velocity, and variability of the data. Volume refers to the size of the data set, which typically could exceed the storage capacity of the analyst’s personal computer. Variety indicates that the data could reside in diverse domains within distinct data repositories while consisting of diverse data types (e.g., continuous, discrete, categorical, ordinal, images, and so forth). Velocity is tied to the rate of generation and transmission of the data that could be expanding at a rate beyond the analyst’s ability to process and to assess the data in a timely fashion. Finally, variability suggests that the previous three “V’s” (volume, variety and velocity) could be deviating over time.
- URL: <https://www.iso.org/standard/75373.html?browse=tc>
- [ISO 23347]                      Statistics – Big Data Analytics – Data Science Life Cycle
- URL: <https://www.iso.org/standard/75289.html?browse=tc>
- [ISO 23348]                      Statistics – Big Data Analytics – Model Validation
- URL: <https://www.iso.org/standard/75290.html?browse=tc>

### I.4 W3C references and associated summaries

- [W3C MVTD]                      Metadata Vocabulary for Tabular Data
- Validation, conversion, display, and search of tabular data on the web requires additional metadata that describes how the data should be interpreted. This document defines a vocabulary for metadata that annotates tabular data. This can be used to provide metadata at various levels, from groups of tables and how they relate to each other down to individual cells within a table.
- URL: <http://www.w3.org/TR/tabular-metadata/>
- [W3C MTDM]                      Model for Tabular Data and Metadata on the web
- This document outlines a data model, or infoset, for tabular data and metadata about that tabular data that can be used as a basis for validation, display, or creating other formats. It also contains some non-normative guidance for publishing tabular data as CSV and how that maps into the tabular data model.
- URL: <http://www.w3.org/TR/2015/REC-tabular-data-model-20151217/>
- [W3C DCAT2]                      Data Catalog Vocabulary (DCAT) – Version 2
- This revision of DCAT has extended the previous version to support further use cases and requirements. These include the possibility of cataloging other resources in addition to datasets, such as data

services. The revision also supports describing relationships between datasets as well as between datasets and other cataloged resources. Guidance on how to document licenses and rights statements associated with the cataloged items is provided.

URL: <http://www.w3.org/TR/vocab-dcat-2/>

[W3C OO]

### The Organization Ontology

This document describes a core ontology for organizational structures, aimed at supporting linked data publishing of organizational information across a number of domains. It is designed to allow domain-specific extensions to add classification of organizations and roles, as well as extensions to support neighboring information such as organizational activities.

URL: <http://www.w3.org/TR/vocab-org/>

[W3C LDP 1.0]

### Linked Data Platform 1.0

LDP defines a set of rules for Hypertext Transfer Protocol (HTTP) operations on web resources, some based on RDF, to provide an architecture for read-write Linked Data on the web.

URL: <http://www.w3.org/TR/ldp/>

[W3C JSON-LD 1.1]

### A JSON-based Serialization for Linked Data

Linked Data is a way to create a network of standards-based machine interpretable data across different documents and Web sites. It allows an application to start at one piece of Linked Data, and follow embedded links to other pieces of Linked Data that are hosted on different sites across the Web. JSON-LD is a lightweight syntax to serialize Linked Data in JSON. Its design allows existing JSON to be interpreted as Linked Data with minimal changes. JSON-LD is primarily intended to be a way to use Linked Data in Web-based programming environments, to build interoperable Web services, and to store Linked Data in JSON-based storage engines. Since JSON-LD is 100% compatible with JSON, the large number of JSON parsers and libraries available today can be reused. In addition to all the features JSON provides, JSON-LD introduces:

- a universal identifier mechanism for JSON objects via the use of IRIs,
- a way to disambiguate keys shared among different JSON documents by mapping them to IRIs via a context,
- a mechanism in which a value in a JSON object may refer to a JSON object on a different site on the Web,
- the ability to annotate strings with their language,
- a way to associate datatypes with values such as dates and times,
- and a facility to express one or more directed graphs, such as a social network, in a single document.

URL: <https://www.w3.org/2018/jsonld-cg-reports/json-ld/#introduction>

[W3C RDF 1.1]

### Resource Description Framework 1.1

The Resource Description Framework (RDF) is a framework for representing information in the Web. This document defines an

abstract syntax (a data model) which serves to link all RDF-based languages and specifications. The abstract syntax has two key data structures: RDF graphs are sets of subject-predicate-object triples, where the elements may be IRIs, blank nodes, or datatyped literals. They are used to express descriptions of resources. RDF datasets are used to organize collections of RDF graphs, and comprise a default graph and zero or more named graphs. RDF 1.1 Concepts and Abstract Syntax also introduces key concepts and terminology, and discusses datatyping and the handling of fragment identifiers in IRIs within RDF graphs.

URL: <https://www.w3.org/TR/rdf11-concepts/>

[W3C ODRL V&E 2.2]

The Open Digital Rights Language Vocabulary & Expression 2.2

The Open Digital Rights Language (ODRL) is a policy expression language that provides a flexible and interoperable information model, vocabulary, and encoding mechanisms for representing statements about the usage of content and services. The ODRL Vocabulary and Expression describes the terms used in ODRL policies and how to encode them.

URL: <https://www.w3.org/TR/2018/REC-odrl-vocab-20180215/>

## I.5 OASIS references and associated summaries

[OASIS AMQP 1.0]

Advanced Message Queuing Protocol Version 1.0

The AMQP is an open Internet protocol for business messaging. It defines a binary wire-level protocol that allows for the reliable exchange of business messages between two parties. AMQP has a layered architecture and the specification is organized as a set of parts that reflects that architecture.

URL: <http://docs.oasis-open.org/amqp/core/v1.0/os/amqp-core-overview-v1.0-os.html>

[OASIS MQTT 3.1.1]

Message Queuing Telemetry Transport Version 3.1.1

MQTT is a client server publish/subscribe messaging transport protocol. It is light weight, open, simple, and designed so as to be easy to implement. These characteristics make it ideal for use in many situations, including constrained environments such as for communication in M2M and IoT contexts where a small code footprint is required and/or network bandwidth is at a premium.

URL: <http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html>

## I.6 Data Mining Group references and associated summaries

[DMG PMML 4.4]

Predictive Model Markup Language 4.4

PMML is XML-based file format to provide a way for applications to describe and exchange models produced by data mining and machine learning algorithms. It supports common models such as logistic regression and feed forward neural networks.

URL: <http://dmg.org/pmml/pmml-v4-4.html>

[DMG PFA]

Portable Format for Analytics

PFA is a mini-language for mathematical calculations that is usually generated programmatically, rather than by hand. A PFA document is a string of JSON-formatted text that describes an executable called a scoring engine. Each engine has a well-defined input, a well-defined output, and functions for combining inputs to construct the output in an expression-centric syntax tree.

URL: <http://dmg.org/pfa/index.html>

## **I.7 TM Forum references and associated summaries**

[TMF BDAG]

### The Big Data Analytics Guidebook

The guidebook provides guidance to a communication service provider on the major components that are needed for the implementation of real-life big data analytics use cases. It defines a reference model, use cases, business value roadmap, building blocks and the analytics big data repository for big data analytics. It also includes addendums, which are;

- Big data analytics use cases – Best practice;
- Big data analytics building blocks – Best practice;
- Big data analytics privacy risk score details – Best practice;
- Big data analytics big data repository – Best practice.

URL: <https://www.tmforum.org/resources/collection/gb979-big-data-analytics-solution-suite-r15-5-1/>

## Bibliography

- [b-ITU-T DDN] ITU-T Study Group 13 work in progress Y.bDDN-req, *Requirements of big data-driven networking*.  
<[http://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=10967](http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=10967)>
- [b-ITU-T DPI] ITU-T Study Group 13 work in progress Y.DPI-ReqFN, *Functional Requirements of Deep Packet Inspection for Future Networks*.  
<[http://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=10546](http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=10546)>
- [b-ITU-T TSAG] ITU-T Telecommunication Standardization Advisory Group, *Outcome of the workshop session on big data in telecommunications sector*.  
<<http://www.itu.int/en/ITU-T/Workshops-and-Seminars/bigdata/Documents/workshop%20outcome.pdf>>
- [b-ITU-T TWR] ITU-T Technology Watch Report, November 2013, [http://www.itu.int/dms\\_pub/itu-t/oth/23/01/T23010000220001PDFE.pdf](http://www.itu.int/dms_pub/itu-t/oth/23/01/T23010000220001PDFE.pdf) *Big data: Big today, normal tomorrow*.  
<[http://www.itu.int/ITU-T/workprog/wp\\_item.aspx?isn=10546](http://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=10546)>
- [b-Chen] Min Chen, Shiwen Mao, Ying Zhang, *et al.* (2014), *Big Data: Related Technologies, Challenges and Future Prospects*, Springer.
- [b-CRA] Challenges and Opportunities with Big Data (2012), *Computing Research Association*, November.
- [b-CSA] Cloud Security Alliance (2012), *Top Ten Big Data Security and Privacy Challenges*.
- [b-IBM] IBM Data Security and Protection, *What are data security and data privacy?*  
<<http://www-01.ibm.com/software/data/security-privacy/>>
- [b-JTC 1] ISO/IEC JTC 1 N12395 (2014), *Resolution Adopted at the 29th Meeting of ISO/IEC JTC 1 15-20 November 2014 in Abu Dhabi*.
- [b-TC 69] ISO/TC 69  
<<https://www.iso.org/committee/49742.html>>
- [b-NEC] NEC (2015), *Software Defined Networking Goes Big: What Big Data Means for SDN*.  
<<http://www.nec.com/en/global/ad/insite/article/bigdata02.html>>
- [b-OASIS] OASIS Committee Categories: *Big Data*.  
<[https://www.oasis-open.org/committees/tc\\_cat.php?cat=bigdata](https://www.oasis-open.org/committees/tc_cat.php?cat=bigdata)>
- [b-theguardian] The Guardian, *Big data and open data: what's what and why does it matter?*  
<<http://www.theguardian.com/public-leaders-network/2014/apr/15/big-data-open-data-transform-government>>
- [b-W3C-BDCG] W3C, *Big Data Community Group*.  
<<http://www.w3.org/community/bigdata/>>
- [b-W3C-OGCG] W3C, *Open Government Community Group*.  
<<http://www.w3.org/community/opengov/>>
- [b-W3C-DPVCG] W3C, *Data Privacy Vocabularies and Controls Community Group*.  
<<http://www.w3.org/community/dpvcg/>>
- [b-FG-DPM] ITU-T Focus Group on Data Processing and Management to support IoT and Smart Cities & Communities  
<<https://www.itu.int/en/ITU-T/focusgroups/dpm/Pages/default.aspx>>
-