



INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.3605

(09/2020)

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

Big Data

Big data – Reference architecture

CAUTION !

PREPUBLISHED RECOMMENDATION

This prepublication is an unedited version of a recently approved Recommendation. It will be replaced by the published version after editing. Therefore, there will be differences between this prepublication and the published version.

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

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

© ITU 2020

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Draft new Recommendation ITU-T Y.3605 (formerly Y.BD-arch)

Big data – Reference architecture

Summary

This Recommendation defines Big Data Reference Architecture (BDRA) that can serve as a fundamental reference point for big data standardization and which provides an overall framework for the basic concepts and principles of big data. The Recommendation provides a description of reference architecture concepts, two distinct viewpoints including user view and functional view, and also cross cutting aspects. Furthermore, the Recommendation addresses layering framework, functional components within framework and detailed functional descriptions for big data.

Keywords

big data, cross cutting, functional component, functional view, reference architecture, user view

CONTENTS

	Page
1 Scope.....	3
2 References.....	3
3 Definitions	3
3.1 Terms defined elsewhere	3
3.2 Terms defined in this Recommendation.....	4
4 Abbreviations and acronyms	4
5 Conventions	5
6 Reference architecture concepts	5
6.1 BDRA architectural views.....	5
6.2 Relationship between the user view and the functional view.....	7
7 User view	8
8 Functional view	9
8.1 Layering framework for big data.....	9
8.2 Functional components.....	10
9 Cross cutting aspects.....	17
9.1 Data redundancy	17
9.2 Performance.....	17
9.3 Data quality	17
10 Security considerations	18
Bibliography.....	19

Draft new Recommendation ITU-T Y.3605 (formerly Y.BD-arch):

Big data - Reference architecture

1 Scope

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.

2 References

The following ITU-T Recommendations and other references contain provisions, which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- [ITU-T Y.3502] Recommendation ITU-T Y.3502 (2014), *Cloud computing - Reference Architecture*
- [ITU-T Y.3519] Recommendation ITU-T Y.3519 (2018), *Cloud computing - Functional architecture of big data as a service*
- [ITU-T Y.3600] Recommendation ITU-T Y.3600 (2015), *Big data - cloud computing based requirements and capabilities*
- [ITU-T Y.3601] Recommendation ITU-T Y.3601 (2018), *Big data – Framework and requirements for data exchange*
- [ITU-T Y.3602] Recommendation ITU-T Y.3602 (2018), *Big data – Functional requirements for data provenance*
- [ITU-T Y.3603] Recommendation ITU-T Y.3603 (2020), *Big data - Requirements and conceptual model of metadata for data catalogue*
- [ITU-T Y.3604] Recommendation ITU-T Y.3604 (2020), *Big data - Overview and requirements for data preservation*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 activity [ITU-T Y.3502]: A specified pursuit or set of tasks.

3.1.2 big data [ITU-T Y.3600]: A category of technologies and services where the capabilities provided to collect, store, search, share, analyze and visualize data which have the characteristics of high-volume, high-velocity and high-variety.

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

3.1.3 big data provenance [ITU-T Y.3602]: Information that records the historical path of data according to the data lifecycle operations in big data ecosystem.

NOTE 1 – Data lifecycle operations include data generation, transmission, storage, use, and deletion.

NOTE 2 – The provenance information provides the details about data source, such as a person responsible for the provision of data, functions applied to data, and information about the computing environment for data processing (e.g. OS, H/W description, locale settings, and time zone etc.).

3.1.4 data catalogue [ITU-T Y.3600]: A listing of all metadata which a data broker makes available.

3.1.5 functional component [ITU-T Y.3502]: A functional building block needed to engage in an activity backed by an implementation.

3.1.6 metadata [b-ITU-T H.752]: Structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment and management of the described entities.

3.1.7 party [b-ITU-T Y.3500]: Natural person or legal person, whether or not incorporated, or a group of either.

3.1.8 personally identifiable information [b-ITU-T X.1361]: Any information that a) can be used to identify the PII principal to whom such information relates, or b) is or might be directly or indirectly linked to a PII principal.

NOTE – To determine whether a PII principal is identifiable, account should be taken of all the means which can reasonably be used by the privacy stakeholder holding the data, or by any other party, to identify that natural person.

3.1.9 provenance [b-ITU-T X.1255]: Information pertaining to any source of information including the party or parties involved in generating it, introducing it and/or vouching for it.

3.1.10 reference architecture [b-ISO/IEC 26550]: Core architecture that captures the high-level design of a software and systems product line including the architectural structure and texture (e.g. common rules and constraints) that constrains all member products within a software and systems product line.

NOTE – Application architectures of the member products included in the product line reuse (possibly with modifications) the common parts and bind variable parts of the domain architecture. Application architectures of the member products may (but do not need to) provide variability.

3.1.11 role [ITU-T Y.3502]: A set of activities that serves a common purpose.

3.1.12 sub-role [ITU-T Y.3502]: A subset of the activities of a given role.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

BDC	Big Data service Customer
BDRA	Big Data Reference Architecture
BDSP	Big Data Service Provider
ETL	Extract, Transform, Load

HTML	HyperText Markup Language
NoSQL	Not only SQL
PII	Personally Identifiable Information
SLA	Service Level Agreement
XML	Extensible Markup Language

5 Conventions

The following conventions apply:

Diagrams are used throughout this Recommendation to help illustrate the big data reference architecture (BDRA). Figure 5-1 provides the conventions used regarding the content of the diagrams.

NOTE – In Figure 5-1, "Aspect" is to be understood as referring to "Cross-cutting aspect".

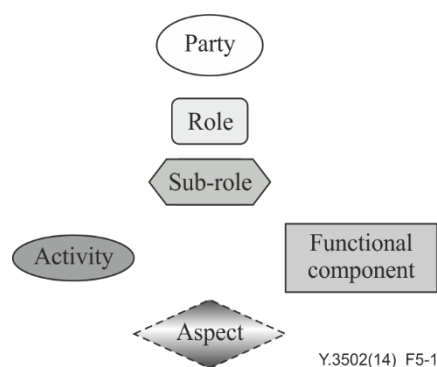


Figure 5-1 – Legend to the diagrams used throughout this Recommendation [Refer to ITU-T Y.3502]

6 Reference architecture concepts

This clause provides an overview of the architectural concepts that are used in this Recommendation.

6.1 BDRA architectural views

Big data systems can be described using a viewpoint approach.

Four distinct viewpoints are used in the BDRA (see Figure 6-1):

- user view;
- functional view;
- implementation view; and
- deployment view.