INTERNATIONAL STANDARD

ISO/IEC 22624

First edition
2020-02

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

Reference number
ISO/IEC 22624:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 38, *Cloud Computing and Distributed Platforms*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Many of the policies and practices in place for handling data in the cloud computing ecosystem need to be described based on the category of the data they address. For instance, personally identifiable information (PII) impose specific data management requirements not only in terms of security but also with regard to mechanisms that allow cloud service users to whom such data relate to exercise control on the usage and transfer of such data. Organisational data such as cloud service usage information and telemetry data from cloud services, which can be used for operational purposes such as improvement of service quality, may have to fulfil specific quality requirements to be useful for a given application.

Customer content data can be related to intellectual property rights and possibly needs appropriate protection by the cloud service provider (CSP). Certain data can be transferred from one jurisdiction to another. Depending on their data category, different instruments (multi-national laws, corporate binding rules, bilateral agreements) are applicable to enable such transfers.

When such policies and practices are to be described, it is helpful to do so in a structured and consistent way so that they can be better expressed, evaluated, analysed, and compared by the stakeholders in the cloud computing ecosystem. ISO/IEC 19944 provides a comprehensive taxonomy defining a fine-grained system of data categories that can be applied to various domains of policies for the handling of data in a cloud computing ecosystem such as cross border data transfer, data geolocation, data usage, data access and data portability, data management including data quality management and data security, or data governance, and provides guidelines on how to describe data handling policies and practices within codes of conduct (CoC).

This document describes such a structured and common approach to express any desired data handling policies and practices. It is important to emphasize that the policies and practices themselves are out of the scope of this document. This document describes a common structure and approach to express any desired data handling policies and practices. It is important to emphasize that the policies and practices are out of the scope of this document. A set of examples from data handling domains are provided in the document as guidance to understand how to use ISO/IEC 19944 regarding application of policies and analysis of policy requirements to such domains.

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

1 Scope

This document:

— describes a framework for the structured expression of data-related policies and practices in the cloud computing environment, based on the data taxonomy in ISO/IEC 19944;

— provides guidelines on application of the taxonomy for handling of data based on data subcategory and classification;

— covers expression of data-related policies and practices including, but not limited to data geolocation, cross border flow of data, data access and data portability, data use, data management, and data governance;

— describes how the framework can be used in codes of conduct for practices regarding data at rest and in transit, including cross border data transfer, as well as remote access to data;

— provides use cases for data handling challenges, i.e. control, access and location of data according to ISO/IEC 19944 data categories.

This document is applicable primarily to cloud service providers, cloud service customers (CSCs) and cloud service users, but also to any person or organization involved in legal, policy, technical or other implications of taxonomy-based data management in cloud services.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 17788, *Information technology — Cloud computing — Overview and vocabulary*

ISO/IEC 19944, *Information technology — Cloud computing — Cloud services and devices: Data flow, data categories and data use*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 17788, ISO/IEC 19944 and the following apply:

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

3.1
codes of conduct
CoC
agreed set of behaviours between organisations to enhance customer and/or partner outcomes and experiences

3.2
confidentiality
property that information is not made available or disclosed to unauthorized individuals, entities, or processes

[SOURCE: ISO 24534-5:2011, 3.11]

3.3
integrity
property of being designed such that any modification of the electronically stored information, without proper authorization, is not possible

3.4
availability
property of being accessible and useable upon demand by an authorized entity

[SOURCE: ISO 22600-1:2014, 3.7]

3.5
data access
process by which a system can read published data on another system

Note 1 to entry: This data access happens over a network connection and the data typically does not persist after the connection is terminated.

3.6
data transfer
copying or moving data from one system to another

3.7
data geolocation
geographic location of a data object at rest

4 Symbols and abbreviated terms

APEC	Asia-Pacific Economic Cooperation
BCR	Binding Corporate Rules
CBPR	Cross-border Privacy Rules
CPTPP	Comprehensive and Progressive Agreement for Trans-Pacific Partnership
CoC	Codes of Conduct
CSC	Cloud Service Customer
CSN	Cloud Service partner
CSP	Cloud Service Provider
DRM	Digital Rights Management
EU	European Union
EUII	End User Identifiable Information
GDPR	General Data Protection Regulation
GPS	Global Positioning System

HBI	High Business Impact
IaaS	Infrastructure as a Service
IPR	Intellectual Property Rights
IRM	Information Rights Management
IT	Information Technology
LBI	Low Business Impact
MBI	Medium Business Impact
NASPO	National Associations of State Procurement Officials
OII	Organization Identifiable Information
PII	Personally Identifiable Information

5 Overview: The need for a structured expression of data policies and practices based on a common data taxonomy

Data policies and practices, at corporate or government level, need to be crisply expressed with the desired degree of precision and clarity. The need for varying degree of precision, along with the need to compare and analyse various policies in an efficient manner, calls for a common and structured approach to the expression of these policies and practices, an approach that is based on a common data taxonomy.

ISO/IEC 19944 provides a comprehensive set of elements which can be used to

a) assign a data category to a given data set (e.g. personally identifiable information (PII), organisational identifiable information, customer content data),

b) provide classes of actions applied to such data (e.g. use to provide a service, to optimize it, to provide marketing information),

c) include scopes explaining on what level the use of data happens (e.g. service level vs. enterprise/ organisational level vs. use by 3rd parties), and

d) define the level of de-identification (or anonymization) applied to a data set (qualifiers such as "identified", "anonymized", "aggregated").

These elements are referred to in the document as “data categories” or “data taxonomy”, “actions”, “scopes”, and “qualifiers” without explicitly referencing ISO/IEC 19944. Clause 6 provides a comprehensive overview of the elements. The framework described in this document references the framework in ISO/IEC 19944.

In order to define application specific data handling policies and practices, these elements need to be applied to the application domain at hand. This includes data classifications with regards to security or risk levels that apply to data, as well as technical and organisational qualifications of data. Hence, the approach described in this document requires the considerations of data categories as described in ISO/IEC 19944 as well as orthogonal information dependent on the concrete application under consideration. Examples which are used to explain this approach therefore employ a tabular representation format emphasizing the orthogonal character of generic data categorization (rows) and application specific elements (columns). Therefore, for a person who is concerned with the development of, for example, enterprise policies for data use by a set of cloud services, all relevant cases which need to be considered are visible.

Implicitly, ISO/IEC 19944 focuses on personal data and PII, and does not explicitly cover non-personal data, or mixed sets of data that contain both PII and non-personal data. Non-personal data is defined as any data that is not personal and is not covered under PII, e.g. scientific data, sales data. Mixed data sets contain both PII and non-personal data such as human resource data that contains both organizational structures and personal employee data. It is important to recognize these different sets as different policies and regulations could apply to each. For example, the EU GDPR[9] regulates aspects of PII and the free-flow of non-personal data regulation[10] sets policies concerning the geo-location and movement of non-personal data. In line with ISO/IEC 19944, this document focuses on PII and does not delve deeper into aspects explicitly related to non-personal or mixed sets of data.

The document is structured as follows:

— Clause 6 describes the framework for the structured expression of data related policies and practices including elements of the framework building on ISO/IEC 19944. It then expands discussion on data classification (6.2.6).

— Clause 7 discusses guidance for using the framework defined in Clause 6.

— Clause 8 covers use of framework in specific areas of concern.

— Clause 9 describes the application of the framework to codes of conduct.

Examples for data handling challenges are provided throughout the document.

6 Framework for the structured expression of data related policies and practices

6.1 General

This document uses the taxonomy and data use expression structure specified in ISO/IEC 19944. Any policy or practice that conforms to this document and uses the taxonomy or data use expression shall meet the requirements of ISO/IEC 19944 as appropriate.

To handle key data management topics, Clause 6 describes a harmonized structure to express a desired policy for data management based on various data types, using data taxonomy in ISO/IEC 19944. The data management policies based on a common structure specified by this document can be expressed, compared and negotiated.

It is important to point out that this document does not define one or more data policies, rather it offers a common structure and framework for others to use in order to express their policy of choice.

Moreover, this document does not stipulate any specific format or syntax to be used to express policies and practices related to a categorization of data. Although tables are frequently employed throughout this document to illustrate the usage of the framework, the use of tabular formats is not normative or mandatory but serves for the presentation of examples only.

6.2 Framework elements

6.2.1 General

ISO/IEC 19944 defines a number of elements to express statements that describe the use of data by a CSP, namely a data categorization hierarchy, a set of qualifiers indicating the level of de-identification of data, and a hierarchy of scopes that describe at which level data are collected and processed by the CSP, a number of actions used to process data, and on which level the result of data processing is used. This clause provides an overview of the elements that are described in detail in ISO/IEC 19944.

6.2.2 Data categories

6.2.2.1 General

The data taxonomy described in ISO/IEC 19944:2017,A.1 as shown in Figure 1 below defines four main data categories, namely customer content data, derived data, CSP data, and account data

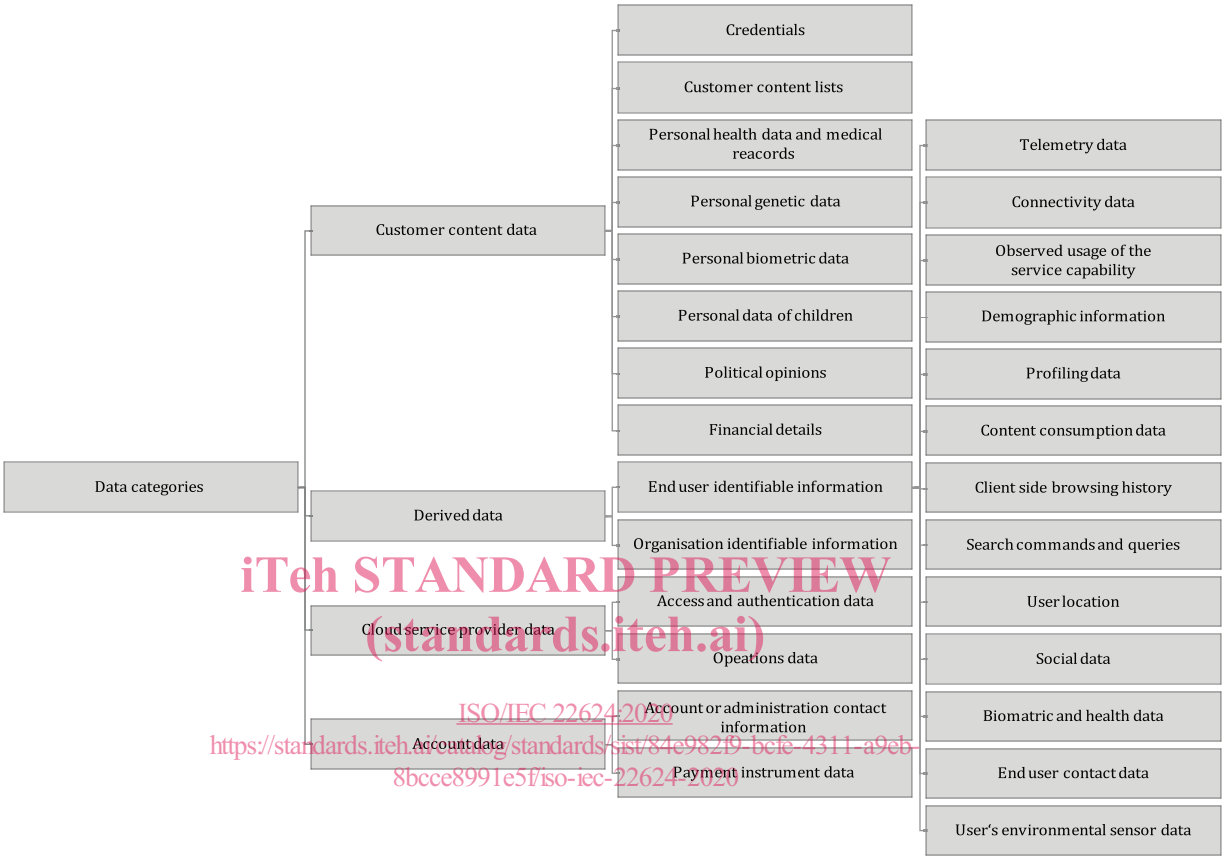

Figure 1 — Data categorization hierarchy according to ISO/IEC 19944:2017, A.1

6.2.2.2 Customer content data

Customer content data is cloud service customer (CSC) data extended to include similar data objects provided to applications executing locally on the device. This includes content directly created by customers and their users and all data that customers provide to the cloud service, or are provided to the cloud service on behalf of customers, through the capabilities of the service or application. This also includes data that the user intentionally creates through the use of the app or cloud service. This data category contains a large variety of sub-categories. The reader is referred to ISO/IEC 19944:2017, 8.2.2 for details.

6.2.2.3 Derived data

6.2.2.3.1 General

Derived data is cloud service derived data extended to include similar data objects derived as a user exercises the capabilities of an application executing locally on the device.

6.2.2.3.2 End user identifiable information

End user identifiable information (EUII) is defined as data associated with a user that are captured or generated from the use of the service by that user; EUII is linkable to that user but is not customer content data. This data category contains a large variety of sub-categories. The reader is referred to ISO/IEC 19944:2017, 8.2.3.2 for details.

6.2.2.3.3 Organization identifiable information

Organization identifiable information (OII) is the data that can be used to identify a particular tenant (general configuration or usage data), is not linkable to a user and does not contain customer content data. This also includes data aggregated from the users of a tenant that is not linkable to the individual user.

6.2.2.4 CSP data

6.2.2.4.1 General

This category includes data that is exclusively under the control of the CSP. It is unique to the system and under the control of the CSP.

6.2.2.4.2 Access and authentication data

Access and authentication data is the data used within the cloud service to manage access to other categories of data or capabilities within the service.

6.2.2.4.3 Operations data

Operations data is data which is used for supporting the operation of CSPs and system maintenance, such as service logs, technical information about a subscription (e.g. service topology), technical information about a tenant (e.g. customer role name), configuration settings/files.

6.2.2.5 Account data

Account data is a class of data specific to each CSC that is required to sign up for, purchase or administer the cloud service. This data includes information such as names, addresses, payment information. Account data is generally under the control of the CSP although each CSC usually has the capability to input, read and edit their own account data but not the records of other CSCs.

6.2.3 Data identification qualifiers

Data in any category can provide or contribute to information that identifies or can be linked to an individual. The extent to which individuals are directly identified in the data, and how easy it is to associate a set of characteristics in the data to an individual is described by the following set of qualifiers (see Figure 2):

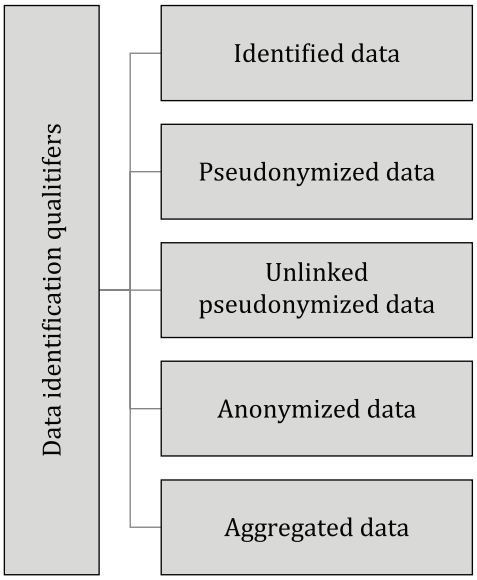

Figure 2 — Data identification qualifiers according to ISO/IEC 19944:2017, A.2

— **Identified data**. Data that can unambiguously be associated with a specific person because PII is observable in the information.

— **Pseudonymized data**. Data for which all identifiers are substituted by aliases for which the alias assignment is such that it cannot be reversed by reasonable efforts of anyone other than the party that performed them.

— **Unlinked pseudonymized data**. Data for which all identifiers are erased or substituted by aliases for which the assignment function is erased or irreversible, such that the linkage cannot be re-established by reasonable efforts of anyone including the party that performed them.

— **Anonymized data**. Data that is unlinked and which attributes are altered in such a way that there is a reasonable level of confidence that a person cannot be identified, directly or indirectly, by the data alone or in combination with other data.

— **Aggregated data**. Statistical data that does not contain individual-level entries and is combined from information about enough different persons that individual-level attributes are not identifiable.

6.2.4 Data usage scopes

ISO/IEC 19944:2017, 9.4.1 defines that "scope" *provides a way to clearly describe the boundaries of collection and use of data in the devices and cloud services ecosystem*. These scopes can be used to describe the applications and services associated with data use (see Figure 3). The definitions are listed in increasing breadth of scope and the wider scopes include the narrower scopes, except for "third party" items which exist in an independent scope. Capabilities are parts of an application or a cloud service which could be one of the services listed in the service agreement. These are parts of the cloud services that a CSP provides, and are a subset of the CSPs overall product and service palette.