INTERNATIONAL STANDARD

ISO/IEC 19944-2

First edition
2022-04

Cloud computing and distributed platforms — Data flow, data categories and data use —

Part 2: Guidance on application and extensibility

Reference number
ISO/IEC 19944-2:2022(E)

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

Published in Switzerland

Contents

Page

Foreword ..... **v**
Introduction ..... **vi**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 Abbreviated terms **2**
5 Presentation of ISO/IEC 19944-1 ..... **2**
6 How to apply ISO/IEC 19944-1 ..... **3**
6.1 General 3
6.2 Generic eCommerce example 3
6.2.1 General 3
6.2.2 Customer content data 3
6.2.3 Derived data 3
6.2.4 Data identification qualifiers 4
6.2.5 Orthogonal facets 4
6.2.6 Data processing categories 5
6.2.7 Data use categories 6
6.2.8 Scopes 7
6.2.9 Data use statements 7
6.3 Privacy examples 8
6.3.1 General 8
6.3.2 Describing the purpose of the processing of PII 8
6.3.3 Using data identification qualifiers with PII 9
6.4 Organization identifiable data examples 9
6.4.1 General 9
6.4.2 Organization identifiable data location requirement examples 10
6.4.3 Organization identifiable data sharing requirement examples 10
6.5 AI example 11
6.5.1 General 11
6.5.2 Facial recognition — Privacy-centric AI example 12
6.6 IoT example 14
6.6.1 General 14
6.6.2 Electrical vehicles 14
7 How to extend ISO/IEC 19944-1 ..... **15**
7.1 General 15
7.2 Data taxonomy 15
7.2.1 General 15
7.2.2 Guidelines for extending the data categories defined in ISO/IEC 19944-1 15
7.2.3 Example of extending the cloud service provider (CSP) and customer content data categories 16
7.2.4 Example of extending the demographic information sub-type 16
7.2.5 Example of extending the financial details sub-type 17
7.3 Custom data facets 17
7.3.1 General 17
7.3.2 Guidance on creating custom data facets 17
7.3.3 Example custom data facet 18
7.4 Data processing 18
7.4.1 General 18
7.4.2 Guidelines for extending data processing categories 18
7.4.3 Examples for extending data processing categories 18
7.5 Data use categories 19
7.5.1 General 19

7.5.2 Guidelines for extending the data use categories 19
7.5.3 Example for AI 19
7.5.4 Facial recognition — Privacy-centric AI example for extending the taxonomy 20
7.5.5 Automotive application — Intellectual property-centric AI/IoT example for extending the taxonomy 21

Bibliography **23**

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards
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 or www.iec.ch/members_experts/refdocs).

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 https://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. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Technical Committee, ISO/IEC JTC 1, *Information technology*, Subcommittee SC 38, *Cloud computing and distributed platforms.*

A list of all parts in the ISO/IEC 19944 series can be found on the ISO and IEC websites.

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 and www.iec.ch/national-committees.

Introduction

ISO/IEC 19944-1 provides a data taxonomy, data processing and use categories and other descriptive facets that can be applied to data. All aspects of ISO/IEC 19944-1 are extensible to meet the needs of diverse users. The standardized ability to categorize data, describe uses of data and apply other facets is useful in several scenarios including the application of policy to data and in describing the use of data to stakeholders.

The aim of this document is to assist users of ISO/IEC 19944-1 by providing examples and guidance for its use across several domains. Additionally, this document provides users who need to extend ISO/IEC 19944-1 with examples and guidance.

Cloud computing and distributed platforms — Data flow, data categories and data use —

Part 2: Guidance on application and extensibility

1 Scope

This document provides guidance on the application of the taxonomy and use statements from ISO/IEC 19944-1 in real world scenarios, and how to develop extensions to the data taxonomy, data processing and use categories and data use statements.

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 19944-1:2020, *Cloud computing and distributed platforms — Data flow, data categories and data use — Part 1: Fundamentals*

ISO/IEC 22123-1, *Information technology — Cloud computing — Part 1: Vocabulary*

ISO/IEC 22989,[1)] *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 19944-1, ISO/IEC 22123-1, ISO/IEC 22989 and the following apply.

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

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

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

3.1
Internet of Things
IoT
infrastructure of interconnected entities, people, systems and information resources together with services which processes and reacts to information from the physical world and virtual world

[SOURCE: ISO/IEC 20924:2021, 3.2.1]

3.2
PII principal
natural person to whom the personally identifiable information (PII) relates

Note 1 to entry: Depending on the jurisdiction and the particular data protection and privacy legislation, the synonym “data subject” can also be used instead of the term “PII principal”.

1) Under preparation. Stage at the time of publication: ISO/IEC FDIS 22989.

[SOURCE: ISO/IEC 29100:2011, 2.11]

4 Abbreviated terms

AI	Artificial Intelligence
CSA	Cloud Service Agreement
CSC	Cloud Service Customer
CSP	Cloud Service Provider
DNN	Deep Neural Network
EV	Electric Vehicle
IaaS	Infrastructure as a Service
IoT	Internet of Things
IPR	Intellectual Property Rights
OPD	Organizational Protected Data
PaaS	Platform as a Service
PII	Personally Identifiable Information
SKU	Stock Keeping Unit

5 Presentation of ISO/IEC 19944-1

To improve transparency and guidance about data flows and data use, ISO/IEC 19944-1 names and describes the flows of data between a device and a supporting cloud service, and how to describe the use of different categories of data by the CSP.

ISO/IEC 19944-1 provides a comprehensive set of elements which can be used to:

— assign a data category to a given dataset, e.g. personally identifiable information, organizational protected data, customer content data;

— provide a set of actions applied to data, e.g. use to provide a service, to optimize it, to provide marketing information;

— define the qualifiers for the level of de-identification applied to a dataset, e.g. "identified", "anonymized", "aggregated";

— describe a use of a certain category of data for a specified purpose with a scope of its usage.

To maintain trust between the stakeholders in the ecosystem of cloud services and to meet the requirements of laws and regulations, it is necessary for service providers to be transparent about their use of the various data types that flow within the ecosystem. ISO/IEC 19944-1 also deals with organizational data and the need to treat some organizational data in particular ways in order to ensure properties such as confidentiality and integrity are maintained.

ISO/IEC 19944-1 introduces the concept of data facets, and data facets are used to extend the expressiveness of data use statements, including adding the concept of which individuals or organizations have control over data.

6 How to apply ISO/IEC 19944-1

6.1 General

This clause provides examples of ISO/IEC 19944-1 applied to several different scenarios. Readers of this document can use the examples in this clause to then apply the data categories, data use statements and other facets to their situation.

6.2 Generic eCommerce example

6.2.1 General

Modern eCommerce applications use and generate many of the data categories specified in ISO/IEC 19944-1. Additionally, the data in eCommerce applications involves other components of ISO/IEC 19944-1 including data identification qualifiers, orthogonal facets, data processing and use categories, scopes, and data use statements. eCommerce applications ordinarily involve organizational protected data such as sales volumes, pricing, customer lists and inventory data.

This clause provides descriptions and examples of how the components of ISO/IEC 19944-1 can be used in a generic retail eCommerce application.

NOTE 1 Per the definition of CSP in ISO/IEC 22123-1:2021, the operator of an eCommerce application is a CSP to its CSCs, even if the application is run on third-party services of infrastructure or platform capabilities types.

NOTE 2 Unless otherwise indicated, all references to CSC and CSP in this document refer to the party, not to the corresponding roles.

6.2.2 Customer content data

credentials: CSCs can provide data objects such as username, password, certificates and biometrics, to access eCommerce services.

financial details: CSCs can provide financial details such as credit card numbers, bank information or other payment information, to pay for purchased products and services.

6.2.3 Derived data

6.2.3.1 End user identifiable information

telemetry data: The CSP can collect data about the use of purchased products and services.

connectivity data: The CSP can collect data about the CSC's connectivity to provide the eCommerce service. For example, the CSP can use IP address information to determine the CSC's location for shipping or pickup information.

observed usage of the service capability: The CSP can collect data about the CSCs preferences and settings for the eCommerce application along with commands issued to the application (e.g. clicking the Submit button).

demographic information: The CSP can collect demographic data (e.g. age, gender) about the CSC.

profiling data: The CSP can use the various categories and instances of data to build a user profile that helps the CSP understand the CSCs interests and preferences.

content consumption data: In cases where the eCommerce application offers content, the CSP can collect data on the CSC's content consumption.

client-side browsing history: The CSP can collect client-side browsing history to help build a profile of the CSC's interests and preferences.

search commands and queries: The CSP can collect search commands and queries to improve the eCommerce application and to understand the CSC's interests and preferences.

user location: The CSP can collect the CSC's location in order to tailor offers made to the CSC.

social data: The CSP can collect the CSC's social data to better understand the CSC's interests and preferences.

6.2.3.2 Cloud service provider (CSP) data

access and authentication data: The CSP can use access and authentication data to allow CSCs to access particular aspects of the eCommerce application.

operations data: The CSP can collect data regarding operation of the eCommerce application including information about individual CSCs' use of the eCommerce application.

6.2.3.3 Account data

account or administration contact information: The CSP will ordinarily collect account data for CSCs of the eCommerce application in order to perform sales transactions.

payment instrument data: The CSP will ordinarily collect payment information from CSCs to perform sales transactions.

6.2.3.4 Organizational protected data:

price lists and pricing algorithms: The CSP can designate pricing information as organizational protected data.

sales data: The CSP can designate sales data as organizational protected data.

customer lists: The CSP can designate its customer lists as organizational protected data.

inventories: The CSP can designate its inventory data as organizational protected data.

6.2.4 Data identification qualifiers

identified data: CSPs operating eCommerce applications will ordinarily collect identified data such as account data, transaction data and profiles linked to account data that contains PII.

pseudonymized data: The CSP can substitute aliases for PII to protect individual privacy when processing of the data does not require PII but there is a potential need to link data back to specific CSCs.

unlinked pseudonymized data: The CSP can delete PII or use aliases when processing data where there is no desire to link the data back to specific CSCs.

anonymized data: The CSP can un-link and alter attributes of the data in a way that does not alter the meaning of the data but makes it reasonably impossible to identify individual CSCs directly or indirectly even if the data is combined with other data.

aggregated data: In some cases, CSPs operating eCommerce applications can aggregate transactional data for reporting and analysis where the PII in each transaction is not relevant to the task.

6.2.5 Orthogonal facets

classification: The data used or generated by an eCommerce application can be classified as High Business Impact, Medium Business Impact or Low Business Impact or using an alternate scheme that describes the significance of the data.

categorization: Data can be categorized according to what the data describes. Examples of eCommerce data categories can include CSC account information, SKU information, price information, sales transactions and inventory information.

operational control: This facet includes the basic actions that can be taken on the data. These actions ordinarily include create, read, update, delete, copy and move.

legal entity: Legal entities that can control eCommerce data include individuals, organizations and public institutions.

legal means: eCommerce data can be protected by several legal means according to local laws and regulations.

6.2.6 Data processing categories

horizontal partitioning or sharding: eCommerce data can be horizontally partitioned based on defined attributes. For example, a dataset of sales transactions can be partitioned by SKU or CSC.

vertical partitioning: eCommerce data can also be vertically partitioned by keeping only a subset of attributes. For example, an inventory dataset can be vertically partitioned by removing the item weight attribute.

data association: eCommerce data can be stored in different datasets which are then linked together. For example, a dataset of transactions can link to other datasets that contain CSC information and to datasets that have detailed information about the products or services purchased.

data aggregation/consolidation: Data analysis can require that different eCommerce datasets be aggregated or consolidated. For example, sales transactions from different geographies can be stored in different datasets which are then aggregated to see global sales information.

data accumulation: Datasets containing eCommerce data ordinarily accumulate new records over time. For example, a dataset containing eCommerce transactions can grow by millions of records each day. These datasets can then be used for time-series analysis to identify trends.

data fusion: eCommerce data from multiple datasets can be combined and then reduced to obtain an improved dataset.

data standardization: Entries in eCommerce datasets can include data recorded in the wrong field. For example, the first and last name of a CSC recorded in a field labelled “First Name” or a postal code entered in a filed labelled “Street Address”. Data standardization means placing the data entries into the correct fields across the entire dataset.

data validation and correction: Data in an eCommerce dataset can be the wrong type or format or simply be incorrect. For example, if an entry of “Yes” is in a field labelled “Age”, the entry is probably incorrect. If an entry of “2020” is in an “Age” field, it can be the correct type and format but is clearly wrong. The process of data validation and correction is the process of correcting the type, format and values of data entries.

data enrichment: Entries in eCommerce datasets can be missing which can interfere with transaction calculations or analysis algorithms. For example, if the price field is blank for a purchase, the amount charged for the sale will be incorrect as will any downstream analysis of the dataset. Data enrichment means filling in missing data entries using some imputation process.

encryption: eCommerce data can be encrypted in flight and at rest to prevent unauthorized access and use.

replication: Copies of eCommerce data can be made and stored in different locations for the purpose of business continuity or recovery in the event of a disaster.

data deletion: CSPs operating eCommerce applications can delete data from time to time. Data deletion is often done when the data is no longer relevant to the CSP, CSC or other interested parties.

secure data deletion: In some cases, the CSP can delete eCommerce data in a way that prevents any party from ever recovering it.

re-identification: An eCommerce CSP can have the need to re-identify data that had previously been de-identified.

6.2.7 Data use categories

provide: eCommerce CSPs will ordinarily use many data categories from multiple scopes to provide the eCommerce service to its CSCs. For example, account data can be used to populate the billing and shipping address fields of a purchase transaction. The postal code portion of account data can be used to calculate shipping costs and time.

improve: Data can be used to improve the eCommerce service. For example, the postal code portion of account data can be used to identify clusters of CSCs which can then be used to site warehouses and reduce shipping times. Data from failed transactions can be used to improve the purchase process and in turn improve CSC satisfaction.

personalize: Data categories such as observed usage of the service capability, demographic information, profiling data, content consumption data and search commands and queries can be used to personalize the eCommerce service for a CSC or group of CSCs. For example, if a CSC or group of CSCs regularly use a search function to find products in a particular category, the service's home page can feature products that correlate to the CSC's search history.

offer upgrades or upsell: eCommerce CSPs can use transaction histories and profiling data to offer upgrades or upsells during the purchase transaction process or through other means. For example, if a CSC selects a quarter inch drill bit, the CSP can offer to upsell a drill bit set.

market/advertise/promote: Similar to "offer upgrades or upsell", transaction data and profile data can be used to promote specified products to a CSC or groups of CSCs.

promote based on contextual information: eCommerce CSPs can use data based on the use of the current capability or on the services and application scope to promote products to the CSC. This data use category does not make use of the CSC's previous use of the service.

promote based on personalization: The eCommerce CSP can use data to change the content of a promotion for a CSC or group of CSCs.

share: eCommerce CSPs can share data with third parties to operate the service. For example, the CSP can share the postal code portion of account data with a third-party shipper to get a shipping cost and delivery time estimate. The CSP can share data with third parties when outsourcing functions such as analytics.

share when required to provide the service: The CSP can be required by law, or by contractual obligations to share eCommerce data.

collect: eCommerce CSPs can collect data from many different categories such as financial details, observed usage of the service, demographic information, profiling data, content consumption data and others. Additionally, the CSP will ordinarily store, prepare and pre-process the collected data for down-stream uses including AI.

train: eCommerce CSPs can use training data from one or more categories to train machine learning models in AI systems. For example, the CSP can use data from the observed usage of the service capability to train a product recommendation model that is presented to CSCs during their use of the eCommerce system.