

Recommendation

ITU-T X.1771 (04/2024)

SERIES X: Data networks, open system communications
and security

Data security – Data protection (II)

Security guidelines for combining de-identified data using trusted third party

ITU-T X-SERIES RECOMMENDATIONS

Data networks, open system communications and security

PUBLIC DATA NETWORKS	X.1-X.199
OPEN SYSTEMS INTERCONNECTION	X.200-X.299
INTERWORKING BETWEEN NETWORKS	X.300-X.399
MESSAGE HANDLING SYSTEMS	X.400-X.499
DIRECTORY	X.500-X.599
OSI NETWORKING AND SYSTEM ASPECTS	X.600-X.699
OSI MANAGEMENT	X.700-X.799
SECURITY	X.800-X.849
OSI APPLICATIONS	X.850-X.899
OPEN DISTRIBUTED PROCESSING	X.900-X.999
INFORMATION AND NETWORK SECURITY	X.1000-X.1099
SECURE APPLICATIONS AND SERVICES (I)	X.1100-X.1199
CYBERSPACE SECURITY	X.1200-X.1299
SECURE APPLICATIONS AND SERVICES (II)	X.1300-X.1499
CYBERSECURITY INFORMATION EXCHANGE	X.1500-X.1599
CLOUD COMPUTING SECURITY	X.1600-X.1699
QUANTUM COMMUNICATION	X.1700-X.1729
DATA SECURITY	X.1750-X.1799
Big Data Security	X.1750-X.1759
Data protection (II)	X.1770-X.1799
INTERNATIONAL MOBILE TELECOMMUNICATIONS (IMT) SECURITY	X.1800-X.1839
METAVEVERSE AND DIGITAL TWIN SECURITY	X.2000-X.2199
SOFTWARE SUPPLY CHAIN SECURITY	X.2150-X.2199
ARTIFICIAL INTELLIGENCE (AI) / MACHINE LEARNING (ML) SECURITY	X.2200-X.2249

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T X.1771

Security guidelines for combining de-identified data using trusted third party

Summary

De-identification is a process to remove the association between a set of identifiable attributes and the data principal. Appropriate de-identification techniques should be used by an organization to balance between its desires to protect data as the data controller and to use the data in new and innovative ways.

Recommendation ITU-T X.1771 provides security guidelines for information and communication technology (ICT) service providers acting as data controllers and for trusted third parties fulfilling the role of combining de-identified datasets. The purpose of these security guidelines is to reduce risks, protect data and promote the utility of the dataset.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T X.1771	2024-04-29	17	11.1002/1000/15886

Keywords

Combining de-identified datasets, combining key, controls, de-identification, de-identified dataset, model for combining de-identified datasets, procedures for combining de-identified datasets.

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, 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 appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2024

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

Table of Contents

		Page
1	Scope	1
2	References.....	1
3	Definitions	1
	3.1 Terms defined elsewhere	1
	3.2 Terms defined in this Recommendation.....	2
4	Abbreviations and acronyms	3
5	Conventions	3
6	Overview	3
	6.1 Combining de-identified datasets	4
	6.2 Generation of a combining key	5
7	Use cases for combining de-identified datasets from organizations	5
8	Models for combining de-identified datasets	6
	8.1 Model with one trusted third party: combining organization.....	6
	8.2 Model with two trusted third parties: key management authority and combining organization	7
9	Procedures to combine de-identified datasets.....	8
	9.1 Procedures for the model with one trusted third party	9
	9.2 Procedures for the model with two trusted third parties.....	12
10	Controls for entities	15
	10.1 Common technical controls applicable to all entities.....	15
	10.2 Controls for trusted third party	16
	Appendix I – Examples of typical de-identification techniques	17
	Bibliography.....	18

Recommendation ITU-T X.1771

Security guidelines for combining de-identified data using trusted third party

1 Scope

De-identification can serve as a solution for an organization to balance between its desires to protect data and to harness the data for innovative purposes. The de-identification techniques described in [ITU-T X.1148] can be used to produce a de-identified dataset from a dataset having an identifying attribute. In addition, there is a need for combined de-identified datasets from different organizations, which can be used for example for statistical calculation, scientific analysis, and other purposes.

This Recommendation provides security guidelines for securely combining de-identified datasets from different organizations. It describes use cases and two models for combining de-identified datasets, identifies related entities, and specifies procedures to combine de-identified datasets submitted from different organizations acting as data controllers or data processors. In addition, this Recommendation proposes organizational and technical controls for trusted third parties which are responsible for combining de-identified datasets for information and communication technology (ICT) service providers.

This Recommendation does not address issues related to regulation.

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 X.1148] Recommendation ITU-T X.1148 (2020), *Framework of de-identification process for telecommunication service providers*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 aggregated data [b-ISO/IEC 20889]: Data representing a group of data principals, such as a collection of statistical properties of that group.

3.1.2 attribute [b-ISO/IEC 20889]: Inherent characteristic.

3.1.3 control [b-ISO/IEC 27000]: A measure that is modifying risk.

NOTE 1 – Controls include any process, policy, device, practice, or other actions which modify risk.

NOTE 2 – It is possible that controls do not always exert the intended or assumed modifying effect.

3.1.4 data controller [ITU-T X.1148]: Stakeholder (or privacy stakeholder) that determines the purposes and means for processing data other than natural persons who use data for personal purposes.

3.1.5 data processor [ITU-T X.1148]: Stakeholder that processes data on behalf of and in accordance with the instructions of a data controller.



3.1.6 data principal [b-ISO/IEC 20889]: Entity to which data relates.

NOTE – The term "data principal" is broader than "PII principal" (or "data subject" as used elsewhere), and is able to denote any entity such as a person, an organization, a device, or a software application.

3.1.7 dataset [b-ISO/IEC 20889]: Collection of data.

3.1.8 de-identification process [b-ISO/IEC 20889]: Process of removing the association between a set of identifying attributes and the data principal.

3.1.9 de-identification technique [b-ISO/IEC 20889]: Method for transforming a dataset with the objective of reducing the extent to which information is able to be associated with individual data principals.

3.1.10 de-identified dataset [b-ISO/IEC 20889]: Dataset resulting from the application of a de-identification process.

3.1.11 direct identifier [b-ISO/IEC 20889]: Attribute that alone enables unique identification of a data principal within a specific operational context.

NOTE – Here and throughout, the operational context includes the information that the entity processing (e.g., de-identifying) the data possesses, together with information that third parties and potential attackers can possess or that is in the public domain.

3.1.12 identifier [b-ISO/IEC 20889]: Set of attributes in a dataset that enables unique identification of a data principal within a specific operational context.

NOTE – See Annex B of [b-ISO/IEC 20889] for a discussion of how this definition relates to those given in other standards.

3.1.13 indirect identifier [b-ISO/IEC 20889]: Attribute that, together with other attributes that can be in the dataset or external to it, enables unique identification of a data principal within a specific operational context.

3.1.14 personally identifiable information [b-ISO/IEC 29100]: Any information that (a) can be used to identify the personally identifiable information (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.

3.1.15 pseudonymization [b-ISO/IEC 20889]: De-identification technique that replaces an identifier (or identifiers) for a data principal with a pseudonym in order to hide the identity of that data principal.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 combined de-identified dataset: Dataset resulting from the implementation of a combining process with the input of de-identified datasets submitted from different organizations.

3.2.2 combining organization: A trusted third party responsible for combining de-identified datasets submitted from different organizations.

3.2.3 combining process: Process of combining de-identified datasets from organizations to produce extensive data in the characteristic of volume that requires a scalable technology for efficient storage, manipulation, and analysis, and to remove the association between a set of identifying attributes and the data principal.

3.2.4 combining key: Secret information shared by relevant organizations for combining de-identified datasets.