INTERNATIONAL STANDARD

ISO/IEC 27559

First edition
2022-11

Information security, cybersecurity and privacy protection – Privacy enhancing data de-identification framework

Sécurité de l'information, cybersécurité et protection de la vie privée — Cadre pour la dé-identification de données pour la protection de la vie privée

Reference number
ISO/IEC 27559:2022(E)

COPYRIGHT PROTECTED DOCUMENT

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 Symbols and abbreviated terms **3**
5 Overview **3**
6 Context assessment **4**
6.1 General 4
6.2 Threat modelling 4
6.2.1 General 4
6.2.2 Security and privacy practices 5
6.2.3 Motives and capacity to re-identify 5
6.3 Transparency and impact assessment 6
6.3.1 General 6
6.3.2 Transparency of actions and stakeholder engagement 6
6.3.3 Privacy-related harms 6
7 Data assessment **7**
7.1 General 7
7.2 Data features 7
7.2.1 General 7
7.2.2 Data principals 7
7.2.3 Data type 7
7.2.4 Attribute types 8
7.2.5 Dataset properties 8
7.3 Attack modelling 8
7.3.1 General 8
7.3.2 Maximum or average risk 9
7.3.3 Population or sample-based attack 9
7.3.4 Data privacy models 9
8 Identifiability assessment and mitigation **10**
8.1 General 10
8.2 Assessing identifiability 10
8.2.1 General 10
8.2.2 Quantifying identifiability 10
8.2.3 Adversarial testing 11
8.3 Mitigation 12
8.3.1 General 12
8.3.2 Reconfiguring the environment 12
8.3.3 Transforming the data 12
8.3.4 Re-evaluation 13
9 De-identification governance **13**
9.1 General 13
9.2 Before data are made available 13
9.2.1 General 13
9.2.2 Assigning roles and responsibilities 13
9.2.3 Establishing principles, policies and procedures 14
9.2.4 Identifying and managing a data disclosure 14
9.2.5 Communicating with stakeholders 15
9.3 After data are made available 15
9.3.1 General 15

9.3.2 Monitoring the data environment 15
9.4 Mitigation in case of incident 15
Annex A (informative) **Example identifiers** **17**
Annex B (informative) **Example threshold identifiability benchmarks** **19**
Bibliography **21**

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 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 27, *Information security, cybersecurity and privacy protection.*

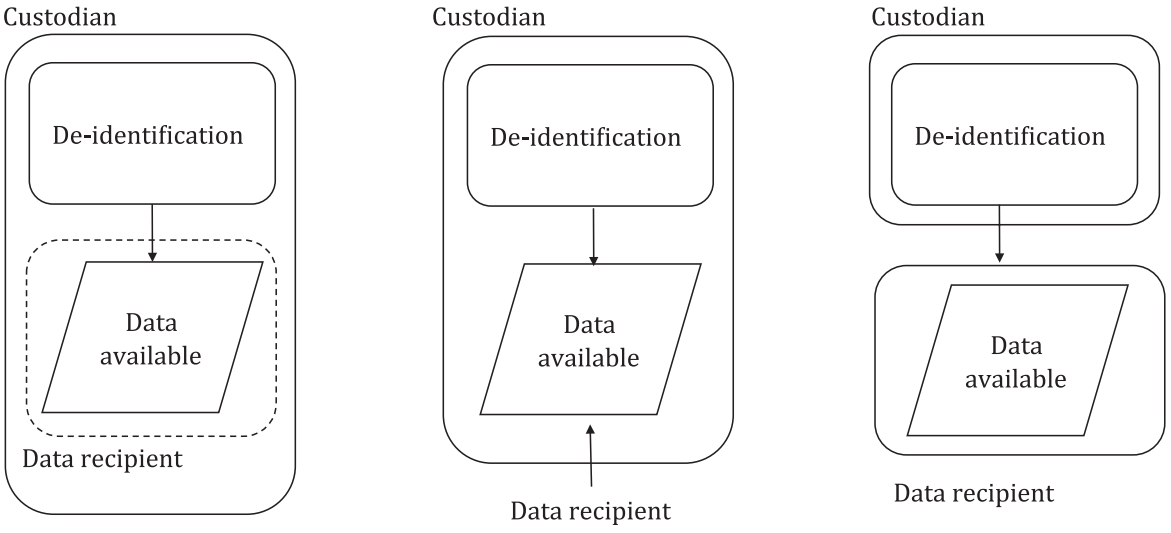
Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

De-identification is one potential means for facilitating the use of personally identifiable information (PII) in a way that does not identify or otherwise compromise the privacy of an individual or a group of individuals. The appropriate use of de-identification techniques can support compliance with regulatory requirements and relevant privacy principles. However, the term "data principal" used in this document is broader than "PII principal" and, for example, includes organizations and computers.

In almost all cases de-identification requires, at the very least, an evaluation of the additional information available to an individual or group that can inappropriately reveal or uncover PII (which is referred to as an adversary, whether a data principal is identified intentionally or not), and how they can combine it to reveal or uncover PII. In short, de-identification requires an assessment of the environment and the circumstances in which the data are made available to data recipients. This considers what additional information is available to an adversary and the possibility of attacks and motivation to re-identify. De-identification also requires an assessment of the data. This determines how the additional information available to an adversary can be used to reveal or uncover PII and the possibility of re-identification, or identity disclosure, by itself or attacks of inference.

This document provides organizations with an implementation framework to govern the appropriate use of data de-identification techniques described in ISO/IEC 20889. This de-identification framework can be applied at any point in the data lifecycle: from designing the means of data collection, the internal reuse of that data, making data available to external partners, or archival. The data recipients can therefore be internal or external to the data custodian that is implementing procedures and practices in accordance with this de-identification framework. As shown in Figure 1 a), use and reuse implies the custodian maintains oversight over the de-identified data while making it available to an internal department or functional group. Figure 1 b) shows external sharing, which implies the custodian maintains oversight over the de-identified data while making it available to an external data recipient (e.g. through a virtual access portal, or a physical data centre). Figure 1 c) shows external release, which implies the custodian transfers oversight over the de-identified data to an external data recipient. In each of these cases, the process of de-identification itself can be transferred to a third party, separate from the custodian or recipient. Written agreements with the recipient determine how data made available after de-identification can be used, in accordance with applicable laws.



a) Use and reuse **b) External sharing** **c) External release**

Figure 1 — Data availability

Information security, cybersecurity and privacy protection – Privacy enhancing data de-identification framework

1 Scope

This document provides a framework for identifying and mitigating re-identification risks and risks associated with the lifecycle of de-identified data.

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.

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 27000, *Information technology — Security techniques — Information security management systems — Overview and vocabulary*

ISO/IEC 29100, *Information technology — Security techniques — Privacy framework*

ISO/IEC 20889, *Privacy enhancing data de-identification terminology and classification of techniques*

ISO 31000, *Risk Management — Guidelines*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 27000, ISO/IEC 29100, ISO/IEC 20889, ISO 31000 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
custodian
person or entity that has custody, control or possession of electronically stored information

[SOURCE: ISO/IEC 27050-1:2019, 3.2]

3.2
data recipient
person or organization by, with or to whom data is accessed, shared or released

3.3
adversary
individual or unit that can, whether intentionally or not, exploit potential vulnerabilities

Note 1 to entry: Adversary, attacker, intruder, snooper, and other similar terms are often used interchangeably in the de-identification literature.

3.4
threat modelling
systematic exploration technique to expose any circumstance or event having the potential to cause harm to a system in the form of destruction, *disclosure* (3.8), modification of data, or denial of service

[SOURCE: ISO/IEC/IEEE 24765:2017, 3.4290, modified — Note 1 to entry has been deleted.]

3.5
privacy impact assessment
PIA
overall process of identifying, analysing, evaluating, consulting, communicating and planning the treatment of potential privacy impacts with regard to the processing of personally identifiable information, framed within an organization's broader risk management framework

[SOURCE: ISO/IEC 29134:2017, 3.7, modified — Note 1 to entry has been deleted.]

3.6
defined population
set of elements that a dataset is drawn from that contributes to the *adversary's* (3.3) ability to identify a data principal

3.7
sample
dataset that is only a proportion of the *defined population* (3.6), such that an *adversary* (3.3) cannot be certain that any particular entity was in it

3.8
disclosure
revealing confidential or personally identifiable information from a dataset based on a vulnerability that is found or exploited

3.9
shared data
dataset in which a fixed set of entities have been granted access to the data by the custodian

3.10
released data
dataset in which the custodian no longer directly controls who has access to the data

3.11
data privacy model
approach to the application of data de-identification techniques that enables the calculation of identifiability

[SOURCE: ISO/IEC 20889:2018, 3.3, modified — The word "formal" and "measurement" have been deleted from the term and "data" added, and "re-identification risk" has been replaced by "identifiability" in the definition.]

3.12
written agreement
data sharing agreement, memorandum of understanding, data access request, contract and any other formally documented agreement

3.13
data transformation
modification of the data

3.14
de-identification governance
system of directing and controlling the de-identification process

[SOURCE: ISO/IEC 38500:2015, 2.8, modified — "the de-identification process" has been added to the definition.]

4 Symbols and abbreviated terms

PIA privacy impact assessment

PII personally identifiable information

P probability function

5 Overview

The goal of this document is to provide a principles-based framework to approach de-identification, which considers procedures, risks, and harms. A principles-based approach to de-identification is intended to be neutral on the specifics of implementation and technologies. The framework is presented in four main parts:

— Context (Clause 6): An assessment of the environment and circumstances in which the data are made available to data recipients, to determine what external information can be available to an adversary. This implies that risk can be managed through contextual controls as well (meaning the IT security controls, obligations described in written agreements, and policy and governance measures).

— Data (Clause 7): An assessment of the data, to determine how the additional information available to an adversary can be used to reveal or uncover PII. Risk can be managed by limiting what data are made available, and in what form that data will be made available (by transforming the data).

— Identifiability (Clause 8): The method of assessing identifiability is a function of context risk (the probability of an attack) and data risk (the probability of disclosure given that there is an attack). An appropriate tolerance shall be defined to ensure the identifiability is below a pre-defined tolerance level.

— Governance (Clause 9): Documented procedures and practices for the custodian to ensure the above are done consistently and effectively, now and in the future, and the preparations that are required before, during, and after de-identified data are made available.

It can be necessary to repeat the process if the resulting de-identified data does not meet acceptance criteria (by the custodian or intended data recipient), in an effort to find solutions that are acceptable to both parties. For example, the privacy and security practices for the data environment can be improved in an effort to reduce threats and improve the utility of data that are made available to the data recipients based on their operational context. The entire approach can be thought of in a somewhat linear fashion, with de-identification governance by the custodian embodying the overall processes of context assessment, data assessment, identifiability assessment and mitigation, and acceptance criteria, as shown in Figure 2. It is, however, possible to reorder elements based on implementation needs (e.g. improving efficiency and scalability for specific data flows). ISO 31000 contains guidelines on managing risk faced by organizations.

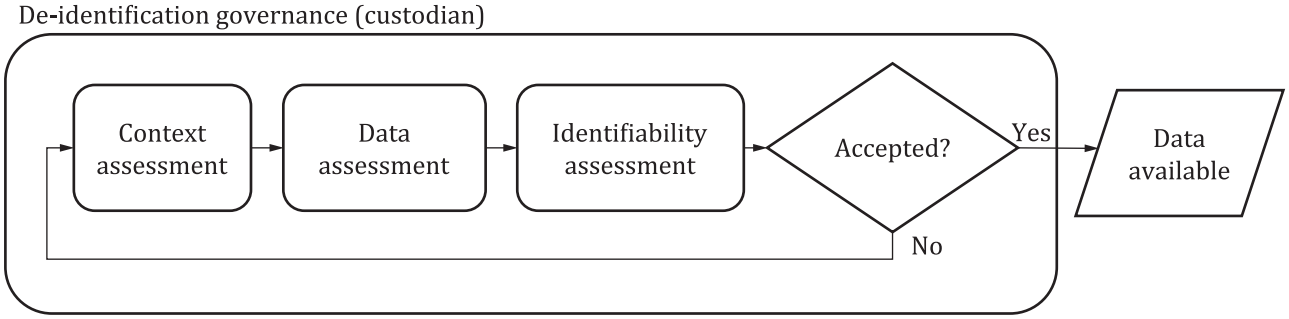


Figure 2 — De-identification framework in practice

6 Context assessment

6.1 General

The custodian shall evaluate the context in which data are being made available to a data recipient (either by providing shared access or by giving a copy of the dataset), to help properly scope the de-identification process.

NOTE Legal requirements can apply.

Determining this context involves a detailed assessment of the environment in which the data are accessed, shared or released, where the data come from and their intended use, and the circumstances under which the data are made available to a data recipient (such as levels of transparency). These elements shall be factored, in one form or another, into the method of assessing risk. For example, detailed checklists can be used to categorize risk in the data environment, and be factored into a standard risk matrix used to compare the possibility of an attack against the impact of identifying a data principal in a given context. It should be noted, however, that disclosures only occur if the attack is successful.

6.2 Threat modelling

6.2.1 General

The custodian doing the sharing or releasing, or a third party doing an assessment, shall use an objective and structured process to evaluate the environment in which the data will be accessed, shared or released. This environment includes persons and their motives (organizational and individual), other data they have access to, and their infrastructure and governance structures (including IT security controls such as those described in ISO/IEC 27002 and ISO/IEC 27701, access policies, etc.). An IT audit or assessment, even self-administered, can capture a great deal of information regarding the release environment, to help frame potential risks (in particular, potential threats).

A structured approach, often known as threat modelling, shall be used to assess the risk of an attack that would reveal or uncover PII. This includes examining what other external data sources can be available and sketching out the who, why and how of a potential disclosure. Potential threats can be:

— Deliberate: A targeted attempt to reveal or uncover PII in the data that are made available to them by an insider to the group or organization that is the data recipient.

— Accidental: A disclosure can also be unintentional, for example a data principal being recognized while a data recipient is working with the shared or released data.

— Environmental: The data can also be lost or stolen in the case where all the controls put in place have failed to prevent a data disclosure.

6.2.2 Security and privacy practices

The security and privacy practices of the data recipient will have an impact on the likelihood of a rogue employee at the data recipient's site being able to re-identify the shared data. A rogue employee can choose not to abide by a contract in the absence of strong mitigating controls. The security and privacy practices can also determine the likelihood of an outsider gaining access to the shared data, either directly or by compromising an insider's credentials.

An evaluation of mitigating controls shall be detailed and evidence based.

NOTE Professional, international, and government regulations, standards, and policies can apply, including ISO/IEC 27002 and ISO/IEC 27701, where appropriate.

Using a standardized approach also ensures consistency, not only for a single organization that is sharing data, but across organizations.

In order to avoid inappropriate or excessive burdens on the data provider or recipient, the evaluation can take into account third party audits and relevant certifications, as well as re-using prior analyses.

6.2.3 Motives and capacity to re-identify

A recipient's motives to reveal or uncover PII can be controlled in part by training, awareness, and obligations described in written agreements, including processes and terms described in ISO/IEC 23751, provided they are enforceable (through legal mechanisms and by refusing to share or release additional data). Obligations in written agreements can include:

— delivery of training on disclosure risks to individuals with access to de-identified data;

— regular reminders of their obligations to uphold data privacy and security policies;

— prohibiting attempts to identify data principals in the data made available to data recipients, or linking data that would extend the profiles of data principals (thereby increasing the risk of disclosing PII) without express permission;

— allowing for spot checks or full audits (possibly by a third party) that ensure compliance with the stated terms of the agreement;

— prohibiting the sharing with any other data recipient without express permission;

— defining the environment in which the data are accessed, shared and released (the infrastructure and governance structures that are expected to be in place);

— defining acceptable use cases and any expectations of how the data are used, or how use cases are evaluated to ensure its use is for purposes that are deemed appropriate.

NOTE It is possible that the enforcement of agreement obligations is not as effective when publishing as open data.

In planning the above obligations, the custodian should know where the data have come from, where they are going, and what the intended uses are for the data being accessed, shared or released. This will help situate the data release and the environment in which they will be made available to data recipients, so that appropriate measures can be taken to reduce the risk of potential disclosures of PII.

6.3 Transparency and impact assessment

6.3.1 General

The custodian shall assess the impact of disclosure, which decides tolerance when identifiability is evaluated, and can include:

— Being as transparent as possible and engaging with stakeholders where practicable. This way expectations of privacy are better understood by the entities involved.

— Whether the data are highly sensitive and intimate, come from vulnerable populations, or can reveal sensitive attributes or be stigmatizing.

— Potential injury or harm to data principals that can arise due to inappropriate processing.

— The trustworthiness of the data recipients (e.g. data sharing agreement in place, history of partnership, incentives to remain an ongoing partner).

— How the purposes for sharing and intended uses are in line with the interest of data principals.

— The extent to which data recipients will thoroughly consider their potential uses of data (for example, the use of privacy impact assessments) and act accordingly.

— Any potential breach of legal principles or fundamental rights.

6.3.2 Transparency of actions and stakeholder engagement

The custodian should explain simply and clearly its data collection practices, how it reuses data with a description of its rationale, and be open to feedback or seek the views of stakeholders on its data sharing activities with the goal of understanding, and where appropriate, addressing their concerns. Trust requires openness. Meaningful discussions with stakeholders can help to establish a reasonable balance of risk and benefits.

The custodian can consider comparing how similar organizations in its sector are releasing or sharing data, and whether any concerns have been raised about their practices. Surveys and focus group work on the data principals' views of data release, sharing and reuse (e.g., by industry associations) can also be considered to help inform decision making.

6.3.3 Privacy-related harms

The custodian should consider the potential privacy-related harms that can result from data being made available, including potential injury or harm due to inappropriate processing. For example, the custodian may:

a) evaluate whether the data are highly sensitive and personal; and also

b) consider use cases that may:

    1) reveal sensitive attributes,

    2) be stigmatizing, or

    3) support decisions that adversely affect data principals.

A custodian should conduct privacy impact assessments (PIA) in order to identify privacy-related risks and the proper controls to mitigate them. The PIA process may be shared and explained to increase transparency. It is a process that begins at the earliest possible stages of an initiative, when there are

still opportunities to influence its outcome and thereby ensure privacy by design. It is a process that continues until, and even after, the project has been deployed.

NOTE The guidelines for the process of conducting a PIA and the report's structure and content is described in ISO/IEC 29134.

Oversight mechanisms should be considered that avoid or meaningfully reduce the risk of harm and thereby engender the data principal's trust in the use of data, and some of these should be transferred as obligations to data recipients.

7 Data assessment

7.1 General

The goal of de-identification is to produce data that meets the requirements of the use case, while protecting the privacy of data principals. A data assessment is used to understand features and attack modelling exploits these features to evaluate potential vulnerabilities (e.g. singling out, linking, and inference) which are described in ISO/IEC 20889. The identifiability assessment in Clause 8 defines a threshold or protection objective based on established benchmarks. The default assumption should be that if an element or feature of the data is not needed for the use case (sensitive information, complex data structures that increase risk, etc.), the custodian should plan to remove it. Anything that the custodian can do to reduce the complexity of the data will, in turn, reduce the complexity of the technical analysis that the custodian shall conduct.

7.2 Data features

7.2.1 General

When thinking about whether, and how, to share or release data safely, a key consideration is the data themselves. The custodian shall evaluate features of the data that are central to identifiability.

7.2.2 Data principals

The custodian should use the following to gain insight into the external background information available to an adversary:

— the kind of data principals represented in the dataset;

— the nature and uniqueness of the data in the dataset and its potential value to an adversary;

— other data about a data principal available to an adversary;

— what a motivated adversary can learn about them through targeted searches.

The custodian should also consider whether there are any vulnerable groups in the data and take extra care in deciding whether or not they are needed in the data the custodian wants to make available and the processes required to protect them.

7.2.3 Data type

Data can be:

— Unit-level records, in which the rows represent a single population unit (person, household, etc.) and the columns represent the information (attributes) collected about them.

— Aggregated, in which cells represent the collection of population units that share categorized information collected about them.

— Unstructured, such as free-form text in which there are no restrictions on content or format, or a combination of the above.

NOTE Each form of data carries different risks and different approaches are needed to protect them.

7.2.4 Attribute types

The custodian shall determine the constituent parts and structured content of the dataset, where feasible doing this at record level granularity and by field type. For instance, ascertaining whether data elements can be:

— direct identifiers;

— indirect identifiers;

— of other concern (e.g. not widely available and targeted due to the confidential or personal nature of the data, possibly warranting additional protection).

Annex A provides examples of direct and indirect identifiers.

NOTE 1 When dealing with sensitive data, the risk is higher both in terms of the likelihood of a deliberate attempt to reveal or uncover PII and the impact. The impact on the organization in terms of public trust and reputation is likely to be greater the more sensitive the data. Identifying target content informs the organization about possible harm that can arise from disclosure.

NOTE 2 Attributes that are indirect identifiers can be the attributes of interest to the data recipients. They are considered in the identifiability assessment and can be removed

7.2.5 Dataset properties

The custodian should consider how properties of a dataset can potentially increase or decrease identifiability as a precursor to the requirement of careful data analysis and can include:

— Data quality: All data contains some level of error which offers some degree of protection. Although the custodian is likely to want to minimize this error for the sake of providing the most useful data possible.

NOTE Further information on data quality models is provided in ISO/IEC 25012.

— Group structure data: This is data that contains information for members of a group who are linked with one another. The data provide (more) information that can make a data principal unique in a defined population.

— Longitudinal data: This is data about a defined population which is collected over time and linked. Potentially unique changes in information over time can be captured and stand out among other longitudinal patterns.

— Population or sample data: Population data are about all persons in a particular identifiable group (as opposed to a random sample). There will be little uncertainty as to who is represented in population data.

7.3 Attack modelling

7.3.1 General

The custodian shall identify a set of potential attacks and associated metrics, which are used for the attack modelling, in conjunction with the threat modelling (see 6.2).

The custodian may objectively consider the following attacks to determine if they are credible possibilities, along with other significant domain-specific or dataset-specific concerns:

— attack where the adversary knows that a target individual entity is in the data;

— attack where the adversary does not, or cannot, know if a target individual entity is in the data;

— attack on all entities (rather than a target individual entity) that can be in the data.

NOTE These are sometimes referred to as prosecutor, journalist, or marketer risk, respectively, in the academic literature.

Metrics can also be averaged across all data principals (average risk), or the maximum taken across all data principals (maximum risk).

7.3.2 Maximum or average risk

— Maximum risk: For this metric, the maximum level of identifiability for a single data principal is taken, measured across all data principals in the shared or released data. This metric is used when there are no controls in place (e.g. public data sharing or release) to prevent the most identifiable data principal from being targeted in an attack. The custodian should assume a high possibility of attack because of the public nature of the sharing or release of data — someone will likely be motivated to attack the dataset if only to demonstrate potential vulnerabilities in it (called a demonstration attack).

— Average risk: For this metric, the average level of identifiability for a single data principal is taken, measured across all data principals in the shared or released data. This metric is used when there are controls in place (e.g. non-public data sharing or release) to prevent any individual data principal from being targeted in an attack, but not necessarily limited to the most identifiable data principal. The custodian should assume an honest but curious adversary when suitable controls are in place to prevent a demonstration attack.

7.3.3 Population or sample-based attack

The adversary has some background information about a specific data principal they are targeting and uses this background information to search for a matching record in the de-identified data made available to recipients. There are three types of attacks, two of which the custodian may evaluate based on whether the adversary knows if a targeted data principal is in the data made available, and a third type when the attack is not targeted:

— Attack of entity in population: In this scenario, an adversary knows the targeted entity is in a defined population in the data made available. The targeted entity will, by definition, be a data principal.

— Attack of entity in sample: In this scenario, an adversary does not, or cannot, know if the targeted entity is in the data being made available, most likely because the targeted entity is in a sample of a larger defined population. The targeted entity can be a data principal, or not.

— Dataset attack: In this scenario, an adversary wants to disclose the identity of as many data principals as possible in a dataset that has been made available.

7.3.4 Data privacy models

The custodian should choose data privacy models to help quantify identifiability in the data by forecasting a measure of data privacy.

NOTE For a description of a taxonomy of technical privacy metrics, see D Wagner I and Eckhoff[21].

Models make assumptions about the adversary and the context in which the data are made available. They can be useful in assessing the likelihood that data principals can be identified directly or indirectly in the data and making informed decisions about what transformations to the data are needed to reduce identifiability (for example, generalization, suppression, aggregation, noise addition), as described in

ISO/IEC 20889. However, the custodian shall consider the context of the data access, sharing or release as well as the data privacy model when making its decision.

There are many models to choose from. One class of models considers linking an entity to a record, to an attribute, or to the made available dataset itself. Another class of models considers what information will be gained in sharing or releasing data. Regardless of what model is used, understanding the assumptions and parameters is critical to the determination of what privacy measures are needed to prevent identity disclosure and how to factor the context of the data release into that assessment.

8 Identifiability assessment and mitigation

8.1 General

The custodian shall factor the context assessment and data assessment into the method of assessing re-identification risks which determine the overall identifiability of the data. Identifiability is a function of the probability of an attack (context assessment) and the probability of successfully identifying a data principal given that there is an attack (data assessment).

8.2 Assessing identifiability

8.2.1 General

In the context of data sharing or release, the custodian shall consider disclosure scenarios to ground the custodian's assessment of identifiability in a framework of plausible events. In constructing these scenarios, the custodian shall consider all the sources of the data to which the would-be adversary can realistically have access. If other custodians have been releasing similar data for a while without any apparent problems, the resources that the custodian shall devote to this can be more modest. The custodian should look for strong precedents, however, from data sharing or releases by reputable organizations such as national statistical organizations with decades of experience in de-identification.

The custodian should avoid focusing too closely on apparent vulnerabilities in the data. For example, a record is unique if no other records share its combination of values on the identifying attributes considered in a disclosure scenario (called uniqueness). Uniqueness does indicate vulnerability, but if there is no well-formed scenario through which that uniqueness can be exploited, no identity disclosure can happen. The custodian shall confirm that a match against a unique record in the custodian's dataset of the selected attributes in the scenario is correct, either through simulation or modelling. This takes into account that the unique record can have a statistical twin in the defined population that is not represented in the scenario sample.

NOTE Attribute disclosures can help enable identity disclosures and can be factored into data privacy models, as described in ISO/IEC 20889.

8.2.2 Quantifying identifiability

The custodian shall use objective methods and well-established benchmarks for selecting an identifiability threshold. The threshold shall work with the data privacy model that is being used and capture the likelihood and impact of identifying a data principal in a given operational context. For public data, the threshold can be based on maximum risk; for non-public data, the threshold can be based on average risk.

Identifiability can be quantified following a standard risk model where the probability of identification (the likelihood of a risk) is given by the probability of identification given a threat times the probability of a threat being realized. That is,

$$P(\text{identification}) = P(\text{identification} \mid \text{threat}) \times P(\text{threat}).$$

NOTE The conditional probability, $P(A \mid B)$, is the probability of A given B, or the probability that A will occur on the condition that B occurs.

Annex B summarizes well-established identifiability thresholds.

A subjective and objective criterion may be used by the custodian to influence the evaluation of impact, which may be based on the items listed in 6.3. If the impact is deemed to be high, that should skew the decision more toward a lower threshold. On the other hand, if the impact is deemed to be low, a higher threshold would be acceptable.

The following elements may be taken into account when assessing identifiability:

— What are the legitimate benefits to data principals or society from analysing the shared data?

— Are the data highly detailed, are they highly sensitive and personal in nature?

— What is the potential injury to data principals from an inappropriate processing of the data?

— What is the appropriateness of approval by data principals for disclosing the data?

The threshold can also be adjusted based on the benefits to entities from analysis of the data.

EXAMPLE A higher threshold, can be justified for the internal reuse, within the custodian, of data by a government department improving services for their citizens and residents, or for external reuse, outside the custodian, of medical trials data by the sponsor for the purposes of developing or improving treatments of related diseases.

In both cases mentioned in the example, data principals are likely to expect compatible or legitimate reuses of data. It is, however, important to meaningfully demonstrate that there are potential benefits to entities, and that expectations are aligned. Otherwise, where suitable for the appropriate lawful basis, pseudonymization can be used while maintaining the identifiable nature of data instead of de-identification which removes the identifiable nature of data.

The adversary, however, does not know which names are correctly matched when considering possible re-identifications. When the adversary attempts to verify the names that were matched, an additional error factor can be necessary to capture this uncertainty.

Whatever identifiability measures the custodian uses, they should be well-established and tested. A standardized approach, adopted widely, is more likely to have assumptions and justifications that have been scrutinized by experts.

8.2.3 Adversarial testing

The custodian should use adversarial or penetration testing to validate assumptions made by simulating attacks using "friendly" adversaries, as part of a practical assessment of a dataset's identifiability. This can be both informative and good practice, but takes skill and expertise as well as time and resources. It is especially important, however, in high risk data sharing or release scenarios such as open data.

A motivated adversary can be characterized as someone who is relatively competent, who has access to external data resources such as the internet and public documents, and is actively willing to make enquiries to uncover information. In basic analyses, they are not typically assumed to have specialist knowledge or advanced computer skills, or to resort to criminality. The custodian can of course use a different set of assumptions about the type of knowledge skills and resources that an adversary may bring to bear if they are well justified and to do so makes sense within the custodian's own disclosure scenarios (see References [16] and [19]).

Adversarial testing mimics more precisely what a motivated adversary can do, explicitly takes into account errors in data and matching, and is based on real data gathering and real external data. However, it is tied very tightly to one particular exercise and therefore doesn't necessarily represent all of the things that can happen. In practice, the custodian should combine data analytical techniques with adversarial testing rather than relying solely on either one.

8.3 Mitigation

8.3.1 General

De-identification essentially attends to either or both of i) the environment in which the data will be made available, or ii) the data to be made available. If the custodian's identifiability analysis (context and data) suggests that the data recipient needs stronger controls, the custodian shall suggest to the data recipient to reconfigure the data environment, or change the data. Also, in deciding on what approaches to apply, note that apparent disclosures (in which there seems to be a disclosure, even if that's not the case) can sometimes be just as harmful to a custodian as actual disclosures. Once mitigation measures have been applied, identifiability should be re-evaluated by comparing with the threshold.

8.3.2 Reconfiguring the environment

Reconfiguring the environment essentially involves controlling who has access, how they access the data and for what purposes. For example:

— allowing access only within the custodian's own secure environment (virtual or physical data enclaves);

— specifying the requisite level of security for the data;

— specifying that all analytical outputs be checked and sanctioned by the custodian before they are published;

— specifying persons who may access the data;

— specifying other organizational mitigation measures such as required training and contractual obligations.

Placing or tightening controls on the environment tends to have quite significant effects on the risk, often ruling out particular forms of attack, for example, and so if the data are sensitive these controls are certainly worth considering.

8.3.3 Transforming the data

Usually the custodian starts from a fairly fixed proposal of what the access, sharing or release environment is, and shall work on changing the data to be made available to the data recipient with the custodian's use case in mind. The custodian should start with the attribute types:

— Direct identifiers should be eliminated by suppression, or by replacing them with random values or pseudonyms that are not based on the identifying information (and irreversible in high-risk data access, sharing or release scenarios, such as public data releases).

— Indirect identifiers should be eliminated by suppression if the information is not needed, or transformed using techniques described below if required to reduce the risk of disclosure to an appropriate level (based on the custodian's scenario).

— Sensitive information should be eliminated by suppression if the information is not needed, or transformed using techniques described below if required to reduce the risk of injury or harms to an appropriate level.

In general, the custodian should be able to reduce identifiability to an appropriate level through generalization, suppression, and sampling as described in ISO/IEC 20889 because these methods are easier to understand and can have meaningful impact on identifiability if used correctly.

— Generalization reduces the level of detail of the information provided, by increasing the size of numerical intervals or merging categories of information.

— Suppression excludes certain attributes from the dataset made available to the data recipient, eliminating that information entirely from consideration.

— Sampling removes entities from the dataset being made available, which creates uncertainty that a particular population unit is actually represented in the data.

The main alternatives as described in ISO/IEC 20889 are various forms of data distortion, techniques that manipulate the data in order to foil disclosure strategies so that an adversary cannot be certain that an attack was successful. Applying data distortion can affect data utility in an unpredictable and non-transparent manner and leaves the custodian with the difficult question about whether or not to share or release information about the distortion. Sophisticated approaches exist (e.g. synthetic data) but are outside the scope of this document.

8.3.4 Re-evaluation

The custodian can rerun their identifiability measurements in order to see what impact reconfiguring the environment or changing the data has on identifiability. Many data privacy models also guide the custodian through the selection of data transformations, provided the custodian has already taken into consideration the context of the custodian's data sharing or release.

9 De-identification governance

9.1 General

De-identification governance comprises three pillars: persons, process and technology. Persons and processes of a de-identification governance framework are supported and enabled by the technology used. The custodian shall formalize and periodically review governance related to its data processing activities and its principles, policies, and procedures for data security, handling, management and storage, and sharing or release. These determine how users' relationships with the data are managed before and after data are made available and the steps needed in the event of a disclosure.

The principles, policies and procedures related to data sharing should be integrated with, and incorporated into, the custodian's wider information and data security practices.

9.2 Before data are made available

9.2.1 General

Organizational structures should be in place before data are made available to ensure risks remain negligible going forward. The necessity and the proportionality of the sharing, and the legal premises to do so should be assessed.

9.2.2 Assigning roles and responsibilities

Individual roles and responsibilities within a custodian should be determined.

— Identify a person who is responsible for authorizing and overseeing the disclosure control process and ensure that they have the necessary skills and knowledge to do this.

— Ensure that all relevant staff are suitably trained and understand their responsibilities for data handling, management, sharing and releasing. This can take the form of:

  — in-house training on the principles and procedures of the custodian's data processing activities;

  — external training on core factors such as disclosure control issues and techniques, data security, data protection law etc.;

  — implementing a staff non-disclosure agreement.

9.2.3 Establishing principles, policies and procedures

A custodian's internal structures for making data available should include principles, policies and procedures that cover:

— The whole de-identification process.

— Monitoring future risk implications for each data release.

— Maintaining a comprehensive record-keeping system across all the custodian's operational activities related to their data protection policies and procedures to ensure there is a clear audit trail.

— Undertaking a privacy impact assessment (PIA) for all the custodian's data products or across the organization as a whole. A PIA can:

  — help the custodian be aware of and address any particular privacy issues,

  — ensure the transparency of the custodian's activities,

  — promote trust in what the custodian does, and

  — help the custodian to comply with relevant privacy legislation.

— Identifying and dealing with cases where disclosure control can be problematic. The custodian should also consider at what point in the process (in dealing with a difficult case) the custodian should seek external help and advice from bodies such as a regulator or de-identification experts.

— Dealing with unintended disclosures. Depending on the custodian's particular needs, it should consider developing separate policies related to different potential unintended disclosures, or develop a single policy. Whichever is chosen, the custodian should consider how an unintended disclosure can occur and how it will respond.

9.2.4 Identifying and managing a data disclosure

A custodian should understand what a data disclosure is and identify ways to reduce the likelihood of such an occurrence. The custodian should:

— Define a data disclosure. Making a distinction between context and data can support a more nuanced evaluation of whether an incident investigation is required, and the possibility of a data disclosure (see Table 1).

— Identify the types of data disclosure relevant to the custodian's data situation.

— Identify those factors likely to lead to a disclosure, such as the loss of an unencrypted disc taken out of the workplace or the accidental emailing of data to the wrong person. Thinking through a range of possible disclosure scenarios can be very useful in helping the custodian identify how a disclosure can arise from their usual processing activities, as well as what errors, procedural violations or malicious intent can also occur.

— Establish measures to limit/avert those factors likely to lead to/facilitate a disclosure.

— Establish how the custodian will address violations of these measures.

Table 1 — Incident and breach scenarios for de-identified data

	Incident	Breach
Context	Evaluate likelihood of context breach	Evaluate likelihood of disclosure
Data	Evaluate likelihood of data disclosure	Evaluate impact of disclosure

9.2.5 Communicating with stakeholders

The custodian should engage and communicate with stakeholders to determine what they would like to know about the custodian's processing activities and to determine what constitutes an appropriate level of information to share. It is likely that stakeholders want to know the "what" of the custodian's processing activities, such as what data, or in which environments. They are also likely to want to know the "how" of the custodian's processing activities, such as how disclosures are avoided or how the custodian determines an environment to be safe. Such explanations should not expose security-related information that can assist an adversary in attacking the system and should be done a way that does not expose trade secrets, or commercially or technically sensitive details.

9.3 After data are made available

9.3.1 General

While the custodian's data can be considered safe at the time of its being made available, it is possible that this is not the case in the medium term.

9.3.2 Monitoring the data environment

Continuing advancements in IT capabilities, supporting ever-greater access to data and capacity for their analysis, and an ever-increasing amount of available data, mean that the custodian should consider the potential for change in the data environment in which the custodian has made data available. Possible actions include:

— Keeping a register of all the data the custodian has made available.

— Keeping track of direct recipients of shared data.

— Comparing proposed share and release activities to past shares and releases to take account of the possibility of linkage between releases leading to a disclosure.

— Maintaining awareness of changes in the data environment and how these can impact their accessed, shared or released data. For example:

  — keeping abreast of developments in new technologies and security that can affect the custodian's data situation by, for example, reading technology journals/blogs, listening to relevant podcasts or attending relevant events;

  — monitoring changes in the law or guidance on data sharing, releasing and dissemination by engaging with relevant custodians such as industry groups and the relevant data protection or privacy authority; and

  — keeping track of current and new public data sources of potentially identifying attributes by, for example, reviewing the information available on the internet and through more traditional sources such as public registers, local community records, estate agents' lists, professional registers, the library, etc.

— Performing periodic adversarial testing (see 8.2.3)

9.4 Mitigation in case of incident

To reduce the impact if a disclosure were to occur, custodians shall be able to respond to a data disclosure and put their plans into action, including:

— containing a breach;

— assessing and dealing with any ongoing risk;

— reporting the breach to the competent authorities and, if required, to the data principals;

— reviewing and learning lessons;

— as applicable, implementing mitigation measures for any injury or harm to a data principal.

Custodians shall put in place mechanisms that can help deal with an unintended disclosure. This includes having effective governance policies and procedures in place which essentially identify who does what, when and how, and generally support a culture of transparency.

— The custodian shall ensure it has a robust audit trail: This demonstrates that the custodian has followed all correct procedures, and identifies where, if at all, in their processing activities it can be necessary to make changes to prevent a similar occurrence.

— The custodian shall have a crisis management policy: This policy should identify key roles and responsibilities and detail an action plan.

The custodian shall ensure it undertakes a periodic review of their processing activities.

Annex A
(informative)

Example identifiers

A direct identifier can be any identifying number, characteristic or code that uniquely identifies a data principal in a given operational context. Common direct identifiers include, but are not limited to:

a) Name

b) Civil identification number

c) Passport number

d) Driver's license number

e) Address details

f) Email address

g) Phone number

h) Fax number

i) Bank account

j) Vehicle identifiers and serial numbers, including license plate numbers

k) Social security number

l) Health card number

m) Medical record number

n) Device identifier and serial number

o) Biometric identification codes, including fingerprints and voice prints, etc.

p) Full face picture images and any other comparable pairs of images

q) Account number, certificate number or license number

r) Internet protocol (IP) address number

s) Web universal resource locators (URLs)

An indirect identifier can be any attribute that does not uniquely identify the data principal alone in the operational context, but in combination with other information can uniquely identify the data principal. Common indirect identifiers include, but are not limited to:

a) Gender

b) Date of birth or age

c) Date of event (e.g. admission, surgery, discharge, visit-related date)

d) Geographic range (e.g. zip code, building name, region)

e) Ethnic origin

f) Nationality, place of origin

g) Language

h) Aboriginal identity

i) Visible minority status

j) Job title, work unit, department and other occupational information

k) Marital status

l) Education level

m) Years of schooling

n) Total revenue

o) Religious beliefs

p) Power consumption of a household

Annex B
(informative)

Example threshold identifiability benchmarks

B.1 Identifiable uniques

The subset of indirect identifiers that are used to model a specific threat, for scenarios of any reasonable complexity, result in a measure of identifiability that is non-zero. This is due to the nature of estimation, and thresholds should therefore be considered a tolerance on uncertainty from a scientific perspective. The following are examples of established thresholds in the estimation of uncertainty as applied to identifiability measures in an operational context.

Table B.1 summarizes well-established thresholds for identifiable uniques estimated from only a sample (see References [14] and [20]). These particular thresholds should not be interpreted as the proportion of uniques in the population. Rather, these thresholds represent an acceptable degree of uncertainty in the estimation of identifiability for uniques in the sample (they can be interpreted as the probability of a correct match from the population given a unique match in the sample). If there are true population uniques, risk goes to 1 as sampling fraction goes to 1.

Table B.1 — Benchmark thresholds for uniques (attack of entity in sample)

Scenario	Context (matrix)	Threshold
Public	High possibility of attack, low impact	Avg 0,005
	High possibility of attack, medium impact	Avg 0,001
	High possibility of attack, high impact	Avg 0,000 5
Non-public	Low-med possibility of attack, low-medium impact	Avg 0,1
	Medium possibility of attack, medium impact	Avg 0,05
	Medium-high possibility of attack, medium-high impact	Avg 0,005

B.2 Identifiable groups

The following categories of identifiability are based on the size of groups with the same set of identifying values in the dataset (an identifiable group). For public data, the probability is based on the minimum size of any one identifiable group. Typically, the identifiable group is in the range of 10-20 data principals that match on the indirect identifiers in a disclosure scenario. If an adversary attempts to randomly assign the names of known data principals to an identifiable group, the probability of the adversary being correct for any one data principal is the inverse of the group size, which gives the probability 0,1 (1/10) to 0,05 (1/20), following best practice for maximum identifiability. For non-public data, the probability is an average across all identifiable groups (average identifiability) and includes an assessment of context (and thereby threats). Table B.2 summarizes well-established identifiability thresholds (see References [15], [17] and [18]).

Table B.2 — Benchmark thresholds for identifiable groups (attack of entity in population or sample)

Scenario	Content (Matrix)	Threshold
Public	High possibility of attack, low impact	Max 0,1
	High possibility of attack, medium impact	Max 0,075
	High possibility of attack, high impact	Max 0,05
Non-public	Low-med possibility of attack, low-medium impact	Avg 0,1
	Medium possibility of attack, medium impact	Avg 0,075
	Medium-high possibility of attack, medium-high impact	Avg 0,05

For identifiable groups, the uncertainty of attempts to verify the names that were matched produces an additional error factor estimated to be in the range of 0,1, if not lower depending on the operational context. For example, for an identifiable group of 20 data principals, the probability the adversary can correctly match an identity is 0,05 (1/20), but to verify requires the estimated verification factor of 0,1, which results in a disclosure risk of 0,005 (see Reference [13]).

Bibliography

[1] ISO/IEC 23751, *Information technology — Cloud computing and distributed platforms — Data sharing agreement (DSA) framework*

[2] ISO/IEC/IEEE 24765:2017, *Systems and software engineering — Vocabulary*

[3] ISO/IEC 25012:2008, *Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model*

[4] ISO/IEC 27002, *Information technology — Security techniques — Code of practice for information security controls*

[5] ISO/IEC 27050-1:2019, *Information technology — Electronic discovery — Part 1: Overview and concepts*

[6] ISO/IEC 27701, *Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines*

[7] ISO 28004-1:2007, *Security management systems for the supply chain — Guidelines for the implementation of ISO 28000 — Part 1: General principles*

[8] ISO/IEC 29134:2017, *Information technology — Security techniques — Guidelines for privacy impact assessment*

[9] ISO/IEC 38500:2015, *Information technology — Governance of IT for the organization*

[10] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, De-Identification of Personal Information, NISTIR 8053. NIST, 2015

[11] ALLIANCE H.I.T.R.U.S.T. *De-Identification Framework*. HITRUST, 2015

[12] Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, April 2014

[13] BRANSON J et al., Evaluating the Re-Identification Risk of a Clinical Study Report Anonymized Under EMA Policy 0070 and Health Canada Regulations' (2020) 21 Trials 200

[14] Data61 and the Office of the Australian Information Commissioner, *De-identification Decision-making Framework*, CSIRO Reports EP173122 and EP175702, 2017.

[15] European Medicines Agency, External Guidance on the Implementation of the European Medicines Agency Policy on the Publication of Clinical Data for Medicinal Products for Human Use' (2018) EMA/90915/2016 Version 1.4</unknown>

[16] INFORMATION COMMISSIONER'S OFFICE *Anonymisation: Managing Data Protection Risk Code of Practice*. Information Commissioner's Office, 2012

[17] Information and Privacy Commissioner of Ontario *De-identification Guidelines for Structured Data*, IPC of Ontario, 2016.

[18] APPENDIX B FROM INSTITUTE OF MEDICINE *Sharing Clinical Trial Data: Maximizing Benefits*. Minimizing Risk, 2015[1)]

[19] OFFICE OF NATIONAL STATISTICS, Guidance on Intruder Testing ((undated))[2)]

1) https://www.nap.edu/catalog/18998/sharing-clinical-trial-data-maximizing-benefits-minimizing-risk

2) https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/disclosurecontrol/guida nceonintrudertesting

[20] UK ANONYMISATION NETWORK, Anonymisation Decision-making Framework. UKAN Publications, 2016[3)]

[21] WAGNER I, ECKHOFF D, Technical Privacy Metrics: A Systematic Survey' (2018) 51 ACM Computing Surveys (CSUR) 57:1

3) https://msrbcel.files.wordpress.com/2020/11/adf-2nd-edition-1.pdf

ICS 35.030

Price based on 22 pages