

Roundtable of G7 Data Protection and Privacy Authorities

Reducing identifiability in cross-national perspective: Statutory and policy definitions for anonymization, pseudonymization, and de-identification in G7 jurisdictions

11 October 2024

Interest in the use and application of processes and technologies for reducing the identifiability of individuals from their personal information¹ has accelerated in recent years. This includes technologies for de-identifying, pseudonymizing, and anonymizing personal information.

When applied appropriately, these processes and technologies can facilitate innovative uses of data, help to minimize privacy risks, and support the protection of the fundamental right to privacy.

At the same time, there is considerable cross-national variation in how these processes are integrated into policy frameworks for privacy and data protection. In G7 jurisdictions, this includes differences in how de-identification, pseudonymization, and anonymization are defined in privacy and data protection statutes, and how these terms interact with privacy requirements. This can lead to less certainty on the part of organizations with respect to their responsibilities.

This report is issued by the G7 data protection and privacy authorities to help promote a consistent understanding of approaches to de-identification, pseudonymization, and anonymization across policy frameworks. The report summarizes policy and legal definitions for each term and identifies key areas of overlap and divergence. It includes, as an annex, an overview of statutory and select non-statutory definitions in each jurisdiction.

This document is not intended to provide compliance guidance with respect to organizations' obligations under specific privacy or data protection laws. Organizations seeking advice on privacy compliance should seek guidance from the appropriate regulatory authority in their jurisdiction.

¹ Except where otherwise noted, this report uses the terms "personal information" and "personal data" interchangeably.

Comparative Overview

G7 jurisdictions have integrated definitions for de-identification, pseudonymization, and anonymization into policy frameworks for privacy and data protection. Broadly, these definitions establish specific legal meanings for these terms, including thresholds at which information is considered to be in scope for a given term. In some cases, definitions also include or imply specific processes or additional requirements that must be met.

Depending on the jurisdiction, information that falls under one definition or another may or may not be subject to fewer restrictions on use and disclosure, and can in some cases fall outside the scope of data protection law. How terms are defined, including parameters set for inclusion and exclusion as well as related requirements and thresholds, therefore has significant consequences for the protection of personal information in a given jurisdiction.

The following sections summarize areas of overlap and divergence in these definitions across G7 policy frameworks. Where relevant, they include consideration of sub-national jurisdictions, as well as the regional framework in place in the EU, in addition to national frameworks. The EU framework, being a regulation, applies therefore also at national level in France, Germany and Italy.

Key areas of overlap and divergence include:

- The extent to which identifiability must be reduced
- The extent to which information can be used to identify a person
- Prescribed processes and techniques for reducing identifiability
- Whether the resulting information is considered personal information

To an extent, and depending on the jurisdiction, the meaning and interpretation of definitions is based also on the definition of “personal information” within a given framework.

De-identification

Definitions of de-identification are included in Canada’s proposed *Consumer Privacy Protection Act (CPPA)*², the UK’s *Data Protection Act 2018 (DPA)*, and the US *Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. Sub-nationally, definitions are included in California’s *California Consumer Privacy Act (CCPA)*, Quebec’s *Act respecting the protection of personal information in the private sector (private sector law)* and *Act*

² Bill C-27, the *Digital Charter Implementation Act, 2022*, would, among other things, repeal Part 1 of the *Personal Information Protection and Electronic Documents Act* and enact the *Consumer Privacy Protection Act*. The Bill is currently being considered for adoption by Parliament.

*respecting Access to documents held by public bodies and the Protection of personal information (public sector law), and Ontario's Personal Health Information Protection Act (PHIPA)*³.

These definitions vary significantly in the extent to which identifiability must be reduced for information to be considered de-identified. Canada's proposed CPPA and Quebec's private and public sector laws require that individuals cannot be *directly* identified from the information. Similarly, the UK DPA defines de-identification in the same terms as pseudonymization, which requires that the information can no longer be attributed to a specific data subject without the use of additional information that is kept separately.⁴ By contrast, the US HIPAA Privacy Rule, California's CCPA and Ontario's PHIPA establish higher thresholds for considering information de-identified by requiring the minimization, to some extent, of the possibility for *indirect* identification as well: the CCPA requires that the information cannot reasonably be used to *infer* information about an individual, PHIPA requires that the information not identify individuals in "reasonably foreseeable circumstances", and the HIPAA Privacy Rule provides that the health information does not identify an individual and that there is no "reasonable basis to believe" that the information can be used to identify an individual.⁵

Definitions for de-identification also vary in whether de-identified information is considered personal information for the purposes of data protection law. Under the US HIPAA Privacy Rule and Ontario's PHIPA – both sectoral statutes addressing privacy in the context of health data – de-identified information is not considered personal (health) information and is therefore not subject to protections applicable to such information. Under Canada's proposed CPPA, the UK's DPA, and Quebec's public and private sector laws, de-identified information is considered to be personal information, and therefore remains within the ambit of those statutes. This variation corresponds, to an extent, with variation in thresholds for considering information to be de-identified: where de-identified information is not considered personal information, thresholds for considering information de-identified are generally higher.

In several jurisdictions, de-identification is considered to be a reversible process. Under the US HIPAA Privacy Rule, the possibility of re-identification

³ In the official French wording of PHIPA, note that the term "de-identify" is translated as "anonymiser", which can also be translated in English as "anonymize".

⁴ The information must also be subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

⁵ The HIPAA Privacy Rule permits protected health information to be de-identified in one of two ways. First, if an expert (a person with appropriate knowledge and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable) determines that the risk of identification is very small, that the health information could be used alone, or in combination with other reasonably available information to identify an individual. (See 45 CFR 164.514(a)(1)). Second, by removing specified identifiers, and any other unique identifying number, characteristic, or code; and there is no actual knowledge that the information could be used alone or in combination with other information to identify an individual. (See 45 CFR 164.514(a)(2))

by means of a code or record designed for that purpose is explicitly contemplated, and the law provides protection for such re-identified health information.⁶ Ontario's PHIPA and Canada's proposed CPPA include provisions concerning limited scenarios in which it is legally permissible to re-identify (or seek to re-identify) individuals from information that has been de-identified.

Pseudonymization

Definitions for pseudonymization are included in the EU and UK's *General Data Protection Regulation* (GDPR) and Japan's *Act on the Protection of Personal Information* (APPI). Sub-nationally, pseudonymization is defined in California's *California Consumer Privacy Act* (CCPA).

For the most part, definitions for pseudonymization are consistent in terms of the extent to which identifiability must be reduced under the applicable framework. Under Japan's APPI, California's CCPA, and the UK and EU GDPR, the threshold for identifiability requires that individuals cannot be re-identified from information that has been pseudonymized unless it is combined with other information.

Under the UK and EU GDPR and the California CCPA, 'pseudonymisation' means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person. In addition, the threshold for identifiability takes into account all the means reasonably likely to be used such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. With respect to the processes involved, Japan's APPI specifies that pseudonymization take place through either the removal of identifiers or their replacement with other identifiers without following patterns that enable restoration of their original state. Pseudonymization under the UK and EU GDPR is generally understood to involve similar processes. It is also included as an example of safeguards that can be implemented to support, for instance, data protection by design and by default, and security of processing.

Across these frameworks, information that has been pseudonymized is considered personal information. Additional measures are also required to prevent individuals from being re-identified, including that information that has undergone pseudonymization be kept separately from additional information that could be used for re-identification. Given that additional

⁶45 CFR 164.502(d)(2)(i).

information can, by definition, lead to re-identification of individuals from pseudonymized information, pseudonymization is not considered to be irreversible under these frameworks.

Anonymization

Definitions for anonymization are included in Canada's proposed *Consumer Privacy Protection Act* (CPPA), the EU and UK's *General Data Protection Regulation* (GDPR)⁷, and Japan's *Act on the Protection of Personal Information* (APPI). Sub-nationally, anonymization is defined in Quebec's private and public sector laws.

Definitions are generally consistent in requiring that anonymized information cannot identify an individual either directly or indirectly. However, there is variation in the threshold for determining identifiability in this context. Under Quebec's private and public sector laws, the threshold is whether it is "reasonably foreseeable in the circumstances" that the information can identify an individual directly or indirectly; the process used for anonymization must also follow criteria and terms established by regulation as well as "generally accepted best practices". Under the UK and EU GDPR, anonymized information must not meet the threshold for being considered personal data, which requires consideration of all the means "reasonably likely to be used" to identify a person, either by the controller or by another person. In Japan, the information must not identify a specific individual, and the process must be done in accordance with regulatory standards. In each case, these thresholds establish a contextual standard whereby factors that could lead to re-identification must be anticipated and addressed.

In some jurisdictions, the process for anonymizing information must be irreversible. This includes Quebec's private and public sector laws and Japan's APPI, which specify that individuals be "irreversibly" no longer identifiable from anonymized information (in Quebec: directly or indirectly). The UK and EU GDPR do not specifically require that the process be irreversible. The EU GDPR specifies that anonymous information is information which does not relate to an identified or identifiable natural person or information where an individual is no longer identifiable. The EU and UK GDPR also acknowledge that technological progress may render anonymization techniques ineffective over time, and that it is the responsibility of data controllers to ensure that techniques remain effective. Some data protection authorities have interpreted irreversibility to be a component of anonymization to some extent.⁸

⁷ The meaning of "anonymous data" is set out by Recital 26 in the UK and EU GDPR.

⁸ See for example, the CNIL: [L'anonymisation de données personnelles](#), May 2020.

At the same time, several regulatory bodies have acknowledged that in practice it may be difficult or impossible to fully eliminate all possibility of re-identification. Requirements for anonymization to be “irreversible” must therefore be understood as irreversibility *for the purposes of data protection law*, rather than a necessarily absolute state. For example, guidance issued jointly by the European Data Protection Supervisor and the Agencia Espanola Proteccion Datos notes that it may not always be possible to reduce the “probability of re-identification of a dataset to zero”, and that “a robust anonymization process aims to reduce the re-identification risk below a certain threshold” rather than necessarily achieve 100% anonymization.⁹ In Quebec, the Regulation respecting the anonymization of personal information specifies that “it is not necessary to demonstrate that zero risk exists” when conducting an analysis of re-identification risks.¹⁰ Similarly, guidance issued by the UK Information Commissioner’s Office notes that “data protection law does not require anonymization to be completely risk-free”, and that entities must only be able to mitigate risks until they are sufficiently remote that the information is *effectively* anonymized when the context is considered.¹¹ Guidelines published by the PPC Japan state that it is not necessarily required to eliminate all technical possibilities to restore the anonymized personal information back to the original data by any measures, but it is required to at least process the information to a state where the business operator handling personal information and the business operator handling anonymized personal information cannot restore the information by ordinary measures based on the ordinary persons’ and ordinary business operators’ abilities and measures, and other factors.¹²

Some regulatory bodies have also considered the question of whether information is anonymous to all parties, or whether the same information can be anonymous to one party and not another, depending on other information available to each. In the United Kingdom, the Information Commissioner’s Office has indicated that, in their view, “the same information can be personal data to one organisation, but anonymous information in the hands of another organisation”, with its status depending on the circumstances and context of its disclosure.¹³ While not a regulatory body, the International Organization for Standardization defines anonymization as preventing identification both by the controller of the information alone or in collaboration with any other party.¹⁴

⁹ [10 Misunderstandings related to anonymization.](#)

¹⁰ Gazette Officielle du Québec, [May 15, 2024, Vol 156, No. 20, 1758, Anonymization of personal information](#), Article 7.

¹¹ [Introduction to anonymization: Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#), May 2021, p.10.

¹² [Guidelines on the Act on the Protection of Personal Information \(Volume on Pseudonymized personal information and Anonymized personal information\)](#), p.30.

¹³ [Introduction to anonymization: Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#), May 2021, p.9.

¹⁴ [ISO/IEC 29100:2024.](#)

Across jurisdictions, anonymous information is considered not to be personal information, and therefore falls outside the scope of legal requirements applicable to such information under data protection law. In some jurisdictions, certain legal requirements are applicable to anonymized information, for example where it is prohibited to re-identify or attempt to re-identify such information.

Conclusion

There are important similarities and differences across jurisdictions in how key terms for reducing identifiability are defined and integrated into frameworks for privacy and data protection.

These include differences in the terms that are used, the circumstances under which information is considered to fall in scope for a given term, and in the privacy requirements applicable to that information. In some jurisdictions, definitions and corresponding requirements for anonymization and de-identification are nearly interchangeable, and the information concerned falls outside the scope of privacy requirements for personal information under data protection law. In other jurisdictions, definitions for de-identification correspond more closely with the concept of pseudonymization, and the information remains personal information under data protection law. This information can be exempted from specific privacy requirements in some jurisdictions, while in others it is not.

These areas of overlap and divergence in policy frameworks create opportunities for greater cross-national harmonization in approaches to de-identification, pseudonymization, and anonymization. This includes possibilities for regulatory cooperation in interpreting the application of frameworks to processes and techniques for reducing identifiability, particularly where these are novel and/or where there is uncertainty as to how they fit within established frameworks. Across jurisdictions, it seems the terms pseudonymization and anonymization are generally less overlapping and capture more granularity than the term de-identification.

Annex

Definitions for key terms

This annex outlines definitions for de-identification, pseudonymization, and anonymization put forward by state bodies in G7 jurisdictions. It includes statutory definitions for these terms, where they exist. It also includes non-statutory definitions, put forward in regulatory and policy documents, where these provide insight into common policy and regulatory uses of those terms, particularly in jurisdictions where no statutory definition is in effect.

De-identification

Canada

While current federal privacy legislation does not define this term, the *Consumer Privacy Protection Act* (CPPA), if adopted,¹⁵ would define 'de-identify' as the action of "modify[ing] personal information so that an individual cannot be directly identified from it, though a risk of the individual being identified remains".¹⁶ Existing documentation from the Government of Canada describes "directly identify" as "the act of establishing a person's identity by means of a direct identifier", i.e. "an attribute that can be used, by itself, to identify a person".¹⁷ The CPPA would require that any technical and administrative measures applied to de-identified information are proportionate to the purpose for which the information is de-identified and the sensitivity of the information.¹⁸

In the province of Quebec, the *Act respecting the protection of personal information in the private sector* and the *Act respecting Access to documents held by public bodies and the Protection of personal information* define de-identified information as information that "no longer allows the person concerned to be directly identified".¹⁹

In the province of Ontario, the *Personal Health Information Protection Act* defines "de-identify" in relation to the personal health information of an individual as "to remove any information that identifies the individual or for which it is reasonably foreseeable in the circumstances that it could be

¹⁵ The *Consumer Privacy Protection Act* has been proposed as part of Bill C-27, the *Digital Charter Implementation Act, 2022*, and is currently being considered for adoption by Parliament.

¹⁶ *Consumer Privacy Protection Act* (CPPA), Section 2(1).

¹⁷ Government of Canada, [Personal Information and Privacy Glossary](#).

¹⁸ *Consumer Privacy Protection Act* (CPPA), Section 74.

¹⁹ [Act respecting the protection of personal information in the private sector](#), Section 12; [Act respecting Access to documents held by public bodies and the Protection of personal information](#), Section 65.1.

utilized, either alone or with other information, to identify the individual”.²⁰ Note that in the official French wording of PHIPA, the term “de-identify” is translated as “anonymiser”, which can also be translated in English as “anonymize”. Guidance issued in 2016 by the Information and Privacy Commissioner of Ontario notes that “de-identification” is “the general term for the process of removing personal information from a record or data set”, and “once de-identified, a data set is considered to no longer contain personal information”.²¹

Non-statutory definitions

In a privacy implementation notice, the Treasury Board of Canada Secretariat, which provides guidance on the interpretation and application of the *Privacy Act*, defines de-identified information for the purposes of the notice as “personal information which has been modified through a process to remove or alter identifiers to a degree that is appropriate in the circumstances”.²²

In a recent investigation into the use of mobility data by a federal public health authority, the Office of the Privacy Commissioner of Canada defined de-identification for the purposes of the report as “a process whereby any personal identifiers, such as names, phone numbers, or device IDs in a mobility data context, are stripped from the data about a specific individual (often replaced with a randomly assigned identifier)”.²³ In an online glossary of terms, the Government of Canada defines de-identification as a “modification of personal information so that the person concerned can no longer be identified”.²⁴

United Kingdom

The *Data Protection Act 2018* (DPA) defines de-identified personal data, in relation to two criminal offences concerning the re-identification of de-identified personal data, as data that “has been processed in such a manner that it can no longer be attributed, without more, to a specific data subject”.²⁵ The DPA’s explanatory notes add that this provision “...defines the meaning of ‘de-identification’ and ‘re-identification’ for the purposes of the [criminal]

²⁰ [Personal Health Information Protection Act](#), Section 2.

²¹ Information and Privacy Commissioner of Ontario, [De-identification Guidelines for Structured Data](#), June 2016.

²² Treasury Board of Canada Secretariat, [Privacy Implementation Notice 2023-01: De-identification](#), March 2023.

²³ Office of the Privacy Commissioner of Canada, [Investigation into the collection and use of de-identified mobility data in the course of the COVID-19 pandemic](#), May 2023, paragraph 26.

²⁴ Government of Canada, [Personal Information and Privacy Glossary](#).

²⁵ *Data Protection Act*, 171(2)(a).

offence and reflects the definition of pseudonymisation in Article 4(5) of the (UK) GDPR”.

For the purposes of data protection law, the term de-identified data is used only in connection with Section 171 of the DPA 2018, which states that the re-identification of ‘de-identified personal data’ is a criminal offence. In this context, ‘de-identified’ personal data is pseudonymised data or data that was considered anonymised but can be re-identified considering all means reasonably likely to be used.

United States

The *Health Insurance Portability and Accountability Act* (HIPAA) defines “de-identified information” as “health information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual”.²⁶ For information to be considered de-identified, it must meet certain implementation specifications.

Sub-nationally, the *California Consumer Privacy Act* defines de-identified information as “information that cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer”, provided a business possessing the information take certain steps with respect to its protection.²⁷

Non-statutory definitions

In a technical paper, the National Institute of Standards and Technology (NIST) has defined de-identification as “a general term for any process of removing the association between a set of identifying data and the data subject”.²⁸

²⁶ 45 CFR [§ 164.514\(a\)](#).

²⁷ [California Consumer Privacy Act](#), 1798.140(m).

²⁸ National Institute of Standards and Technology, [De-Identifying Government Datasets: Techniques and Governance](#), September 2023.

Pseudonymization

Canada

Non-statutory definitions

In a privacy implementation notice, the Treasury Board of Canada Secretariat, which provides guidance on the interpretation and application of the *Privacy Act*, describes pseudonymization as “a process of masking direct identifiers” and notes that pseudonymization is a popular form of de-identification. The notice further suggests that pseudonymization occurs when direct identifiers are replaced with aliases and that same alias is used consistently across datasets.²⁹

The Government of Canada’s Personal Information and Privacy Glossary defines pseudonymization as “a de-identification technique in which the attributes that allow the direct identification of a person are replaced by pseudonyms”, noting that “in pseudonymization, attributes that allow the indirect identification of a person are not modified”.³⁰

The Department of Justice Canada, in a discussion paper, has defined pseudonymization as “a special form of de-identification where new data elements are substituted for identifying information”.³¹

European Union and United Kingdom

The UK *General Data Protection Regulation* (UK GDPR) and the EU *General Data Protection Regulation* (EU GDPR) define pseudonymization as “the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person”. This means that the use of ‘additional information’ can lead to the attribution to individuals.³² In practice, pseudonymization is commonly understood to involve replacing directly identifying data attributes in a dataset (for example, name) with a form of indirectly identifying data attributes (for example, alias, sequential numbering, key-coded data, and many implementations of hashing).³³

²⁹ Treasury Board of Canada Secretariat, [Privacy Implementation Notice 2023-01: De-identification](#), March 2023.

³⁰ Government of Canada, [Personal Information and Privacy Glossary](#).

³¹ Department of Justice Canada, [Privacy Act Modernization: A Discussion Paper](#), p.4.

³² See, for example, EDPS and AEPD, [10 Misunderstandings Related to Anonymisation](#), p.3.

³³ CNIL, May 2020, [L’anonymisation de données personnelles](#).

The UK Information Commissioner's Office has noted that pseudonymization refers to "techniques that replace, remove, or transform information that identifies an individual", including for example "replacing one or more identifiers which are easily attributed to individuals (such as names) with a pseudonym (such as a reference number)".³⁴

According to the UK GDPR and the EU GDPR, personal data which have undergone pseudonymization, which could be attributed to a natural person by the use of additional information, should be considered to be information on an identifiable natural person. In other words, pseudonymized information that has gone through a pseudonymization process is still personal data.³⁵

Japan

The *Act on the Protection of Personal Information* defines pseudonymized personal information as information relating to an individual prepared in a way that makes it not possible to identify a specific individual unless collated with other information, by either deleting all individual identification codes contained in the personal information (in the case of personal information containing such codes) or deleting a part of the identifiers or their equivalent contained in the personal information (in the case of all other personal information).³⁶ In both cases, it is permissible to replace all individual identification codes, or a part of the identifiers or their equivalent, with other identifiers or their equivalent (rather than deleting them) without following patterns that enable restoration of their original state.

In general, 'identifiers' refers to all items that can be used to identify a specific individual, including any information that can be easily collated with other information and thereby used to identify that specific individual.³⁷ 'Individual identification codes' generally consist of codes converted for use by computers to identify a specific individual by a distinguishing feature, or codes

³⁴ Information Commissioner's Office, [Chapter 3: pseudonymisation Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#), p.3.

³⁵ EU GDPR, [Recital 26](#).

³⁶ [Act on the Protection of Personal Information](#), Article 2.

³⁷ More specifically, see the definition of "identifiers or the equivalent" in Article 2(1) of Japan's APPI: Article 2(1) "Personal information" in this Act means information relating to a living individual which falls under any of the following items:

(i) information containing a name, date of birth, or other **identifier or the equivalent** (meaning all items (excluding individual identification codes) made by writing, recording, sound or motion, or other means, in a document, drawing, or electronic or magnetic record (this includes a record created in electronic or magnetic form (meaning electronic form, magnetic form, or any other form that cannot be perceived with the human senses; the same applies in item (ii) of the following paragraph); hereinafter the same); hereinafter the same) which can be used to identify a specific individual (this includes any information that can be easily collated with other information and thereby used to identify that specific individual);

(ii) those containing an individual identification code.

used to identify a specific user, purchaser, or recipient that are assigned or recorded differently for each of them.³⁸

United States

At the sub-national level, the *California Consumer Privacy Act* defines pseudonymization as “the processing of personal information in a manner that renders the personal information no longer attributable to a specific consumer without the use of additional information, provided that the additional information is kept separately and is subject to technical and organizational measures to ensure that the personal information is not attributed to an identified or identifiable consumer”.³⁹

Non-statutory definitions

The National Institute of Standards and Technology (NIST) has defined pseudonymization as “a particular type of de-identification that both removes the association with a data subject and adds an association between a particular set of characteristics relating to the data subject and one or more pseudonyms”.⁴⁰ NIST adds that the technique is commonly used “so that multiple observations of an individual over time can be matched and so that an individual can be re-identified if there is a policy reason to do so.”

Anonymization

Canada

The *Consumer Privacy Protection Act* (CPPA), if adopted, would define ‘anonymize’ as “to irreversibly and permanently modify personal information, in accordance with generally accepted best practices,⁴¹ to ensure that no individual can be identified from the information, whether directly or

³⁸ See the definition of “identifiers or the equivalent” in Article 2(2) of Japan’s APPI: Article 2(2) “**Individual identification code**” in this Act means one prescribed by Cabinet Order which consists of any character, letter, number, symbol or other codes falling under any of the following items: (i) characters, letters, numbers, symbols or other codes converted in order to be provided for use by computers, used to identify a specific individual by a distinguishing physical feature of theirs; (ii) characters, letters, numbers, symbols or other codes used to identify a specific user, purchaser, or recipient, which are assigned differently for each of them regarding the use of services for an individual or the purchase of goods for an individual, or which are stated or recorded in an electronic or magnetic means differently for each of them in a card or other document issued to an individual.

³⁹ [California Consumer Privacy Act](#), 1798.140(aa).

⁴⁰ National Institute of Standards and Technology, [De-Identifying Government Datasets: Techniques and Governance](#), September 2023.

⁴¹ At the committee stage, a motion has been adopted to replace “generally accepted best practices” with a standard of “no reasonably foreseeable risk in the circumstances”.

indirectly, by any means”.⁴² The CPPA would not apply to personal information that has been anonymized.

Sub-nationally, the province of Quebec’s *Act respecting the protection of personal information in the private sector* and *Act respecting Access to documents held by public bodies and the Protection of personal information* define information as being anonymized if it is “at all times, reasonably foreseeable in the circumstances that it irreversibly no longer allows the person to be identified directly or indirectly”.⁴³ These laws further specify that information must be anonymized according to generally accepted best practices and criteria and terms determined by regulation. The corresponding regulations establish various steps and criteria that must be followed when undertaking anonymization.⁴⁴ Under these laws, anonymized information may only be used for “serious and legitimate purposes” by private sector organizations, and for “public interest purposes” by public sector organizations.

Non-statutory definitions

In a privacy implementation notice, the Treasury Board of Canada Secretariat, which provides guidance on the interpretation and application of the *Privacy Act*, defines anonymized information for the purposes of the notice as “personal information that has been de-identified to the point that there is no serious possibility of re-identification, by any person or body using any additional data or technology at this point in time”.⁴⁵

The Government of Canada’s Personal Information and Privacy Glossary defines anonymization as “a de-identification technique that consists of irreversibly altering personal information so that the person concerned cannot be reidentified”.⁴⁶ In a discussion paper, the Department of Justice Canada has noted that “generally speaking, “anonymized” information has been irreversibly stripped of personal identifiers”.⁴⁷ In an investigatory report of findings, the Office of the Privacy Commissioner of Canada has noted that de-identified information does not qualify as anonymous if it is still possible to link the de-identified data back to an identifiable individual.⁴⁸

⁴² The [Consumer Privacy Protection Act](#), Section 2(1).

⁴³ [Act respecting the protection of personal information in the private sector](#), Section 23; [Act respecting Access to documents held by public bodies and the Protection of personal information](#), Section 73.

⁴⁴ Gazette Officielle du Québec, [May 15, 2024, Vol 156, No. 20, 1758, Anonymization of personal information](#).

⁴⁵ Treasury Board of Canada Secretariat, [Privacy Implementation Notice 2023-01: De-identification](#), March 2023.

⁴⁶ Government of Canada, [Personal Information and Privacy Glossary](#).

⁴⁷ Department of Justice Canada, [Privacy Act Modernization: A Discussion Paper](#).

⁴⁸ PIPEDA Case Summary [#2009-018](#).

European Union and United Kingdom

The UK *General Data Protection Regulation* (UK GDPR) and the EU *General Data Protection Regulation* (EU GDPR) specify that the principles of data protection, in other words data protection legislation, should not apply (i) to “anonymous information, namely information which does not relate to an identified or identifiable natural person” (ii) or to “personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable”.⁴⁹ These statutes note that, to determine whether a natural person is identifiable, “account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly.” Further, to ascertain whether means are reasonably likely to be used to identify the natural person, “account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments”.

The UK Information Commissioner’s Office has issued draft guidance (applicable within the UK) noting that anonymous information can be understood as the end result of a process that converts personal data into information that data protection legislation no longer applies to, as the information is no longer personal data.⁵⁰ In France, the French data protection authority (CNIL) has issued guidance that defines anonymization as “a process that consists of using a set of techniques in a way that renders impossible, in practice, the identification of a person by any means, irreversibly”.⁵¹ The European Data Protection Supervisor has issued a joint document with the Spanish data protection authority noting that, unlike pseudonymized personal data, “anonymous data cannot be associated to specific individuals”.⁵² In Germany, some data protection laws of the federal states define anonymization, with deviations in detail, as the alteration of personal data in such a way that the individual details of personal or factual circumstances can no longer, or only with a disproportionate amount of time, cost, and labour, be assigned to an identified or identifiable natural person.⁵³

Japan

The *Act on the Protection of Personal Information* (APPI) defines ‘anonymized personal information’ as information relating to an individual that can be

⁴⁹ EU GDPR, [Recital 26](#).

⁵⁰ [Introduction to anonymization: Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#), May 2021.

⁵¹ [L’anonymisation de données personnelles](#), May 2020.

⁵² EDPS and AEPD, [10 Misunderstandings Related to Anonymisation](#), p.3.

⁵³ BfDI, [Positionspapier zur Anonymisierung unter der DSGVO unter besonderer Berücksichtigung der TK-Branche](#), June 2020, p.3.

prepared in a way that makes it irreversibly not possible to identify a specific individual, by either deleting all individual identification codes contained in the personal information (in the case of personal information containing such codes) or deleting a part of the identifiers or their equivalent contained in the personal information (in the case of all other personal information).⁵⁴ In both cases, it is permissible to replace all individual identification codes, or a part of the identifiers or their equivalent, with other identifiers or the equivalent (rather than deleting them) without following patterns that enable restoration of their original state.

The APPI further requires that, when preparing anonymized personal information in an anonymized personal information database or equivalent, businesses handling personal information must process personal information in accordance with standards prescribed by Order of the Personal Information Protection Commission as those necessary to make it impossible to identify a specific individual and restore the information to its original state.⁵⁵

United States

Non-statutory definitions

The National Institute of Standards and Technology has defined anonymization as "a process that removes the association between the identifying dataset and the data subject", and notes that the term anonymization "is reserved for de-identification processes that cannot be reversed".⁵⁶

Acknowledgements

This document has been co-developed by the data protection authorities of the G7, as part of the G7 Emerging Technologies Working Group. Its development was led by the Office of the Privacy Commissioner of Canada.

The G7 data protection authorities thank the following organizations for their review and feedback on sections of the document:

- The California Privacy Protection Agency
- The Commission d'accès à l'information du Québec
- The Information and Privacy Commissioner of Ontario
- The US Department of Health and Human Services

⁵⁴ [Act on the Protection of Personal Information](#) (APPI), Article 2(6).

⁵⁵ [Act on the Protection of Personal Information](#) (APPI), Article 43(1).

⁵⁶ National Institute of Standards and Technology, [De-Identifying Government Datasets: Techniques and Governance](#), September 2023, p.13.