APPENDIX 9: The lawful pathways explained

While there are differences between the sets of privacy principles that operate in Australian jurisdictions, there are many commonalities, including how the principles both facilitate and restrict the sharing of personal information.

Generally, organisations are permitted to collect personal information where the information is necessary and relevant for their functions or activities (and in some instances, with the individual's consent); this is known as the 'primary purpose' for the collection. Organisations are then permitted to use and / or disclose that personal information for the primary purpose.

If the organisation then wishes to use or disclose the information for a different (secondary) purpose, it will need to satisfy at least one of a limited number of exceptions. Thus, only *some* secondary purposes are allowed.

Under privacy law, using and sharing health data about health consumers for synthetic health data projects will always be for a *secondary* purpose, as the primary purpose for collecting this information was to provide health care services to individuals (see 'Step 1: Assess the use case' above).

synthetic health data requests will generally involve handling personal information (specifically health information, which is a special subset of personal information) at two possible stages, each of which will require a lawful privacy pathway in order to proceed:

- When selecting and handling the source data from which synthetic health data will be generated (i.e. the source data will be considered personal information).
- When the re-identification risk associated with a synthetic health dataset is more than very low (i.e. in these circumstances, the synthetic health dataset will be considered personal information).

Multi-party projects

Where an organisation needs to collect personal information from other organisations in order to generate synthetic health data, the privacy and legal risks associated with these activities must first be identified and appropriately managed. This could include, for example, a scenario where one health department will disclose health information to another health department, which will create a linked health dataset for synthetic health data generation. Or it could include a scenario where a health department will disclose health information to a university research team, who will use it to generate a synthetic health dataset for research-related purposes.

In these cases, a Privacy Impact Assessment (PIA) must be completed in order to ensure that each organisation in these scenarios can *disclose* and/or *collect* the personal information under their own privacy obligations. The PIA should identify the most appropriate lawful privacy pathway/s for *each* participating organisation, and *each* part of the data journey.

The possible lawful pathways for using and disclosing personal information for a secondary purpose are explained below. Where organisations need to assess the possible lawful pathways that apply to their use case, they should also refer to the text of the Use and Disclosure privacy principles in the privacy laws that apply to them (see Appendix 3, The policy and legal framework underpinning this Framework, for a description of and links to the different privacy laws and privacy principles that could apply to an organisation).

'Directly related' and 'within reasonable expectations'

Privacy laws allow for health information to be used and shared outside of an organisation for a secondary purpose, if that secondary purpose is *directly related* to the primary purpose for which the information was collected. The secondary purpose must also be within the reasonable expectations of the individual, and in some cases, the organisation must have no reason to believe that the individual concerned would object to the use or disclosure.

For example, if information is collected in order to provide a health service to the individual, the use of the information to send an appointment reminder to the individual is for a secondary purpose that is directly related to the primary purpose, which an individual should reasonably expect.

The NSW Privacy Commissioner has advised that a directly related purpose "would be the type of situation that people would quite reasonably expect to occur with their personal information".[49]

Examples of disclosures of health information that the NSW Privacy Commissioner considers appropriate[50] under this test include:

- providing information to a person or organisation involved in the ongoing care of the patient
- investigating and managing adverse incidents or complaints about care or patient safety
- monitoring, evaluating, and auditing the provision of a particular product or service that the organisation has provided, or
- managing a legal claim made by the person.

[49] Privacy NSW, *A Guide to the Information Protection Principles*, 1999, p.35.
[50] NSW IPC, *Statutory guidelines on the management of health services*, p. 6.

If the individual has not been made aware (such as through a collection notice included on a patient form) that their personal information will be used or disclosed for the secondary purpose, there is a greater risk that they would not 'reasonably expect' the disclosure to take place, and this pathway may not be available.

Using or processing personal information for the purpose of de-identifying it to the extent it is no longer personal information may, in some circumstances, be considered a 'normal business practice' that is incidental or directly related to the primary purpose of collection.[51]

From a practical perspective, using health consumer information to generate synthetic health data can meet the 'directly related' and 'within reasonable expectations' where:

5. the use case is for a clear 'public benefit' purpose related to providing health services, and where the expected benefits from the use case are related to consumer health or health system outcomes;
6. the aim in creating and managing the synthetic health dataset is to achieve a 'de-identified' dataset for the use case, that significantly minimises the risk to individuals compared to if the source dataset was used for that use case; and
7. steps have been taken to set expectations with health consumers about how their health information will be used.

These requirements are reflected in the Use Case Assessment checklist in Appendix 4.

These conditions are explained in more detail above under 'Step 1', and form the basis for determining whether a particular synthetic health data use case can proceed under this Framework.

With consent

Personal information can be used or disclosed for a secondary purpose if an organisation has the consent of the individual/s to whom the information relates.

To be a valid consent under privacy law, such as to authorise a use or disclosure of personal information that would not otherwise be allowed, consent must be:

8. Voluntary (i.e. the individual opted in, had the opportunity to change their mind later, and has not since withdrawn their consent)
9. Informed (i.e. the individual was told about this data use proposal in a comprehensible manner before they chose to participate)
10. Specific (i.e. the consent was specific to this data handling proposal, not bundled in with other topics)
11. Current (e.g. given within the last two years), and

[51] See the OAIC's guidance, *De-identification and the Privacy Act*, March 2018, available at: https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/handling-personal-information/de-identification-and-the-privacy-act

12. Given by a person with capacity (e.g. parents can consent on behalf of children).

To rely on consent as a lawful pathway, the individual's consent must specifically cover the use or disclosure for the purpose being proposed, with transparency around which organisations will be handling their information.

From a privacy risk management perspective, express consent is typically easier to establish and evidence than implied consent (which relies on non-action or silence). This means an individual has been given the option to consent, and responds affirmatively – for example, by ticking a box, signing a consent form, or saying 'yes'. Organisations seeking to rely on 'with consent' as their lawful pathway also need mechanisms in place to manage consent – for example, to record consent and to have processes in place where an individual wishes to withdraw their consent.

Many data assets containing (or derived from) consumer health information are *collected* with the consent of the individual. However, if the individual did not consent at that time for their information to later be *used* or *disclosed* for a specified purpose, the organisation wishing to use or share the information for a synthetic health data project will need to either obtain the individual's specific consent to use and / or share their information or will need to rely on an alternative lawful pathway to authorise the use or disclosure.

From a practical perspective, obtaining and managing valid consents is not a suitable lawful pathway to pursue for large-scale synthetic health data projects, particularly where organisations wish to use or share data that has already been collected from individuals.

Seeking ethics approval: management of health services & research projects

Most privacy laws carve out special exceptions for using and sharing health information for purposes related to the management of health services and / or research. Although not well defined, privacy laws draw a distinction between 'research', and other activities, including management of health services, although the legal tests for each purpose are often similar, and both commonly require obtaining approval by an appropriate Human Research Ethics Committee (HREC) (depending on the jurisdiction).

When seeking approval from an HREC, organisations must ensure all aspects of the proposed information flows associated with a synthetic health data project are approved appropriately by the HREC. This includes any *disclosure* of information by one organisation to another organisation that is *collecting* it to prepare source data for synthetic health data generation, as well as the subsequent *use* of the information to generate the synthetic health data.

Sharing of data across jurisdictions and sectors with HREC approval relies on meeting tests set out under multiple pieces of privacy law and different statutory guidelines. In order to

seek a *single* ethics review for a project that involves data sharing across multiple jurisdictions and sectors, the organisations should apply to a registered HREC that has been nationally certified, and is recognised under the National Mutual Acceptance (NMA) scheme. This is because the NMA scheme (in which all state and territory jurisdictions participate) supports a single ethical review for multi-centre projects across state and territory public sector organisations.

Where organisations need to *share* real health data in order to prepare a synthetic health dataset, or where a synthetic health dataset is still considered 'personal information' due to the level of re-identification risk, seeking approval and a waiver of consent from an HREC will most likely be the most appropriate and practical lawful pathway to support the use and sharing of this data (although this will depend on the jurisdiction and the other lawful pathway options that may be available to an organisation).

This is the same lawful pathway that many organisations would most likely follow when seeking to collect and handle *real* data about people for their projects (e.g. for a clinical research project).

Required or authorised under another law

A collection, use or disclosure of personal information may be required or authorised under another law. If such an action is 'required' under law, the organisation handling the personal information would have no choice but to handle the information as directed under the applicable legislation. (An example is when an investigative or law enforcement body uses its compulsion powers to compel production of documents.)

If such an action is instead 'authorised' under another law, the organisation would be permitted to handle the information as set out in the legislation but would have a choice as to whether or not they do so.

The OAIC also describes the need for clear and direct language when seeking to rely on this exception as a lawful pathway:

> "An act or practice is not 'authorised' solely because there is no law or court/tribunal order prohibiting it. Nor can an act or practice rely solely on a general or incidental authority conferred by statute upon an agency to do anything necessary or convenient for, or incidental to or consequential upon, the specific functions and powers of the agency. The reason is that the purpose of the APPs is to protect the privacy of individuals by imposing obligations on APP entities in handling personal information. A law will not authorise an exception to those requirements unless it does so by clear and direct language.[52]"[53]

[52] See *Coco v The Queen* (1994) 179 CLR 427.
[53] OAIC, *Australian Privacy Principles (APP) Guidelines*, December 2022, B.135.

This pathway could be relevant where the principal laws that govern the functions or services of an organisation require or authorise a particular collection, use or disclosure of information that is applicable for a synthetic health data project. Other laws may also authorise using and sharing data for secondary purposes under certain circumstances, for example, the *Data Availability and Transparency Act 2022* (Cth) (DAT Act).

Effectively de-identified to be safe for sharing

Only robustly and effectively de-identified data will be considered safe to share under this pathway, which is one of the key aims in creating and handling synthetic health datasets with a very low re-identification risk.

In the context of privacy law, the term 'de-identification' must be understood by reference to the meaning of 'personal information'. 'De-identified' data therefore means that a person's identity is no longer apparent, or cannot be reasonably ascertained, following the application of one or more de-identification techniques to 'personal information'.[54] Examples of de-identification techniques are discussed in more detail in Appendix 7.

In theory, this means that 'de-identified' data is no longer 'personal information' for the purposes of regulation by privacy laws.

However even if direct identifiers such as name and address, or individual healthcare identifiers or unique reference numbers, are removed from a data set, a person's identity may still be 'reasonably ascertainable'. In fact, the Australian Privacy Commissioner has warned that de-identification "can be effective in preventing re-identification of an individual, but may not remove that risk altogether", for example if "another dataset or other information could be matched with the de-identified information".[55]

This means that if the surrounding context, and other available information used in combination with the data to be shared, could be used to ascertain a person's identity, the data should be assumed to be re-identifiable. In other words, the data should still be considered 'personal information', and privacy protections applied accordingly.

Further, 'identifiability' in law does not necessarily imply that a person's name or legal identity can be established from the information. The Australian Privacy Commissioner has said that:

[54] See the NSW Information & Privacy Commission Fact Sheet 'De-identification of personal information', 2020, https://www.ipc.nsw.gov.au/fact-sheet-de-identification-personal-information

[55] OAIC, *De-identification of data and information*, April 2015; was previously available at https://www.oaic.gov.au/information-policy/information-policy-resources/information-policy-agency-resource-1-de-identification-of-data-and-information but has since been replaced by OAIC, *De-identification and the Privacy Act*, 21 March 2018, available at https://www.oaic.gov.au/privacy/guidance-and-advice/de-identification-and-the-privacy-act/

> "Generally speaking, an individual is 'identified' when, within a group of persons, he or she is 'distinguished' from all other members of a group."[56]

De-identification should be seen as a methodology to manage 'personal information', in order to *reduce* the likelihood of any individual being identifiable. Data should not be considered completely 'de-identified' (and thus outside the regulation of privacy laws) unless it has been tested for re-identification risk, and found not to pose such a risk. It is also important to note that some privacy laws (e.g. the WA PRIS Act) still regulate some aspects related to the handling of de-identified data (such as limiting the transfer of de-identified information outside of Australia, ensuring the security of de-identified information, and prohibiting the re-identification of the data).

Only data that has been de-identified to the point that there is only a very low chance of re-identification or 'singling out' will it be considered 'safe' to use or share under this lawful pathway.

It should also be noted that even where health information has been effectively de-identified to be safe for use and / or sharing, where it is being used for a research project, additional ethical considerations may need to be made even if the research may be eligible for 'lower risk research' ethics review pathways on the basis no personal information will be used.[57]

This Framework seeks to rely on this pathway to support various use cases involving synthetic health data on the basis it has been de-identified to the point of no longer being 'personal information'. If synthetic health data has not been robustly and effectively de-identified, this pathway will not be suitable to support the use or sharing of synthetic health data, and an alternative pathway must first be settled (e.g. seeking ethics approval and a waiver of consent from an HREC).

Organisations should be aware that data considered 'de-identified' in one context may not remain de-identified in another. For example, unit record synthetic health data in a restricted, protected environment (such as a secure data enclave) and subject to specific governance controls may be considered 'effectively de-identified' within that environment and context. However, if the same data were published and made publicly available – or if it was subject to a data breach – the risk of re-identification may increase, and the data may become 'personal information' again (and so will become subject to privacy obligations again).

Other factors may also impact or heighten the risk of re-identification, such as if the source dataset (or a closely related dataset) and / or the trained model used to generate the synthetic health dataset were exposed (e.g. were made available on the dark web) or made available to users who have access to the synthetic health dataset. As such, re-

[56] Office of the Australian Information Commissioner, *What is personal information?*, May 2018. p.8, available at https://www.oaic.gov.au/agencies-and-organisations/guides/what-is-personal-information

[57] See National Health and Medical Research Council, *National Statement on Ethical Conduct in Human Research* (2025) at 5.1.15 – 5.1.18. Available at: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2025

identification risk should not be considered 'static'. Organisations will need to re-assess re-identification risk where there are changes relating to the data, the environment, or other relevant factors.

Other exceptions

Privacy laws may provide other exceptions which support handling personal information for synthetic health data projects. For example, the NSW Privacy Commissioner can issue a temporary Public Interest Direction (PID), or the NSW Minister of Health can issue a permanent Health Privacy Code of Practice (HPCOP) that may also authorise organisations subject to the NSW HRIP Act to collect, use or disclose health information in a manner not otherwise permitted under the NSW HPPs. Similar schemes are available under privacy laws in Queensland, the Northern Territory, Tasmania and Western Australia, as well as the Privacy Act. However, Victoria (with respect to health information) and the ACT do not have these mechanisms.

Privacy laws also provide other exceptions which can authorise the collection, use and disclosure of personal information, although they will not be relevant for synthetic health data projects. These exceptions include, for example, for law enforcement purposes, where there is a serious threat to life, health or safety, and in the cases of missing persons. Specifically for the handling of health information, there are also exceptions which pertain to the provision of healthcare services to the individual about whom the information relates. For example, there are exceptions that allow a health service provider to share the patient's health information amongst a treatment team, or with family members in certain circumstances.

From a practical perspective, the range of other exceptions which can authorise organisations to collect, use or disclose health information in a manner not otherwise permitted under their privacy obligations are not relevant for synthetic health data projects – particularly where privacy laws already have special exceptions designed for research projects and management of healthcare activities.

Further Resources

- OAIC Australian Privacy Principles Guidelines (see in particular 'Chapter 6: APP 6 – Use or disclosure of personal information')
- See also discussion in Appendix 5 (Impact Assessment) in relation to Aboriginal HRECs
- NHMRC Guidelines approved under Section 95A of the Privacy Act 1988
- NHMRC National Statement on Ethical Conduct in Human Research 2025
- NSW IPC Statutory Guidelines on Research – HRIP Act
- NSW IPC Statutory Guidelines on the management of health services – HRIP Act

- See Appendix 3, *The policy and legal framework underpinning this Framework*