

Data & Research Governance Policy

Version 5.1 | 31 May 2024

| | | | | |
|-----------------------------------|---|--------------------------|----------------------|---|
| Document Title: | Data & Research Governance Policy | | | |
| Summary: | This policy provides the Cancer Institute NSW (the Institute) with an instrument to formally manage its data assets in a collaborative, consistent and co-ordinated manner. | | | |
| Date of Issue: | 23-Nov-21 | Next Review Date: | 29-May-26 | |
| Status: | Active | Review Cycle: | Every 2 Years | |
| Contact Officer: | Chief Data Officer | | | |
| Applies To: | All CINSW staff including contingent workers and contractors. | | | |
| References: | The following are related policies, guides, frameworks that support this policy: NSW Open Data Policy NSW Cyber Security Policy NSW Data and Information Custodianship Policy NSW Information Management Framework NSW Information Classification, Labelling and Handling Guidelines | | | |
| Version and Change History | Version | Who | Date | What |
| | 5.1 | Kim Cameron | 31/05/2024 | Update to new template |
| | 5 | Faith Papuni | 14/10/2021 | Updates following review meeting on 13/10/21 |
| | 5 | Faith Papuni | 24/09/2021 | Extensive review of policy to ensure alignment to NSW Health Data Governance-initial draft |
| | 4 | Jamie Gabriel | 25/06/2019 | Amendments to Section 3 to reflect roles & responsibilities outlined in NSW Health Data Governance Framework. Plus, minor admin changes |
| | 3 | Phil Siddens | 24/05/2017 | Minor formatting and edits |
| | 2 | Narelle Grayson | 09/09/2015 | Review and updated policy |
| | 1 | Stephen James | 03/01/2010 | Initial draft |
| Approvals | Version | Who | Date | Record |
| | | | | |
| | 5 | Maria Mury | 23/11/2021 | Approved and recorded in HPRM Notes E10/13359 |
| | 4 | Lisa Cox | 03/07/2019 | Approved E19/19299 |
| | 3 | David Currow | 23/06/2017 | Approved |
| | 2 | Beth Macauley | 19/10/2015 | Email approval |
| | 1 | Beth Macauley | 27/07/2010 | Email approval |

Table of Contents

| | | |
|-----------|--|-----------|
| 1. | Introduction..... | 4 |
| 1.1. | Overview | 4 |
| 1.2. | Purpose | 4 |
| 1.3. | Scope | 4 |
| 2. | Five safes framework..... | 5 |
| 3. | Policy statement..... | 7 |
| 4. | Roles and responsibilities..... | 9 |
| 5. | Legal and regulatory compliance | 10 |
| 6. | Policy review | 11 |
| 7. | Terms and definitions | 11 |
| 8. | Related documents | 13 |

1. Introduction

1.1. Overview

The Cancer Institute NSW (the Institute) was established in July 2003 through the Cancer Institute (NSW) Act 2003 to lessen the impact of cancer in NSW. The Institute is funded by the NSW State Government and governed by the Cancer Institute NSW Board. The Institute's objectives are to:

- reduce the incidence of cancer in the community
- increase the survival rate for people diagnosed with cancer
- improve the quality of life of people diagnosed with cancer and their carers
- provide a source of expertise on cancer control for the government, health service providers, medical researchers, and the general community.

The Institute's data assets are used to support these objectives.

1.2. Purpose

The goal of this policy is to ensure that the Institute's data assets are well-maintained and utilised to achieve its objectives. It also outlines the principles the Institute will follow to maintain data integrity, security, quality, and usability throughout its lifecycle.

This policy was created in accordance with the:

- Cancer Institute NSW Data Strategy, which outlines the Institute's operations and data investments to improve cancer outcomes for NSW residents; and
- aligns with both the broader NSW Health Data Governance Framework (GL2019 002), which outlines the roles and responsibilities related to data governance across NSW Health, as well as
- the 5-Safes Framework, a multi-faceted approach to effectively managing data in line with relevant legislative and other compliance obligations.

1.3. Scope




This policy covers all areas of data asset management for the Institute data assets and specifies the Institute's obligations for ensuring that data use for research and health-services management is conducted responsibly and ethically. This policy establishes safeguards to ensure that:

- there is a strategic alignment of the Institute's data assets to the Cancer Institute NSW Cancer Plan and the NSW's Data Strategy
- there is compliance with relevant legislation, policies and procedures and standards
- there are mechanisms in place to ensure there is confidence in the data used to inform decisions
- there are effective assurances and controls in all aspects of the data management life cycle
- there is clarity on roles and responsibilities and accountability relating to data and research governance

2. Five safes framework

The Institute is committed to effective compliance with its privacy and confidentiality obligations and in managing the risks associated with the use and release of data.

The Five Safes framework¹ is a globally recognised risk management model that will be used to identify and manage data sharing risks across five “safety” dimensions: people, projects, settings, output and data.

| Function | Description | Potential Risk |
|---|---|---|
| People  Share data only with authorised users | Safe people ensure users can be trusted to use data in an appropriate manner. <i>All Data users must have the requisite knowledge and skills and act in accordance with required standards of behaviour.</i> | Users of the data: <ul style="list-style-type: none"> • lack of relevant expertise • have insufficient statistical skills to analyse the data effectively • and/or their organisation are unlikely to be able to manage data breach risks effectively • and/or their organisation have a history of breaching terms and conditions. • are subject to a conflict of interest |
| Settings  Using data only in a safe and secure environment | Safe Setting ensure that access facility is safe against unauthorised access. <i>This refers to the way in which data are accessed and is subject to practical controls, both electronically and in terms of the physical environment.</i> | Data are: <ul style="list-style-type: none"> • misplaced, intercepted, or revealed during transmission to the setting (data/privacy breach) • vulnerable to unauthorised access at the setting (data/privacy violation) • used for purposes other than those authorised (including linking to other data) • removed from the approved setting; and • not destroyed at the end of the project |
| Projects  Sharing data only for appropriate and authorised purposes | Safe Projects ensures that the use of data is appropriate. <i>Data must be used in a legal and ethical manner, and the project is designed to benefit the public.</i> | <ul style="list-style-type: none"> • the project is unlikely to yield public benefits proportionate to risk • the project design is unlikely to achieve stated goals • consent arrangements are illegal |

¹ <https://www.aihw.gov.au/about-our-data/data-governance/the-five-safes-framework>

Outputs



Ensuring public outputs from data sharing projects to not identify people or organisations

Safe Outputs ensures that statistical findings remain confidential where applicable

Safeguards must be in place when disseminating the project's findings to reduce risk.

- outputs do not meet confidentiality requirements
- outputs are provided without data supplier approval; and
- output treatments differ from previously revealed data.

Data



Applying appropriate protections





Safe data ensures there is minimum risk of disclosure.


Data must be properly processed to ensure that it cannot be used to identify individuals or organisations.

- no identifiers are deleted
- data variables not required for the project are included
- data analysis is insufficient to prevent disclosure and reidentification of individuals

3. Policy statement

The Data Lifecycle for the management and use of data within the Institute is underpinned by the data management functions outlined below. When accessing data, all staff, including contractors, and consultants must ensure the following:

| Function | Statement |
|--|---|
| Policies, Guidelines & Standards  Data Ownership | <ul style="list-style-type: none"> roles of data sponsor, data custodian, and data steward must be been clearly defined for all data assets there must be a legislative basis for the collection and storage of the data asset the Institute must comply with all with applicable state and federal laws, regulations regarding data exchange, and standards |
| Information Architecture & Integrations  Reliable Flow of Information | <ul style="list-style-type: none"> there must be a data dictionary for all data assets using an Institute provided template, which clearly outlines data elements, master data, reference data, and this must be kept centrally under version control and in appropriate storage data assets must be stored in secure environments that can be that can be monitored to ensure access control and allow anomaly detection and technologies used for storage must be approved by the Data Governance Steering Committee data must be periodically backed up on an agreed schedule to prevent loss, and appropriately disposed of when it is no longer needed |
| Data Quality Framework  Data Stewardship | <ul style="list-style-type: none"> a Data Quality Framework (DQF), created using an Institute template, must be in place for all Data Assets documenting all rules used in the data engineering process, and any caveats of use the DQF must be held centrally, made available to all staff (and external parties where appropriate) and held under a review cycle the data custodian and data users must ensure appropriate procedures are followed to uphold the quality and integrity of the data they access data records must be kept up to date throughout every stage of the business workflow and in an auditable and traceable manner. Data should only be collected, extracted, manipulated, and reported for legitimate uses and to add value to the Institute. where appropriate, before any data (other than publicly available data) is used or shared outside the Institute, verification with the Data Custodian is required to ensure the quality, integrity and security of data will not be comprised data shall be retained and disposed of appropriately under the relevant legislation, standards and policies issued by the State Records Authority, and the Institute's Records and Information Management Policy (E06/10927) |
| Reporting & Analytics  | <ul style="list-style-type: none"> all use cases for a Data Asset must be logged the creation of interactive analytics for Data Assets is limited to the agreed upon technologies |

| | |
|--|--|
| <p>Knowledge from Information</p> | <ul style="list-style-type: none"> all sharing of data via interactive analytics must be monitored and regularly reviewed to ensure it is fit for purpose all research projects conducted using Institute Data Assets shall be monitored by the Data Governance Team using agreed upon metrics and must comply with the relevant data and research governance procedures data and data analytics are valuable resources of the Institute. The Institute adheres to the recommendations in the <i>NSW Government Intellectual Property Framework (2020)</i>² with regards to ownership of data and data analytics outputs |
| <p>Privacy, Compliance & Security</p>  <p>Role Definition & Accessibility</p> | <ul style="list-style-type: none"> all staff working with data in the Institute must have appropriate training, and training requirements will be maintained by the Data and Research Governance Team any data breach must be reported to the Data and Research Governance Team to protect the Institute's data against internal and external threats (e.g., breach of privacy and confidentiality, or security breach) a privacy threshold assessment shall be undertaken prior to the commencement of any new activity or change to an existing activity that involves the collection, storage, use or disclosure of personal or health information to identify whether a Privacy Impact Assessment (PIA) is required |

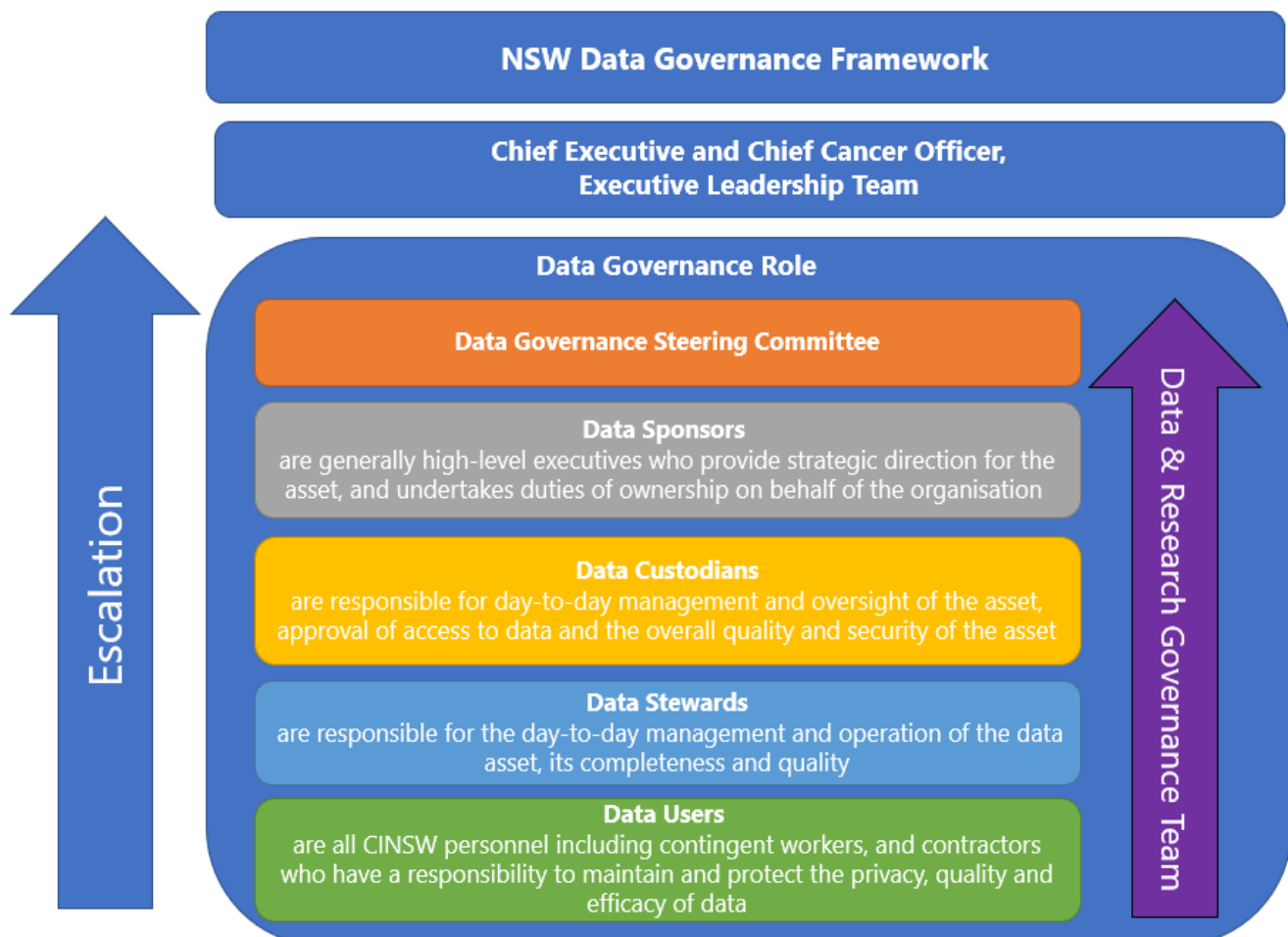
² <https://arp.nsw.gov.au/c2021-11-nsw-government-intellectual-property-framework/>

4. Roles and responsibilities

The [NSW Health Data Governance Framework](#) (GL2019_002) stipulates that each state-wide data asset must have a designated Data Sponsor, Data Custodian and Data Steward. The Institute follows best practices by nominating a data sponsor, custodian, and steward for each of its data assets. The roles and responsibilities of these positions and Data Users are defined below:

Data & Research Governance Roles and Responsibilities

Effective data and governance is based on clear roles and responsibilities, which ensures that data is managed efficiently throughout the data lifecycle.



Adapted from: <https://data.nsw.gov.au/data-governance-toolkit-0>

5. Legal and regulatory compliance

The Institute adheres to all applicable regulations, policies, and procedures governing data collection, management, and preservation, confidentiality, data sharing, data linkage, and public disclosure detailed in the [NSW Health Data Governance Framework](#) (GL2019_002) which are listed below:

The legislative context for each Institute Data Asset will be detailed in the Institute's Data Asset Framework.

| Legislation | Description |
|---|--|
| Cancer Institute (NSW) Act 2003 | the Cancer Institute NSW was established in July 2003 through the Cancer Institute NSW Act 2003 as a response to the need to further decrease the devastating impact of cancer on our society. Its functions, therefore, support the overall mission of lessening the impact of cancer on members of the public in NSW |
| Data Sharing (Government Sector) Act 2015 | authorises public sector agencies to share data with another public sector agency for specific purposes |
| Government Information (Public Access) Act 2009 (NSW) | the GIPA Act facilitates public access to NSW Government information. It does this by authorising and encouraging the release of information by NSW Government agencies, giving members of the public the right to request access to government information, and by ensuring government information is only restricted where there is an overriding public interest against disclosing the information |
| Health Services Act 1997 | regulates the governance and management of the public health system in NSW |
| Health Administration Act 1982 | sets out the broad roles of the Minister and Secretary, NSW Health in relation to the health portfolio generally |
| Health Records and Information Privacy Act 2002 | establishes controls and obligations related to personal health information. NSW Health can, without consent, use and disclose both identifiable and de-identifiable personal health information when it is necessary to use (health) information to manage health services |
| Human Tissue Act 1983 | contains provisions relating to the use of tissue, including blood, from either a living or dead body for medical, scientific, or therapeutic purposes |
| Public Health Act 2010 | contains a range of legislative provisions relating to the protection of public health in NSW. |
| Privacy and Personal Information Protection Act 1998 | establishes controls and obligations related to personal information |
| State Records Act 1998 | sets out the records management responsibilities of public offices |

6. Policy review

This Policy will be reviewed and updated every two (2) years from the approval date, or more frequently if appropriate. In this regard, any staff members who wish to make any comments about the policy may forward their suggestions to the Manager, Data Governance.

7. Terms and definitions

The below list is the Institute's common terms and definitions used for Data & Research Governance and is reviewed periodically to ensure they align with the wider definitions used within NSW Health and NSW Government.

| Term | Definition |
|-------------------------------|--|
| Data | Any facts, statistics, instructions, concepts, or other information in a form that is capable of being communicated, analysed, or processed (whether by an individual or by a computer or other automated means). For the purposes of this Policy Directive, the terms data and information are used interchangeably |
| Data Asset | a data asset is a structured collection of data developed for a broad purpose. An enduring data asset (or enduring linked data asset) is a subset of this category, denoting the linkage of a larger range of data that is designed for potentially many purposes and users. A data asset could also include models, methodologies, and algorithms |
| Data Dictionary | a resource that provides standard data collection descriptions and definitions. It may describe the collection's content, format, and structure, as well as any relationships between data elements or data collections. Both data dictionaries and metadata serve the same purpose: they provide a centralised resource to help users better understand the data with which they are working |
| Data Quality Framework | for all Data Asset's held in the institute, there should have an accompanying comprehensive data asset quality statement created using a standard Institute Template (REF) template, which describes in detail the caveats on data elements and existing use cases made available across NSW Health, and includes quality and coverage information that is updated with each data release |
| Data Type | <p><u>Clinical Data</u></p> <p>Aggregated data: consolidated data relating to multiple individuals</p> <p>Unit record data: records of data that relate to the health of an individual</p> <p>De-identified data: data about a person whose identity is not apparent and cannot be reasonably ascertained</p> <p>Identifiable data: personal, unique identifying data e.g., name, address, age, date of birth, ethnicity, or diagnosis. It can also be a combination of data that allows a person's identity to be "reasonably ascertained"</p> <p><u>Corporate Data</u></p> <p>data about how the organisation is managed including its Finance, Human Resources and Records Management. This may or may not be of a sensitive nature</p> |
| Data Use | the 'use' of health data generally refers to how data is communicated or handled. When using CINSW data, it is assumed that the quality will be evaluated using data quality guidelines |

| | |
|--------------------------------------|---|
| Information | is data that have been processed and analysed into a form (physical, oral, or electronic) that is meaningful to the recipient. For the purposes of this Policy Directive, the terms information and data are used interchangeably |
| Health Information | has the same meaning as in the <i>Health Records and Information Privacy Act 2002</i> (HRIP Act), i.e., personal information or an opinion about a person's physical or mental health or disability, or a person's express wishes about the future provision of health services for themselves, or a health service provided, or to be provided, to a person |
| Management of Health Services | refer to activities which fall within the 'funding, management, planning, or evaluation of health services', i.e., activities for the purpose of undertaking, monitoring, assessing, or improving clinical practices and health service delivery |
| Metadata | the structured description of data characteristics such as content, quality, and format. It provides a common meaning, allows for comparisons, and facilitates the retrieval, use, and management of data resources |
| Personal Information | is information or an opinion (including information or an opinion forming part of a database and whether recorded in a material form) about an individual whose identity is apparent or can reasonably be ascertained from the information or opinion as defined in the <i>Privacy and Personal Information Protection Act 1998</i> (PPIP Act) |
| Research | can be defined as a systematic investigation to establish facts, principles or knowledge or a study to obtain or confirm knowledge. A defining feature of research is the validity of its results. The knowledge that is generated by research is valid in the sense that what is discovered about the particular facts investigated can be justifiably claimed to be true for all like facts |

8. Related documents

The following is a list of related documents that are either referenced and/or have been adapted to develop the Institute's Data & Research Governance Policy V5.0, the framework and the data lifecycle management standard operating procedure.

| Document | Description |
|--|---|
| Australian Institute of Health and Welfare Data Governance Framework | describes the data governance arrangements for the custody of AIHW data |
| Cancer Institute (NSW) Act 2003 | an act relating to the constitution, objectives, and functions of the Cancer Institute NSW |
| Cancer Institute NSW Cancer Plan | sets out a coordinated and collaborative approach to cancer control for NSW, to ensure we are working together to accelerate better outcomes for people affected by cancer |
| Cancer Institute NSW Data Strategy | outlines the Institute's operations and data investments to improve cancer outcomes for NSW residents |
| Data NSW Data Governance Toolkit | a resource designed to help agencies improve their ability to govern their data |
| Five Safes Framework | a multi-faceted approach to effectively managing data in line with relevant legislative and other compliance obligations |
| NSW Information Management Framework | practical tool that outlines a shared direction for information management in the NSW public sector |
| NSW Data Governance Framework | which outlines the roles and responsibilities related to data governance across NSW Health |
| NSW Government Data Strategy | sets a collaborative, coordinate consistent and safe approach to using and sharing data and insights across government to inform decisions and actions to achieve the best possible outcomes for the people and businesses of NSW |