Guidelines for the access to and use of the Human Services Dataset

Document approval

The Guidelines for the access to and use of the Human Services Dataset has been endorsed and approved by:

|  |  |
| --- | --- |
| Jessica Stewart  Executive Director, Family and Community Services Insights Analysis and Research (FACSIAR)  Approved: 1 July 2021 |  |

Document version control

|  |  |
| --- | --- |
| Distribution: | Public |
| Document name: | Guidelines for the access to and use of the Human Services Dataset |
| Trim Reference | D20/1997906 |
| Version: | Version 3.0 – July 2021 – Changes following governance transition to FACSIAR |
| This document replaces | Version 2.0 – February 2020 – Approved by Data Custodian, endorsed by the HSDGAC – Changes to content and forms  Version 1.0 – August 2019 – Approved by Executive Director, TFM. |
| Document status: | Revised |
| Authoring unit: | FACSIAR |
| Date: | 1 July 2021 |
| Next Review Date: | 12 months from version date, or earlier if there are legislative or significant operational changes. |

Table of contents

Contents

[1 About these guidelines 4](#_Toc67476366)

[2 The Human Services Dataset 5](#_Toc67476367)

[3 Access to and use of the dataset 6](#_Toc67476369)

[3.2 Approved Purposes 6](#_Toc67476370)

[3.3 Prohibited Uses 7](#_Toc67476371)

[3.4 Community Impact and Ethics Approval 7](#_Toc67476372)

[4 Approved Analysts 8](#_Toc67476373)

[4.1 Process for Approving Analysts 8](#_Toc67476374)

[4.2 Approved Analyst Training 9](#_Toc67476375)

[5 Data and analytic requests 10](#_Toc67476377)

[6 Checking of Analytic Outputs 12](#_Toc67476378)

[6.1 Output checking process 12](#_Toc67476379)

[6.2 Five Safes Framework 12](#_Toc67476380)

[7 Contact the FACSIAR Project Team 13](#_Toc67476381)

[APPENDIX A: Data Items Inventory for the Human Services Dataset 14](#_Toc67476382)

[APPENDIX B: Request process map 15](#_Toc67476383)

[APPENDIX C: Data request form 16](#_Toc67476384)

[APPENDIX D: Analytics Proposal Form 17](#_Toc67476385)

1 About these guidelines

These guidelines provide details of the conditions and requirements for obtaining data from the Human Services Dataset (HSDS). These guidelines should be read in conjunction with the [Public Interest Direction](https://www.ipc.nsw.gov.au/sites/default/files/file_manager/PID%20under%20PPIP%20Act%20-%20Their%20Futures%20Matter.pdf)[[1]](#footnote-1) and [Health Public Interest Direction](https://www.ipc.nsw.gov.au/sites/default/files/file_manager/PID%20under%20HRIP%20Act%20-%20Their%20Futures%20Matter.pdf)[[2]](#footnote-2) (the PIDs) made by the NSW Privacy Commissioner for the Their Futures Matter (TFM) Project.

Effective oversight of the HSDS involves proactive management of security and privacy risks such as data breaches and re-identification. The FACS Insights, Analysis and Research (FACSIAR)[[3]](#footnote-3) governance approach is informed by the NSW Government’s [Data Sharing Principles](https://data.nsw.gov.au/data-sharing-principles), which are based on the Five Safes Framework[[4]](#footnote-4).

The Data Sharing Principles are:

* Safe Projects: Share data only for safe and authorised purposes.
* Safe People: Share data only with authorised users.
* Safe Settings: Share data only in a safe and secure environment to minimise the risk of unauthorised use or disclosure.
* Safe Data: Apply appropriate and proportionate protections to the data.
* Safe Output: Ensure data outputs do not identify people before any further sharing or release.

The Human Services Dataset Governance Advisory Committee (HSDGAC) has been established to provide guidance to the FACSIAR HSDS Governance and Privacy Project Team on the management of the HSDS. The Committee is chaired by FACSIAR and is represented by the following NSW Government agencies:

* NSW Department of Premier and Cabinet
* NSW Department of Communities and Justice, including NSW Bureau of Crime Statistics and Research
* NSW Department of Education
* Centre for Health Record Linkage, NSW Ministry of Health (CHeReL)
* Centre for Epidemiology and Evidence, NSW Ministry of Health
* NSW Data Analytics Centre, NSW Department of Customer Service
* Aboriginal Housing Office
* Legal Aid NSW
* NSW Police.

2 The Human Services Dataset

The Department of Communities and Justice (DCJ), in collaboration with NSW government departments and agencies, has created a cross-agency linked dataset under the TFM Project. The dataset has brought together data from four key NSW clusters including Education; Health; Stronger Communities and Customer Service.

The metadata document, providing a list of the data items included in the HSDS, is shown at **Appendix A**.

The data of the HSDS is grouped into three tiers, and access permissions for each tier are set out in the PIDs:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type of data | Description of data | Provided for | Availability for access | Who can access |
| Tier One | Unprocessed, unit-record level data provided by agencies.  Unique identifiers are applied to this data. | For linkage | **Closed**  (access by linkage agency only for an Approved Purpose) | Data Linkage Centre (NSW data)  Australian Government Linkage Agency (Commonwealth data) |
| Tier Two | De-identified, unit-record level data pertaining to individuals.  De-identification checks are applied. | For analysis, modelling, research or evaluation | **Shared**  (controlled access to approved data users for an Approved Purpose) | Data Analytics Entity  Approved Analysts |
| Tier Three | Aggregated (and, if needed, confidentialised) outputs from analysis of Tier Two data.  Does not contain identifying information. | For release | **Open**  (unrestricted access – for an Approved Purpose) | Government agencies (generally)  Research sector (generally)  NGOs (e.g. funded service providers)  General public |



3 Access to and use of the dataset

The PIDs have confined ‘Analytical Services’ to the study, analysis, modelling, research or evaluation of Tier Two Data. Only Approved Analysts and the Data Analytics Entity can provide Analytical Services for and on behalf of the FACSIAR HSDS Governance and Privacy Project Team[[5]](#footnote-5).

The PIDs permit Approved Analysts to access and analyse Tier Two data for the Approved Purposes.

There are two access points to the HSDS:

* NSW Data Analytics Centre’s (DAC) Secure Analytics Lab (SAL) - for linked NSW datasets; and
* ABS DataLab – for data linked with the Commonwealth Multi-agency Data Integration Project (MADIP) data.

The PIDs allow for linkage with Commonwealth data. The Commonwealth data will be available from the ABS DataLab. To access the Commonwealth linked data, approval is required from the HSDS Data Custodian and the MADIP Custodians.

3.1 Approved Purposes

Two threshold criteria inform considerations of whether a proposal meets the Approved Purposes:

1. Conformance to two guiding principles:
2. the proposed activity must facilitate or enable the Project Objectives, which is to ensure that effort and funding across government is focussed on interventions that will **improve long-term outcomes for Vulnerable Children or Young Persons and their families** at the earliest opportunity
3. the data will be used to **design and deliver better government services for Vulnerable Children or Young Persons and their families**.
4. Under the overarching rubric of these guiding principles, the data that is collected and used will:

* provide specific identifications of trends and gaps in government service usage and delivery;
* facilitate services that are better tailored to the needs of Vulnerable Children or Young Persons and their families both now and in the future;
* deliver clear evidence on service, support and program effectiveness and a detailed view of resource allocation and gaps;
* provide valuable information for research and planning of government supports and services; and
* enable Participating Agencies and other government agencies to meet the Project Objectives by implementing new policy and program development directed to improving outcomes for Vulnerable Children or Young Persons and their families.

3.2 Prohibited Uses

The following uses of the HSDS are not supported:

* Attempting to re-identify an individual;
* Targeting an individual to provide a service;
* Use for a compliance or enforcement purpose; and
* Use for a commercial purpose.

3.3 Community Impact and Ethics Approval

Applicants are responsible for determining if they need ethics approval for their projects. In particular, applicants are asked to consider Aboriginal impact resulting from their use of data, and to seek ethics review if relevant. If ethics review is required, approval must be obtained before the proposal will be considered by the Data Custodian.

Projects need to be conducted in line with the National Statement on Ethical Conduct in Human Research from the National Health and Medical Research Council (NHMRC) Act 1992. In Australia, Human Research Ethics Committees (HRECs) review research proposals that involve human participants to ensure that they meet ethical standards and guidelines.

A list of registered HRECs can be found at the [NHMRC website](https://www.nhmrc.gov.au/sites/default/files/documents/attachments/november_2020_list_of_human_research_ethics_committees_registered.pdf).

4 Approved Analysts

All approved analytic activities or projects that interact with Tier Two data must be performed by an Approved Analyst.

The PIDs outline obligations for Approved Analysts who use the HSDS. Approved Analysts must:

* have been approved to use Tier Two data in the HSDS for analysis, modelling, research or evaluation
* be under a contractual obligation to comply with the *Personal Information Protection Act 1998* (PPIP Act) and the *Health Records and Information Privacy Act 2002 (*HRIP Act), to the extent modified by the PIDs
* undertake privacy verification checks (‘Information Protection Gates’):
* before commencing analysis of the requested Tier Two data, to ensure that the data does not contain any personal information;
* before analytics output, results or findings are externally released, to ensure that only de-identified information is disclosed to third parties (including the requesting agency), and individuals cannot be re-identified in any externally released data
* adopt best practice privacy, security and de-identification practices to minimise disclosure risk, such as the [De-Identification Decision Making Framework](https://publications.csiro.au/rpr/download?pid=csiro:EP173122&dsid=DS3)[[6]](#footnote-6) and the Five Safes Framework.
* keep Tier Two data confidential, secure and protected from unauthorised use and disclosure as well as loss.

Approved Analysts must not:

* use Tier Two or Tier Three data to re-identify, or cause any data to be re-identified (except as part of the Information Protection Gates)
* use any data that is derived from the Human Services Dataset other than for the Approved Purposes
* use any data from the Human Services Dataset to target an individual person.

Approved Analysts may also be subject to audits to ensure that any access and use of Tier Two data is consistent with the privacy and security requirements outlined in the PIDs.

4.1 Process for Approving Analysts

An Approved Analyst can be engaged to access the HSDS, provided that:

* the analytics being undertaken has been approved by the Data Custodian; and,
* access to data is on a need-to-know basis for the duration of the approved project period.

To be eligible for approval, the analyst must:

* demonstrate evidence of technical ability in data or statistical analysis
* sign a legally binding data privacy and confidentiality agreement
* confirm that they have read and understood DCJ’s data breach policy
* have a current Working with Children Check (WWCC)
* have a current National Police Check (Criminal Records Check)
* be willing and available to undertake training on appropriate use of data, and privacy, confidentiality and security obligations

Applicants who may not have access to staff with the relevant skill set may speak to the Project Team about seeking the services of a FACSIAR Approved Analyst or alternatives.

The Project Team generally considers applications for individuals to become Approved Analysts as part of the analytic proposal – please see Analytics Proposal Form at **Appendix D**.

Applications to become Approved Analysts are collated and put to the HSDGAC for consideration and endorsement monthly (first week of each month). The HSDGAC may seek further information from the applicant or provide feedback.

Following endorsement for approval by the HSDGAC, the Data Custodian will be informed of the recommendation and make a final decision.

Approved Analysts can be approved to work on more than one project. If an Approved Analyst is required to assist on subsequent projects, they must receive approval from the Data Custodian for the subsequent projects before they can commence accessing relevant data.

It is the responsibility of Approved Analysts to notify the Project Team of any updates to personal details or change in circumstances, e.g. new expiry date and number of WWCC records.

4.2 Approved Analyst Training

Approved Analysts must undertake training before accessing data. Approved Analysts receive training on:

* privacy obligations (such as accessing data on a need-to-know basis and reporting and responding to data breaches)
* security obligations (including ICT and physical security)
* conditions of data use associated with the HSDS.

5 Data and analytic requests

Any analytic activities that makes use of the Tier Two data contained within the HSDS must adhere to the requirements specified within the PIDs.

There are two broad types of requests for Tier Two data:

1. Data requests
2. Proposals for in-depth analytics.

The processes involved in both types of requests are explained below, and shown in the flowchart at **Appendix B**.

1. **Data requests** typically involve providing readily available information from Tier Three data for reporting purposes, includes the LGA and district [insights](https://www.theirfuturesmatter.nsw.gov.au/investment-approach/district-and-lga-insights).

Data requests must be made using the form at **Appendix C**.

1. **Proposals for in-depth analytics** will typically involve the extraction and assembly of customised data from Tier Two data for analytics. In some cases, the proposal may involve linkage with additional Tier One data provided by the requesting party (enduring or one-off linkage).

Access to Tier Two data will require an Approved Analyst to access the requested data via the NSW DAC’s SAL platform or the ABS DataLab. Applicants may provide Approved Analysts for the analytics project, nominate staff for approval, or they may seek the services of a FACSIAR Approved Analyst. Any costs associated with accessing the HSDS will be borne by the requesting party.

Careful consideration should be given as to what data is being requested, as any changes occurring after the data extraction and preparation activities have commenced may incur extra fees and could extend delivery timeframes.

Analytics proposals must be made using the form at **Appendix D**.

Data requests and proposals for in-depth analytics must demonstrate compliance with the conditions and requirements of the PIDs, in order to ensure that sensitive data will not be used or disclosed in a manner that could identify or target individuals, or compromise data security. In addition, all information obtained from the HSDS must be used in accordance with the Approved Purposes of the PIDs.

Analyst requests seeking to use postcode data must include justification for need and details of intended use. The Project Team will assess such requests on its merits, separate to the proposal itself – the postcodes must substantively support the analysis.

FACSIAR will work with the requesting party to ensure that the specifications of the request are appropriate and relevant requirements in the PIDs are met. It is recommended that applicants contact FACSIAR prior to making any request for data. FACSIAR will inform the applicant as soon as possible in the event that a request requires further clarification.

5.1 Approval Process for Analytics Proposals

The application process can be summarised as follows:

1. **Proposals outline project and any nominated analysts -** Application forms (Appendix C or D) should be completed outlining the project and any nominated analysts. Forms should be sent to the FACSIAR HSDS Governance and Privacy Project Team <[HSDS@facs.nsw.gov.au](mailto:HSDS@facs.nsw.gov.au)>.
2. **Project team intake, triage and assessment of proposal -** The Project team will assess the proposal and engage with the applicant to discuss the project, provide feedback and seek further information.
3. **Human Services Dataset Governance Advisory Committee (HSDGAC) endorsement sought -** All proposals for in-depth analytics which appear to align with the Approved Purposes are considered by the HSDGAC. In assessing any analytics proposal, consideration will be given to:
   1. whether the project aligned with the Approved Purposes
   2. the availability of suitable data
   3. timeframes and resourcing
   4. proposed Analysts
   5. technical feasibility of the project - this may involve consultation with the Data Linkage Centre (NSW CHeReL), Australian Government Linkage Agency (ABS) and Data Analytics Entity (NSW DAC) as appropriate.
4. **Data Custodian approval sought -** Where the Committee is satisfied that the proposal meets the requirements under the PIDs, the Data Custodian will be informed of recommendations for approval from the Committee and make a final decision.
5. **Data access -** Access to the data is via the NSW SAL or ABS Data Lab for analysis of data by Approved Analysts.
6. **Information Protection Gates -** A series of output checks for identifiable information, publication risks and data interpretation are undertaken before data release.

The timeframe from request to data release can vary widely, depending on the complexity of the project and availability of the data.

A flow chart illustrating the approval process in **Appendix B.**

6 Checking of Analytic Outputs

6.1 Output checking process

To minimise disclosure risk, there are analytic output checkpoints that must be passed prior to external release as Tier 3 data (including any release to the requesting agency or party):

1. All analytical outputs must be aggregated, and where necessary, confidentialised, before it can be taken out of the DAC’s SAL or ABS DataLab analytic environment. As part of this process, the Approved Analyst must check the outputs to ensure these do not contain any personal information and that individuals cannot be re-identified.
2. Aggregated outputs are checked by the DAC and FACSIAR (as the governance lead) to ensure that these outputs do not contain personal information and individuals cannot be re-identified.
3. Authorised Reviewers from relevant agencies review the aggregated outputs for disclosure and publication risks. Members of the HSDGAC are involved if they are also the nominated Authorised Reviewer for their agency.
4. The Data Custodian’s approval may be sought for release of aggregated outputs that relate to significant projects or sensitive topics.

Use of linked Commonwealth data will also need to go through NSW and ABS disclosure vetting processes prior to release.

6.2 Five Safes Framework

In applying the Five Safes, the following steps can be taken to manage and minimise disclosure risk:

|  |  |
| --- | --- |
| Safe Project | Data is used for an approved project that serves the Approved Purposes, with a valid analytics aim, for public benefit, and not for compliance or regulatory purposes (such as targeting a specific individual). This is underpinned by multiple layers of oversight and clear lines of accountability. |
| Safe People | The Approved Analysts are appropriately authorised to access and use the data. (see Process for Approving Analysts in section 4). |
| Safe Settings | The HSDS can be accessed by Approved Analysts in secure analytic environments, via the NSW DAC’s SAL or the ABS DataLab. |
| Safe Data | Personal identifiers are removed from project data and controls are in place to prevent misuse and re-identification. This could involve supervisors of Approved Analysts manually checking Tier Two data for identifying information, before it is released to Approved Analysts for analytics work. |
| Safe Outputs | Outputs are produced at an aggregate level and do not contain unit-level data.  Supervisors should check outputs for disclosure risk, including inadvertent disclosure, before outputs are externally released or made public. Steps could include:   * aggregating outputs to a minimum cell size of ten * suppressing the cell so that the output at risk is not disclosed * expanding the output to include additional data, therefore increasing the sample size * considering any additional contextual information that would place the output at greater risk of disclosure * adopting the ABS’s Statistical Area Levels (for location-based statistics) or Confidentialisation (where a high disclosure risk exists due to uncommon or unique traits)   Additional checks are undertaken by the Data Analytics Entity (DAC, ABS) to ensure individuals cannot be re-identified, and relevant agencies for disclosure and publication risks. |

7 Contact the FACSIAR Project Team

Proposals or queries can be made to the Project Team: [HSDS@facs.nsw.gov.au](mailto:HSDS@facs.nsw.gov.au)

APPENDIX A: Metadata document for the Human Services Dataset



This metadata document provides a list of the data items included in the HSDS. This version contains data collected up to 30 June 2017.

APPENDIX B: Request process map



APPENDIX C: Data request form



**This document contains the form for Data Requests using the Human Services Dataset. This request form is for the extraction of readily available information from Tier Three data for reporting purposes.**

**Before filling out this form, please refer to ‘Guidelines for access to and use of the Human Services Dataset’ as well as** [**Public Interest Direction**](https://www.ipc.nsw.gov.au/sites/default/files/file_manager/PID%20under%20PPIP%20Act%20-%20Their%20Futures%20Matter.pdf) **and** [**Health Public Interest Direction**](https://www.ipc.nsw.gov.au/sites/default/files/file_manager/PID%20under%20HRIP%20Act%20-%20Their%20Futures%20Matter.pdf) **made by the NSW Privacy Commissioner.**

**Submissions should be made to:** [HSDS@facs.nsw.gov.au](mailto:HSDS@facs.nsw.gov.au)

|  |
| --- |
| **Data Request Details** |
| **Name:** |
| **Position:** |
| **Agency, employer or organisation:** |
| **Purpose of request:** |
| **Details of data being requested:** |
| **Audience for the final product and intentions to publish:** |

|  |
| --- |
| **Office Use** |
| **Meets criteria for Data Request (Yes/No)** |
| Reasons if no: |
| **Extracted in-house or by external analyst:** |

APPENDIX D: Analytics Proposal Form



**This document contains the form for Analytics Proposals using the Human Services Dataset. This request form is for the extraction and assembly of customised data from Tier Two data for analytics.**

**Before filling out this form, please refer to ‘Guidelines for access to and use of the Human Services Dataset’ as well as** [**Public Interest Direction**](https://www.ipc.nsw.gov.au/sites/default/files/file_manager/PID%20under%20PPIP%20Act%20-%20Their%20Futures%20Matter.pdf) **and** [**Health Public Interest Direction**](https://www.ipc.nsw.gov.au/sites/default/files/file_manager/PID%20under%20HRIP%20Act%20-%20Their%20Futures%20Matter.pdf) **made by the NSW Privacy Commissioner.**

**For information on specific data available, see Appendix A for the metadata document which is updated annually when the dataset is refreshed.**

**Submissions should be made to:** [**HSDS@facs.nsw.gov.au**](mailto:HSDS@facs.nsw.gov.au)

|  |
| --- |
| **Analytics Proposal Details** |
| **Name:** |
| **Position:** |
| **Agency, employer or organisation:** |
| **Purpose of proposal (including how the proposal meets the criteria for the Approved Purposes under the PID or Health PID):** |
| **Details of the use of the Human Services Dataset (include specific items of interest, study cohorts of interest, dates of interest and age ranges etc):**  **If you are interested in using postcode data, please provide reason(s) why and how it would support your research or analysis:** |
| **Research methods to be used, and if applicable, any hypothesis testing:** |
| **Name, position and agency/employer/organisation of persons undertaking the research:** |
| **Name, position, experience and agency/employer/organisation of any proposed analysts requiring approval to use the Human Services Dataset, including:**   * **whether they have Working with Children Checks, National Police Checks or approval and intent to obtain these** * **whether they have demonstrated experience in data or statistical analysis** |
| **Audience for the final product and intentions to publish:** |
| **Timeframes and resources required to undertake the research:** |
| **Consideration of Aboriginal impact:** |
| **Are you providing data for linkage with the Human Services Dataset? (please indicate Yes or No)**  **If Yes, is this data linkage a one-off arrangement or an ongoing arrangement?** |

|  |
| --- |
| **Office Use** |
| **Supported by the HSDGAC (Yes/No/Further work required)** |
| Reasons if no/further work required: |
| **Approved by Data Custodian (Yes/No)** |
| Reasons if no: |

1. Made under section 41(1) of the *Privacy and Personal Information Protection Act 1998* (NSW). [↑](#footnote-ref-1)
2. Made under section 62(1) of the *Health Records and Information Privacy Act 2002* (NSW). [↑](#footnote-ref-2)
3. FACS Insights, Analysis and Research, within the NSW Department of Communities and Justice. [↑](#footnote-ref-3)
4. The Five Safes Framework has also been adopted by the Australian Government, the UK Government, the New Zealand Government, the South Australian Government, AIHW and ABS, to manage data sharing and disclosure risks. [↑](#footnote-ref-4)
5. The PIDs refer to the Project Team as the team established by the NSW Government to coordinate and lead the TFM Project. This Project Team is led by FACSIAR. Further details on the TFM Project are described in paragraphs 4 to 6 of the PIDs. [↑](#footnote-ref-5)
6. Developed by the Office of the Australian Information Commissioner and CSIRO’s Data61. [↑](#footnote-ref-6)