



Research Data Collection
Data Management Plan

<Dataset Title>

<Authors/Submitters>

<Logo, if required>

This document records the whole process of the data collections, including pre-preparation data, data ingest, catalogue registry, license, data quality assurance, roles specified, the checklist of publishing data, issues identified, and the final status of collections on the National Computational Infrastructure.

Revision history

Date	Version	Description	Author(s)
2014/03/04	0.1b	Draft release of data management plan incorporating NCI and GA	Jingbo Wang, Irina Bastrakova (GA)

Release history

Date	Version	Status	Audience	Approval
2014/03/14	0.1c	Limited released consultation	Core Partner Data managers	Ben Evans
2014/05/30	0.1d	Feedback from users		Jingbo Wang/Irina Bastrakova (GA)
2014/07/15	0.1d	The online form is available https://datamgt.nci.org.au		Jingbo Wang
2015/03/20	0.2a,b,c	Updated by Lesley Wyborn/Jingbo Wang/Fabiana Santana	Data Managers	Ben Evans/Irina Bastrakova (GA)
2015/06/18	0.2d,e,f	Comments from Lesley Wyborn, Mingfang Wu, Julia Martin, Aaron Sedgmen, Helen Glaves (British Geological Survey) /Irina Bastrakova	NCI research data collections technical advisory group	NCI research data collections technical advisory group
2015/08/31	1.0	The online form is updated		Jingbo Wang

1. Data Collection Overview

Data Description	
Hierarchy Level*	<p>Please specify whether it is a collection or series.</p> <p>A series is a subset of one collection. If a collection includes several subsets, the top level will be called a collection, and each subset will be called a series.</p> <p>A DMP is needed at both collection and series level.</p>
Dataset Title*	Please specify name/title of the data collection or series.
NCI Data Collection Code*	NCI will provide the project code.
Dataset abstract*	<p>Provide an adequate summary of the content including:</p> <ul style="list-style-type: none"> ● The method that has been used to generate the data. ● If the data will contribute to a pre-existing collection at NCI? ● Whether the data are a unique, unrepeatable measurement of the environment? ● If the data has a broad geographical or temporal extent that makes it useful to others? ● If the data includes changes in the processing methods, new standards or set any precedents? ● If the data supports current projects or new trends in science? ● Are the data likely to be cited or referenced in a publication?
Dataset Ownership*	Please state which organization(s) own(s) the data.

	<p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p>
Dataset Custodianship*	<p>Please state which organization(s) is (are) custodian(s) of the data.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p>
Does the dataset have a Persistent Identifier (PID), such as Digital Objective Identifier (DOI)*	<p>If yes, please specify the exact citation.</p> <p>If not, will the data provider mint a DOI and if so when in the publication workflow? Or will the data provider agree for NCI to mint the DOI? (Please refer to NCI's data policy on how DOIs will be minted).</p>

Other Dataset Identifiers*	Please provide any other identifiers that the dataset is known by.
Cross Reference	Please provide any published journal articles data citation using the data.
Source of Data*	Please state where the data come from if it is provided by the third party. For example, who collected the data.
Dataset Provenance*	Please provide the derivation history of the data or versioning since the data has been received from the third party. If relevant, include PIDs of other datasets it is linked to (e.g., derived from, a subset of, transformed, etc.). If known, include PIDs of software, services or instruments, used to produce the dataset.
Dataset Purpose*	Please provide a description of why the dataset was developed and how it was intended that it would be used.
Dataset Status*	Status of the dataset. Please select from the pick list: completed historicalArchival obsolete ongoing planned required underdevelopment
Alternative Title*	Please provide other names the dataset is known by, including any acronyms.
Topic Category	Please select from the pick list:

	<p>farming</p> <p>biota</p> <p>boundries</p> <p>climatologyMeteorologyAtmosphere</p> <p>economy</p> <p>elevation</p> <p>environment</p> <p>geoscientificInformation</p> <p>health</p> <p>imageryBaseMapsEarthCover</p> <p>intelligenceMilitary</p> <p>inlandWaters</p> <p>location</p> <p>oceans</p> <p>planningCadastre</p> <p>society</p> <p>structure</p> <p>transportation</p> <p>utilitiesCommunication</p>
Activities*	If the dataset was collected as part of a specific grant program, or project, etc., please specify the name.
Roles/Contacts from the Provider Organization	
Principle Investigator	Please provide the following information of the person who is the

responsible for the data*	Principle Investigator (i.e., in charge) of the development of the data. Name: Organization name: Position name: Telephone number: Delivery Point: City: Postal Code: Country: Email:
Content Expert *	Please provide the following information of the person who is responsible for the content and maintenance of the content in the data Name: Organization name: Position name: Telephone number: Delivery Point: City: Postal Code: Country: Email:
Contact for Media Communication*	Please provide the following information of the media person who is responsible for media communications such as writing reports, making videos, broadcasting, press articles, etc.

	<p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p>
Data transfer/Ingestion	<p>Please provide the following information of the person who is responsible for the transfer of the dataset to NCI and its ingestion at NCI.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p>
Quality Assurance/ Quality Control	<p>Please provide the following information of the person who can provide information on the Quality Assurance (QA)/Quality Control (QC) procedures that have been applied to the dataset prior to its</p>

	<p>transfer to NCI.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p> <p>Please note: it is the data provider's responsibility to undertake QA/QC of the data and metadata prior to transfer.</p>
Publishing	<p>Please provide the following information of the person in the organization providing the data who is authorized for data publishing, defining how the data should be cited and verifying that the licensing and other legal attributes are in accordance with the policies of the organization.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p>

	<p>Email:</p> <p>Please note: Data publishing is a joint effort by the data provider and NCI and this role should be different to any of the other roles described in the DMP.</p>
Maintenance	<p>Please provide the following information of the person who will provide information on any changes/updates to the dataset and/or its metadata.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p>
User Support	<p>Please provide the following information of the person who can provide user support on the content and other technical issues specific to the dataset.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p>

	Postal Code: Country: Email:
Data Service Development	Please provide the following information of the person who will provide technical support to develop services, such as web portal, data registry and query system, etc., at NCI. Name: Organization name: Position name: Telephone number: Delivery Point: City: Postal Code: Country: Email:
Data Service Operation	Please provide the following information of the person who will assist in maintaining and keeping the data services running on a regular basis. Name: Organization name: Position name: Telephone number: Delivery Point: City: Postal Code:

	<p>Country:</p> <p>Email:</p>
Data Life-cycle Control	<p>Please provide the following information of the person who makes higher level data life cycle decisions.</p> <p>Name:</p> <p>Organization name:</p> <p>Position name:</p> <p>Telephone number:</p> <p>Delivery Point:</p> <p>City:</p> <p>Postal Code:</p> <p>Country:</p> <p>Email:</p>
Constraints on the Data	
Limitations on using the data*	<p>Please provide information on the fitness for purpose including any limitations on use for the dataset and specify any special requirements such as 'Not suitable for navigation purposes'; 'Research and data collection in indigenous communities' and 'Ethnographic data'</p>
License Type For Access Constraints *	<p>Please provide information on additional constraints such copyright, patent, trademark, etc.</p> <p>copyright</p> <p>patent</p> <p>patentPending</p>

	<p>trademark</p> <p>license</p> <p>intellectualPropertyRights</p> <p>restricted</p> <p>otherRestrictions</p>
Copy of Access Constraints*	Please provide physical copies of the license and other information that pertain to constraining access to the data.
License Type For Constraints on Using*	<p>Please specify the license that applies to the data from the pick list:</p> <p>copyright</p> <p>patent</p> <p>patentPending</p> <p>trademark</p> <p>license</p> <p>intellectualPropertyRights</p> <p>restricted</p> <p>otherRestrictions</p>
Copy of Constraints on Using*	<p>Please provide physical copies of the license and other information that pertain to constraining use the data.</p> <p>Please note here two type of constraints: access and use. "Access" means that the user can browse the data, read the metadata information, view the images, but not work with the data. "Use" means that the user can use the data to generate some results or work with the data for his projects. In many cases, the constraints of accessing and using the data can be the same.</p>

Security Classification*	<p>Please select the level of restriction from the pick list:</p> <ul style="list-style-type: none"> unclassified restricted confidential secret topSecret
---------------------------------	---

2. Detailed Information on the Dataset

Extent	
Spatial Extent*	<p>Please provide spatial extent for the dataset (lat/long) in the format of [-90 90]/[-180 180] and datum/projection.</p> <p>West Longitude : East Longitude : South Latitude : North Latitude :</p>
Temporal Extent*	<p>Please provide date/time (e.g. start date/ end date) when the dataset was acquired. The end date may be the expected date of the project completion in the case of an ongoing project.</p> <p>Begin Position: (YYYY-mm-dd) End Position :</p> <p>“now” can be used for “End Position” if the data collection is an ongoing process.</p>
Formats	

Format Name	Please provide the name of the data format (e.g. NetCDF)
Format version*	Please provide the version of the data format (e.g. 1.0, 2.0, CF) The data needs to be in an appropriate format for ingestion. Please specify if any specialist software is required to access the data, and/or transform the data to an appropriate format.
Usage	
Tools*	Please specify what tools are needed to access, use, and display the data.
Software*	Please specify what software (including which version) was used to generate the data.
Algorithms/Libraries*	Please specify what algorithms/libraries were used to generate this data.
Flexibility*	Please specify whether multiple users, on multiple systems, can use the data for multiple purposes. For example, does the data need to be converted to another format for users from different communities, or can it be used as is, on different operating platforms? Does the data have to be scaled up or down for different use cases?

3. Workflow

Pre-publishing Preparation	
Ingest Process*	NCI strongly recommends that the network is used to transfer the data. Please specify the tools/environment requirement to make this happen.

Quality Assurance (QA)/ Quality Control (QC) Processes*	<p>All the data and metadata should be finalized before transferring the data. The data provider needs to supply documentation on the QA/QC processes that were applied to the data and metadata prior to transfer to NCI. The QA/QC documentation should include information on how the data and metadata were cleaned and validated, and including how the integrity of data structures or file formats were verified. The QA/QC process should confirm that the data are in a suitable condition to be added to the collections at NCI, i.e. readable, undamaged, uncorrupted, and robust enough to be handled by high performance systems at NCI</p>
Maintenance Process*	<p>Please select the maintenance frequency from the pick list:</p> <ul style="list-style-type: none"> continual daily weekly fornightly monthly quarterly biannually asNeeded irregular notPlanned unknown
Backup and Recovery plan*	<p>Please specify whether NCI needs to host an archival copy of the data collection. Specify if the data are required to be kept in secure conditions and how and where this is to happen.</p>
Publishing Process	

Services URL*	Please provide the URL of the data services.
Services Description*	<p>Please select the preferred service type (including service interface and technology) from the following pick list:</p> <p>ESRI: ArcIMS</p> <p>ESRI: ArcGIS</p> <p>W3C: HTML: INFORMATION</p> <p>W3C: HTML: DOWNLOAD</p> <p>OGC: Web Feature Services</p> <p>OGC: Web Mapping Services</p> <p>OGC: Web Processing Services</p> <p>OGC: Catalogue Services for the Web</p> <p>OGC: Web Coverage Services</p> <p>OGC: Sensor Observation Services</p> <p>OPeNDAP: OPeNDAP</p> <p>Please also specify the version and standards of the service type requested.</p> <p>The URL should be provided if available.</p> <p>Please note: NCI provides generic and latest version of data services, provided there are no special requirements. However, if the data provider needs a customized version or other data services that are not immediately available, please send the request to datacollections@nci.org.au so that NCI can take further action based on the available resources.</p>
Local file path*	NCI will populate the local storage file path of the data on /g/data at NCI.
Catalogue Registry*	NCI will provide the web link of the catalogue registry for the data. The cross-reference and synchronization of the metadata catalogue are required if both the data provider and NCI hosts dual copies of the metadata catalogue.

	<p>However, if the data provider hosts the metadata catalogue relevant to this data and that catalogue is not maintained at NCI, please provide the relevant links.</p> <p>Please refer to NCI's data policy (will be available in mid-2015) or discuss with NCI's data collection team datacollections@nci.org.au for more details.</p>
Federated Catalogues*	<p>Please list all catalogues that includes information on your data other than NCI's catalogues, such as ANDS Research Data Australia (http://researchdata.ands.org.au/), FIND (http://find.ga.gov.au/) data.gov.au (http://data.gov.au), or CMIP5 (http://esgf.llnl.gov)</p>
Search Words/ keywords*	<p>We require the ABS Field Research List as keywords. http://www.abs.gov.au/AUSSTATS/abs@.nsf/0/6BB427AB9696C225CA2574180004463E?opendocument</p> <p>If additional keywords are needed please choose words from a governed vocabulary that is in a controlled namespace (and provide links).</p>
Website*	<p>Please provide details of any known website which accesses this data. For example, the VHIRL accesses natural hazards data, VGL accesses geophysical data, BCCVL accesses climate model data, and CWSlab accesses ACCESS data.</p>
Readme and Additional Documentation * (eg., Data Description and Data specification)	<p>NCI will provide the link pointing to the readme file served within NCI's OpenDAP data services.</p> <p>Please provide a link to any additional document that describes the dataset and its specifications. An example of such a document is: http://www.ga.gov.au/_data/assets/pdf_file/0018/11871/Australian-Reflectance-Grid-ARG25-Product-Information-Beta-Release-PDF-2.pdf</p>
Feedback	<p>Please provide feedback or comments to datacollections@nci.org.au</p>
Retention of Data*	<p>Please specify the time frame the data is to be maintained at NCI.</p>

Others

How-to Guide*	Provide a high level document stating how to access and use the data in the home directory. If possible provide an example that demonstrates how to use the data, or a paper about how the data has been analyzed. This part aims to make the user friendly.
----------------------	--

NOTE: A collection description file and details of the data license needs to accompany all data collections. These two files should be saved in the home directory under each collection allocation. The templates for the collection description file and license file are available at NCI documentation repository (<https://datamgt.nci.org.au>).

4. Time Frame for ingestion and publication of the data

Data Ingest	Please provide the time frame when the data will be available for ingestion onto a staging area on the NCI file system.
Metadata catalogue	Please provide the time frame when the metadata catalogue entries will be finalised.
Data and Metadata quality assurance after ingestion*	Please provide the time frame you think will be required for QA/QC'ing the data and metadata catalogue after the ingestion is finalized at NCI.

Please note: NCI's data policy is in progress. Please send email to the help desk datacollections@nf.nci.org.au, should you have any questions.

Source Documents used to develop the NCI Data Management Plans:

1. Australia National Data Service <http://ands.org.au>
2. Bastrakova, I.V., Ardlie, N., Regan, J. May 2013. Geoscience Australia Community Metadata Profile of ISO 19115:2005, Version 1.0 http://www.ga.gov.au/corporate_data/76066/Rec2013_017.pdf
3. British Atmospheric Data Centre <http://badc.nerc.ac.uk/data>
4. Nature Environment Research Council <http://www.nerc.ac.uk/research/sites/data/policy.asp>
5. Procedures and Guidelines for the Management of Research Data and Records, prepared by Information Management, University of Melbourne Library. http://researchdata.unimelb.edu.au/documents/2011-06_Procedures_for_the_management_of_RRD2011_v10.pdf
6. Research data management plan, University of Oxford <https://www.admin.ox.ac.uk/rdm/dmp/plans/>
7. Digital Curation Centre data management plan examples and check-lists
<http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples>
http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP_Checklist_2013.pdf
8. Data management plan, University of Western Australia Information services November 2012.
9. Framework for creating a data management plan, University of Michigan.
<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html>