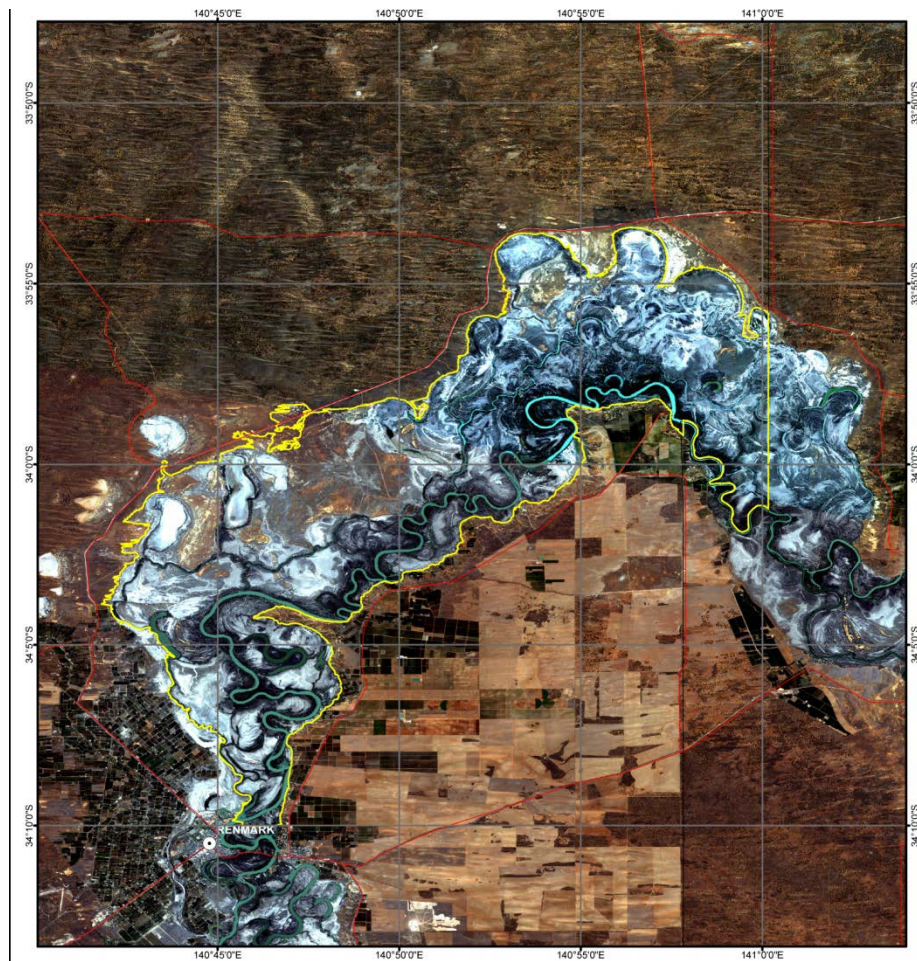




National Environmental Science Programme

Data Management and Accessibility Guidelines



VERSION CONTROL REVISION HISTORY			
Version	Date revised	Reviewed by (Name, Position)	Comment (review/amendment type)
V1.0	11/12/2014	Dave Johnson, A/g Assistant Secretary, ERIN Science and Monitoring Branch	First draft based on NERP Data and Accessibility Guidelines
V2.0	09/03/2016	Dave Johnson, Director, Science Partnerships Section	Revised draft of NESP Data Management and Accessibility Guidelines
V3.0	25/01/2017	Naomie Johnson, A/g Director, Science Partnerships Section	Final NESP Data Management and Accessibility Guidelines

Cover image: Riverland map produced by Environmental Resources Information Network, Australian Government Department of the Environment and Energy

NESP Data Management and Accessibility Guidelines v3.0
 National Environmental Science Programme
 Australian Government Department of the Environment and Energy
 Email: research@environment.gov.au

National Environmental Science Programme

Citing this publication



Department of the Environment and Energy 2017, National Environmental Science Programme Data Management and Accessibility Guidelines v3.0, Australian Government Department of the Environment and Energy, Canberra, Australia, pp. 15.

With the exception of logos or where otherwise noted, this report is licensed under the Creative Commons Australia Attribution 4.0 Licence. The terms and conditions of the licence are at: <http://creativecommons.org/licenses/by/4.0/deed.en>

Contents

Introduction	4
Purpose	4
Scope	4
Definitions	5
Planning for discoverable, accessible and re-usable data	6
Data management guidelines	6
Licence of Research Outputs	6
Publication of Research Outputs	7
Exceptions and requirements regarding the open access policy	9
Peer-Reviewed Publications	10
Intellectual Property	10
Communications related to the publication of research products	10
Compliance with the Data Management and Accessibility Guidelines	10
Appendix A: Best practice publication and management of grey literature.....	12
Appendix B: Peer-reviewed publications.....	13
Peer reviewed publications	13
Understanding pre-print, post-print and re-prints: Which should be made publicly available?	13
Tools for understanding journal copyright arrangements	13
Appendix C: Research Product Submission Form (copy).....	15

Introduction

Government agencies and research communities have identified the need to promote open access to public sector and publicly funded information. The Australian Government's position is that information funded by the Government is a national resource that should be managed for public purposes. Open access to Government funded information is the default position of the Department with exception only for privacy, security or confidentiality reasons.

These Guidelines are consistent with the Prime Minister's announcement of the Australian Government Public Data Policy Statement on 7 December 2015, and with national and international open access principles and practices which are underpinned by legislative instruments such as the *Australian Information Commissioner Act 2010*, *Freedom of Information Amendment Reform Act 2010* and the *Public Governance, Performance and Accountability Act 2013*. The Department of the Environment Information Strategy 2013-2017 states:

'Open access to information is the default position of the Department, with exception only if required for privacy, security or confidentiality reasons.'

The Australian Government Digital Continuity Principle 4 encapsulates effective open access where digital information is discoverable, accessible and useable.

'Digital information is discoverable when it can be easily found. It is accessible when it can be easily retrieved and read in context and it is usable when it can be easily evaluated or understood, edited, updated, shared and reused as appropriate by those who need it' ([Digital Continuity Principles, National Archives of Australia](#)).

Providing discoverable, accessible and re-usable data and information products derived under the National Environmental Science Programme (NESP) will provide up-to-date, high quality data and information to decision-makers, environmental managers, other scientists and the general public. This will increase the capacity of the programme to support a more collaborative, informed approach to managing Australia's environment.

Purpose

This document provides guidance on the expectations of the Department of the Environment and Energy (Department) to ensure outputs from the NESP are publicly and freely accessible and available on the internet, for use by all persons, as required by the NESP Guidelines. While it is expected that all non-sensitive NESP research products are openly and freely available on the internet, these guidelines are not intended to prescribe specific standards or methods of making data and information openly available. But rather, they are intended to provide more detail on expectations and best practice procedures to identify solutions to achieve discoverable, accessible and reusable NESP research products. We recognise that discipline-specific standards of data management apply and encourage NESP researchers to apply these standards wherever possible.

This document should be read in conjunction with the NESP Guidelines and requirements of NESP funding agreements.

Scope

These Guidelines apply to all research products generated from NESP-funded research. A broad range of research products are expected to be generated throughout the life of the program and may include the following:

- raw data sets including spatial data
- analysis and data products such as GIS-derived maps

- publications including scientific papers, reviews, books, book chapters
- grey literature including fact sheets, project profiles and technical reports
- images, maps, photos, videos, animations
- models and other tools (e.g. Decision Support Tools) such as software created by the research process - including value added components developed for off the shelf or open source software.
- websites
- mobile or tablet apps
- unspecified emerging technology

Definitions

In this document, except where otherwise expressed, the following definitions are applicable:

Attribution	Means giving appropriate credit, providing a link to the licence, and indicating if changes were made, but not in any way that suggests the licensor endorses you or your use. (Source: AusGOAL Creative Commons Attribution v4.0 International, 8 December 2014)
Data	Individual pieces of information
Embargo	In the context of this document, refers to the period of time before which a journal article or data sets can be made publicly available on the internet. Please note that this is different from a media embargo.
Exceptions to the open access policy	Data or other research products for which open access is not desirable. This includes information that may be culturally, environmentally or socially sensitive. Exceptions must be reported to the Department in prescribed milestone reporting.
Grey literature	Literature produced and disseminated outside of commercial publishing. In the NESP context, this includes fact sheets, project profiles and reports
Metadata	Contextual information associated with research outputs that supports: <ul style="list-style-type: none"> • discovery • assessment • access • re-use • verification and provenance • integration, synthesis and aggregation • curation and preservation
Openly available	Refers to the making of information available at no cost under licensing terms and in formats that allow users to re-purpose the information from its original form. This is consistent with the Australian Government Principles on open public sector information developed by the Office of the Australian Information Commissioner.
Open access	
Open format	A specification for storing and manipulating content that is usually maintained by a standards organisation. In contrast, a proprietary format is usually maintained by a company, with a view to exploiting the format by incorporating it into other products they sell, such as software. Open formats are critical to the

effectiveness of the 'open access' concept. Information and data published using an open format ensures that users, regardless of their operating system or platform will be able to access information (Source: [AusGOAL](#), 3 March 2014)

Publicly available

Means placed on an internet site which is accessible to the public and discoverable by internet search engines, such as Google Scholar

Research products

For the purposes of this report, research products include all research outputs and existing material incorporated into research outputs including publications, reports, data, software, models, algorithms, software metadata, models and other tools, websites, mobile or tablet applications, and unspecified emerging technology.

Planning for discoverable, accessible and re-usable data

Data management planning is essential to achieve successful delivery of open access research. A data and information management plan is required for each NESP project before the project commences. Ideally a data management plan should consider the following:

- description of the data to be collected, methods and procedures applied, and timing of collection
- a commitment to the publication of data and a date by which this will occur
- the expected repository or other location where the data will be published and available to the public
- description of the metadata standards that will be used
- a description of information products likely to be generated and how these will be made openly accessible
- the details of a contact person that is responsible for ensuring the data management plan is published on the internet, implemented, and kept up to date. The Department should be advised as soon as practicable of changes to this contact.

Data management plans need to be supported by infrastructure, such as:

- allocation of resources for data management – from the initial data capture through to ongoing delivery and curation
- IT Infrastructure– the hardware, software and other facilities which underpin data-related activities
- support services– resources allocated to support implementation of data management plans
- metadata management – so that data records can be used for both internal and external purposes.

Data management guidelines

Researchers are required to make all NESP research outputs publicly available on websites with a persistent and enduring link. Researchers should be guided by their applicable NESP Hub Long Form Funding Agreement, particularly [Sections 11.7 Licence of Research Outputs](#), [11.8 Publication of Research Outputs](#) and [11.9 Peer-Reviewed Publications](#).

Licence of Research Outputs

NESP hub funding agreements require all research outputs to be made publicly available under the latest [Creative Commons framework](#) (Creative Commons Version 4.0 International as at the date of this publication) using a [Creative Commons Attribution licence \(CC BY4.0 at the date of publication\)](#). This

should allow for research outputs to be deposited into an appropriate subject and/or institutional repository.

The Office of the Australian Information Commissioner and the Council of Australian University Librarians endorsed the Australian Government Open Access and Licensing Framework ([AusGOAL](#)) framework for the licensing of research data. AusGOAL establishes a common approach to data licensing across research and government and is consistent with the Department's Information Licensing Policy.

Publication of Research Outputs

NESP hub funding agreements require that hubs ensure all research outputs are publically available for no charge on appropriate institutional repositories and/or websites. The funding agreement allows for instances where research products cannot be made available publically, and these are covered below in the *'Exceptions and Requirements regarding the Open Access Policy'*. Providing open-access to the data and information products derived under the NESP will provide up-to-date, high quality data and information to decision-makers, environmental managers, other scientists, and to the community. Research products should be discoverable and available without cost.

Except in cases of sensitive data or information, a requirement that interested parties send an email to obtain access to an otherwise closed data product is not considered to represent open access and is therefore not an acceptable practice for NESP research products. See *'Exceptions and Requirements regarding the Open Access Policy' p.9*.

Data publishing

It is expected that researchers will take all reasonable steps to deposit research data in an appropriate subject and/or institutional repository within/upon 12 months of collection of the full data set. Where a publication (e.g. journal article) is being produced prior to the collection of a full data set, the portion of the data set on which the publication is based is required to be deposited in an appropriate subject and/or institutional repository within/upon 12 months of collection of that portion; or where it's past 12 months, as soon as it is apparent that a publication based on that portion of the data set is going to be produced. Metadata and the output should be stored in an open format together, in a way that clearly shows how they are linked. Outputs can also be stored in proprietary formats and AusGOAL identifies a range of [open formats](#) that facilitate reuse and value-adding.

All Australian universities have repositories with potential for providing access to research outputs. Researchers with institutional affiliations can typically contact their university library for more information and assistance on how and what to deposit.

National and international infrastructure also exists in specific disciplinary domains and there are a number of significant repositories for research information in Australia including:

- [Atlas of Living Australia](#)
- [Australian Urban Research Infrastructure Network](#)
- [Integrated Marine Observation System](#)
- [Terrestrial Ecosystem Research Network](#)
- [Australian Ocean Data Network](#)

The publishing of significant datasets as a data paper published within a data journal is encouraged and recognised for the benefit of obtaining a peer-review of the dataset and for the chance for researchers to

have their dataset contribute to their citation index. A non-exhaustive list of data journals is provided by the [University of Edinburgh](#).

Metadata

It is expected that researchers will take all reasonable steps to attach high quality metadata to all products resulting from NESP funding. High quality metadata can facilitate the discovery of data and information products and provides the contextual information needed to enable the end-user to utilise the data in an appropriate and consistent fashion. Use of metadata structure standards allows for more successful data integration and increases data value.

The use of relevant metadata standards can assist in meeting these requirements and allows for more successful data integration, and increases the value of data. Metadata standards can vary according to discipline but the standards must be an accepted and best practice standard. The Reference Architecture for the National Environmental Information Infrastructure provides some preferred general standards for a range of data (<http://www.neii.gov.au/system/files/filedepot/1/NEII%20Reference%20Architecture.pdf>). For spatial metadata, researchers are encouraged to conform to the Australian standard ANZLIC Metadata Profile or, at a minimum, record the mandatory elements of the standard. Variations on the Australian standard are acceptable, for example ISO 19115/19139 MCP (Marine Community Profile). [ISO 19115-1](#) is the international standard that underpins the ANZLIC spatial metadata standard.

The [Open Geospatial Consortium](#) (OGC) provides the most widely accepted formats and protocols for spatial data delivery. Although the OGC offer comprehensive protocols and standards, this does not preclude the use of proprietary formats in particular areas or institutions. In terms of output format, the spatial data shapefile format is acceptable although it does not meet all of the requirements of an OGC format. Shapefile is proprietary format published by the Environmental Systems Research Institute (ESRI). It has become a widely used format and there are readily accessible translators to enable data in this format to be translated to geographic information systems requiring a different format.

These Guidelines encourage the contribution of metadata to the Australian National Data Service (ANDS) data discovery portal, [Research Data Australia](#) (RDA). RDA harvests information from data repositories around Australia. The metadata for the associated data should contain a link to related publications and to the NESP funding grant programme. Best practice for linking to publications and grants from data (or the data's metadata) are available from ANDS (contact@ands.org.au) and from the Council of Australian University Librarians [Research Advisory Committee](#).

At a minimum, requirements for effective metadata are:

- available online, linked to the data (also online) and accessible by web search engines or discovery facilities.
- identifies the attributes, methods and procedures used for determining the attributes of all values within the dataset/stream.
- defines or links to online definitions of all terms used in the data
- provides contacts and access locations for the data
- provides provenance for any data that is used in generating the product.

Persistent Identifiers

Persistent identifiers are globally unique numeric and/or character strings that reference a digital object and are intended to function for the long term. Assigning a Digital Object Identifier (DOI) to data and other research products facilitates data citation and is considered best practice. A DOI is a type of persistent

identifier that indicates a dataset will be well managed and accessible for long term use. It is now routine practice for publishers to assign DOIs to journal articles and for authors to include them in article citations.

Treatment of non peer-reviewed research products

Organisations are increasingly able to produce and make available research and information in a range of formats using digital technologies and online platforms for dissemination. This provides for a level of flexibility, relevance and timeliness often lacking in formal publications such as journal articles and books. However many of the social and economic benefits of these research products are lost because of a lack of bibliographic and production standards, evaluation criteria, systematic collection and preservation strategies and links to the specific underlying data.

Research outputs (peer reviewed or not) which will be sold, including published books, must be made available through key, relevant libraries, at a minimum those libraries participating in the Commonwealth Library Deposit and Free Issue Scheme. www.finance.gov.au/librarydeposit/. This scheme ensures that Commonwealth government-funded publications are deposited in selected libraries and are freely available to library users. There are currently twenty seven institutions subscribing to this scheme, listed at <http://www.finance.gov.au/sites/default/files/library-mailing-list.pdf>.

Many of the guidelines listed here for publications, such as journals and data sets, also apply to a range of other valuable documents and resources produced by NESP projects, such as reports, fact sheets (grey literature), images, photographs and videos, models and software, websites, and mobile and tablet applications. All data products should be openly and freely available using best practice specific to the research product type. Images, photographs and videos should be considered to be part of the research data in the case they have contributed in a significant manner to the development of a dataset. It is not expected that all audio/visual outputs produced across the research program would be included. Grey literature may require additional metadata, and further terminology definition to describe its production, formats and source.

Best practices for the management of grey literature are given in *Appendix A*.

Exceptions and requirements regarding the open access policy

The Department recognises that open access to information may not be suitable in cases where that information is culturally, environmentally or socially sensitive, or could contravene privacy laws. Decisions to restrict access to sensitive research products should be justified, and made by those closest to the source (i.e., the data steward). In cases where restricted access applies, the data steward should keep an enduring copy of the unaltered data and make freely available metadata which describes the data and why it has not been released. Sensitive data may include, but is not limited to:

- location information for highly desirable or collectable species
- location information for rare species
- culturally significant site data
- social data restricted by privacy law or considerations, and
- other heritage or Indigenous matters.

It is the Hub Leader's responsibility to collate instances of exceptions to these guidelines. Collated lists are generally at project level and include basic information on the data collected and the justification for its non-release. Proposed exceptions to the open access policy must be reported to the hub Steering Committees and to the Department, as an attachment to annual Progress Reports, and in Research Plans. Owners of the data that constitute the exception should be open to discussions with the Department

regarding that data, the potential utility and relevance of the data to the Department's business and how relevant portions of that data could be provided so as to assist decision-making.

Peer-Reviewed Publications

Open access to peer-reviewed publications, and well described data arising from the NESP, will be an enduring legacy of quality assured data and information that will assist both decision makers and the wider research community. NESP hub funding agreements allow for the hub to grant a licence or an assignment of Intellectual Property (IP) to the publisher of a peer-reviewed journal, book or similar publication, provided that a copy of the article (as accepted for publication after peer-review) is made publicly available for no charge on an appropriate repository or website, within/upon 12 months of publication of the article. Publication of only an abstract or citation, or the requirement for interested parties send an email to obtain access to an otherwise closed publication, is not considered to represent open access and is therefore not an acceptable practice for NESP publications.

It is the responsibility of the hub to understand the copyright and licencing arrangements being entered into when signing agreements with publication houses. Most journal copyright arrangements distinguish between three versions of a published peer-reviewed article: pre-print (the 'submitted' version), post-print (as accepted for publication after peer-review) and re-print (publisher's final .pdf version of the publications). The Department's preference is for the re-print to be placed on a website at the time of publication. However, in cases where this is not possible, a copy of the post-print placed on a website within/upon 12 months of publication fulfils the NESP funding requirements. Post-prints should include a text header advising where the article has been published, a link to the article's DOI, and should be saved in a .pdf format to enhance discoverability by Google Scholar.

Further guidance and a list of tools to aid understanding of journal copyright and licencing arrangements are given in *Appendix B*.

Intellectual Property

It is recognised that Intellectual Property typically resides with PhD students and that this ownership is included in the contracts held between PhD students and their institution. These Guidelines do not seek to contravene these contracts.

It is expected that all products such as journal articles produced by PhD students will adhere to the open access intent of these Guidelines and be discoverable and accessible free of charge within/upon 12 months of publication. Data collected by PhD students may be considered an exception to the open access policy [see '*Exceptions and Requirements regarding the Open Access Policy*', p.9]. Hubs should retain a register of likely dates by which students' data will be made publically accessible (i.e. following submission of their thesis and publication of related papers). Earlier publication of data by PhD students is encouraged where possible and metadata should be made freely available in line with the guidance at '*Metadata*', p.8.

Communications related to the publication of research products

Summaries of all released research products should be provided to the Department a minimum of five working days before the release, and be accompanied by the NESP research product submission form [Appendix C]. The Department's [Environmental Resources Information Network](#) (ERIN) has requested that staff be deemed eligible and able to subscribe to an RSS feed or otherwise be notified as new data sets become available.

Compliance with the Data Management and Accessibility Guidelines

The information contained in these Guidelines is adapted from relevant sections in the NESP Long Form Funding Agreement (Sections 11.7 Licence of Research Outputs, 11.8 Publication of Research Outputs and

11.9 Peer-Reviewed Publications) and therefore represents a contractual requirement for the Hubs. Consistent with the concept of earned autonomy, the Department will monitor compliance to these guidelines as part of our assessment of the Hub's annual progress report. Instances of non-compliance to the guidelines may result in the relevant milestone payment under the hub funding agreement being delayed or reduced.

Appendix A: Best practice publication and management of grey literature

Adapted from: Amanda Lawrence, Grey literature Strategies ARC Linkage Project Swinburne University, DRAFT 1.0 12 March 2014

Grey literature, such as factsheets, project profiles and reports, should meet minimum bibliographic standards, including:

Minimum essentials:

- Title
- Year of publication (or more specific date if you wish)
- Author(s) – optional for corporate author
- Producing organisation
- Page number on each page
- URL to document or abstract page or DOI or link to repository location
- Copyright or Creative Commons licence – it is a government recommendation that material is published CC By unless there are grounds to do otherwise.

Recommended:

- Header/Footer on every page with title or some other way of identifying the document.
- Location of producing organisation – City and Country
- Contact details of producing organisation and/or URL
- Identifier: this could be an ISBN (International Standard Book Number), ISSN, an internal identification number, a handle or a Digital Object Identifier (DOI).
- Email contact of author
- Author institution (if different from Producing organisation)

Publishing formats

- Main options: PDF, Word, HTML
Providing all three is the optimal situation. HTML is preferred by many for accessibility reasons. PDF format is not desirable, though is preferred to publications not being provided.
- Aim for optimal accessibility: all formats are made more accessible by the use of headings, bullets, tables. More on PDF and web publishing accessibility guidelines is available from AGIMO
 - <http://www.finance.gov.au/blog/2012/09/20/pdf-accessibility-becomes-iso-standard/>
 - <http://www.w3.org/TR/WCAG20-TECHS/>

Publishing options

- Consideration should be given to whether the literature could be posted or archived to a repository or database and then linked to from your website. Some options are:
 - Policy Online (<http://apo.org.au> or editor@apo.org.au)
 - Pandora, NLA archive
 - Internet Archive

Appendix B: Peer-reviewed publications

Peer reviewed publications

Peer-reviewed publications include: peer reviewed scientific papers; books; and other peer-reviewed published material. An electronic copy of all peer-reviewed articles (as accepted for publication after peer-review) must be made openly and freely available on the internet, if not immediately, then within/upon 12 months of publication. Researchers are not permitted to grant a licence for NESP funded articles to publishers if that licence does not allow the hub to place a complete copy of the article (not just an abstract or citation) on its website within twelve months of publication.

Understanding pre-print, post-print and re-prints: Which should be made publicly available?

Journal copyright arrangements often distinguish between three versions of a published peer-reviewed article - pre-print, post-print and re-print. A pre-print refers to a paper that has been submitted for publication but which has not yet undergone peer-review. The manuscript is termed a post-print after the author has re-written the manuscript based on the comments of reviewers or editors. The re-print refers to the publisher's final PDF version of the paper encompassing the journal-specific formatting and which contains the same figures and text exactly as the post-print.

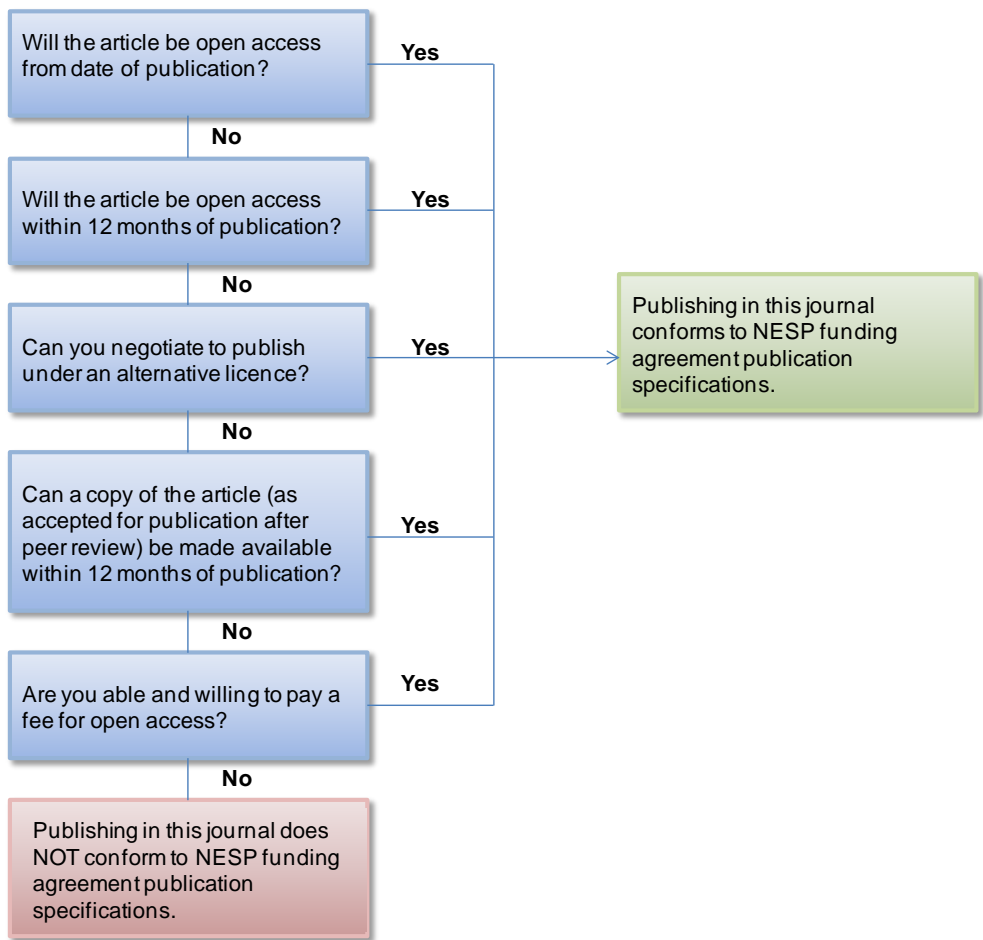
Tools for understanding journal copyright arrangements

It is the responsibility of the research hubs to understand the copyright and associated licensing arrangements being entered into when signing agreements with publication houses. Internet tools such as SHERPA/RoMEO can be used to determine the default position of the publication house.

Where the default position of a publication house does not conform to the NESP requirements for open access, hubs may be in a position to negotiate licensing arrangements by stipulating that their funding body has a requirement for all research products to be open access. Many journals will negotiate the licence conditions applying to individual journal articles in cases where research funding organisation has an open access policy. Some publications allow researchers to pay a fee to make the article open access. Payment of open access fees is the responsibility of the research hub. The Department's preference is for research funding to be used for research, in preference to publication fees.

Alternatively, it is often possible to negotiate with the publisher to publish under an alternative licence scheme that meets the Australian Government's open access policy. AusGOAL provides licences for a range of information access situations that are designed to manage legal risks associated with making information available for reuse and are endorsed by the Office of the Australian Information Commissioner. These include a non-commercial licence which may be acceptable to the publisher. Figure 1 is a flow diagram which may aid in determining whether publishing in a specific journal conforms to the NESP funding agreement specifications.

Figure 1. Choosing a journal that conforms to the NESP funding agreement publication specifications.



Appendix C: Research Product Submission Form (copy)



Research Product Submission Form

The Australian Government will use the research product and information provided in this form to share the knowledge generated under the NESP, support its adoption and identify opportunities to promote NESP research.

Hold the cursor over each answer field for instructions on completing this form.

Project Information

Hub name	Select from the options below
Project number	
Project title	
Product lead author/researcher	
Program funding (NESP/NERP/CERF/Combination)	

Research Product Information

Select the type of research product	Select from the options below
Other	
Title	
Journal or location name	
DOI	
Citation	
Synopsis	
Media/other interest	
Keywords	

Research Product Accessibility

Date of publication	
Date of product availability	
Location of product availability	
Version of product attached	

I certify that I have attached the copy of the publication that can be provided to and used within the Australian Government in accordance with the specific publisher copyright restrictions that relate to this product.

Email Form

Functionality may be limited unless opened in Adobe. If the instructions do not appear or the above button does not work, visit www.environment.gov.au/nesp for instructions and email the form to your Hub Communications or Administration Officer.