The Research Council of Norway

Open Access to Research Data

Policy for
The Research Council of Norway

The Research Council of Norway

Open Access to Research Data

Policy for
The Research Council of Norway

The Research Council of Norway
P.O.Box 2700 St. Hanshaugen
N–0131 OSLO
Telephone: +47 22 03 70 00
Telefax: +47 22 03 70 01
bibliotek@rcn.no
www.rcn.no/english

The report can be ordered at:
www.forskningsradet.no/publikasjoner

or green number telefax: +47 800 83 001

English translation: Carol B. Eckmann
Design omslag: Design et cetera AS
Photo/illustration: Shutterstock

Oslo, November 2014

ISBN 978-82-12-03382-5 (pdf)

Foreword

Transparency and knowledge sharing are critical for research. Rapid technological development is changing how research is conducted and results are shared. Advancement in research is increasingly being driven by access to new, large amounts of data. The results of publicly funded research comprise a public good that is valuable to the research community as well as to society at large. Better access to research data will enhance the quality of research in that results can be validated and verified in a more effective manner and datasets can be used in new ways and in combination with other datasets. Creative, expanded use of research data will pave the way for more wide-ranging interdisciplinary research, as well as innovation in the development of society and industry. In addition, data sharing will result in a more rational use of resources within the research sector because efforts will not be duplicated and because datasets may be used in other contexts than originally planned. Open access to results and data will also be valuable for education and give research a higher profile within society at large.

This policy sets out how the Research Council will use its instruments to promote open access to research data. The principles and guidelines presented in this document are intended to serve as a guide for researchers when planning their research projects. The policy will be revised on the basis of the experience gained from their implementation, as well as international developments, changes to the legal framework, and new technological potential.

1. Introduction

Objectives

The Research Council has drawn up this policy to lay the foundation for improved quality assurance, storage, accessibility to and sharing of research data by researchers. This will lead to:

- Enhanced research quality through better opportunities for expanding on previous studies and combining data in new ways;
- Transparency in the research process and better opportunities for verifying scientific results;
- Increased cooperation and less duplication of research efforts;
- More innovation in the private and public sectors;
- Greater efficiency and better utilisation of public funding.

The role of the Research Council

Deposit of and accessibility to research data for further use must take place in a cooperative effort between researchers, research institutions, scientific journals, infrastructure providers and government authorities. This will require new ways of thinking among the actors involved, as well as the development of new competency and training, and new infrastructures and tools. The Research

Council seeks to be a driving force for the storage and sharing of research data. This means that the Research Council will facilitate cooperation between the actors, finance relevant activities, and provide guidance to the research communities through:

- investing in infrastructure that supports quality assurance, long-term storage, sharing and reuse of research data;
- promoting an effective division of tasks between public infrastructure and service providers by organising meeting places and offering advisory services;
- implementing procedures in the application assessment process that ensure relevant grant applications include data management plans;
- implementing procedures in project follow-up activities that ensure the data management plans are being followed by projects granted funding;
- continuing the practice in Research Council contracts that requires research data to be stored in a safe and secure manner for a minimum of 10 years;
- providing an up-to-date list of approved, national (and any international) infrastructures for storing, archiving and sharing data;
- encouraging the institutions to devise their own policies and guidelines to improve the accessibility of research data.

Delimitation

Input data are data that already exist, regardless of the research to be conducted. These may comprise information collected for a different purpose (e.g. public administrative data, clinical data or weather data) or they may be physical or digital collections of objects and texts (e.g. libraries, text corpuses or other scientific collections). Information on the Internet may also be defined as input data in this context, and such information is highly heterogeneous. Data which are used as input data in research, but which have been collected, generated or processed by other researchers or research institutions than those conducting the research, will normally not be encompassed by the guidelines in this policy.

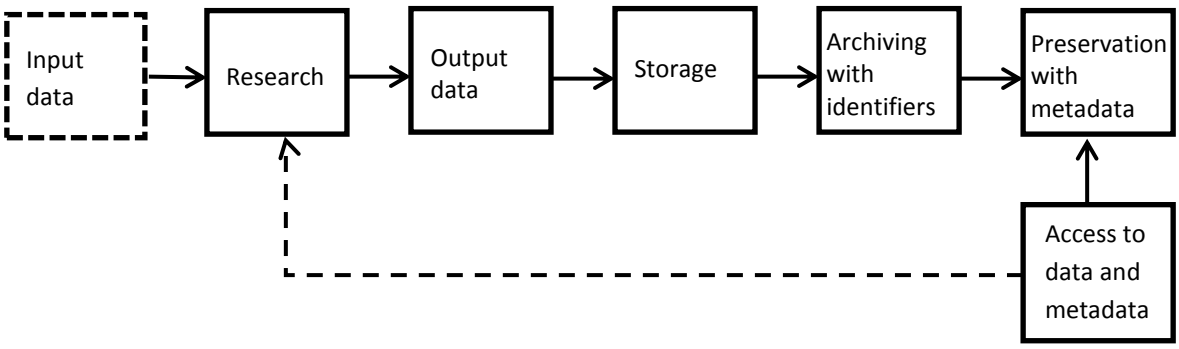
Output data are data generated through research. These may be data that are generated through new analysis or by combining existing input data, but they may also comprise entirely new data that are generated through new data collection. Such data will typically come from experiments, simulations, fieldwork or interviews. Output data are always a direct result of research activity, regardless of whether the data are based on input data or whether they are collected from scratch. Thus, output data from previously conducted research which are used as input to new research are designated as input data based on the definition in the paragraph above.

The principles and guidelines in this policy apply only to output data that are generated through research activity funded wholly or partly by the Research Council.

Use of data in the research process

It is common to use many data sources in research, and output data from research are reused, combined with other data, and distributed both directly and through the use of data by others. To

enable the reuse of output data, they must be processed after they are generated. This may be illustrated with the following process, which is divided into phases:[1]



The first phase following generation is storage in a physical infrastructure. When the data are stored in a safe and secure manner in the form of reserve copies, they must be archived. This phase entails, among other things, ensuring that the data cannot be changed and making the data retrievable through the use of identifiers. The next phase consists of enhancing the quality of the data by classifying them and making them more usable. This phase may be called the preservation phase, and it is where the data are enriched with metadata, which is descriptive information about the data. This may include descriptions of the experiment or simulation from which the data were generated, references to standards used in the work with and categorisation of the data. This may also involve descriptions of the quality of the data and descriptions which make it possible to search for the data, such as time period, geographic region, scientific field and research topic. Without good metadata it is difficult to search for data, and the data will not be usable by others due to uncertainty related to what they actually represent.

After the preservation phase, the data may be used by others, but it may be necessary/desirable to include a final phase consisting of data curation (not shown in the figure). Curation refers to the maintenance and updating of the data and ensuring the data are relevant for further use over time.

In the view of the Research Council, storage and archiving are critical for the capacity to retrieve scientific results data. Furthermore, metadata must be generated to ensure that output data have long-term value and can be made accessible to others. As a consequence, a policy for open access to research data must also encompass the storage, archiving and generation of metadata.

2. Principles and guidelines in the Research Council's policy

Main principle

Open access

The Research Council's policy follows an open by default principle with regard to access to research data. The Research Council will therefore help to ensure that research data in general are made openly accessible, but that an exception is made for data that cannot or should not be disseminated.

[1] The phases following the generation of results-based data are based on http://www.clir.org/initiatives-partnerships/data-curation

The Horizon 2020 programme's definition of open access[2] states that access to scientific information must also be free of charge. The Research Council has decided to base its policy on the premise that the user should cover the actual costs incurred from data retrieval. This is closer to the OECD's definition of open access,[3] which states that access must be provided at the lowest possible cost, preferably at no more than the marginal cost of dissemination.

Exceptions

There are a variety of challenges associated with making some datasets openly accessible. The reasons for restricting access include:

Security concerns

When open access to the data may threaten personal or national security, the datasets ***must not*** be made openly accessible.

Personally sensitive data

When open access to the data conflicts with the applicable statutory framework regarding the protection of personal privacy, the datasets ***must not*** be made openly accessible.

Other legal factors

When open access to the data conflicts with other legal provisions, the datasets ***must not*** be made openly accessible.

Commercial factors

Data that have commercial value and are generated in projects in which a company is the contractual partner with the Research Council ***may*** be exempted from the general principle of open access. In these cases, it is recommended that the data are made available after a certain period of time, preferably after three or five years.

Other factors

When open access to data will have major financial or practical implications for those who have generated/collected the data, the datasets ***may*** be exempted from the general principle on open access if a satisfactory argument is made for this.

The Research Council's guidelines

The Research Council has drawn up a set of guidelines for the archiving, dissemination and sharing of research data. The guidelines will be followed up through the Research Council's research funding instruments and through the National Financing Initiative for Research Infrastructure.

1.0 Research data should be stored/archived in a safe and secure manner.

- 1.1 The data should be stored in secure archives, either in a central repository at the relevant institution or in national archives.

2.0 Research data should be made accessible for reuse.

- 2.1 Research data should be made accessible to all relevant users, on equal terms, as long as there are no legal, ethical or security-related reasons to preclude this.

---

[2] European Commission: "Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020", Version 1.0, 11. December 2013. Available here in PDF. [Downloaded 16.04.2014].

[3] OECD: "Principles and Guidelines for Access to Research Data from Public Funding", April 2007. Available here in PDF. [Downloaded 16.04.2014].

3.0 Research data should be made accessible at an early stage.

3.1 The data used as the basis for scientific articles should be made accessible as soon as possible, and never later than at the time of publication.

3.2 Other data that may be of interest for other research should be made accessible within a reasonable amount of time, and never later than three years after the project has concluded.

4.0 Research data should be accompanied by standardised metadata.

4.1 The metadata should enable others to search for and use the data.

4.2 The metadata should follow international standards.

4.3 The metadata should describe the quality of the data.

5.0 Research data should be provided with a license for access, reuse and redistribution.

5.1 The license should be internationally recognised.

5.2 The license should set as few restrictions as possible on the access, reuse and redistribution of the data.

6.0 Research data should be made accessible at the lowest possible cost.

6.1 Metadata should be made accessible at no charge and be published in a way so that they may be harvested and used in research data searches.

6.2 Research data should preferably be made accessible at no charge.

6.3 The price of access to research data should never be higher than the actual costs of making them available.

7.0 Research data should be provided with a long-term plan.

7.1 A plan should be prepared for the management of data that are considered to have long-term value.

7.2 The projects should specifically address how research data that are not considered to have long-term value should be managed, or if relevant destroyed, after a certain period of time.

FACTS

Clarification of terms

Some of the key terms in the area of open access to research data can be ambiguous. Consequently, it is important to specify what the Research Council means by these terms in order to clarify the conceptual framework underlying these principles and guidelines.

Research data	*Research data is understood in this policy to mean the registration/recording/reporting of numerical scores, textual records, images and sounds that are generated by or arise during research projects.*

Publicly funded	*All projects and activities that are wholly or partly funded with government allocations are considered to be publicly funded.*
Open access	*Open access is understood as the principle that research data should be accessible to relevant users, on equal terms, and at the lowest possible cost. Access should be easy, user-friendly and, if possible, Internet-based.*

The Research Council of Norway
Drammensveien 288
P.O. Box 564
N-1327 Lysaker

Telephone: +47 22 03 70 00
Telefax: +47 22 03 70 01
post@rcn.no
www.rcn.no

Cover design: Design et cetera AS
Cover photo: Digital Vision

Oslo, December 2014

ISBN 978-82-12-03382-5 (pdf)