University of Helsinki research data policy: Background memorandum

Table of Contents

UNIVERSITY OF HELSINKI RESEARCH DATA POLICY: BACKGROUND MEMORANDUM................................................ 1

1 BACKGROUND .................................................................................................................................................. 2


2 THE OPERATIONAL STARTING POINTS, SCOPE AND WORK OF THE RESEARCH DATA POLICY COMMITTEE IN 2014 7

2.1 APPOINTMENT AND DUTIES OF THE COMMITTEE .................................................................................. 7

2.2 OPERATIONAL SCOPE OF THE COMMITTEE ......................................................................................... 8

2.3 WORK OF THE COMMITTEE ................................................................................................................ 9

3 VISION, OBJECTIVES AND ADDED VALUE OF THE RESEARCH DATA POLICY ................................................ 9

3.1 VISION ........................................................................................................................................................ 9

3.2 OBJECTIVES AND ADDED VALUE ........................................................................................................ 10

4 RESOURCES AND RESPONSIBILITIES REQUIRED BY THE RESEARCH DATA POLICY .............................. 12

5 DEFINITIONS ................................................................................................................................................ 14
1 Background

The significance of online research corpora, i.e., research data, to scientific research has increased along with the development of digital communications. The digital revolution has enabled researchers to distribute, open up and use large datasets in novel ways. The wide availability and usability of datasets greatly enhance and improve research and international cooperation, enable new discoveries, and promote equal opportunities for the utilisation of data among researchers.¹

Research data management is subject to numerous national and international recommendations and obligations. According to a Government resolution dated 3 March 2011, public datasets must be openly available and reusable. In March 2014, the Ministry of Education and Culture responded by launching the Open Science and Research 2013–2017 initiative, which aims for Finland to become a leading country in the openness of science and research by the year 2017.² The project pertains to research publications, research data and research methods. The European Commission has resolved that publicly funded online research corpora must be freely available to researchers, companies and individuals alike.³ Many research funding providers require systematic data management and the opening up of research data. The Academy of Finland, for instance, recommends that research corpora be made public and requires that all research plans include a data management plan.⁴ The Horizon 2020 EU Research and Innovation Programme also includes a pilot project pertaining to open access to research data.

The demands for the efficient use of research data support efforts to open up data and ensure opportunities for secondary research. Opening up data to the public requires systematic data management. In addition to external demands, certain norms within academia, such as repeatability, verifiability and controllability, steer researchers towards documenting their data appropriately and producing and managing their research knowledge in systematic ways.⁵

The League of European Research Universities, LERU, has published a set of guidelines for research data management entitled The LERU Roadmap for Research Data in December 2013.⁶ LERU recommends that each member university draft an organisation-specific research data policy, which contains agreed-upon principles and protocols for the collection, storage, use and management of research data at the level of the university.

The University of Helsinki began preparing its research data policy in spring 2014. The preparation process included charting previous research data management projects, outlining the objectives of the policy and forming a committee for preparing the research data policy for the rector’s approval.

---

² http://openscience.fi/.
³ European Commission: Towards better access to scientific information: Boosting the benefits of public investments in research. 7/2012.
⁴ The Academy of Finland promotes open access, 26 June 2014. (in Finnish).
⁵ Sami Borg: University research data policy. 27.8.2014. (in Finnish).
1.1 The research data committee of the University of Helsinki 2008–2011 and the Research Data Pilot Project at the Helsinki University Library 2011–2012

In 2008–2011, a research data committee appointed by Vice Rector Hannele Niemi and chaired by Docent Marjut Salokannel operated at the University of Helsinki. The committee’s mission was to analyse how different fields within the University of Helsinki could better make their research data available to the research community, primarily for the purposes of verifying research results and, depending on the field, for further utilisation in subsequent research projects. The final report of the committee was published in May 2011. The Helsinki University Library implemented a research data pilot project in 2011–2012 with the aim of tackling problems pertaining to the management, storage and use of research data through an international research group operating at the University of Helsinki. The goal was to investigate various research data-related practices, questions and problems arising in different phases of the research project. The final report of the pilot project was published 8 March 2012.

The final reports of both the University’s research data committee and the Library’s pilot project presented a number of proposed measures pertaining to the collection, storage and use of research data. Tables 1 and 2 present the measures and an estimate of how they were implemented by December 2014. The estimates show that some of the proposed measures are no longer relevant or that national projects such as the Open Science and Research Initiative and related solutions must be taken into consideration in their implementation. It should also be noted that not all measures have been implemented systematically throughout the University. University-level research data management services fail to reach all researchers, thus failing to support researchers and the University’s objectives in an optimal way.

Table 1: Proposals of the research data committee (2011) and their implementation in 2014

<table>
<thead>
<tr>
<th>Proposal</th>
<th>Phase</th>
<th>Responsible organisation</th>
<th>NB!</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making research data openly available in order to verify research results</td>
<td>Partially implemented</td>
<td>Principal investigators, researchers, research groups</td>
<td>Differences between disciplines and special fields are considerable. Communications have a significant impact. No centralised services are provided.</td>
</tr>
<tr>
<td>Creating mechanisms to support researchers in the preparation of funding applications and the</td>
<td>Partially implemented</td>
<td>Research Affairs (Research Funding Services and Research)</td>
<td>Cooperation between Research Affairs, the Helsinki University Library and the IT Centre will be continued.</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Proposal</th>
<th>Phase</th>
<th>Responsible organisation</th>
<th>NB!</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Library promotes and supports the accessibility and availability of research data as well as opportunities for secondary research.</td>
<td>Partially implemented</td>
<td>The Helsinki University Library</td>
<td>Some attention has been paid to metadata and ontology-related competence. The Library has both campus-specific research data committees and a joint</td>
</tr>
<tr>
<td>Confirming regulations/principles regarding the ownership of research data</td>
<td>Partially implemented</td>
<td>Research Affairs (legal counsels for research), principal investigators</td>
<td>Project-specific agreements will be concluded. No principles have been drafted. Ownership issues have not been discussed within the University.</td>
</tr>
<tr>
<td>Outlining field-specific deadlines for opening up research data</td>
<td>Partially implemented</td>
<td>Research groups</td>
<td>Can this be done centrally? Field-specific differences must be taken into consideration.</td>
</tr>
<tr>
<td>Creating an integrated information environment (metadata standards, interfaces)</td>
<td>Not started</td>
<td></td>
<td>Relevant national projects (the Open Science and Research Initiative and services provided by the CSC) should be taken into consideration.</td>
</tr>
<tr>
<td>Implementing (campus-specific) training on drafting research data management plans</td>
<td>Implemented</td>
<td>The Helsinki University Library</td>
<td>Implemented as part of data management training for doctoral schools (one-credit module that includes other aspects)</td>
</tr>
<tr>
<td>Acquiring more storage space (close to researchers to resolve immediate storage-related issues)</td>
<td>Partially implemented</td>
<td>The IT Centre</td>
<td>Taking into consideration relevant national projects (services provided by the CSC)</td>
</tr>
<tr>
<td>Contributing to national long-term data preservation efforts</td>
<td>Implemented</td>
<td>Research Affairs (Research Administration)</td>
<td>Participating in the CSC’s Long-Term Preservation Solution for Research Data (PAS) committee. National long-term preservation solutions are important in fields with no established international data banks (including preservation solutions for specific types of data).</td>
</tr>
</tbody>
</table>

Table 2: Proposals arising from the Research Data Pilot Project at the Helsinki University Library (2012) and their implementation in 2014
- The Library invests in metadata and ontology-related competence.
- The research data committee or equivalent will continue operating at the Library until support for data management is established as part of Library operations.

The Library provides support for researchers in research data management-related issues. Research practices vary between disciplines, which is why the need for support for research data management varies as well.

- The competence of Library staff is developed in such a way that each campus has at least one information expert who can support researchers in research data management. One of these experts, the data librarian, is an ICT expert.
- In the future, all Library support staff will possess general research data management competence and knowledge. Customer service staff are familiar with research data-related sources and is able to direct customers to appropriate sources.

In the future, the Library will increase cooperation with researchers/research groups within the University (the IT Centre, Research Affairs, the National Library of Finland), nationally (the Open Science and Research Initiative coordinated by the CSC) and internationally (OpenAIREplus) to enhance research data management.

The Library is active on international forums and is the National Open Access Helpdesk (NOAD) of the OpenAIREplus project in Finland.

The Library implements the EU recommendation to increase the openness of research data produced with public funding.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Status</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Library invests in metadata and ontology-related competence.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The research data committee or equivalent will continue operating at the Library until support for data management is established as part of Library operations.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Library provides support for researchers in research data management-related issues. Research practices vary between disciplines, which is why the need for support for research data management varies as well.</td>
<td>Implemented (Experts are employed on campuses.) / Partially implemented (Extensive piloting and creation of advisory service packages is ongoing.)</td>
<td></td>
</tr>
<tr>
<td>In the future, the Library will increase cooperation with researchers/research groups within the University (the IT Centre, Research Affairs, the National Library of Finland), nationally (the Open Science and Research Initiative coordinated by the CSC) and internationally (OpenAIREplus) to enhance research data management.</td>
<td>Implemented</td>
<td>Network Focus Group meetings have been planned.</td>
</tr>
<tr>
<td>The Library is active on international forums and is the National Open Access Helpdesk (NOAD) of the OpenAIREplus project in Finland.</td>
<td>Implemented</td>
<td></td>
</tr>
<tr>
<td>The Library implements the EU recommendation to increase the openness of research data produced with public funding.</td>
<td>Implemented</td>
<td>Access services The Library opens up some of its data to the public.</td>
</tr>
</tbody>
</table>
In the future, the University of Helsinki will have a data catalogue or information system where research metadata can be stored and from which it can be subsequently harvested into a national metadata catalogue. This catalogue can be linked to the Research Database TUHAT and other systems.

- The Library will submit a proposal to the IT Centre regarding a project to investigate different possibilities of creating an information system or adopting an existing one for data management at the University of Helsinki.

The University of Helsinki and the Helsinki University Library will further the adoption of permanent identifiers for projects, publications, researchers and data.

- The TUHAT research database should have a more obvious function for publishing the location of the data. This would increase the visibility of data also among researchers.
- The Library should have more influence on the development of the TUHAT database.

The University of Helsinki will enable its researchers and research groups to store their research data on University servers and in discipline-specific, national or international data archives such as IDA, FSD and Pangaea.

The University should have clear guidelines for determining the ownership of data in different fields and the persons responsible for acquiring necessary permissions from copyright holders. The University should also have clear guidelines for acquiring research permits and drafting descriptions of personal data files in accordance with relevant legislation.

Every research plan should include a data management plan outlining the

<table>
<thead>
<tr>
<th>Action</th>
<th>Status</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the future, the University of Helsinki will have a data catalogue or information system where research metadata can be stored and from which it can be subsequently harvested into a national metadata catalogue. This catalogue can be linked to the Research Database TUHAT and other systems.</td>
<td>Not started</td>
<td>The TUHAT research database enables adding a link to data (one link per publication).</td>
</tr>
<tr>
<td>The Library will submit a proposal to the IT Centre regarding a project to investigate different possibilities of creating an information system or adopting an existing one for data management at the University of Helsinki.</td>
<td>Partially implemented</td>
<td>ORCID and ResearcherID are marketed as part of the training provided by the Library. The TUHAT research database has been developed further.</td>
</tr>
<tr>
<td>The University of Helsinki and the Helsinki University Library will further the adoption of permanent identifiers for projects, publications, researchers and data.</td>
<td>Partially implemented</td>
<td>Advice on available storage options is provided.</td>
</tr>
<tr>
<td>The TUHAT research database should have a more obvious function for publishing the location of the data. This would increase the visibility of data also among researchers. The Library should have more influence on the development of the TUHAT database.</td>
<td>Partially implemented</td>
<td>Project-specific agreements are being drafted and advisory services provided. The results of the Open Science and Research Initiative must be taken into consideration.</td>
</tr>
<tr>
<td>The University of Helsinki will enable its researchers and research groups to store their research data on University servers and in discipline-specific, national or international data archives such as IDA, FSD and Pangaea.</td>
<td>Partially implemented</td>
<td></td>
</tr>
<tr>
<td>The University should have clear guidelines for determining the ownership of data in different fields and the persons responsible for acquiring necessary permissions from copyright holders. The University should also have clear guidelines for acquiring research permits and drafting descriptions of personal data files in accordance with relevant legislation.</td>
<td>Partially implemented</td>
<td></td>
</tr>
<tr>
<td>Every research plan should include a data management plan outlining the</td>
<td>Partially implemented</td>
<td></td>
</tr>
</tbody>
</table>
acquisition, use, preservation and protection of the data, including primary materials, as well as any plans as to how its further use will be enabled. The University should ensure that researchers have at their disposal in writing clear practices for drafting these plans. This requires cooperation between several parties.

(Support and training are provided.)

2 The operational starting points, scope and work of the research data policy committee in 2014

2.1 Appointment and duties of the committee

The Strategic Plan of the University of Helsinki 2013−2016 includes the following development areas: “research and teaching infrastructure of an international standard” (1.1) and “effective structures and clear practices” (3.2). The goal is to streamline operational practices and create services that support the core operations of the University. A University-level research data policy supports the implementation of the strategic development areas.

The data policy also supports the University’s strategic objective of ranking among the top fifty universities in the world. To achieve this goal, the University of Helsinki must have common principles for handling research data and clear processes for research data management. By outlining a research data policy the University is able to meet national and international expectations concerning research data management.

On 14 May 2014, the rector launched a research data policy project to achieve the aforementioned objectives and appointed a committee to carry out the task (Rector’s Decision 71/2014, in Finnish). The duties of the committee included

- determining objectives for the research data policy and investigating the added value thus created,
- preparing a proposal for the research data policy of the University of Helsinki to the rector,
- outlining plans to implement the policy, and
- distributing information and reports about the progress of the work to various interest groups.

The members of the committee included Vice Rector Keijo Hämäläinen (chair), IT Manager Minna Harjuniemi, Professor Sampsa Hautaniemi, Professor Timo Honkela, Campus Library Director Pälvi Kaiponen, Copyright Attorney Maria Rehbinder (Aalto University) and University Researcher, Docent Reijo Sund. In addition, Senior Advisor in Research Administration Eeva Nyrövära was appointed secretary of the committee.

The preparatory work of the committee was carried out in cooperation with the scientific community, faculties, departments, independent institutes and the University of Helsinki Central Administration. Experts from the Helsinki University Library, IT Centre and the University of Helsinki Central Archives were consulted during the process.
The collection, use and distribution of research data has been extensively discussed both in Finland and on international forums. The research data policy of the University of Helsinki is closely linked to this discussion and the committee strove to enable various Finnish stakeholders to participate in the drafting of the policy and to influence other relevant national efforts. Particular emphasis was placed on cooperation between the University of Helsinki and Aalto University, and Copyright Attorney Maria Rehbinder acted as a liaison between the universities in the project. Committee chair Keijo Hämäläinen and secretary Eeva Nyrövaara participated actively in the Open Science and Research Initiative coordinated by the Ministry of Education and Culture.

Several universities belonging to the League of European Research Universities (LERU) have prepared detailed data policies. The research data policy committee decided to model its work on the following research data policy documents of LERU member universities: the Research Data Management Policy of the University of Edinburgh and the Policy on the Management of Research Data and Records of the University of Oxford.

2.2 Operational scope of the committee

The committee outlined its mission as follows: the research data policy pertains to the online research corpora used, generated and edited in connection with research projects, i.e., research data. Physical materials (such as printed copies) on which the research data are based and biological materials were left outside the scope of the policy.

However, the committee points out that the management of physical and biological materials is a key aspect of good scientific practice. The use of biological materials in research is governed by specific legislation and ethical guidelines. The use of physical research materials is governed, for example, by instructions for document preservation or the 2013 archiving plan (AMS) as well as the criteria for the preservation of research materials outlined cooperatively by Finnish universities. If online materials have been documented and processed for long-term preservation, hard copies usually need not be preserved. However, the preservation of printed materials must be determined on a case-by-case basis, taking into consideration the fact that permanent electronic storage requires permission from the Finnish National Archives Service. When determining the need for archiving, issues such as the following must also be taken into consideration: the expected cost-benefit ratio of preserving the data, the need (if any) to verify

---

14 Legislation and guidelines for research on Flamma: [https://flamma.helsinki.fi/portal/home/sisalto?_nfpb=true&_windowLabel=contentviewer&contentId=HY303218&lang=en&_pageLabel=content_view](https://flamma.helsinki.fi/portal/home/sisalto?_nfpb=true&_windowLabel=contentviewer&contentId=HY303218&lang=en&_pageLabel=content_view).
15 Archiving plan 2013 on Flamma (in Finnish): [https://flamma.helsinki.fi/portal/home/sisalto?_nfpb=true&_pageLabel=content_view&_nfls=true&placeId=HY055745&contentId=HY306953](https://flamma.helsinki.fi/portal/home/sisalto?_nfpb=true&_pageLabel=content_view&_nfls=true&placeId=HY055745&contentId=HY306953).
data from hard copies or other materials, and data protection issues. The Central Archives of the University of Helsinki provides instructions and guidance on the preservation of printed materials. Materials used in teaching and administrative documents were left outside the scope of the policy. The research data policy takes into consideration open access issues. The openness of scientific publications was also left outside the scope of the policy (for further information on the openness of research publications, see Rector's Decision 126/2008, in Finnish).

2.3 Work of the committee

The committee appointed by the rector convened four times: 12 June 2014, 1 September 2014, 29 October 2014 and 12 December 2014. The committee used a wiki (https://wiki.helsinki.fi/display/Tutkimusdatapolitiikka/Tutkimusdatapolitiikka) for collecting background materials and document drafts.

The committee produced the following documents: The University of Helsinki Data Policy, The University of Helsinki research data policy: Background memorandum, and The University of Helsinki research data policy: Implementation plan 2015–2017.

The preparation of the research data policy was discussed at the Research Council on 4 June 2014, 12 November 2014 and 15 December 2014, as well as by the management group of the University on 17 April 2014 and 13 January 2015. The committee secretary gave presentations on the preparation of the research data policy at the following meetings and events: the summer seminar of the Enterprise Architecture Committee of the University of Helsinki on 17 June 2014, a meeting of the research and postgraduate education network on 1 October 2014, the national ICT seminar of Finnish universities (IT-päivät) on 5 November 2014, and a seminar organised by the Finnish Social Science Data Archive (FSD) on 1 December 2014.

No separate hearings were organised for the preparation of the research data policy. The committee submitted its background memorandum, the research data policy and the implementation plan to the academic community for commenting to allow various stakeholders to influence the preparation of the policy. In addition, all documents were uploaded to the wiki for commenting (http://wiki.helsinki.fi/display/Tutkimusdatapolitiikka/Tutkimusdatapolitiikka). Commenting in the wiki required logging in with a valid Haka user ID. The commenting period was 4–28 November 2014.

3 Vision, objectives and added value of the research data policy

3.1 Vision

Good data management practice is a part of good scientific practice and enhances the research process. According to the data management handbook published by the Finnish Social Science Data Archive, good data management entails utilising scientifically and ethically sound data collection, research and evaluation methods and preserving any data thus generated in
accordance with the criteria for scientific information. The University-level research data policy enables various actors to follow unified principles for the use and submission of online research materials. By determining the principles of management, the policy also indirectly improves the quality of research materials.

The vision of the research data policy is for online research corpora (research data) generated at the University of Helsinki to be

- authentic and reliable;
- intelligible, available and accessible;
- safely and securely preserved;
- openly distributable (taking into account ethical and legal issues);

and for researchers operating at the University to

- receive the necessary training, guidance and support in data management; as well as
- acknowledge that data management is a key aspect of research.

3.2 Objectives and added value

The objective of the University of Helsinki research data policy is to agree on common principles for collecting, storing, using and maintaining research materials, i.e., research data, at the University. A University-level policy ensures that research projects carried out at the University comply with good data management practice.

The detailed objectives of the policy and the added value thus generated include:

- improving the quality and transparency of research by increasing openness and complying with good data management practice;
- creating unified practices for the storage, maintenance and systematic disposal of research data, thus ensuring its accessibility, availability and distributability; facilitating both research cooperation and the usability of research data in teaching through its good accessibility, availability and distributability;
- creating unified research data services, thus saving financial and personnel resources at the University and increasing the time available for research;
- enabling researchers to open up research data to the public after its primary use, thus promoting a more diverse use of materials;

---

The Finnish Social Science Data Archive: Aineistonhallinnan käsikirja. Miksi aineiston hallintaa ja jatkokäyttöä? (in Finnish). Data management entails that research materials (data) and related descriptive information (metadata) have been created, stored and organised in such a way that the data remains usable and reliable and that data protection is maintained throughout the lifespan of the data.
preparing for the long-term preservation of digital materials (http://www.kdk.fi/en/long-term-preservation) and the long-term preservation solution (Tutkimus-PAS) drafted as part of the Open Science and Research Initiative; 

- obligating researchers and other relevant parties to agree upon the ownership of research data and its access rights and terms of use before the start of research projects, thus avoiding disputes; 

- determining the necessary responsibilities and resources and agreeing on the distribution of labour between various University bodies; and 

- raising discussion on the ownership of research data generated under the auspices of the University of Helsinki. Ownership can be realized in a number of ways, such as: a) the University owns and holds other relevant rights to databases containing research data centrally and grants parallel access rights to researchers, other universities or interested parties with, for example, the CC BY 4.0 licence; or, b) the ownership of relevant rights is decentralised, so researchers retain rights to databases and any data therein as well as related responsibilities, either to their own data or jointly to the entire database. In the latter option, the University does not participate in data-related agreements, but the researchers in possession of the data must agree upon the rights and permissions for further use separately.

In its research data policy, the University of Helsinki takes into account both the national objective to increase the openness of data in Finland and the recommendations of several research funding providers that favour the openness and transparency of research data. The openness of research data is based on general scientific principles and protocols, the requirements and interests of research funding providers (including societal impact and the taxpayers' perspective), the recommendations or requirements of scientific publishers as well as the objectives of research groups and individual researchers.

The research data policy enables and enhances the active and systematic use of research materials generated and managed at the University after the original compilers of the data stop actively utilising them (further use in secondary research). Published datasets can be referred to in the same way as research publications. Opening up research data can be considered beneficial in the following ways:

- Open data enables research of high quality (through optimal knowledge of existing information in terms of research materials).
- Open access increases the shared use of research data and the number of users, thus improving efficiency and productivity.
- Openness enables minimising redundancy in data collection and research. The societal impact of research increases once research data can be utilised by private citizens.
- The openness of research supports innovation and the commercialisation of research results.

The committee recommends that research data be opened to the public using Creative Commons (CC) licences. The Creative Commons Nimeä (BY) 4.0 licence is the most frequently used Creative Commons licences.
licence for databases. The *Creative Commons 4.0* licence family takes into account the sui generis protection of databases. As for licensing metadata, the committee recommends forfeiting any rights with the *Creative Commons 0* licence. This means forfeiting the requirement to indicate the name of the author(s). Other licences may be used for weighty reasons.

The date of opening up research data must be considered field and research specifically. Sometimes data cannot be made public due to ethical or legal reasons.

### 4 Resources and responsibilities required by the research data policy

The research data policy pertains to students and staff of the University of Helsinki, grant-funded researchers associated with the University as well as online research materials collected in connection with projects coordinated by the aforementioned persons or other core operations of the University. The use and management of materials collected as part of collaborative projects are agreed upon in writing at the start of the project. The use of research materials owned by third parties (such as in register-based research) is governed by the rules and regulations issued by the owner.

The implementation of the policy will incur equipment, storage and personnel costs. The costs and resources necessary for its implementation and ways to take them into consideration in the planning of operations for the coming years will be evaluated in 2015. Every University unit must take into consideration the measures and responsibilities outlined in the research data policy. Below is a brief description of the distribution of responsibilities regarding the implementation of the research data policy. More detailed descriptions of the necessary measures and responsibilities can be found in the implementation plan.

#### Responsibilities of the University leadership

- Ensuring sufficient operational resources
- Rewarding successful research data management and identifying means to gain academic merits
- Defining openness strategies

#### Responsibilities of the Central Administration (Research Affairs in particular)

- Coordinating operations (maintaining an internal networking model where Research Administration, IT Management, researchers and representatives from the Library and the Central Archives jointly develop data management practices)
- Taking into consideration the enterprise architecture (drafting service descriptions and service production processes)
- Identifying and drafting openness policies, practices and related good contract management
• Drawing up clear instructions and providing assistance in the publication and licensing of research data and the drafting of authorship and ownership agreements
• Developing the TUHAT research database to support the objectives of the research data policy
• Supporting University units (faculties, departments, independent institutes, doctoral schools) in the identification of measures required by the research data policy
• Cooperating in national development work (the Open Science and Research Initiative, and the Tutkimus-PAS long-term preservation solution)
• Monitoring international development trends
• Organising staff training in research data management
• Providing orientation materials and training for instructors

Responsibilities of the Helsinki University Library, Central Archives, the National Library of Finland and the IT Centre
• Identifying measures required by the research data policy, and planning operations and finances accordingly
• Charting technological and service-related needs
• Constructing and maintaining a research data and service infrastructure
• Building a research management tool and providing user training
• Providing training and guidance
• Implementing metadata, ontology and long-term preservation services and guidance
• Providing support and guidance in general archiving principles and procedures at the University; providing archiving instructions
• Archiving administrative documents generated in the course of research
• Agreeing on mutual cooperation

Responsibilities of faculties, departments and independent institutes
• Identifying measures required by the research data policy and planning operations and finances accordingly
• Familiarising new students and staff with good scientific practice and research data management
• Organising research data management training together with the Library and other experts
• Managing research infrastructures in accordance with the University of Helsinki Research Infrastructure Programme
• Identifying and sharing good research data management practices

Responsibilities of doctoral schools
Organising research data management training together with the Library and other experts

**Responsibilities of research groups and principal investigators**
- Observing good data management practices
- Improving the repeatability of research (ensuring appropriate referencing and metadata production)
- Familiarising new students and staff with good scientific practice and research data management
- Identifying and sharing good research data management practices

**Responsibilities of individual researchers**
- Observing good data management practices
- Drafting data management plans
- Drafting and implementing plans pertaining to the quality and lifespan of research results
- Allocating resources to the publication and re-use of research materials
- Agreeing on authorship and ownership in accordance with University guidelines
- Licencing results in accordance with University guidelines.
- Enhancing opportunities for the repeatability of research (ensuring appropriate referencing and metadata production)
- Deciding on access rights to research data
- Publishing and opening up research data, when possible
- Identifying and sharing good research data management practices

**5 Definitions**

Below is a list of definitions related to the research data policy. They will be clarified further during the Open Science and Research Initiative. The clarifications will affect the contents of research data policy-related websites and intranet pages.

**Open data**

The research data are available in a usable and intelligible format online with no restrictions as to use or editing. The research data must always be traceable to be reliable.

**Open Science**

The notion that all research methods, results and data must be carefully documented and available to all users.

**Creative Commons licence**
Creative Commons licences are used for forfeiting parts of copyrights and granting viewers or other users certain rights to use the work in question.

**Digital information**
Materials stored in or converted into digital format (such as scanned documents) and related metadata.

**Digitisation**
Converting analogue materials (such as printed documents) into digital format.

**Holding period**
A clause regulating that open access to data begins only after a set waiting period (such as six months from publication).

**Primary use**
Primary use refers to the original purpose of the data indicated in the research plan or otherwise.

**Secondary research/re-use**
Secondary research refers to cases where primary or other users utilise the research data for a research purpose other than the one indicated in, e.g., the original research plan.

**Data curation**
Data curation refers to measures related to maintaining and updating the contents of research materials.

**Reference information**
Descriptive metadata appended to the materials by the party handing them over for preservation.

**Metadata**
Metadata refers to descriptive information on the context, content and structure of data, as well as their management and processing, throughout the lifespan of the data. This information can be used, e.g., for searching for, locating and identifying materials. Metadata is essential in terms of finding, cataloguing and using materials. Metadata contains both descriptive information and technical system metadata.

**Long-term preservation**
Long-term preservation refers to preserving the structure, content and integrity of online materials in the long term. In Finland, the National Digital Library is responsible for long-term preservation.

**Raw data**
Raw data refers to unprocessed data generated in multiple formats that can undergo several transformations and rounds of processing in the course of their lifespan. In the natural sciences, raw data are often generated by research instruments such as telescopes, satellites, synchrotrons and, increasingly, computer simulations. In the humanities and social sciences, raw data are generated through, e.g., interviews and questionnaires.

**The fourth paradigm**
The fourth paradigm refers to data-intensive research enabled by modern computers, software and research methods. It provides researchers with tools for conducting experiments and making calculations in greater magnitudes and numbers and on a greater scale than ever before.

Data management plan

A data management plan is both a plan and mechanism for guiding the collection, quality assurance, analysis, publication and storage of digital information. The Academy of Finland, for example, requires that the data management plan determine how research materials are acquired, used and stored and how they are made available for re-use after the end of the research project.

Data infrastructure

Data infrastructure is the part of e-infrastructure containing all mutually compatible basic services and tools for the production, storage, distribution and utilisation of data.

Data mining

Data mining refers to a number of methods for finding relevant information from large datasets. Data mining can be applied very extensively, as it can be done using raw data. Typically, the data being mined can include industrial process measurements, customer database extracts or web server logs. The definition of data mining does not pose any limitations as to the mining methods used.

Subsequent use of data

Subsequent utilisation refers to private and legal persons being able to copy, edit, publish and distribute data for different purposes. Much of the raw data and research materials collected with public funding can be and is utilised in both research and the development of different everyday products and services.

Privacy protection

Privacy protection refers to the protection of personal details from unlawful use when handling personally identifiable information.

Data protection

Data protection refers to administrative and technical measures for ensuring that information is only available to authorised persons or parties, that the data cannot be altered by unauthorised persons or parties, and that data and information systems can be utilised by authorised persons or parties. Related concepts include access control, confidentiality, privacy, authentication, integrity and security.

Research data management

Good research data management is systematic. It ensures that research data are handled in accordance with good scientific practice and that any data generated through research are preserved in accordance with the criteria for scientific information.

Research materials

Research materials include all materials produced, edited and used in scientific research as a basis for research results.
Research data

Research data refers to research materials in digital format (requires systematic electronic storage). They can be stored in different file formats.