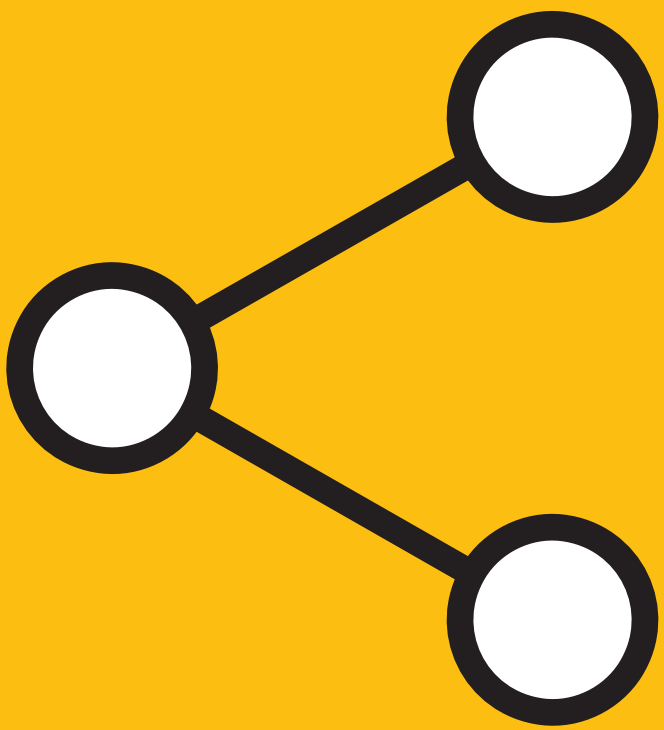
Open Access to Research Data – Status, Issues and Outlook



norden
NordForsk

POLICY PAPER 1 | 2016

norden

NordForsk

NordForsk Policy Paper 1 –2016

Open Access to Research Data
– Status, Issues and Outlook

NordForsk, 2016
Stensberggata 25
N–0170 Oslo
www.nordforsk.org
Org.nr. 971 274 255

Design: jnd.no
Printed by: 07 Group

ISSN 1504-8640

Open Access to Research Data – Status, Issues and Outlook

Table of Contents

1. Introduction **6**

The Impact of Improved Sharing and Re-Use of Research Results 6

Open Science 7

Open Access to Research Data 8

2. Analysis and Issues – Potential Frameworks for Discussions and Actions **14**

The Importance of Structured Data Management 14

Research Data Issues 14

Barriers and Enablers 16

Research Funders and Academia 18

Data Infrastructure 19

The International Arena 20

3. Current status **24**

Denmark 24

Finland 26

Iceland 28

Norway 30

Sweden 32

The EU 34

The UK 35

The US 35

4. Main Findings and Outlook **38**

Main General Findings 38

Main Findings Concerning the Nordic Countries 38

Some Potential Actions at the Nordic Level 39

References **40**

Preface

This report responds to the request by the Nordic Committee of Senior Officials for Education & Research (ÄK-U) within the Nordic Council of Ministers to NordForsk to produce an overview of current knowledge within the area of Open Access to research data and identify possible joint frameworks and the corresponding implications. The full request is formulated as follows: "ÄK-U beslutade att ge NordForsk i uppdrag att initiera en kunskapsöversikt inom Open Access till forskningsdata, som beskriver utvecklingen i de nordiska länderna och EU, och identifierar möjliga gemensamma ramar (juridiska, tekniska, ekonomiska etc.), samt implikationerna av dessa för OA till forskningsdata."

NordForsk has decided to appoint a project leader, the Director of the Nordic eScience Initiative (NeGI) at NordForsk, Sverker Holmgren to handle the request from the Nordic Council of Ministers. The work of the project leader has been supported by a reference group, consisting of Juha Haataja (Finnish Ministry of Education and Culture), Ásdís Jónsdóttir (Icelandic Ministry of Education, Science and Culture), Hanne-Louise Kirkegaard (Danish Ministry of Higher Education and Science), Jarkko Siren (European Commission, Research Executive Agency), Roar Skålin (Research Council of Norway) and Anna Wetterbom (Swedish Research Council).

The project leader has prepared this report in response to the request put forth by the Nordic Council of Ministers to NordForsk. The reference group has commented on draft versions in a physical meeting and by email, providing information on the status within their respective countries and input on the factual parts of the report.

Oslo, June 2016

Gunnel Gustafsson
Director of NordForsk

1. Introduction

1. Introduction

The Impact of Improved Sharing and Re-Use of Research Results

Leaders, policymakers and politicians in the research and innovation field are currently intensely discussing how to further improve the sharing and re-use of research results. The issue as such is not new; similar discussions have been critical to the evolution of modern research since the Renaissance. However, the rapid development of information and communication technology (ICT) has now made it possible for research results to be shared in fundamentally new ways and to an extent which is already having a profound effect on research. New computer-based tools provide powerful means for retrieving information and knowledge from the digitally available resources – often in revolutionary and unexpected ways. Digitalisation is changing the *modus operandi* of research, providing new possibilities for geographically distributed collaboration and the sharing and re-use of research results. This development is also a fundamental enabler for a much faster and potentially even more extensive impact of research on the progress and prosperity of society as a whole.

In many fields of research and innovation the amount of data produced is growing at an exponential rate. Research and innovation in a growing number of areas is becoming data-driven, meaning that data represents the most valuable asset in projects and that analysis of digital datasets is becoming the main tool. Many statements on the importance of this process have already been made. These range from one-liners like "*data is the new oil*", focusing mostly on the innovation aspects and potential impact on society of re-use of data in general, to more elaborate statements focusing on the importance of data for research itself and its internationalisation.

A fundamental early reference, pointing to the importance of the development, is the book "The Fourth Paradigm: Data-Intensive Scientific Discovery". The similarly fundamental reports, "Riding the Wave: How Europe can gain from the rising tide of scientific data" and "Science as an Open Enterprise", both outline a series of policy recommendations for important actors in the field. Many important international organisations have pointed to the need for structuring and opening up the research data system. Here, the early work by OECD presented in the report, "Principles and Guidelines for Access to Research Data from Public Funding", has been an important contribution, recently followed up by the report, "Making Open Science a Reality", by the same organisation. Another important step was taken in the European Commission publication, "Recommendation on Access to and Preservation of Scientific Information", which sets targets for openness in all European research and also provides recommendations to the Member States of the European Union.

Currently, many national governments are actively engaging in the process of achieving better exploitation of research data. Moreover, European and global organisations, e.g. Science Europe, LERU, EARTO, ESFRI, e-IRG, UNESCO, OECD, and the Global Research Council are discussing these issues, and more focussed global efforts such as the Research Data Alliance (RDA) and CODATA are actively exploring the area. Numerous studies and reports with slightly different perspectives and aims are being produced, and most of these documents include recommendations to different stakeholders. These recommendations often have the same or similar goals, but still contain subtle differences that make the policy landscape complex and difficult to navigate. Buzzwords like *Big Data* and *the Data Society* are also widely used in many different settings. The use of such poorly defined concepts might blur serious discussions, but the "Big Data Hype" also means that a discussion on sharing and re-using research results has entered communities where such issues were rarely voiced before. The current status of the discussion, with many actors, initiatives and reports, can be seen as a manifestation of the fact that exploitation of research data is deemed to be very important both within research and for other parts of society, but that the

development is still only in its initial phase. A consolidation process is both needed and foreseen. A recent report, the "ERAC Opinion on Open Research Data", has been produced by a task force within the European Research Area Committee (ERAC). This publication is the result of an extensive effort and contains a detailed survey of the field, including a discussion on terminology, and it presents a set of recommendations covering four different areas: training of stakeholders and awareness raising, data quality and management, sustainability and funding, and legal issues. The ERAC report also includes several annexes with supporting material, providing a solid basis for further discussions on most issues related to Open Access to research data discussed further below.

Open Science

A main goal for a future research system is that research results and recourses are used in an optimal way for society as a whole. In this context, there is currently a strong movement towards the *democratisation of research data*, i.e. research data produced as a result of public funding should be considered to be part of the *common good* of society. The research data should then as far as possible be made openly available to anyone who is interested in it. In essence, two fundamental arguments for openness are commonly brought forward:

- **Improved quality of research:** A transparent presentation of research results and methods facilitates maintenance of a high level of integrity in research by enabling other researchers (and other parts of society) to scrutinise the research process and the results critically, reducing the risk of poor research, unethical behaviour and falsification of results.

- **More "bang for the buck":** Open access to research results and resources enables both other researchers and other parts of society to utilise the results and resources for further research, innovation, commercialisation and other societal improvements. This promotes faster accumulation of the research record and avoids duplication of research efforts.

The insight that an open model for research is preferable both for research itself and society in general has led to the development of the *Open Science paradigm*, promoting an open operating model for all publicly funded or managed research efforts. Here, the key objective is to openly publish research articles, research data, methods, software and educational resources as quickly as possible – for others to use and re-use. Open Science is often presented as more than just a set of tools, methods and recommendations. Instead it is considered to be a mode of thinking where open-mindedness, networking and international collaboration are key elements. Sometimes, the discussion is expanded even further to the concept of *Responsible Research*, including not only Open Science but also issues such as public engagement, gender equality, ethics, and science education.

A comprehensive definition of Open Science, established and maintained in an open and collaborative process in the spirit of the Open Science paradigm itself, can be found on Wikipedia: *Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society, amateur or professional. It encompasses practices such as publishing open research, campaigning for open access, encouraging scientists to practice open notebook science, and generally making it easier to publish and communicate scientific knowledge.* A more elaborate definition of the Open Science concept can be found in e.g. the OECD report "Making Open Science a Reality". The EU, too, is a main actor in the field of Open Science (earlier under the name Science 2.0) and has taken several concrete actions and produced a number of important documents, e.g. the background document, "Public Consultation: 'Science 2.0': Science In Transition".

Open Access to Research Data

To give the necessary background for a discussion on issues and common frameworks for Open Access to research data, it is beneficial to isolate the fundamental concepts and elaborate briefly on the definitions:

[Digital] **Data:** Digital data represents other forms of data using some specific digital encoding. Using advanced encodings, all sorts of potentially very complex input can represented. Already from this it is clear that it is critical to amend a digital dataset with *metadata*, describing at least the encoding. However, normally much more metadata, i.e. data describing the data, is needed. This usually includes technical information, e.g. on data structures, licensing terms, and which standards the dataset conforms to. The metadata may also state why and how the dataset was generated, who created it and when, providing information on the *data provenance*.

Research Data: Often, the discussion on research data is restricted to *result data* (or output data) produced by the research processes. This provides a distinction from *source data* (or input data) which is data that already exists in general even if the research at hand had not been initiated. The source data could have been obtained for another purpose (e.g. data for public administration, clinical data or weather data), it could be digitalised collections of objects and texts, e.g. made available on the internet, or it could be data with its origin in other research processes.

Another distinction is often made between the data (including metadata) directly needed to reproduce and validate the results presented in a scientific publication and other data (including metadata) that is produced as an intermediate result or as a bi-product of the research activity. In this context, data directly produced by different experiments, simulations or observation processes is often referred to as *raw data*. Sometimes, e.g. when discussing issues of ownership and intellectual property rights, there is also a need to separate between the data items as such and the *datasets*, *databases* or *data collections*, referring to the collective properties of the data, including metadata.

Open [Access]**:** It is widely recognised that some research data cannot be made fully openly accessible. For example, research data should not be made accessible if doing so would violate laws and regulations or could pose a threat to the security of individuals or society. Furthermore, some research data must be protected to make sure that the research conforms to statutory frameworks for personal privacy and ethical guidelines. In some research fields the discussions on the access policy for specific research datasets, e.g. containing data for human individuals, quickly turns complex. Here, legal and ethical expertise and truly serious consideration are needed.

The overall goal of Open Access to research data is to provide the freedom for any user to use the data in any way, for any lawful purpose, without introducing any unnecessary legal or technical barriers. To fully specify the conditions for use of the data, a *licencing system* may be used, referring to an appropriate definition of the ownership of a data collection. Here, the owner states which specific licence should govern the usage rights. One standard licence often used in open access settings is the Creative Commons Attribution license (CC-BY). This licence states that any user may copy and distribute the object considered and make derivative objects based on it only if they give the owner credit in the manner specified together with the object itself. Another interesting licence is the Creative Commons Public Domain licence (CC0), attempting to waive ownership rights altogether and stating that the object considered is simply globally available in the public domain without any restrictions.

Research data can be produced in settings where commercialisation issues are of integral importance, e.g. if the data is generated in collaborative and possibly jointly funded research projects where a commercial entity is a contractual partner. Again, legal expertise might be needed to set up the

frameworks for management of intellectual property rights and publication of the research results, including the research data. However, established mechanisms like patents may often be used to protect the commercialised product of the project, normally enabling the publication of the underlying research results, including data, at least after a certain period of time.

In some settings, research data can also be sensitive, e.g. for reasons of personal privacy or societal security. Such data may of course not be made fully openly available, and existing legal and ethical frameworks will restrict access. However, in some cases it may be possible to provide anonymised versions of the data in an Open Access setting. Furthermore, it is often possible to provide open access to all or some of the metadata describing a dataset or a certain data item without revealing any sensitive information. This might still provide very valuable information, enabling the re-use of the actual sensitive research data after taking the necessary actions needed to fulfil the requirements of privacy and ethics.

Many definitions of open access to research data include a statement that the access should not only be open but also free of charge. This can be based on an argument that free access facilitates the democratisation of research and bridges digital divides. Other definitions instead state that the data should be provided at the lowest possible cost, preferably at no more than the marginal cost of the data retrieval.

[Open] **Access:** Merely providing storage for research data and providing basic access services does not lead to a functioning research data system. Several other issues must be resolved to enable re-use by a wide range of users in research and society. Here, it is often stated that the *FAIR principles* must apply, i.e. the data must be Findable, Accessible, Interoperable and Re-usable. These principles include aspects such as digital data identifiers and standardisation efforts and frameworks. The data should also be assessable, meaning that potential users must be able to judge the data quality and understand the data provenance. Sometimes quality is guaranteed by formal adherence to quality and/or procedural standards, but for research data it is more common that quality is assessed by the level of trust the potential user has in the data provider and the procedures used when creating the data. An important aspect of re-usability may be that different sources of data might have to be combined, requiring *data integration* to provide the user with a unified view of several datasets. Digital data identifiers are also a precondition for data citations and provenance, which in turn are requirements for a reward system that gives credit to producers of research data.

Another important aspect of data access is the *data lifetime*. Enabling access to data for a limited period of time may be quite straightforward, especially if the data is only used internally in a research project. Guaranteeing that the data should be *preserved* for a long time (or “indefinitely”) requires that the value of the data is maintained over time – a process that requires persistence and knowledge. Often, the term *data curation* is used to describe the data management activities related to organisation, validation and annotation of data such that the dataset remains archived and preserved. Additional services are then also needed to make the data available and re-usable also in a long-term perspective.

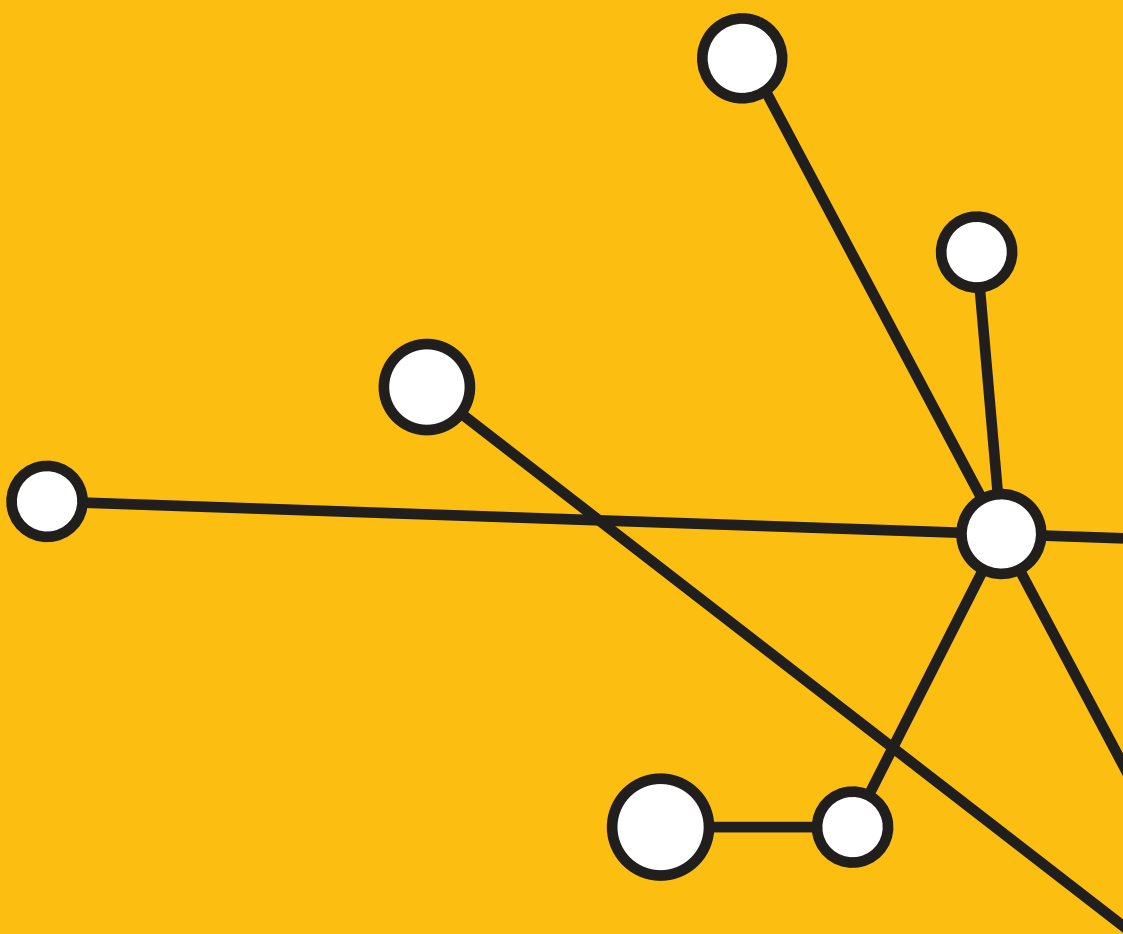
[Data] **Infrastructure:** As a foundation for data sharing and re-use, sustainable services for enabling access to, storing, preserving and curating large amounts of data need to be in place. The data infrastructure also provides the technology platform needed to allow the combined use of digitally enabled resources, across topical and national boundaries. It forms the platform for services for accessing the data, for actually storing and preserving the data and metadata, and for analysing the data to extract useful information. A data infrastructure normally includes internet connectivity services for people and data in a global context, data storage services for storing and preserving data, and computing services for analysis and modelling. A data infrastructure also needs to include relevant user support services and competent staff for providing these services.

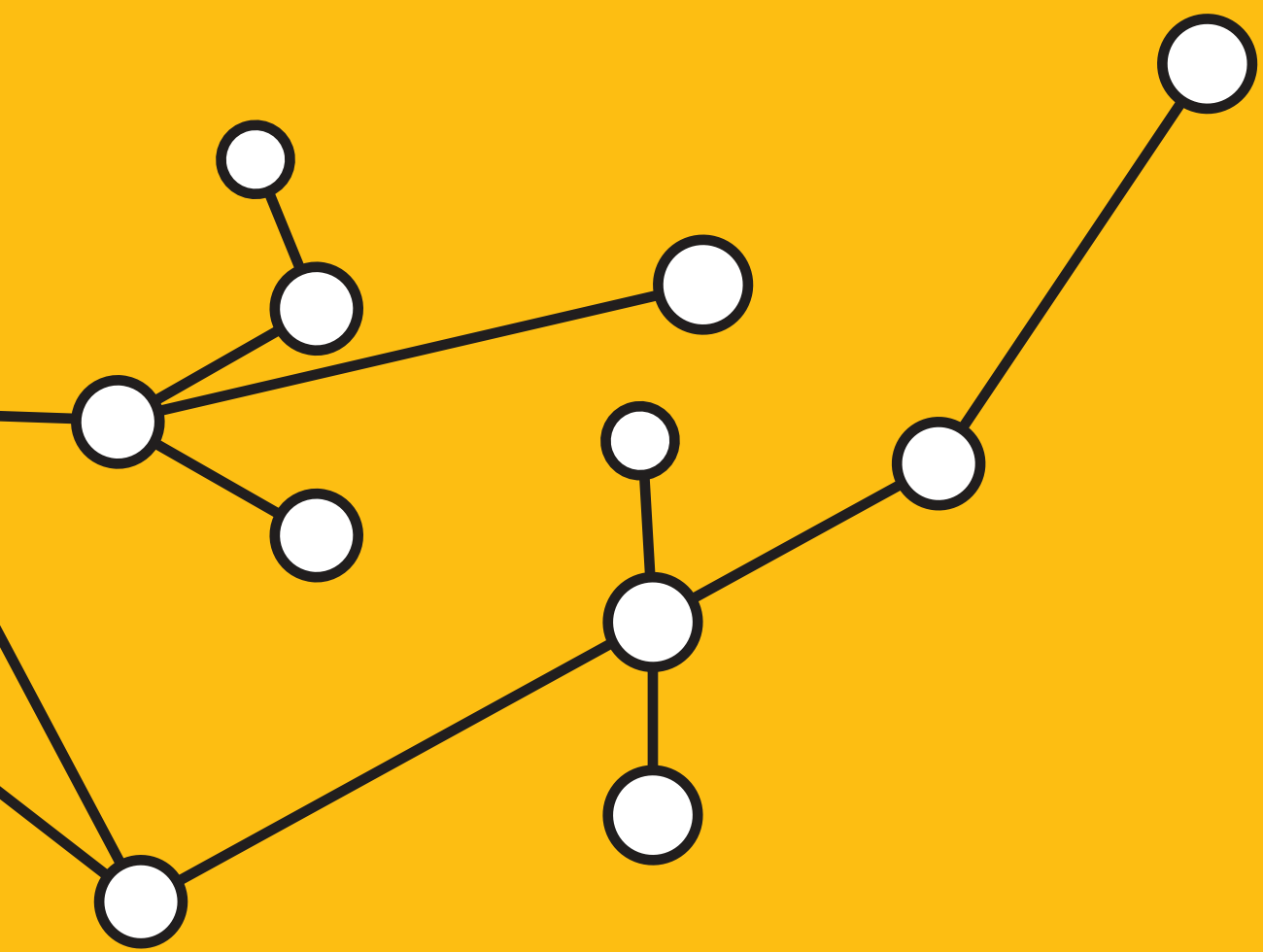
Relation to other Types of Research Results: The classical means of making research results available and re-usable is via publications in text form in peer-reviewed journals, conference proceedings and sometimes also research monographs/books. Here the notion of *Open Access* is well established and often also accepted by research communities and institutions. Many funders and research institutions demand that the publications are made openly available either in journals that provide open and free access to the articles directly or by parallel publication of the articles in open repositories provided by institutions or other actors. Some journals and repositories provide mechanisms for attaching datasets to the publications as supplementary material. Also, some research domains, e.g. Astronomy and Climate Research, have themselves established repositories for storing research data and mechanisms for connecting datasets to publications.

Another example of a well-established and accepted concept is that of *Open Source Software*. Here, the source code for a software or tool is made available by the owner of the software with a license specifying how it may be used. The licence might e.g. state that anyone may study, change, and distribute the software for any purpose. Often, open source software is developed in a collaborative public manner, and it is argued that the open source model generates a more diverse scope of design perspectives and opportunities for quality control than any single organisation is capable of developing and sustaining by itself.

To fully implement the Open Science paradigm, it is argued that *Open Notebook Research* is needed, making the entire primary record of a research project publicly available online as it is performed. This involves making the personal, or laboratory, notebook of the researcher available along with all raw and processed data, and any associated material, as this material is generated. This is the logical extreme of a transparent approach to research. For example, it explicitly includes making failed, less significant and otherwise unpublished experiments and research procedures openly available. Such a system would e.g. reduce the risk of duplication of research efforts, and it would also have to be built on a structured approach to documentation of the research process, including research data management. However, implementing Open Notebook Research would significantly increase the workload of individual researchers and groups, and it can only be implemented based on a fundamental restructuring of the research funding and research merit systems.

2. Analysis and Issues – Potential Frameworks for Discussions and Actions





2. Analysis and Issues – Potential Frameworks for Discussions and Actions

The Importance of Structured Data Management

A structured approach to data management is a prerequisite for modern research and a fundamental building block for enabling the re-use of research data – even if such a framework is not implementing the full Open Access regime. Without a structured approach, research data will easily be lost or become impossible to re-use in other ways. Here, a wider discussion on the re-use of research results and Open Access to research data can be facilitated by bringing forward good examples of data management practices in research areas where such actions have already emerged because of requirements of research.

Many reports and initiatives have concluded that data management and aspects of data re-use should be considered during the entire life cycle of a research project, i.e. during planning, execution and afterlife. The practical implementation of a structured data management framework is greatly facilitated by a mindset among all the involved actors where research data is viewed as an important and valuable asset in the research process. This includes not only the researchers themselves but also policymakers, research funders and organisations hosting the research efforts.

A way of ensuring that a serious approach to data management is followed is to require that all research projects formulate a *data management plan*. This is a structured document that describes how a research activity will handle its data both during the project and after it is completed, presenting how the data will be managed in the present and potentially prepared for preservation and re-use in the future. Developing a data management plan is a critical element in the planning phase of a research project. It then normally needs to be revisited during the implementation and execution of the research project and possibly also after the research project has finished. The plan should describe how the many aspects of data analysis, data management, metadata generation, and data preservation will be dealt with, and it should discuss what data to store, preserve and make available for re-use. It should also discuss the ownership of the data collections, if relevant including intellectual property rights, and include an account for the relevant actors and roles in the data management and data stewardship processes. Furthermore, a description of governance structures and funding streams for the data management should also be considered. Many examples of data management plans exist, and today consolidation at the national and international level is probably needed to avoid confusion among researchers and research organisations.

The *Data Access Policy* is another central document describing the policies and rules that apply for use and re-use of the data generated. Developing such a policy might be part of planning a (usually large) research or research infrastructure initiative, but it may be more common for the policy to be developed and decided on at an institutional/organisational level and thereafter applied to all research activities within that institution/organisation.

Research Data Issues

Diversity. Research data comes in very many different shapes and sizes and needs to be handled on many different time-scales. The data is also used in many different ways. High-level policy initiatives may impose overarching rules, e.g. specifying that a data management plan is required for research projects. However, the practical implementation of such a requirement quickly comes to a level of detail

where the specifics of different research fields and projects must be taken into account. History shows that attempts to set up one-size-fits-all solutions are prone to fail, both at the policy level and when it comes to practical implementation. In some contexts, the concept of the *four Vs*, i.e. Volume, Variety, Velocity and Veracity (truthfulness and provenance) of data is used to structure the diversity issues as well as to bring some clarity into what is actually meant by the term *Big Data*.

Deciding on which Data to Store and Preserve. A basic issue in discussions on research data is deciding on which data should be stored and made available during a project, and then which data should be preserved and made available for re-use after a project has ended. Practical and economic realities almost always make it impossible to store all research data and make it available for re-use, especially for extended periods of time.

A main argument for openness in research is that it facilitates reproducibility and quality control. In this context it is natural to require that the research data that underpins a research article or other publication is openly provided, e.g. as supplementary material to the publication. However, it is often difficult to specify at what level of analysis the data has to be made available for the results to be considered reproducible. Moreover, reproducibility normally requires a specification of the software tools that were used for the analysis (down to a level of version number and possibly also providing a detailed description of the setup of the computer system used) and an exact specification of the procedures that were undertaken when managing and analysing the data. Providing documentation and an open account for this will require that a more complete Open Science framework is put in place.

Much of the data produced in a research project may not be directly attributable to a specific scientific publication. However, it may still be highly valuable as source data for other research projects, as input in innovation processes and for the general public. For research and society to take full advantage of the major investments in research, steps must be taken to expand the ability to re-use a wider set of research data than that which can be directly connected to research publications. Such re-use needs to be enabled both within disciplines and across national borders and domain boundaries. It may be highly relevant to store and preserve both raw data and datasets from experiments that were deemed to be less interesting for a specific publication. In general, it is of course impossible to know which datasets might become valuable in the future and, on the other hand, it is in practice normally not possible to store and preserve all data from a research project. Here, the collective knowledge of the research fields and the competence of the researchers should be combined with discussions among policymakers and research leaders to arrive at solutions that provide sufficiently good coverage of future opportunities within the financial restrictions imposed. Here, evidence-based arguments for the importance of data sharing for research and society are needed. The recent report, "The Data Harvest: How sharing of research data can yield knowledge, jobs and growth", presented by RDA, elaborates further on this issue and presents concrete figures on how innovation based on sharing research data can indeed yield knowledge, jobs and growth in Europe. Here, the vision is of researchers and innovators openly sharing data across technologies, disciplines and countries to address the grand challenges of society.

Defining the Legal Framework. A fundamental issue in discussions on research data is the ownership of the datasets and collections. For written text, there are well-established national (and international) notions of *copyright* governed by legal practice. However, these systems may not be fully applicable or adapted to digital research data. For example, in some cases the data itself cannot be connected to a legal notion of ownership, e.g. if it contains measurements of "facts" in nature. Instead, the concept of ownership may be applicable to the data collections, referring to the collective properties of the data, and how the data is organised in the collections.

The ownership of research data collections is important for several reasons. In general, ownership can indicate that the owner is exclusively entitled in some way to take advantage of the data collection.

However, other statutory frameworks may at the same time indicate that research data should be made openly and freely available for re-use. The ownership of research data may also come with obligations. For example, a legal framework regulating access to information owned by public entities might be seen to be applicable to digital research data as well. This could then imply that the public entities owning the data collections are obliged to provide open access to it, and possibly also to preserve it for the future.

In this context it is important to note that the legal framework for research data may not be understood in the same way by different actors in the research system, and the legal ownership structures may be different from the de-facto ownership established by tradition or community policies in research and research institutions. This may cause difficulties and true conflicts at different levels, especially in the current situation where the importance of research data is growing rapidly.

The existing national frameworks for ownership and access to research data differ widely. In many countries further work on issues connected to ownership of research data is also often needed, e.g. to arrive at clear and generally accepted definitions of ownership and intellectual property rights for research data collections as well as an effective statutory framework for access and re-use. Some intergovernmental organisations, e.g. the EU, have legislative powers and are taking concrete actions towards a more coordinated and consistent situation across national borders. International organisations can also take important steps to clarify relevant issues and point to opportunities for coordination even without formal legislative action.

Barriers and Enablers

The Researcher's Perspective. Several recent reports identify a number of perceived or real barriers to data sharing and re-use among researchers. The same reports also often list enablers of data sharing to provide an indication of where further development might be needed. These reports are often based on surveys among researchers. The following lists of barriers and enablers are extracted from the report, "Sowing the seed: Incentives and motivations for sharing research data, a researcher's perspective", presented by the European Knowledge Exchange Initiative. Here, the researchers interviewed identify the following barriers to data sharing and re-use:

- fear of competition, of being scooped and therefore reduced publication opportunities;
- cost in both time and money to prepare data and documentation for sharing and absence of funding to do so;
- absence of professional rewards for data sharing;
- lack of standards and data infrastructure;
- ethical and legal constraints.

The same report also lists the following enablers for data sharing:

- data sharing expectations of funders and journals;
- peer expectations and sharing practices in the research community;
- availability of data repositories and standards;
- desire to showcase data quality;
- researchers' data management skills;
- organisational support;
- acknowledgement received for data sharing;
- data publication and metrics.

Defining Roles, Mandates and Interfaces. Historically, a single actor (e.g. a research group) has often taken on many, or even all, roles related to a specific dataset. Experience shows that this frequently results in inefficient, costly and/or unsustainable solutions for data management and access. The data does not become easily re-usable – especially not for other fields of research or for society in general. It is not uncommon that the data is in effect not handled at all, which leads to it being lost after or even during a research project.

Actors specialising in different tasks can more easily exploit synergies and set up cost-efficient implementations and operation of services and solutions. Furthermore, specialised actors can maintain the level of competence and commitment that is needed to provide easy re-use of research data within domains and across national and topical boundaries. The development towards a research data system with a set of well-defined roles performed by a rather small number of specialised actors is a maturation process, it can e.g. be compared to the system for research publications with actors for writing, publishing, providing access to and preserving the printed material which has developed over many hundred years. Here, initiatives on Open Access to scientific publications had an earlier start than discussions on research data, and the development has proceeded further.

In the current situation where the importance of research data is growing rapidly, it is natural that many different actors take an active interest in the field. This commonly results in the initiation of overlapping and potentially competing initiatives within organisations, between different organisational levels, and between local, national and international efforts. Here, both leadership among politicians and policymakers and genuine, formalised support in the research system are needed to arrive at consolidated and generally accepted solutions.

A common conclusion is that a well-functioning system for research data should consist of a well-defined set of specialised actors, taking one or more role each. Here, the mandates for the different actors should be specified, and their means of interaction should be sufficiently formalised. A relevant set of roles to be considered for such a system are:

- the data owner – the actor that formally owns the data collection. This is also the actor who is responsible for compliance with the legal and ethical framework applicable to the data collection;
- the data steward – the actor that ensures that the data is of high-quality, accessible, discoverable etc., in a consistent and sustainable manner;
- the data services provider – the actor that provides services for using the data;
- the data access provider – the actor that provides access to the data;
- the data infrastructure provider – the actor that provides eInfrastructure services and resources for storing and/or preserving the data.

Defining Funding Streams and Governance Structures. Providing mechanisms for re-use of research data gives rise to apparent costs in the budgets of research groups, universities and research funders. Such costs obviously need to be considered in an overall perspective, taking into account the benefits that come from re-using the research data. There is an obvious risk that different actors revert to a discussion focussing more on identifying the budgets in which the costs should be visible rather than how the opportunities for improved efficiency and quality of research should be valued. Leadership among politicians and policymakers and an integrated view of the research funding system are needed to arrive at the best possible solutions for research and society as a whole. Again, evidence-based arguments for the importance of data sharing for research and society are called for, and such arguments must be dealt with in a context of how the sharing of research data can improve research and be valuable to society in other ways as well.

Many definitions of Open Access to research data include a statement that the access should not only be open but also free of charge. This can be based on an argument that free access facilitates the democratisation of research and bridges digital divides. Other definitions state instead that the data should be provided at the lowest possible cost, preferably at no more than the marginal cost of the data retrieval.

The definition of roles, actors and interfaces goes hand-in-hand with defining attached governance structures and funding streams to ensure efficiency and sustainability for all parts of the research data system. Here, it is essential that the costs and properties (e.g. relevant quality parameters) for different services and procedures are made transparent so that different options for implementing them can be compared. This could for example lead to one commercial actor being preferred for fulfilling one or more of the roles. The leadership from politicians and policymakers needs to ensure that the research data system and governance structures are built using an integrated perspective, providing the best possible solution for the research system and society as a whole. It is important to note that such a system may not seem optimal from the perspective of some of the individual actors.

Research Funders and Academia

The effect of processes where "money talks" should not be underestimated. Based on the policy frameworks and tasks defined at the political level and within their own organisations, research funding organisations such as national research councils have always played a central role in influencing the research system. This presents both opportunities and responsibilities. Organisations that distribute public funding have an obligation to maximise the benefits of their grants towards goals that have been set by the politicians. Private research funders also have similar obligations, often towards similar goals. Such goals can range from providing support to the best possible curiosity-driven basic research to e.g. directly maximising the short-term impact on innovation and growth in companies. In the context of re-use of research data (and of research results in general), it can be argued that there is a significant benefit to be found in a wide span of such goals. Based on initiatives at the political level, funding organisations have a responsibility to take clear actions to investigate the field further and incorporate the relevant data management and data re-use aspects in their processes for research strategy, evaluation and funding. Furthermore, the research community expects research funders to provide a unifying and strategically supported approach to issues such as research data. This requires research funders to build competence to be able to meet the needs in this field.

Many funding organisations have published a policy for Open Access to research data and also impose a structured approach to research data management by e.g. requiring that a data management plan is provided in research proposals. However, to fully deal with the opportunities and issues related to research data, a more integrated approach and complete set of actions needs to be considered. Such an approach must be based on relevant decisions in relevant political and policy-making bodies and might include:

- raising awareness of the importance of research data both towards society, including the political system, and towards the research community;
- including aspects of data management and Open Access to research data in strategic planning and strategies;
- participating in, and potentially leading, relevant policy activities at national and international levels. This includes discussions and agreements with relevant actors on issues such as data access modes, data infrastructures and long-term preservation;
- adopting a policy for Open Access to research results or Open Science, with specific elements discussing data management and re-use of research data;

- including aspects of data management and Open Access to research data in the planning of calls for proposals. This includes actions towards defining roles and stakeholders, governance structures and funding streams;
- including aspects of data management and Open Access to research data in evaluations of research proposals and individual researchers. This includes ensuring that instructions to scientific staff and external evaluators contain the relevant guidelines and that evaluation panels comprise the relevant expertise;
- including aspects of data management and Open Access to research data in contracts and other agreements on research funding;
- including aspects of data management and Open Access to research data in reporting of research projects and other efforts related to research.

Universities are assumed to be governed by principles established to ensure the independence of the research system and the integrity of the research community. A traditional academic governance system is sometimes seen as conservative and hesitant to embrace new research policy initiatives. In fact, this might be considered to be an essential feature of an independent and collegial governance structure to safeguard the independent evolution of research. However, universities and other organisations performing academic research are also research funders, using public funding to support their research activities. Academia shares an obligation to maximise the benefits of e.g. the basic university grants towards society as a whole. This means that academia, too, has a responsibility to further investigate the field of data management and Open Access to research. Academia should take action to evaluate the benefits of re-use of research results in general, and especially with regard to research data, and incorporate relevant data management and data re-use aspects in the internal processes for strategic research planning, evaluation and funding. In particular, academia has a responsibility to discuss the perceived barriers to data sharing and re-use presented earlier, including the incentive structures for researchers to take part in data management and data publication activities. Academia also bears a large responsibility with regard to providing reasonable career paths for persons involved in data management and for providing the education and training needed both for further research activities and for other parts of society.

Data Infrastructure

The digital data, including associated metadata, needs to be managed, stored and preserved in a cost-efficient and effective manner, with appropriate quality and safety assurances. Access to the data across borders and domain boundaries must also be secured to enable new, potentially unexpected exploitation. For this, a suitable eInfrastructure framework is required, providing versatile services and tools for both data management and access. The short document, “Summary of Policy Recommendations Drawn from the e-IRG Blue Paper on Data Management”, drawn up by a working group established by ESFRI and e-IRG, presents a set of recommendations on data infrastructures with a focus on the needs of large-scale international research infrastructures. However, the recommendations are in general also applicable in other settings and on other scales and provide a starting-point for further discussions.

Some elements of a data infrastructure are especially fundamental. For example, a system with *persistent data identifiers* is needed for referring to different datasets in an unambiguous way over long periods of time. Furthermore, the technologies for implementing appropriate security and data protection policies must be operational at all times, and the corresponding services need to be continuously maintained. In general, the data needs to be managed, stored and preserved in a cost-efficient way, with appropriate quality and safety assurances. Access to the data across borders and domain boundaries must also be secured. The development of data infrastructures must be

complemented with research-policy and coordination efforts to accomplish the above goals in practice. These processes must be driven by the needs of research and society, and many actors have to be orchestrated to complete the build-up and operation of data infrastructures rapidly, while still maintaining sustainable and cost-efficient solutions.

The International Arena

The development in the Nordic countries is very much a part of an international development. The activities in the European arena are particularly important with regard to policy-making and implementation in the field of re-use of research results in general as well as in the specific field of Open Access to research data. It is important for the viewpoints of actors in the Nordic countries to enter the discussions on future policy actions at an early stage to be able to influence the development. Good examples from the Nordic countries should be made visible, and in some cases there is an opportunity to take the lead both in policy discussions and in practical implementation. Here, the Nordic countries are small and have limited resources, and an exchange of knowledge and ideas at the Nordic level can facilitate active participation in initiatives at a wider international level. The presence of Nordic researchers in bottom-up initiatives like the Research Data Alliance is currently limited, and actions may also be taken to increase the awareness of such activities.

3. Current Status

3. Current Status

In general, the Nordic countries started discussions on re-use of research results in general, including Open Access to research data, early on. However, there are large differences between the countries when it comes to the approach taken and the further development of the discussions into policies and operational initiatives. In some of the Nordic countries the progress towards Open Science lags behind major countries globally, like the US and the UK, while in e.g. Finland a major effort has been initiated with the explicit and ambitious goal to become one of the leading countries in openness of science and research by the year 2017.

A main driver of actions in the Nordic countries is the EU Recommendation on Access to and Preservation of Scientific Information presented in 2012, but again in e.g. Finland the implementation of eInfrastructures for research data had already started well before this. The perceived slow development in some of the Nordic countries might at least in part be the result of careful strategic planning processes among stakeholders and serious concerns about the complexity of the issues and the costs. However, this may also be a sign that the issue of Open Access to research data has not been a main focus of important actors. From the Finnish example it is clear that significant progress can be made in a short time based on leadership, allocation of resources, and the engagement of central actors in a coherent process towards a common goal.

Denmark: In 2014, the Danish Ministry of Higher Education and Science adopted a national strategy for Open Access to research articles from publicly funded institutions. The strategy has an ambitious goal, stating that by 2017, 80 per cent of the articles should be freely available via the internet. The Danish Open Access Indicator, which was launched in March 2016, shows that only 18 per cent of scientific publications produced at Danish universities are Open Access today. Even though the current measurement is based on data from 2014, the results show that Denmark is far from reaching its goal. So, the National Steering Committee for Open Access, which coordinates the implementation of the strategy, will have to discuss which measures ought to be taken to make significant progress in the process in the years to come.

The Danish Government has so far not developed a national strategy for Open Access to research data. However, the Ministry of Higher Education and Science has adopted a Danish Code of Conduct for Research Integrity, which stresses the need to ensure that research performed in Denmark is reproducible and that the research results are verifiable by other actors. This implies that research data underlying a research article has to be stored and that such data must be accessible. In 2014, the Danish Rectors' College, the Danish eInfrastructure Collaboration and Denmark's Electronic Research Library established a *Steering Group for National Data Management*. This group presented a strategy on data management in 2015. This document does not have a specific focus on Open Access as such; instead, it advocates a structured approach to data management, data preservation and data infrastructures. It also argues that a top-down approach to data management, policy development and implementation has proven to be prone to failure. Instead a bottom-up process built on stakeholder collaboration should be chosen.

The operational Open Access policy for the Danish public research councils and foundations reflects the Government strategy for research publications by imposing a requirement that, if the journal allows it, the article should be made openly available. No discussion on research data is included in the operational policy, and no requirements on e.g. providing a data management plan are imposed when applying for a grant from these funding agencies.

.dk *The Danish Government has so far not developed a national strategy for Open Access to research data. However, the Ministry of Higher Education and Science has adopted a Danish Code of Conduct for Research Integrity, which stresses the need to ensure that research performed in Denmark is reproducible and that the research results are verifiable by other actors.*

.fi

Finland: The discussion on re-use of research materials had an early start in Finland, and today concrete actions have been initiated within most important core areas. Several of these actions have already resulted in operational strategies, policies and services. Unlike many other countries, an overall view of issues on open access to research results has been adopted, and a major initiative on Open Science, the *Open Science and Research Initiative (ATT)*, is led by the Ministry of Education and Culture and executed in collaboration with research institutions, funding agencies and providers of digital services for research. A central international communication channel and gateway for this initiative is the Openscience.fi website.

The ATT initiative has led to the channelling of substantial resources into policy-making, implementation and monitoring, with the objective of ensuring that Finland becomes one of the leading countries in openness of science and research by the year 2017 and to ensure that the possibilities of open science will be widely utilised in Finnish society. The status of research organisations' openness in the operational cultures was evaluated 2015.[1] According to this analysis, no higher education institution has yet to reach the highest maturity level in openness. The Universities of Helsinki and Jyväskylä have reached the second-highest level. Five institutions were placed at the third level, fourteen at the fourth, and nine at the lowest. Over half of all institutions have been actively promoting openness. When it comes to openness, universities are clearly ahead of polytechnics. In order to monitor progress, a similar analysis will be repeated annually until 2017. In 2016, funding organisations are also being evaluated.

The Finnish *National Research Data Initiative (TTA)* has been an important enabler for the current development in Finland. This project was initiated and funded by the Ministry and was performed during 2011–2013. It was a broad-based co-operative network for the development of research data services and promotion of open knowledge and interoperability. As a result of the initiative, a centralised research data infrastructure using a research data enterprise architecture and metadata models was developed.

In 2014, the Ministry of Education and Culture of Finland released *The Open Science and Research Roadmap 2014–2017* which sets the policy framework for national efforts in the field. This document is also complemented with an Open Science Handbook and a Data Management Guide directed at Finnish researchers. A set of services for data management and access was previously developed under the TTA initiative, including services for data storage and data publishing. Several other services are currently being implemented, including facilities for digital preservation of research materials and more elaborate tools for data management.

The Academy of Finland is currently implementing the practices outlined in the Open Science and Research Roadmap when providing funding for research projects. The Academy requires that academy-funded publications are made openly available. The Academy also requires that applications include a data management plan, describing how the research data in the project will be used and re-used, how the rights of ownership to and usage of the data used and generated by the project will be distributed, and how the data produced will be stored and subsequently made available within and outside the project both during the project and after the project has ended. Finally, the Academy recommends that research projects also make their research data available through major national or international archives or storage services that are of relevance within their own fields.

1 openscience.fi/openculture

.fi *The ATT initiative has led to the channelling of substantial resources into policy-making, implementation and monitoring, with the objective of ensuring that Finland becomes one of the leading countries in openness of science and research by the year 2017 and to ensure that the possibilities of open science will be widely utilised in Finnish society.*

Iceland: Since 2012, legislation requires that all publications based on research funded by the large public competitive funds in Iceland (all of which are administered by the Icelandic Research Centre – Rannís) are made openly available. However, the implementation and monitoring of this legislation is still incomplete, and further efforts are needed. One of the steps being taken is the construction of a repository for publications that will be effective in 2016. Discussions on Open Access to research data have recently been initiated both within the Ministry of Education, Science and Culture and at the National and University Library, and awareness of the importance of issues relating to open access to digital research results, especially for smaller countries, is growing. The Ministry is currently drawing up a plan for Icelandic higher education and research for the years 2017–2021, and the importance of structured data management and open access to research data is likely to be included there. Currently, no requirements on e.g. providing a data management plan are imposed when applying for a grant in the public competitive funds.

.is *Discussions on Open Access to research data have recently been initiated both within the Ministry of Education, Science and Culture and at the National and University Library, and awareness of the importance of issues relating to open access to digital research results, especially for smaller countries, is growing.*

.no

Norway: According to the Research Council of Norway's principles for open access, the results of publicly funded research should be self-archived in an open electronic repository and published in peer-reviewed open access journals. Moreover, by using services within the national research information system CRIStin (Current Research Information System in Norway), the fulfilment of this requirement can be monitored.

Following a process with stakeholder consultations and an extensive survey directed to Norwegian researchers, the Research Council of Norway launched a policy on Open Access to Research Data in 2014. This policy sets out best practices and recommendations rather than strict requirements, and states that research data produced in projects fully or partly funded by the Council should be openly available by default. Here, the data should be made available on equal terms for all users and under internationally recognised licenses, giving as few restrictions as possible for use, re-use and redistribution. The policy recognises that some research data cannot be made openly available due to legal or ethical requirements or privacy issues. Datasets that would be highly impractical or costly to make accessible may also be exempted from the default principle. An interesting feature of the policy is that it states that the data should be made available at the lowest possible cost (i.e. possibly not for free); however, the cost must not exceed the actual cost of disseminating the dataset. The policy also presents a number of guidelines for the archiving, dissemination and sharing of research data, including that the data should be archived at established secure data centres, equipped with a long-term data management plan, and augmented with metadata based on international standards. The policy covers research data generated through research partially funded by the Research Council of Norway. Some of the Norwegian universities have developed local policies based on the Research Council policy.

The Research Council of Norway has taken several actions towards implementing the policy on Open Access to Research Data. The Council accepts data archiving costs as part of the operational expenses of funded projects. Furthermore, within the funding scheme for Centres of Excellence, additional funding can be made available to cover "particularly high operating costs", e.g. for using eInfrastructure for high-performance computing and storage of large amounts of data. The Research Council of Norway seeks to encourage the further development and establishment of well-designed infrastructures for data storage and data management. A significant number of projects that develop data access solutions or make data openly available are funded by the Council. Discussions are currently underway on how to implement data management plans. One option is to require data management plans (DMPs) for the projects that are funded. The preferred format for DMPs will be that of the Digital Curation Centre in the UK (DMPonline).

So far, the Research Council of Norway's strategies for Open Access to publications and to research data are not strongly connected to each other, and there is as of yet no well-established and formalised Norwegian initiative within the more general field of Open Science.

In January 2016, the Norwegian Government appointed a working group to assist in producing national guidelines for open access to publications, aimed at research funding institutions, institutions performing research, and individual researchers. It is likely that the earlier efforts of the Research Council will be valuable in this process.

.no *Following a process with stakeholder consultations and an extensive survey directed to Norwegian researchers, the Research Council of Norway launched a policy on Open Access to Research Data in 2014.*

Sweden: Since 2010, the Swedish Research Council requires that all scientific publications fully or partly funded by the Council must be published with Open Access. From 2015, only articles published with Open Access are formally accepted as a basis for reporting research to the Council. The Council further states that from 2017, researchers funded by the Council will be required to publish their results under a CC-BY-license. However, so far the monitoring efforts in the practical research funding processes have been limited, and it is unclear to what extent these rules are followed in practice.

In 2014, the Swedish Government asked the Swedish Research Council to propose national guidelines for open access to scientific information. After consultations with stakeholders, the Swedish Research Council submitted a proposal to the Government in January 2015. This proposal consists of two parts: Guidelines for Open Access to publications, advocating a transition from a subscription-based system to a system fully based on Open Access for all research publications; and a description of a process towards providing Open Access also to research data. The aim is to make openly available all research data, produced in whole or in part with the support of public funds, as soon as possible. However, the proposal acknowledges that the implementation of such a system is a tedious and time-consuming process that requires further analysis, and that it is also heavily dependent on international developments and the availability of eInfrastructure resources. A main conclusion from the analysis of the legal framework presented in the proposal from the Swedish Research Council is that, according to the national principle of public access to official records, the Swedish universities are responsible for archiving and long-term preservation of research data produced by their researchers. In general, the discussions within universities on how to meet this requirement is not well developed yet, also with regard to which data needs to be preserved.

To start the transition towards an open access system, the Swedish Research Council recommends that selected pilot calls should be implemented in 2015–2020 in which the research data providing the basis for scientific publications should be made openly accessible.

In autumn 2016, a new Research Bill will be presented. It is likely that this will assign the National Library of Sweden a coordinating role for Open Access to research publications while the Swedish Research Council will be assigned a similar role for research data. It is still unclear how these efforts would be coordinated and how they will be integrated with potential future efforts to arrive at a more complete Open Science landscape. It is also unclear how these efforts will be related to future activities initiated by the universities.

.se *A main conclusion from the analysis of the legal framework presented in the proposal from the Swedish Research Council is that, according to the national principle of public access to official records, the Swedish universities are responsible for archiving and long-term preservation of research data produced by their researchers.*

The EU: The European Union can influence the research system through different forms of legislation drafted by the European Commission and adopted by the Parliament and Council, and through policy-making and research funding managed by the Commission. With regard to legislation, a reform of the European copyright directive is being prepared. Consultations and discussions have taken place on allowing public interest research organisations to carry out text and data mining of digital content to which they have lawful access. The EU Database Directive is also of relevance for research data. A future review of this directive would have substantial impact on the openness of research and on data intensive research in Europe.[2]

The European Commission "Recommendation on Access to and Preservation of Scientific Information", presented in 2012, provided an important step towards an Open Access research data landscape in Europe. This document sets targets for openness in European research and also contains a set of recommendations to the Member States of the European Union. As has already been noted, many of the actions on Open Access to research results taken by the Nordic countries during recent years can be attributed to this document. The Recommendation was accompanied by the Commission Communication, "Towards better access to scientific information: Boosting the benefits of public investments in research", in turn followed up by the document, "Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020", describing how the ideas presented in the Communication are implemented in the Commission's current funding schemes. Here, the beneficiaries must ensure that all peer-reviewed scientific journal articles resulting from Horizon 2020 funding are published as Open Access and deposited in an accessible repository. The researchers should at the same time seek to deposit the research data needed to validate the results presented in the publication.

As a result of the Commission Communication, the Open Research Data Pilot has been introduced in the EC Horizon 2020 funding schemes. This initiative aims at making the research data generated by research projects in selected areas accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access. The pilot focusses on the data needed to validate the results presented in scientific publications. Here, the guidelines state that the projects should provide information about tools and instruments necessary for validating the results and, where possible, provide the tools and instruments themselves.The guidelines also state that the data should be provided free of charge, and that the management and potential re-use of all data in the research projects should be discussed in a data management plan. It is mandatory for all projects to set up a data management plan, and rather detailed guidelines for this are provided. However, the plan does not have to be provided at the stage where the project proposals are evaluated; instead, it can be set up in the start-up period of the granted projects.

The Digital Single Market is the EU strategy to make the EU a world leader in the digital economy while the "three O's" of Open Innovation, Open Science and Open to the World are introduced as the strategy to make EU a leader in Research and Innovation. In the crossroads of these two strategies, the concept of the European Open Science Cloud has been introduced to provide the eInfrastructure governance and resources needed for practical implementation at the European level.[3,4] Further steps towards the implementation of a complete Open Science system have been envisioned for the Horizon 2020 Work Programmes 2014–2015 and 2016–2017. The current design of the pilot encourages participation while providing flexibility for projects that might need to restrict access to some data. While this basic approach is defined in the Horizon 2020 Model Grant Agreement the default pilot areas are defined in the Work Programme and could again change in the funding period 2018–2020.

An important aspect of the Horizon 2020 Open Research Data Pilot is the requirement to develop a data management plan. While guidelines and tools are available,[5] many consortia feel the need for

2 ec.europa.eu/internal_market/copyright/prot-databases/index_en.htm

3 ec.europa.eu/digital-single-market/en/cloud

4 europa.eu/rapid/press-release_SPEECH-15-5243_en.htm

5 Horizon 2020 Guidelines on Data Management and the DCC DMPOnline tool

more practical guidance. Guidance is also needed for project reviewers who are required to provide feedback to grantees on the quality of the data management in general and data management plans in particular in their projects. Such guidance represents a significant challenge for the EC and its agencies implementing the Horizon 2020 Work Programmes, due to a lack of useful examples. At the same time it also represents an opportunity in the form of a large collection of data management plans that will be submitted by the consortia. These plans can be analysed to derive knowledge on the state-of-the-art and to provide recommendations for improving research data management in Europe.

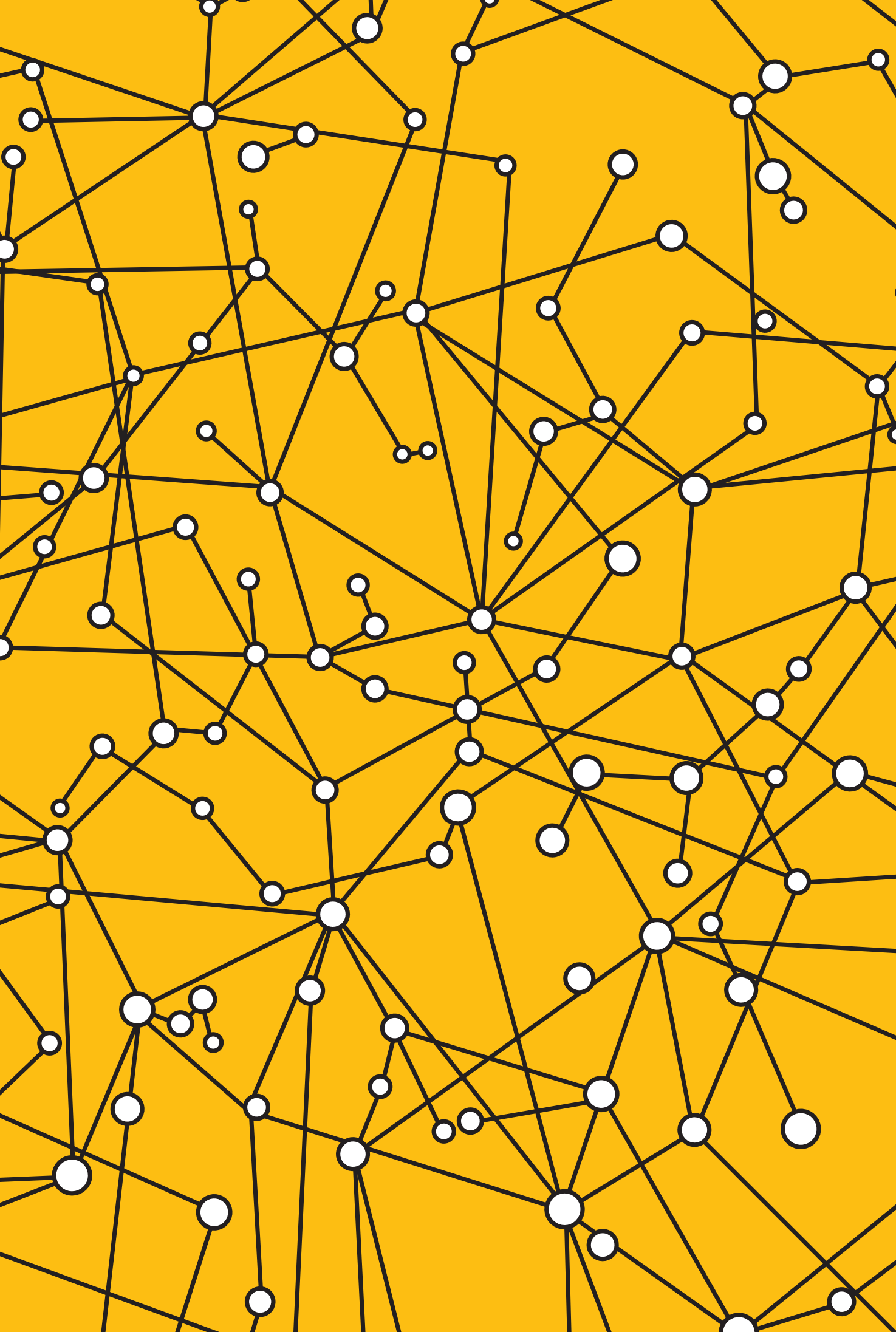
The European Research Council (ERC) has developed an Open Access policy that requires electronic copies of publications supported by ERC funding to be made publicly available no later than six months after the official publication date of the original article. The ERC also considers it essential that primary data, as well as data-related products such as computer code, is deposited in the relevant databases no later than six months after the date of publication. However, no recommendation on providing Open Access to such research results has currently been given.

The UK: The Research Councils UK partnership presented a common data policy in 2011 which was later revised in 2015. This policy is based on the principle that publicly funded research data is a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner. The policy states that institutional and project-specific data management policies and plans should be in accordance with relevant standards and community best practice. Here, the Digital Curation Centre provides examples of data management plans and lists the requirements for such documents. The Research Councils UK policy also states that research data with acknowledged long-term value should be preserved and remain accessible and usable for future research. To enable research data to be discoverable and effectively re-used by others, sufficient metadata should be recorded and made openly available to enable other researchers to understand the research and re-use potential of the data. Published results should always include information on how to access the supporting data. The Research Councils UK recognises that there are legal, ethical and commercial constraints on the release of research data. To ensure that the research process is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these are considered at all stages in the research process.

The US: The White House Office of Science and Technology Policy Directive, "Expanding Public Access to the Results of Federally Funded Research", presented in 2013, states that all major federal research and development agencies are required to store and make their data publicly accessible to search, retrieve, and analyse. Here, the scope of the directive is the data necessary to validate research findings, including datasets used to support scholarly publications. The agencies are currently in the process of implementing this directive, e.g. by providing specifications for data management plans and modifications of funding schemes.

The National Science Foundation (NSF) has required research proposals to provide a data management plan since 2010, and the instructions were recently updated to specify that the plan must address how the project will comply with the agency's Data Sharing Policy covering primary data, samples, physical collections and other supporting materials.A pilot for a National Data Service has also been initiated as a collaborative effort between publishers, university libraries, repositories, computing centres, and discipline-specific federations. Here, a key element is to provide a connection between publications and the underlying data. Furthermore, NSF has initiated an effort on setting up "big data hubs", providing a forum for community building and build-up of knowledge within research data management.

4. Main Findings and Outlook



4. Main Findings and Outlook

Main General Findings:

• It is generally recognised that fully open access to research data can sometimes not be implemented, e.g. for personal data and for data where access is regulated by contracts with third parties. However, well-structured processes and efficient resources for maintaining and sharing research data are still needed to cater to the needs of the distributed and persistent research communities using such more closed datasets.

• To arrive at a functioning system for Open Access to research data, four enabling components have to be present:

– suitable definitions, policies, rules and standards;

– cost-efficient eInfrastructures for data storage and the eScience tools enabling discovery and easy access/use of research data;

– established funding schemes and governance structures providing support for data access and storage, long-term preservation, and curation;

– a widely-accepted system for credit to researchers, providing benefits for researchers that contribute to data sharing.

To fully exploit the possibilities provided by introducing Open Access to research data, a more integrated view of research results is required. Such a discussion ranges from the efficient management of research data and other types of research results, through the sharing of such results within communities, to providing fully Open Access to a variety of different types of research results. In other words, a more complete Open Science landscape needs to be studied.

Currently, a large number of actors are producing reports on sharing and re-use of research data. The reports often have slightly different perspectives and aims, and most of them include recommendations to different stakeholders. The recommendations may be similar but they still contain differences that make the policy landscape complex and difficult to navigate. A consolidation process at the international level, both with regard to concepts and terms and with regard to recommendations and policies would be beneficial for researchers and other actors alike.

Main Findings Concerning the Nordic Countries:

• The discussions on Open Access to research data in the Nordic countries started early. However, the attention to the issue and the amount of resources allocated to policy-making and implementation differs significantly between the countries. The progress in some of the countries can be considered as rather moderate, while e.g. Finland is moving ahead rapidly. Major countries like the US and the UK are also moving ahead quickly and can today be considered to be ahead of several of the Nordic countries.

• The Finnish example shows that rapid progress towards Open Access to research data is possible if a coherent Open Science initiative is initiated at a high level, including a major effort in both policy-making and implementation and involving all relevant actors. Here, national leadership and a combination of top-down and bottom-up processes seem to be a key factor enabling progress. However, in Finland as well, the discussion within the research community on how to give credit to researchers contributing to data sharing has only just started.

• In some of the Nordic countries, the discussion and efforts on e.g. Open Access to research publications, to research data, and to other types of research results can be perceived as disconnected from each other. At the international level, e.g. in Europe, the discussion today has often moved on to include a more overall Open Science perspective.

Some Potential Actions at the Nordic Level:

• **Exchange of information and experience:** The Nordic countries are small actors in an international perspective, and the amount of resources that can be allocated to efforts on Open Access to research results and Open Science is limited. An extended exchange of information and experiences can improve the efficiency of national efforts. Such efforts might be pursued at several levels, including policymakers, research funders, research organisations and providers of data infrastructure. Such exchange could potentially result in more harmonised policies and guidelines. This would be beneficial for pan-Nordic research efforts, and could also result in improved impact on and visibility in the international arenas.

• **Facilitating progress on specific issues:** In the presentation above, a number of specific issues in need of further discussion and actions have been identified. Again, the amount of resources and competence in the Nordic countries is limited in an international perspective, and cooperative efforts might prove to be more efficient, allow for more rapid progress and improved international impact. The scope of Nordic collaboration can be used for setting up time-limited working groups, e.g. in the form of NORIA-net actions. Within certain fields, existing Nordic collaborative efforts can be employed for concrete actions.

• **Implementing pilots:** The Nordic arena for funding of research is rather small. It can also be agile and provide an opportunity for implementing pilot actions on Open Access to research data and Open Science. It might be an interesting perspective to consider Nordic research funding to fully implement an Open Science perspective, from the planning and evaluation of calls for proposals, through evaluation of proposals and contracts, to reporting and final evaluations. Such efforts could provide an interesting pilot for the national actors in the Nordic countries and would enable Nordic collaboration to take a lead within an important area.

References

Danish E-Infrastructure Cooperation (2015): *National strategi for forskningsdatamanagement* [Danish].
www.deic.dk/sites/default/files/uploads/PDF/National%20Strategi%20for%20Forskningsdata%20Management%202015-2018.pdf

Danish Ministry of Higher Education and Science (2014): *The Danish Code of Conduct for Research Integrity.*
ufm.dk/en/publications/2014/files-2014-1/the-danish-code-of-conduct-for-research-integrity.pdf

ERAC (2016): *ERAC Opinion on Open Research Data.*
era.gv.at/object/document/2402/attach/Opinion_on_open_research_data_ST01202_EN16.pdf

European Commission (2016): *Open innovation, open science, open to the world.*
bookshop.europa.eu/en/open-innovation-open-science-open-to-the-world-pbKI0416263/

European Commission (2016): *Guidelines on Data Management in Horizon 2020.*
ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

European Commission (2016): *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020.*
ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

European Commission (2012): *Communication. Towards better access to scientific information: Boosting the benefits of public investments in research.*
ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf

European Commission (2012): *Recommendation of 17.7.2012 on access to and preservation of scientific information.*
ec.europa.eu/digital-single-market/en/news/commission-recommendation-access-and-preservation-scientific-information

European Commission (2010): *Riding the Wave. How Europe can gain from the rising tide of scientific data.*
Final report of the High Level Expert Group on Scientific Data.
ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=6204

European Knowledge Exchange Initiative (2004): *Sowing the seed: Incentives and motivations for sharing research data, a researcher's perspective.*
www.knowledge-exchange.info/projects/project/research-data/sowing-the-seed

European Research Council ERC (2016): *Open Access Guidelines for research results funded by the ERC. Revised February 2016.*
erc.europa.eu/sites/default/files/document/file/ERC_Open_Access_Guidelines-revised_feb_2016.pdf

Finnish Ministry of Education and Culture (2014): *Open science and research leads to surprising discoveries and creative insights. Open science and research roadmap 2014–2017.*
openscience.fi/documents/14273/0/Open+Science+and+Research+Roadmap+2014-2017/e8eb7704-8ea7-48bb-92e6-c6c954d4a2f2

Hey, Anthony J.G. (2009): *The Fourth Paradigm: Data-intensive scientific discovery.* Redmond, Washington: Microsoft Research.

League of European Research Universities (LERU) (2012): *LERU statement on Open Research Data.*
www.leru.org/files/publications/Open_Access_to_Research_Data-FINALdocx.pdf

OECD (2015): "Making Open Science a Reality", *OECD Science, Technology and Industry Policy Papers,* No. 25. OECD Publishing, Paris.
dx.doi.org/10.1787/5jrs2f963zs1-en

OECD (2007): *OECD Principles and Guidelines for Access to Research Data from Public Funding.*
www.oecd.org/sti/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm

The Open Science and Research Initiative (ATT) (2015): *Open Science and Research in Finland. Openness in the operational cultures of higher education institutions.*
openscience.fi/openculture

The Open Science and Research Initiative (ATT) (2014): *Open Science and Research Handbook.*
openscience.fi/handbook

The Open Science and Research Initiative (ATT): *Data Management Guide.*
openscience.fi/data-management-guide

The Research Council of Norway (2014): *Open Access to Research Data.*
www.forskningsradet.no/en/Open_access/1254008537671

Research Councils UK (2013): *RCUK Policy on Open Access and Supporting Guidance.*
www.rcuk.ac.uk/documents/documents/rcukopenaccesspolicy-pdf/

The Dutch EU Presidency Conference 'Open Science – From Vision to Action' (2016). *The Amsterdam Call for Action on Open Science.*
english.eu2016.nl/binaries/eu2016-en/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science/amsterdam-call-for-action-on-open-science.pdf

The Royal Society (2012): *Science as an Open Enterprise.*
royalsociety.org/topics-policy/projects/science-public-enterprise/report/

Science Europe (2016): *Funding Research Data Management and Related Infrastructures.*
www.scienceeurope.org/uploads/PublicDocumentsAndSpeeches/SE-KE_Briefing_Paper_Funding_RDM.pdf

Swedish Research Council (2015): *Proposal for National Guidelines for Open Access to Scientific Information.*
publikationer.vr.se/en/product/proposal-for-national-guidelines-for-open-access-to-scientific-information/

The White House (2013): *Increasing Access to the Results of Federally Funded Scientific Research.*
www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

NordForsk, 2016
Stensberggata 25
N–0170 Oslo
www.nordforsk.org