The Research Council of Norway

How should we share research data?

Report and recommendations relating to licensing and making research data available

Contents

Preface

Society is undergoing a digital transition and data are becoming an increasingly important part of the research society, business and industry, and the public administration. Increased sharing of research results can generate several positive effects such as quicker knowledge development and more efficient research, better quality assurance and greater public trust in research. At the same time, sharing data from publicly funded research is expected to improve utilisation of research results in society.

But what are the legal conditions that apply to sharing of research data from publicly funded research, and how can we encourage more sharing and re-use of research data?

This Committee has worked on these issues for just over a year through internal meetings, public debate and two open consultation rounds. This report contains the Committee's report and recommendations for national guidelines relating to licensing and making research data available.

Sharing of publicly funded research data can be limited by several legal factors such as data protection, protection of sensitive information and matters related to commercialisation. Correct and expedient use of licenses is therefore important to achieve as much data sharing as possible within these limits. At the same time, the Committee has found that other factors, such as incentives in the research system, technical factors and insufficient skills in the research communities, are the most important obstacles to data sharing. We hope this report and the recommendations herein can form a useful knowledge basis and tool for future work on legal and practical aspects of data sharing. Furthermore, we hope it can help to encourage responsible sharing and re-use of research data in a way that also safeguards considerations for research ethics, data protection and other issues described in this report.

More sharing of research data combined with ever increasing technological possibilities will hopefully bring about a range of positive effects in society going forward. However, the future we are stepping into is developing at an immense pace, which can lead to unforeseen effects, including in areas far beyond the research ethics issues and other societal challenges the Committee has highlighted in this report. We therefore advise regular evaluations of and adjustments to data sharing guidelines in the time ahead as we as a society learn more about the effects of sharing data from publicly funded research.

The Committee has comprised representatives from business and industry and the media, as well as representatives covering a broad range of research institutions and scientific disciplines. We would also like to thank the external resource persons who have provided input during this process and who have, among other things, prepared the appendices to this report.

The Research Council's secretariat, represented by Siri Lader Bruhn and Nenitha Dagslott, has done an outstanding job during this process, not least considering the short time horizon given the extensive remit and that the work has been carried out in full during the COVID-19 pandemic. The Committee is grateful for being assigned this particularly interesting assignment and for the valuable input from a broad research sector, public sector and industry. I would personally like to thank the members of the Committee for outstanding, efficient and informative work, and I hope and believe our discussions and recommendations will be useful to society in future work on sharing of research data.

The recommendations in the report are unanimous.

Chair of the Committee,
Jan Magnus Aronsen

Summary

Data generated through publicly funded research should in principle be considered a public good and shared with other researchers, business and industry, and society at large. This will help to strengthen the quality of and trust in research by allowing for better validation and verification of research results. It will also help to speed up knowledge development and, in turn, value creation and innovation in the private and public sectors. Research and innovation are increasingly driven by access to new and large quantities of data. Efficient structures for transparency and sharing of research data are necessary to ensure data quality, efficient utilisation of resources in the research communities and increased use of research findings in society. There is already a great deal of sharing and re-use of research data taking place, but it is still the case that much of the data produced through publicly funded research are not shared and re-used to the desired extent. At the same time, it can be unclear what legislation regulates sharing and re-use of publicly funded research and how such data can and should be licensed. Key questions in this context are who holds the rights to the data, right of use and responsibility for them.

The Committee's remit was to look into issues relating to rights and licensing in connection with the sharing of research data by reviewing relevant legislation, describing different types of licenses and considering the consequences that different licensing choices will have for the researcher, the institutions and society at large. The Committee was also asked to discuss the conditions that must be put in place to achieve the objectives of more sharing and re-use of research data.

The members of the Committee represent a broad range of disciplines in the research sector, including two members with legal expertise in the fields of intellectual property law and data protection law. The members represent the higher education sector and the institute sector, as well as all parts of Norway. The perspectives of business, industry and civil society are also represented. The Committee's work has taken place in the period from June 2020 to October 2021 and has resulted in a memo that was published in November 2020 and this final report. The process has included two consultation rounds and several meetings and seminars with different stakeholders and actors. This report is the Committee's final delivery and contains both a review of legal issues related to ownership and sharing and re-use of data, and recommendations to the Norwegian Government and the ministries, funders and policy implementation systems, and research institutions, respectively. The Committee hopes that the review can form a useful knowledge base in the further work on data sharing and pertaining rights.

The main results of the Committee's report can be summarised as follows:

- Research data must as a rule be considered public information on a par with data from other public activities.
- The term research data covers a broad range of data, and there are notable differences between different disciplines when it comes to both the amount of data generated and how shareable they are. This must be taken into account in policy making in this field going forward, and it is important that the research communities themselves take part in the debate and take ownership of the topic.
- Many laws and regulations address rights, requirements, obligations and protection relating to research data. In principle, they facilitate and require the sharing of research data from publicly funded research.
- Substantial expertise is required to interpret the extensive and somewhat fragmentary regulatory framework and arrive at good solutions when data is to be shared with others.
- The use of licences with as few restrictions as possible both on access and re-use is a good tool for achieving more sharing and re-use of research data.
- In cases where the regulatory framework protects data subjects or there are other considerations that prevent open sharing of research data, it is important to find appropriate, discipline-specific solutions to how data can nonetheless be shared in accordance with the FAIR principles.
- There is currently no optimal licence that contains all relevant terms and conditions for research data. The Committee therefore proposes considering whether a new licence should be devised and suggests that Norway could initiate a process to this end.
- There are many other factors aside from the legal framework that influence researchers' opportunities and willingness to share research data. These include available infrastructure, resources and skills and competence.
- Public research institutions should have an overall responsibility for ensuring that research data produced by their employees are shared subject to expedient licences or dedications to the public domain. The institutions should therefore develop long-term strategies in line with national and international requirements, describing how they will ensure sharing of data in a re-usable form, including institutional infrastructure and expertise.
- The current incentive system should be reviewed and adjusted to more appropriately reward the sharing of research data.
- It is unclear what is considered publicly funded research, for example in cases where research projects are a collaboration between public institutions and business and industry. The Committee recommends that this issue should be reviewed and discussed in a group/forum with relevant stakeholders.

In its recommendations, the Committee proposes measures that can provide conditions more conducive to sharing and re-use of research data, both when it comes to legal aspects and the use of licences and non-legal barriers. The Committee has also devised a set of overall guidelines for the use of licences for research data – the 'Licence etiquette rules'.

The Committee's work has identified a number of questions and issues that have not been possible to answer or discuss in detail within the limits of this work. The Committee has nonetheless wanted to raise them so that, hopefully, the research sector as a whole will continue to work on the topics and find appropriate answers and solutions in the time ahead.

Introduction

Background

Transparency and knowledge sharing are key aspects of research and innovation in many parts of the world. At the same time, society is facing major challenges that must be resolved to ensure sustainable global development and knowledge-based business and industry. The technological development over the past decades has dramatically changed the way in which research is carried out, how the results are shared and how they can be used in new ways. Research and innovation are increasingly driven by access to new and large quantities of data. Efficient structures for transparency and sharing of research data are necessary to ensure data quality, efficient utilisation of resources in the research communities and increased use of research findings in society.

At the international level, the desire for greater openness in research, and not least sharing of data, is nothing new, but we see that it is now even more emphasised than previously. The ongoing coronavirus pandemic has enhanced this trend. OECD works with EU member and non-member states to promote open science and open innovation. It revised its recommendations in 2021 on open sharing of data and digital objects from publicly funded research. UNESCO has announced that it will revise its recommendations on open science. The objective is to define shared values and principles for open science and identify measures that the 193 member states are expected to endorse. The recommendations will build on the UNESCO Strategy on Open Access to Scientific Information and Research and their new Recommendation on Open Educational Resources (OER).

The European Commission is also taking stronger action when it comes to open science. In the EU research programme Horizon Europe, open science will be the rule rather than the exception. Applications will be assessed in relation to their open science practice under the criteria 'Excellence' and 'Implementation' (with the exception of the European Research Council and the European Institute of Science and Technology). Stricter requirements will also be made of data being made available in accordance with the FAIR principles. The EU's establishment of the European Open Science Cloud (EOSC). (EOSC) is a key measure that better facilitates open sharing of data. This joint infrastructure aims to bring together institutional, national and European stakeholders, initiatives and data infrastructure to develop an inclusive ecosystem for open science in Europe. In the Norwegian context, open science is highlighted as a clear ambition in the Research Council's main strategy and policy for open science, among other sources. Open science is also assumed to be prominent in the next version of the Norwegian government's Long-term Plan for Research and Higher Education.

Norway harbours high ambitions to take a strong and leading role in the data economy in order to realise the values inherent in data and thereby create more jobs for the future. This is clearly signalled in Report No 22 to the Storting (2020–2021) *Data som ressurs – Datadrevet økonomi og innovasjon* and the strategy document *Nasjonal strategi for kunstig intelligens* (2020). Furthermore, a number of high expectations of the institutions and services in the university and university college sector (higher education sector) have been set out in the new digitalisation strategy and the white paper on governance of state universities and university colleges (in Norwegian only). To meet these ambitions, it is essential to develop a good framework for efficient sharing and re-use of data, as well as good practices and a culture for such activities. To achieve more efficient and seamless access to data, it must be clarified who can use the data, as well as whether the data are subject to exclusive rights, who can use them and, if so, on what conditions. Correct and expedient use of data licenses and other re-use agreements will be important instruments for achieving the goal of more sharing and re-use of research data.

Data generated through publicly funded research should in principle be considered a public good and shared on a par with other public data in accordance with national and international guidelines. The results from such research are a common good that is valuable to both science and society. Better access to research data will help to strengthen the quality of and trust in research by allowing for better validation and verification of research results. The research system has traditionally shared knowledge by publishing scientific articles, while the data on which the articles are based have not been made public to the same extent. In some fields of research, many findings cannot be reproduced (or disproved) because the data are of low quality or inaccessible. At the same time, it is also a challenge that large quantities of research data are collected and processed without being used, or that they could have been re-used in subsequent research or in other ways in society. More systematic sharing of these data would enable the general public and other researchers to verify the quality and validity of the data on which the research and publications are based.

Sharing of data also contributes to increased value creation in that other researchers and society at large can use the data in new ways and in combination with other datasets. Good infrastructures for data sharing can facilitate far more efficient summaries and meta-analysis of the existing knowledge in different fields. It is also possible to increase society's utilisation of research data for innovation and commercialisation purposes. As such, it will also be important to align policies across sectors and to coordinate the various funding schemes and other policy instruments. More sharing of research data will promote more

interdisciplinary research and innovation related to societal and industrial development. However, it will also contribute to a more nuanced public debate and better democratic processes by giving the general public better insight into the basis for research.

The research sector is facing many of the same challenges as other sectors when it comes to data sharing. Established standard formats, good data infrastructures, sustainable funding models as well as competence raising and incentives at the individual and institutional level are crucial to achieving more sharing and re-use of data. There may also be additional challenges related to the sharing of research data, however, due to uncertainties about legal frameworks for data sharing and a lack of incentives for sharing in the research system. In certain scientific fields, such as the humanities and social sciences, methodological aspects of the data collection can also make it difficult to filter out 'clean' data from the research activity itself.

Licenses are increasingly used to ensure acknowledgement and regulate re-use. Expedient use of licenses can be an important instrument to achieve more sharing and re-use of research data. The use of licenses is not consistent, and the guidelines for Norwegian researchers are inadequate. Licenses with limitations for re-use of the data are also used to some degree, and there is little knowledge about the consequences of such restrictions to further use of the data in society. More knowledge and an awareness of the use of different licenses in the research communities are needed to ensure that research data can become a more beneficial contribution to value creation in society.

Licenses have been the subject of much debate, both in Norway and internationally, in connection with the requirement for open access publishing. The Norwegian Directorate for ICT and Joint Services in Higher Education & Research's (UNIT) report from December 2019 on Rights and licensing issues relating to open access publishing (in Norwegian only), by Associate Professor at UiB Torger Kielland focuses on Creative Commons licenses and management of publications protected by copyright at the research institutions. Kielland's report does not deal with research data, and input from the research sector indicated a need for further studies on rights management and licensing when it comes to research data as well.

As a response to this input, the Research Council, on behalf of the Ministry of Education and Research, appointed this Committee to look at rights and licensing issues in connection with the sharing of datasets from research. This report is the final delivery from the Committee. It provides a systematic review of relevant legislation that regulates sharing and re-use of data, which types of data are generated through research, which restrictions on sharing the different types of data may be subject to, and which licenses can be used for the various types of data without comprising the restrictions on sharing. The Committee finds it important to emphasise that the report does not provide conclusive answers to the challenges we are facing in this field. However, it does provide an overview and some clear recommendations to all parts of the sector about what can be done to achieve more sharing and re-use of research data.

The Committee's remit

The National strategy on access to and sharing of research data (2017) states that research data must be as open as possible, as closed as necessary. It is an objective that the results from publicly funded research should contribute to value creation and benefit the general public. Other important reasons for making research data more accessible are to improve verifiability and quality assurance and reduce duplicate data capture. Expedient and consistent use of licenses, which supports simpler sharing across fields and sectors and is in line with international recommendations and practice, will also be of great importance.

The Committee's remit was to look into issues relating to rights and licenses in connection with the sharing of datasets in research by reviewing relevant legislation, describing different types of licenses and considering the impact of different choices of license for the researcher, the institutions and society at large. The Committee was also asked to discuss the conditions that must be put in place to achieve the objectives of more sharing and re-use of research data.

The assignment was twofold, with two deliveries.

1. In phase one, the Committee made recommendations on which discussions of principles should be addressed, which structural changes should be considered, and, if relevant, which legislation should be reviewed so that Norway can achieve its ambition of more sharing and re-use of research data. The Committee was asked to view its work in the context of other ongoing initiatives that have been implemented to generate knowledge about data and data processing in different segments of society. The first delivery took the form of a memo that the Committee submitted to the Ministry of Education and Research (autumn 2020), which formed the basis for the next phase of the Committee's work.

2. In phase two, the Committee invited relevant stakeholders to participate in an open process with different discussion arenas and other opportunities to present their points of view. The final delivery from the Committee is this report, which summarises its work and a set of national recommendation on the use of licenses for research data (fall 2021).

Composition of the Committee

The Committee represents a broad range of disciplines in the research sector, including two members with legal expertise in the fields of intellectual property law and data protection law. The committee members represent the higher education sector and the institute sector and cover all parts of Norway. The perspectives of business, industry and civil society are also represented. The Committee comprises the following members:

Role	Name	Position	Employer, geographical area
Chair	Magnus Aronsen, Chair of the Committee	Associate professor	University of Oslo, Institute of Basic Medical Sciences/Oslo University Hospital
Member	Koenraad De Smedt	Professor	University of Bergen, Department of Linguistic, Literary and Aesthetic Studies
Member	Ingrid Heggland	Senior research librarian	The Norwegian University of Science and Technology (NTNU), Library Section for Collections and Digital Services
Member	Liv Dingsør	CEO	DigitalNorway
Member	Stein Tronstad	Section leader	Norwegian Polar Institute
Member	Ole Petter Pedersen	Editor	Tech Weekly (Teknisk Ukeblad)
Member	Inger Berg Ørstavik	Professor	University of Oslo, Department of Private Law
Member	Tobias Mahler	Professor	University of Oslo, Department of Private Law
Observer	Nina Karlstrøm	Section leader	UNIT
Observer	Katrine Weisteen Bjerde	Policy adviser	UNIT
Head of the secretariat	Siri Lader Bruhn	Special adviser	The Research Council of Norway
Secretariat	Nenitha Dagslott	Adviser	The Research Council of Norway
Secretariat	Jon Ø. Flæten (until Feb. 2021)	Special adviser	The Research Council of Norway

Other contributors

In its review of the legal framework for sharing and re-use of research data, the Committee has received invaluable help from experts on Norwegian and European legislation in the area. Research fellow Kristina Stenvik at the Department of Private Law (UiO) has reviewed and systematised the Norwegian regulatory framework. This review has formed the basis for the Committee's discussions on the legislation that regulates sharing and re-use of research data in Norway and can be read in full in Appendix 1[1] to this report.

Heather Broomfield, senior adviser at the Norwegian Digitalisation Agency and research fellow at the Department of Public Law (UiO), and Espen Dennis Kristoffersen, head of analysis and evaluation at the Ministry of Local Government and Modernisation, have contributed a review of relevant European legislation that applies to and regulates the sharing of research data in Norway. This review can be found in Appendix 2[2] to this report.

The Committee is grateful for these contributions, which have provided a useful point of departure for discussions and formed the basis for several of the Committee's recommendations in this report.

Process

The Committee started its work in June 2020. Ten committee meetings have been held at regular intervals during the period. In October 2020, the Committee submitted the memo Research data in a data-driven society (in Norwegian only) as a product of the first phase of the Committee's work. The Committee

1 Appendix 1 is available in Norwegian only. The document is an elaborated version of the text in "PART 2 – Legal framework and use of licences and dedication to public domain for research data" in this report.

2 Appendix 2 is available in Norwegian only. The document is a description of the Directive (EU) 2019/1024 on Open Data and Public Sector Information and how it can be translated into a Norwegian context.

initiated an open consultation round after the memo was published. It received 37 written responses that were incorporated into the second phase of its work.

During the first half of 2021, the Committee also had several meetings with relevant parties, including the ministries' research committee, Universities Norway's strategic unit for research UHR Research, the Norwegian node of the Research Data Alliance, Abelia's data forum, the Norwegian Association of Researchers' copyright committee and the management of some universities. The chair of the committee and several committee members have also given presentations at seminars and conferences during the period. In March 2021, the Committee organised an open input conference with oral input and panel debates.

A draft report was published in June 2021. The Committee initiated a further open consultation round with the deadline set to 1 September. 15 written responses were received during this round and the input has been considered and incorporated into this final report by the Committee.

It has been important to the Committee that the process leading up to the final report has been as open and inclusive as possible. The Committee's assignment has involved complex questions and issues that have a bearing on all parts of the research sector as well as other sectors and actors. The open process has led to many interesting and useful discussions, and contact and dialogue has been established across research communities. The Committee is grateful for all contributions it has received and looks forward to further cooperation to achieve the goal of a more open research sector.

Delimitations and structure of the report

The Committee has received a great deal of attention and created engagement, indicating that the topic is widely relevant. It also demonstrates a major need to find good solutions that enable the research sector to comply with policy requirements, and for researchers and others to gain access to good and valuable datasets. The Committee's remit is in principle clearly delimited to apply to research data and pertaining rights and license issues, but it is important to point out that several ongoing processes and initiatives concern the sharing of data in general, and it is therefore important to see this work in context. This report must thus be seen as one of several contributions to achieving more sharing and re-use of publicly funded data in a larger perspective, but with particular emphasis on data produced through research activity.

It became clear through the consultation rounds that many people in the research sector are concerned with topics related to the sharing and re-use of research data, and we have received a great deal of useful input as to what should be discussed in this report. The Committee identified a pressing need for many types of clarification and many important issues that should be resolved. The Committee discusses these issues, but has not aimed to provide answers or conclusions in this report. These issues will require long-term joint efforts by many actors to arrive at good answers.

The report builds to some extent on the Committee's memo from October 2020, but also raises discussions that have taken place within the Committee as well as the extensive and valuable dialogue the Committee has had with different external actors during the period. It also contains a thorough review of the legal framework for sharing and re-use of research data.

The Committee has received many requests to deal with the licensing of research-related results or products, which are not data per se. This could be source code and software, models, methodology protocols etc. These also form an important part of open and reproducible research. There are a number of legal frameworks that could be considered relevant in this context, such as the Patents Act and the Copyright Act, and there are many different licenses of relevance to source code and software beyond those relevant to use in the context of research data. Concepts and terminology related to source code and software can also vary somewhat to those used for research data. The Committee therefore considers this to be an area in need of a dedicated and thorough review. Due to limitations in terms of time and scope, we have chosen to delimit the present report to research data. The topic is nonetheless briefly addressed in Subchapter 7.3.6.

Another question the Committee has been asked during the process is related to the remit's underlying premise that it is publicly funded research the report and recommendations are to address. However, it is simply not completely clear what this entails. It could be claimed that research is publicly funded from the first penny invested in it by the public sector, while it could also be argued that this would be an unreasonable interpretation. We touch on this in Subchapter 6.2 The Open Data Directive, but the Committee has chosen not to give an interpretation of what lies in the term 'publicly funded'. This matter is not for the Committee to define. The recommendations are nonetheless based on the premise of public funding of research, despite the current lack of a clear definition of this concept.

The report comprises three parts. The first part describes what the Committee has included in the term 'research data', what can in some cases distinguish this type of data from other types of data, the life cycle of research data and the difference in meaning between open data and FAIR data. This part shows that the term 'research data' encompasses a broad area, and the Committee has attempted to distinguish between research data that 'qualify' for sharing and research data that do not. This is an important distinction that has a bearing on which expectations and requirements apply to sharing in different cases. The second part of the report deals with the legal framework that is relevant to sharing and re-use of research data, and the Committee's assessment of the need for clarification and, where relevant, amendment. This part also reviews different types of licenses and their function, and the consequences different choices of license may have for re-use

of research data. Part two concludes with the Committee's recommendations as regards legislation and its overall recommendations on the use of licenses (License etiquette rules). The last part presents what the Committee considers the most important factors that influence researchers' and institutions' opportunities and willingness to share research data. Finally, the Committee presents its recommendations to the Norwegian Government and the ministries, funders and policy implementation systems, and research institutions, respectively. The Committee is of the opinion that the recommended measures can contribute to more sharing and re-use of research data in the future.

Part 1

Characteristics of research data

Photo: Shutterstock

Many areas of policy making and solutions developed to encourage more sharing of public data in general will also be relevant to the sharing of research data. However, research data have some distinctive characteristics and features that must be taken into account when preparing guidelines to ensure that their full potential can be utilised. In this part of the report, we describe different variants of research data that could require different assessments and considerations for sharing and re-use. We also describe the research data's life cycle and how different stages of the cycle can require different approaches to such assessments. Finally, we describe the main types of data this report will deal with.

1. WHAT DOES THE TERM 'RESEARCH DATA' MEAN?

The term 'research data' can be interpreted in different ways in different scientific disciplines and across the research sector. There is no one definition, and the term can be applied to different types of data and data-related objects. This can make it challenging to talk about research data across fields and sectors and, not least, to find good joint solutions to how to manage research data. In this chapter, we present some definitions that we find clear and useful. We also provide an overview of different characteristics and qualities of research data that may have a bearing on how the rules and guidelines will apply. In Chapter 5 we have attempted to describe research data that 'qualify' for sharing, first and foremost to highlight that some research activity does not produce shareable data in the traditional sense. This is an important clarification to ensure that such research activities are not made subject to requirements for sharing data when sharing is not possible, or be disadvantaged in cases where sharing of research data is used as an assessment criterion or express goal in more and more contexts.

In the course of the Committee's work, UNIT started the project National terminology for research data management (the follow-up if this project has now been transferred to the Norwegian Directorate for Higher Education and Skills). The Norwegian node of the Research Data Alliance (RDA) is managing the project and will propose a Norwegian set of terminology by the end of 2021 for use in connection with research data management. The product of this important project is likely to make future discussions on research data and data management more precise and expedient, and the Committee looks forward to their delivery.

1.1 Definition of research data

There are many and partly overlapping definitions of research data, but a common denominator is that research data are information that in one way or another is used or produced in connection with research activities. Most definitions point out that data are information used as the background or basis for scientific findings or analyses, and that they can also be used to verify and validate research results. Research data can exist in either digital or physical formats, but it is most relevant to discuss digital data in connection with data sharing.

There are already many definitions of research data and the Committee has therefore not wanted to write its own definition. We instead refer to two existing definitions at two different levels of detail that we believe to be useful, each in its own way.

EOSC DEFINITION OF RESEARCH DATA IS:
Data collected or produced in the course of scientific research activities and used as evidence in the research process, or commonly accepted in the research community as necessary to validate research findings and results.

OPENAIRE DEFINITION OF RESEARCH DATA IS:
Research data are the evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical). These might be quantitative information or qualitative statements collected by researchers in the course of their work by experimentation, observation, modelling, interview or other methods, or information derived from existing evidence. Data may be raw or primary (e.g. direct from measurement or collection) or derived from primary data for subsequent analysis or interpretation (e.g. cleaned up or as an extract from a larger data set), or derived from existing sources where the rights may be held by others. Data may be defined as 'relational' or 'functional' components of research, thus signalling that their identification and value lies in whether and how researchers use them as evidence for claims. They may include, for example, statistics, collections of digital images, sound recordings, transcripts of interviews, survey data and fieldwork observations with appropriate annotations, an interpretation, an artwork, archives, found objects, published texts or a manuscript.

The European Open Science Cloud's (EOSC) definition is brief, which can be useful when discussing general problems related to research data. OpenAire's definition is more detailed, which can provide clarification when looking at problems in more detail and identifying solutions and measures.

Consensual definitions can be useful when discussing topics across sectors and levels of administration. However, since the report's discussion about research data is broad-ranging, the Committee wishes to refer to research data in a way that covers several different characteristics. It is not therefore expedient to have a single general and consensual definition. The Committee finds it important to be clear about different aspects of the data that could have a bearing on how to make sensible decisions about sharing, potential restrictions and choice of licenses. We discuss several of these characteristics in the following, and hope that the discussion will provide an overview of how heterogeneous research data actually are, making it inexpedient to address it as a single thing with specific characteristics when taking steps to make research data more available.

1.2 Raw data vs. processed data

Raw data refers to data in their original form, as they are when collected. When such data are systematised, sorted or analysed, they become processed data. If the data are organised and structured in a system that can be accessed electronically, they are often called a database. In a legal context, databases can be legally protected if certain conditions in the law are met. The situation is different for unstructured raw data (we will discuss this in more detail in Part 2). Under certain circumstances, the processing of data can confer exclusive rights, such as copyright, and these rights can in some cases also apply to databases.

1.3 Source data vs. data generated during the research process

In connection with the sharing and re-use of research data, it will often be relevant to distinguish between existing data that is re-used in research (often called source data) and new data generated during the research process. Source data can be information collected for purposes other than research, such as data from public administration, health services, private enterprises or data in the form of physical collections, literary texts or other artistic production. It can also be data from previous research activities. The distinction between source data and data generated from research is significant in relation to rights to data and the use of licenses. Researchers often use source data for further research and must in such case comply with licenses, agreements and legislation that apply to the source data. This can make further sharing challenging or even impossible.

Research data that is generated without the use of source data can be more freely shared by the responsible researcher, research group or research institution in accordance with applicable laws and guidelines. In many cases, research will entail a combination of the use of source data and data generated during the projects, which serves to complicate the assessment of further sharing. This is discussed in more detail in Part 2.

1.4 Qualitative vs. quantitative data

Another relevant dimension of research data is whether the data are collected using quantitative or qualitative methods. The method used can[3] have an impact on the objectivity of the data and whether it is possible to filter out 'clean' data from the research activities themselves. It can also have an impact on how much metadata, other documentation and information about the context is needed to understand and interpret the data. Quantitative methods are often used to test ideas and hypotheses presented at the start of a project. This means that data are collected based on predetermined parameters. Data from such research can often be quantified in numerical terms and can be used for statistical analysis. When using qualitative methods, data collection is often more integrated into the research activities themselves, which can make it difficult to distinguish clearly between data (that can be analysed or interpreted) and the researcher's intellectual contribution. Such processes are often more hermeneutic than other research processes, and new understandings and elements are drawn in along the way. This can make the distinction less clear.

1.5 Personal data vs. data about the world at large

Within several fields of research, such as medicine, biotechnology, social sciences and the humanities, information about persons will often comprise part of the data basis for the research. Such data must be processed differently than data that provides information about the world at large. Information and data that can be linked to individuals will fall under the scope of the Personal Data Act and require a lawful basis for processing. In research, such a basis for processing can often be obtained in the form of a consent. If personal data fall under a special category (often called sensitive personal data), what is known as explicit consent is required, which means that the consent is given in a particularly clear way. The consent is also required to be specific and informed, which means that the purpose of the consent has been clearly and precisely worded. As a rule, personal data cannot be used for other purposes than that described in the consent. This means that, in principle, personal data cannot be openly shared, but with an explicit consent to sharing, it will be possible based on the terms and conditions described. In such case, it will be relevant to use a license. If an informant withdraws consent during the process, the data cannot be used by the research project or shared with others for further processing.

If the data are anonymised, making it impossible to identify individuals in the dataset, the personal data legislation ceases to apply and the data can be shared with others without

3 It is important to point out that quantitative methodology does not necessarily yield objective data, for example data collected through individual questionnaire surveys in the social sciences.

restriction. It is nonetheless a matter of research ethics whether anonymised data should be shared without the data subjects being informed, when data is collected based on consent. We will take a closer look at the GDPR in Subchapter 6.4 Restrictions on sharing of data.

1.6 Big data vs. 'the long tail of research data'

There is no clear definition of the term 'big data', but in general, it describes collections of data whose size and complexity make them impossible to utilise through traditional data processing. Examples include weather data and transport data. As a rule, the data come from many different sources, but if standardised, they can provide new and important information when compiled and analysed. The term big data is often used in connection with data economy, data mining and artificial intelligence. Data at the other end of the scale can for the sake of simplicity be called small data, although this is not always a fitting description. This category can also encompass relatively large datasets, but they are characterised by often containing more precise and detailed information than big data. They are also less suited to automatic analysis and machine processing. Small data are often heterogeneous and will in many cases need individual consideration and adaptation to be made available for re-use and compilation by others. This kind of data is often referred to as 'the long tail of research data'.[4]

2. OPEN DATA AND FAIR DATA

Sharing the results of research is an important principle to ensure transparency and verifiability. In recent years, open science has become a particularly important policy area, among other things because digitalisation has enabled the re-use of results in completely new ways for researchers, industry and society. The amount of data collected by the research communities and society at large has significantly increased in step with rapid technological development. This has also given rise to new challenges in terms of data management, and essential changes have been made to the policies that govern this area. In the early phase, it was important to bring all the research communities on board to the idea that sharing of data was necessary and the right thing to do in modern research. The term 'open data' and the concept 'open by default' were introduced. This has subsequently been met with a fair amount of criticism, including because it takes little account of the challenges relating to data protection, security, commercialisation and costs associated with making data open. The concept has therefore been toned down somewhat, and we now refer to making data and other results from research 'as open as possible, as closed as necessary'. This is also reflected in the Government's strategy and the Research Council's Policy on Open Science.

At the same time, we are seeing a shift of focus from the process of sharing and making data available to the actual re-use of the datasets. The sharing of data is not an end in itself. It is rather a means of achieving the goal of better utilising publicly funded data for new knowledge development and the development of technology, products and services that are useful to individuals and society at large. The term 'FAIR data' is commonly used to describe data adapted to re-use in accordance with the 'FAIR principles' (Findable, Accessible, Interoperable, Reusable). This shift is clearly shown in the European Commission's Open Science Policy Platform: final report. The report states that there must be a balance between requirements for openness and protection of intellectual property rights (IPR) in line with the principle 'as open as possible, as closed as necessary'. The report also underlines that the FAIR principles should be the goal, not necessarily that the data should be open. The report calls for greater awareness of the value of IPR and the individual contributions researchers make to solve major societal challenges.

Since the terms 'open data' and 'FAIR data' are both used to describe access to and the conditions for re-using data, and are to some extent used interchangeably, it is important to stress that they do not refer to the same thing. In the following, we will take a closer look at what the terms entail and the potential practical significance of these two ways of sharing data.

2.1 Open data

Open data means that the data can be shared without restrictions in terms of who can use them (restricted access) or what the data can be used for (restricted purpose). Since public data is collected through activities or projects that receive public funding, it is a goal that as much data as possible are made open. Certain restrictions can complicate this situation, both in relation to public data in general and research data in particular. When it comes to research data, it is important to distinguish between data subject to legal restrictions on sharing and data that is not shared for other reasons. Legal provisions that restrict what access can be granted to different types of data and how they can be (re-)used are found in many legislations since they are intended to protect different interests. Such provisions are found in legislation that applies to different areas, including the Health Research Act, the Security Act and the Copyright Act. It can be challenging for the researcher, research group or institution tasked with ensuring lawful sharing of data to identify and familiarise themselves with the applicable restrictions. This can give rise to uncertainty and lead to differing practices for sharing and licensing of datasets. We will discuss this in more detail in Part 2 of this report and in Appendix 1.

As regards data that are not subject to statutory restrictions, it is still the case that some data cannot be shared. To explain why, it is necessary to look at the underlying factors that affect practice. In these cases, it is the responsibility of the researcher/institution to assess whether and to what degree the data can be made open. Reasons for not sharing data could be inadequate infrastructures for sharing, research ethics grounds,

4 LongTailOfData2016 (e-irg.eu)

a lack of legal assistance and knowledge, or a lack of resources and incentives. This will be discussed in more detail in Part 3 of the report.

2.2 FAIR data

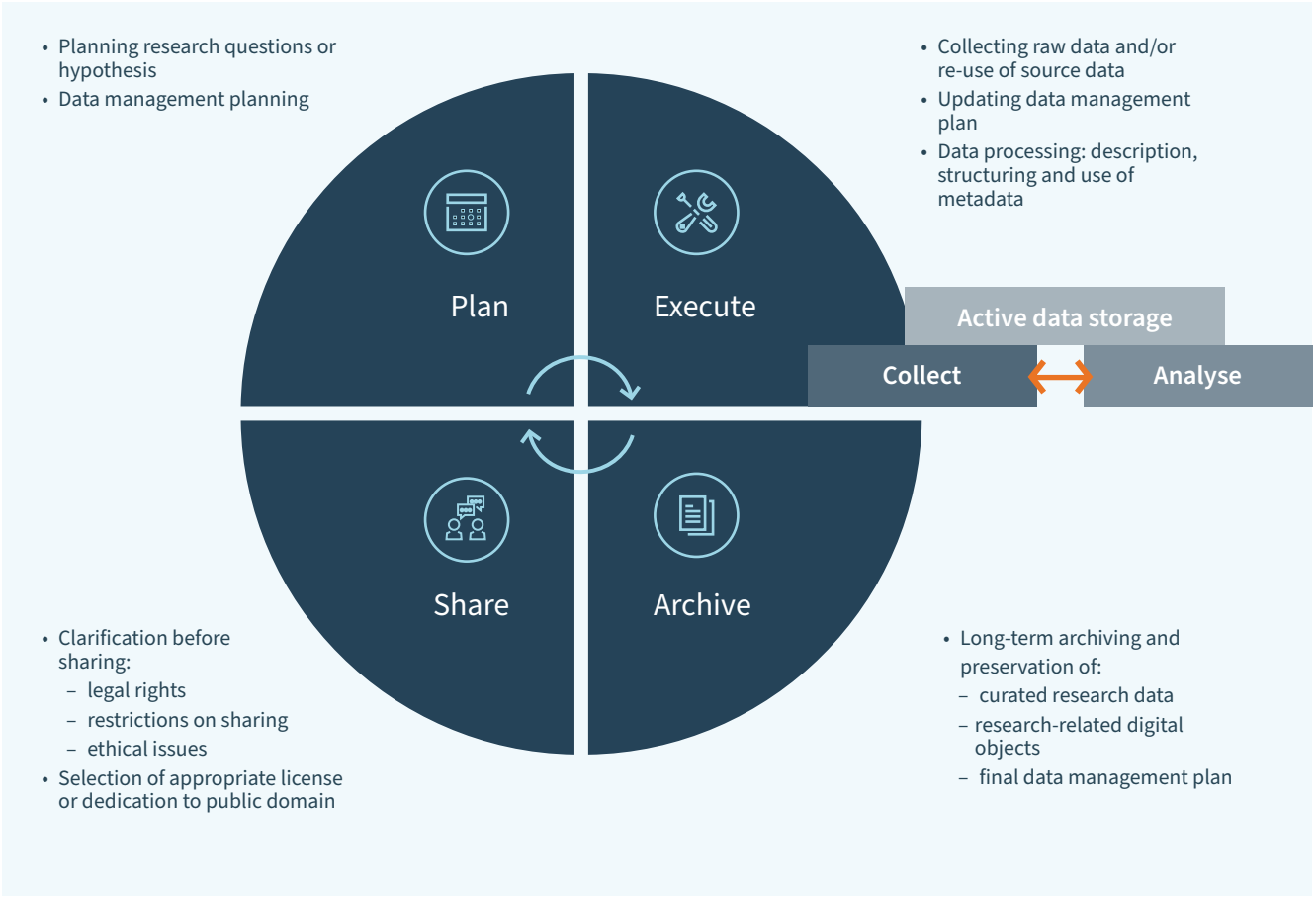
While open data, as a rule, means that the data can be used by everyone without restrictions, the term FAIR data emphasises technical accessibility, which requires good machine-readable metadata including license information. FAIR data must be findable, accessible, interoperable and reusable, but this does not necessarily mean open access for everyone. The costs associated with making research data FAIR exceed the costs of making research data open. To become completely FAIR, data must be standardised to a high degree. This often requires significant investments in terms of both time and resources, since research data are often heterogeneous and unstandardised. However, it is important to underline that the data's FAIRness can be assessed along an axis rather than from an either/or perspective.

The FAIR principles do not necessarily mean that the data should be open, and open data does not necessarily entail that the data should be FAIR, although, as a rule, both are preferable.

3. RESEARCH DATA'S LIFE CYCLE

Research data will typically go through a dynamic life cycle in a research project, from planning and collection to archiving and, where relevant, sharing (cf. Figure 1). In contrast to data from other public sector activities, which are often collected automatically or systematically for public administration purposes, research data are generally collected because the researcher wants to test hypotheses or generate new knowledge about specific topics. Research data must often be linked to rich metadata in order for the data to make sense and be possible for others to re-use. This applies in particular to research data that are not part of large data collections where infrastructure and uniform standards and formats are in place. Research related digital objects with persistent identifiers (PID)[5], associated metadata and, for example, source code and methodology protocol will in many cases also exist beyond the project period together with the data. The storage of data and research-related digital objects requires further resources for maintenance and updating through long-term storage and data curation. This must take place in line with other technological developments to enable re-use and further development of the data later on. Before research data is shared, any legal rights, restrictions on sharing and ethical issues must be clarified before deciding which license or dedications to public domain it will be expedient to use.

Figure 1 The life cycle of research data in research projects



5 What are persistent identifiers (PIDs)? – ORCID

Scientific publications such as articles, books or monographs often form part of the life cycle of a research project. However, research data do not always result in publications, for example if they are seen as negative or uninteresting in light of the hypothesis or research question (often called publication bias). Such data can nonetheless be useful for researchers and service and product developers who see an unexploited potential for knowledge development, re-use or further development. At the same time, such data will be necessary to be able to conduct meta-analysis of the overall existing knowledge in a given field, and a lack of openness in this respect could lead to systematic bias in consequent knowledge summaries. Open project banks with good project descriptions, associated research results and data management plans could be a good tool to prevent publication bias and contribute to robust and reproducible research. OpenAire's project bank is a good example of such a project bank.

The matter of *when* data should be shared in a research project becomes important in a legal sense because research data change status as a legal object[6] depending on where in the project's life cycle the data are. The Committee will come back to this in Part 2.

4. MAKING DATA AVAILABLE AND PUBLICATION

Work methods and processes in the research sector, as in other sectors and industries, have changed in recent years through extensive digitalisation. The way in which data is made available is no exception. Previously, data were generally placed in the institution's internal archives and databases, if they were kept at all after the research had been completed and the scientific articles published. Now, however, data is more often actively made available through APIs and publication of datasets as independent, scientific products with citable titles, attribution and metadata. The introduction of various permanent identifiers, such as DOIs[7], has provided new opportunities for citation, while the use of licenses provides good opportunities to set conditions for re-use. The introduction of these kinds of digital solutions has led to a greater need for standard licenses to replace previous forms of individual agreements making data available manually.

5. RESEARCH DATA THAT QUALIFY FOR SHARING AND RE-USE

As we have shown, research data are a heterogeneous category that encompasses many different types of information and sources. The characteristics of the data can affect how suitable they are for sharing and how much processing is necessary for them to be shared. It is important to underline that a great deal of research data have few or no complicating aspects pertaining to them, and with the right processing, they can be shared without problem. Other data could be protected by law and need more extensive processing to be shared. These can often only be shared in aggregate form or in systems with restricted access. Some data have an even stronger protection in law and may not be lawful to share at all. The latter will by definition not qualify for sharing.

Some disciplines work mainly with interpretation and thus conduct less empirical research that produces results in the form of shareable data. New interpretations of literary or philosophical works or rules of law are examples of elements that cannot be shared in the same way as other types of data. Such interpretations will as a rule form an integral part of a scientific publication to support the answer to the research question, and the data that are the source of the research constitute independent works in themselves. In such cases, the sources must be made available from the infrastructures of those who hold the rights to the content, such as archives, libraries and other online resources. Since this report discusses the sharing of research data, it makes little sense to include such types of information that are generally unsuitable for sharing. Another example of data that are not qualified for sharing is personal data that must be anonymised to such a great extent that, in practice, the quality of the data is degraded, in some cases to the extent that they can no longer be defined as re-usable research data.

When the Committee presents its recommendations in the following, it is important to underline that they apply to research data that qualify for sharing and re-use. The Committee is particularly concerned with such a qualification so that the research communities that are not able to share data to the same extent as others are not disadvantaged as a result of an ambition to increase the sharing of research data, with pertaining measures, instruments and incentive mechanisms.

It will not be possible for a relatively small committee whose members do not represent all aspects of the research areas to define in full what should be considered qualified data, but we believe that the above review can form a good basis for arriving at such a qualification within the different research communities.

[6] Information, documents, databases and publications are examples of legal objects.
[7] Digital Object Identifier System (doi.org)

Part 2

Legal framework and use of licences and dedication to public domain for research data

Photo: Sebastian Duda, Shutterstock

We begin this part of the report with a review of the legal framework that regulates the sharing and re-use of research data. This text is a summary of a more comprehensive text found in appendices 1 and 2 (only available in Norwegian) to the present document. We then go on to discuss different types of licences and dedications to public domain, their function and what consequences the choice of different types of licences and dedications to public domain will have in relation to re-use of research data. Finally, the Committee's overall recommendations concerning the legal framework for sharing of research data and recommendations on the use of licences are presented.

6. LEGAL FRAMEWORK FOR SHARING OF RESEARCH DATA

There are clear socio-political expectations for research data to be shared as much as possible [8,9], but these expectations are not reflected in any coherent legislation regulating the rights to, or sharing and re-use of, such data. It is challenging and difficult for individual researchers, research communities and institutions to comply with political and research ethics expectations for research data to be made available while navigating a fragmented and complex legal landscape.

The purpose of this chapter is to provide a clear and brief overview of the most relevant European and Norwegian rules of law that apply to the sharing of research data. We will discuss both rules that provide a legal basis for sharing and rules that impose restrictions on sharing, within the framework of the national strategy. Since the Committee's work is limited in time and scope, the review does not aim to provide a complete or comprehensive account of all the relevant legislation for specific disciplines. EU legal acts, such as the Open Data Directive[10] are discussed in Subchapter 6.2, while the EU's General Data Protection Regulation,[11] the Database Directive[12] and the Digital Single Market Directive[13] are discussed in connection with the review of relevant national legislation in Subchapter 6.3.

As we will discuss in greater detail in Chapter 7 agreements entered into and licences can also confer rights to, impose restrictions on or stipulate requirements for sharing of research data. The rules of law set the limits for agreements and licences in the field of research. The parties to an agreement can decide to depart from non-mandatory legal provisions. However, some rules cannot be departed from by agreement, including the provisions on right of access in the Freedom of Information Act, research ethics rules, and provisions on personal data, national security and duty of secrecy. In the field of intellectual property law, there is a certain freedom to enter into agreements that depart from the arrangements set out in law, and rights can, in principle, be transferred. The rules on how rights are achieved, i.e. the conditions for intellectual property being protected by the Copyright Act or an invention being patentable, cannot be departed from by agreement. Freedom of contract only applies to the transfer or utilisation of rights that arise or are applied for on the conditions of the law. Agreements concerning rights can also be entered into before the rights in question arise. Although rights can be transferred, the rules of law impose certain restrictions: The originator cannot waive the right to attribution and respect set out in Section 5 of the Copyright Act, and the Act Respecting the Right to Employees' Inventions sets out a framework for agreements on the transfer of inventions. Licensing and dedication to the public domain for research data are elaborated on in Chapter 7.

6.1 The relationship between European and national law

Norway is not a member of the EU, but it is an associated country and is affected significantly by the EU's policy through the EEA Agreement. The Norwegian parliament, the Storting, is the advisory body for the Government in relevant EU/EEA matters through the European Consultative Committee. EU matters of EEA relevance, including changes to Norwegian law, are considered by the Storting.[14]

The EU uses different kinds of legal acts. Norway is obliged to implement binding legal acts that are incorporated into the EEA Agreement, such as regulations and directives, into Norwegian law. The most relevant ones for the present report are the General Data Protection Regulation (GDPR), the Open Data Directive, the Database Directive and the Digital Single Market Directive. A regulation is a binding legal act that must be applied in every detail throughout the EU. A regulation has no direct effect on Norway, but they are often implemented as national regulations and sometimes through legislative amendments or the introduction of new legislation. A directive sets out a goal that the countries must achieve, and is often more general in its wording than a regulation. Directives can be minimum harmonisation, in which case they set minimum standards while countries, including Norway, are free to introduce duties or rights that exceed those outlined in the directive. In Norway, most directives are implemented through national regulations, in the same way as EU regulations, but more comprehensive framework directives may require amendment to existing laws or the introduction of new ones.[15]

[8] https://www.regjeringen.no/en/dokumenter/national-strategy-on-access-to-and-sharing-of-research-data/id2582412/?ch=1
[9] https://www.regjeringen.no/no/dokumenter/meld.-st.-22-20202021/id2841118/?ch=1 (In Norwegian only)
[10] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.172.01.0056.01.ENG
[11] https://eur-lex.europa.eu/eli/reg/2016/679/oj
[12] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31996L0009
[13] https://eur-lex.europa.eu/eli/dir/2019/790/oj
[14] https://www.stortinget.no/no/english/International-delegations/the-delegation-to-efta-and-eea-parliamentary-committees/the-norwegian-parliament-and-the-eea-agreement/
[15] https://www.europalov.no/laer-mer/eu-rettsaktene (In Norwegian only)

6.2 The Open Data Directive

6.2.1 Introduction

The Open Data Directive (EU) 2019/1024[16] is a recast of the previous Public Sector Information Directive (2003/98/EF) with an amending directive (2013/37/EU). This means that the Open Data Directive retains much of the content of the previous directives, in addition to some new regulation. The most relevant point in this context is that the scope of the Open Data Directive has been expanded to include, subject to certain conditions, publicly funded research data. This subchapter is a summary of Appendix 2, which discusses the Directive in more detail.

6.2.2 The Open Data Directive in brief

The 2019 Open Data Directive, which replaced both the two preceding directives, was motivated by the wish to be able to fully utilise the potential inherent in public sector information. The Directive emphasises the following:

- provision of real-time access to dynamic data via adequate technical means;
- increased supply of valuable public data for re-use, including from public undertakings, research performing organisations and research funding organisations;
- handling the emergence of new forms of exclusive arrangements;
- use of exceptions to the principle of charging the marginal cost;
- the relationship between this Directive and certain related legal acts, such as the GDPR[17] and the Database Directive.[18]

The Directive is a minimum harmonisation directive that encourages the member states to go further than what follows from the Directive in making information (data) available for re-use. Important elements include removing obstacles and introducing uniform rules on key issues, for example conditions for re-use, equal treatment and competition rules. This concerns anything from formats, principles for charging, standard licences and exclusive arrangements to special rules for research data.

The Open Data Directive contains no formal definition of open data, but the following is quoted from recital 16 of the preamble:

'Open data as a concept is generally understood to denote data in an open format that can be freely used, re-used and shared by anyone for any purpose.'

Data are not open in the sense of the Directive if any legal restrictions apply to sharing of the data. Examples include if intellectual property rights apply to the data or if the data in question are personal data. A document that is not initially considered open data may be reclassified as such, for example by personal data being rendered anonymous in accordance with the GDPR. The Directive expressly states that there is no obligation to create or adapt public documents or provide extracts of such documents in order to meet user needs if doing so would go beyond a simple operation, cf. Article 5(3). This also applies to publicly funded research data.

6.2.3 Expansion of the scope of the Directive to cover publicly funded research

Sharing and re-use of research data have not as yet been regulated by EU legislation. The grounds cited for expanding the Directive's scope of application in Article 1(1)(c) to include publicly funded research data pursuant to the conditions set out in Article 10 were that the data had already been paid for by society and should therefore be made available whenever possible to contribute to new research and new innovations. Open access policies aim in particular to provide researchers and the public at large with access to research data as early as possible in the dissemination process and to facilitate its use and re-use, cf. recital 27 of the preamble.

A recognition that research data differ somewhat from, e.g., administrative data, is inherent in the conditions that must be met in order for publicly funded research data to fall within the scope of the Directive, cf. Article 10(2). Among other things, publicly funded research data are only covered if they have already been made publicly available through an institutional or subject-based repository. See Subchapter 2.6.2 for further information.

6.2.4 The Directive's definition of research data

The definition of research data is found in Article 2(9):

'"research data" means documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results.'

It follows from recital 27 of the preamble that research data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images, but also meta-data, specifications and other digital objects. Scientific articles differ from data generated from research activities in that they report and comment on findings resulting from scientific research, and are thus exempt.

[16] Directive on Open Data and Public Sector Information ('the Open Data Directive')

[17] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Official Journal of the European Union L 119 of 4 May 2016, p. 1).

[18] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (Official Journal of the European Communities L 77 of 27 March 1996, p. 20).

6.2.5 National policies and measures for publicly funded research data

The Directive obliges member states to adopt national policies, for example in the form of guidelines, and take relevant action to support the availability of publicly funded research data.[19] The purpose is to make publicly funded research openly accessible in line with guidelines for open access. This approach is based on the principle of 'open by default', cf. recital 16 of the preamble. Any restrictions on access to data should therefore be justified. There is a wish to move away from a practice whereby data are closed by default and access is considered on request.

It follows from Article 10(1) second sentence of the Directive that any concerns relating to intellectual property rights, personal data protection and confidentiality, security and legitimate commercial interests, shall be taken into account in accordance with the principle of 'as open as possible, as closed as necessary'. It is therefore emphasised that the Directive does not affect protection of personal data that follows from, e.g., the GDPR, and that means that any re-use must comply with the purpose limitation stipulated there (Article 5(1)(b) and Article 6). Article 10(1) second sentence is thus not to be interpreted as entailing any restriction of the protection, and any flexibility afforded by the GDPR and other relevant legal acts must not be exceeded.

Open access is understood as the practice of providing online access to research outputs free of charge and without restrictions on use and re-use other than the possibility to require acknowledgement of authorship, meaning that the author(s) are credited, cf. recital 27 of the preamble. This does not mean that Article 8 on standard licences does not apply to publicly funded research data. It will be possible to set objective and proportionate conditions for such data too, provided that they are non-discriminatory and justified on grounds of a public interest objective, see Chapter IV of the Directive for details. The Directive encourages the countries to ensure that standard licences are available and encourages their use. However, public sector bodies are not obliged to use them unless the countries make them mandatory. It will be important, in relation to research data as well as other types of data, to establish clear recommendations about which licences should be used for different types of data in order to avoid interoperability challenges etc.

6.2.6 Re-use of publicly funded research data

Article 10(2) deals with more specific cases where publicly funded research data are covered by and can be re-used in accordance with the Directive:

'Without prejudice to point (c) of Article 1(2), research data shall be re-usable for commercial or non-commercial purposes in accordance with Chapters III and IV, insofar as they are publicly funded and researchers, research performing organisations or research funding organisations have already made them publicly available through an institutional or subject-based repository. In that context, legitimate commercial interests, knowledge transfer activities and pre-existing intellectual property rights shall be taken into account.'

It is a condition for re-use that the data must already have been made publicly available through an institutional or subject-based repository by researchers, research performing organisations or research funding organisations in order to fall within the scope of the Directive. Therefore, the Directive does not apply to research data until they have been made publicly available in this way. However, since this is a minimum harmonisation directive, the countries may extend the application of the Directive to research data made publicly available through other data infrastructures than repositories, for example through open access publications, as an attached file to an article or a paper in a data journal, cf. recital 28 of the preamble.

The Directive makes it clear that re-use includes both commercial and non-commercial use. It is thus the re-user's right to decide what to use the publicly funded research data for once access has been granted. This is nevertheless conditional on the use being in accordance with Chapter III on conditions for re-use and Chapter IV on non-discrimination and fair trading – see Chapter 4 of the present document for other provisions of relevance to research data. It also follows from Article 8(1) on the right to set conditions.

The Directive does not define what is considered publicly funded research. This could be interpreted as affording the countries a certain flexibility in interpretation as regards a minimum level of funding. In such case, it would go against the goal of achieving greater alignment of the countries' rules and practices, which is an important objective of the Directive. Other qualifying conditions that must be met in order for research data to fall within the scope of the Directive, for example that data must already have been made publicly available, nonetheless suggest that the countries have limited discretion to construe a certain level of funding. Data produced with more than 'insignificant' public contributions should therefore be assumed to be covered by the designation publicly funded research data.

It also follows from recital 28 of the preamble that certain obligations stemming from the Directive should be extended to research data generated from scientific research activities subsidised by public funding or co-funded by public and private-sector entities. This is not a requirement, and it is thus up to the countries to decide on such extensions. It should be

[19] Norway already has a national policy in this field. However, it falls outside the scope of this chapter's purpose to consider whether the Norwegian policy is in line with the Directive. A draft regulation is under consideration by the EU, and the process going forward may entail changes. https://www.regjeringen.no/en/dokumenter/national-strategy-on-access-to-and-sharing-of-research-data/id2582412/?ch=1

assumed that subsidies in the sense of the Directive does not refer to a general tax break scheme such as the Norwegian Skattefunn scheme. The Directive does not state which, if any, obligations should also apply to data resulting from research co-funded by public and private-sector entities. It will be up to each country to decide this at the national level. In any case, such obligations will be limited to obligations that follow from the Directive, which means that research data will not fall within the scope of the Directive unless they have already been made publicly available. It is thus within the control of researchers and rights holders when and if research falls within the scope of the provisions in the Directive, and scientific articles are exempt, as they are not considered data generated from the research activities. It is also worth mentioning that many journals require research data on which a published work is based to also be made publicly available, thus making the data accessible. Another general consideration in the Directive is to avoid establishing rules that will make companies less likely to invest in public-private partnerships. Member states should strike a balance between the need for data to be open and the financial realities.

6.3 National legislation

Norway has no comprehensive legislation governing research data. Regulation of this area is fragmentary and governed by parts of a number of different laws. The regulatory framework can be divided into three main categories: 1) *intellectual property rules* that confer rights to data, collections of data or inventions based on data, 2) *rules that facilitate, regulate or require sharing of data*, and 3) *rules that restrict sharing of data.*

6.3.1 Intellectual property rules

Research data that constitute works of art or literature (copyright)

Copyright protects works of art or literature that are 'the result of an original and creative intellectual effort', cf. the Copyright Act Section 2 second paragraph, and confers an exclusive right to produce copies of such work and make it available to the general public. This includes, among other things, non-fiction texts and oral lectures, cf. the Copyright Act Section 2 second paragraph letters a) and b). Such texts or lectures may contain research data. Although a person can obtain rights to the text or lecture, only the concrete and unique form of expression can be subject to copyright, and not the actual content. Research data that consist purely of facts or information are not subject to copyright, nor are ideas or scientific theories. Different publications that contain research data may as such be subject to copyright, but the exclusive rights conferred by the copyright do not cover the research data on which the content is based. Copyright can thus give the author exclusive rights to an article or other publication, but presents no obstacle to others making use of the research data.

Within some research disciplines, works of art or literature constitute the actual or empirical basis for research. In such cases, it is natural to characterise them as research data. If the research data contain works of art or literature subject to copyright, further sharing, re-use and further development of research data may only take place insofar as it can be done without infringing on the original rights.

Under the Copyright Act, a database can be protected by copyright if the conditions for protection are met. In order to achieve protection as intellectual property, either the selection of data or the structure and composition of the database must be the expression of an individual creative intellectual effort. It is the compilation of data that may be eligible for copyright protection, and not the data themselves. Scientific databases will rarely qualify as intellectual property, as their arrangement and compilation do not typically entail creative choices. Therefore, it is usually the protection of databases, which is discussed in the subchapter on *Protection of collections of research data (database protection)*, that applies.

Copyright protection is afforded to the author/originator; the person who performed the individual creative effort. Several persons jointly may have performed such efforts as to give rise to copyright. However, supervision, laboratory work, data analysis or funding cannot confer copyright. It is also emphasised that only natural persons may be authors/originators. In order for an institution to have copyright, the copyright must be transferred by means of an agreement. Such agreement can be implicit in an employment relationship, even when this is not explicitly stated in the employment contract.

Protection of collections of research data (database protection)

According to the Norwegian Copyright Act Section 24, a person who creates a 'database' can acquire an exclusive right to control the database if the obtaining, verification or presentation of the contents has entailed 'a substantial investment'. This exclusive right means that the author has exclusive rights to control all or a substantial part of the contents of a database, as well as to copy the database and make it available to the general public.

The Database Directive defines a database as a collection of independent works, data or other material arranged in a systematic or methodical way and individually accessible by electronic or other means.[20] The database protection thus applies to a collection of data, and not raw data or individual data points. It is the *collection* that is subject to protection.

The condition for the protection of databases is that the obtaining, verification or presentation of the contents entailed a substantial investment. It is worth noting that the investment

[20] The Database Directive 96/9/EC, Article 1(2).

RESEARCH WITH COPYRIGHTED SOURCE DATA

In research, the term 'research data' is often not an unambiguous concept, but a collective term that covers data collected by researchers from different disciplines, each with their own distinctive traditions. In some disciplines, the distinction between input data in research projects and intellectual property can be blurry. Artistic research, for example, can be a closely interwoven combination of artistic production and academic reflection. In some cases, the data collected may be protected by copyright. In disciplines that study artistic expressions (language, literature, music, dance, theatre etc.), the research data are often based on copyrighted material. Language corpora, for example, are scientific annotations of text or speech, and each piece of input can be independent intellectual property protected by copyright. A musicological database will often contain sheet music as well as audio and video recordings of different performances. Each performance can have several originators (composer, lyricist, performer, producer, choreographer etc.), and there could be different organisations managing their moral and economic rights. In such cases, further sharing, re-use and development of research data may only take place insofar as it can be done without infringing on the original rights. Overlapping rights makes it more complicated to clarify the copyright issue before the material is published. It is important when clarifying copyright issues relating to research data to safeguard the distinctive nature of the different disciplines and identify discipline-specific challenges.

must be linked to the obtaining, verification or presentation, in other words to the work of structuring and compiling the database. Investments in connection with data production do not entitle a database to protection.

The requirement for a 'substantial investment' as worded in the law is unclear and relative when it comes to who makes the investment and how much it is in monetary terms or time. A common linguistic understanding of what constitutes a substantial investment indicates that a substantial expenditure of resources in terms of time or money is required. Due to the ambiguous meaning of the phrase, what constitutes a substantial investment must therefore be decided on a case-to-case basis.[21]

The holder of the rights to a database is the party that makes the substantial investment. The rights holder will typically be an institution, in contrast to copyright protection, which will often be afforded to the researcher. The institution will also be the rights holder in cases when, for example, an employee has considered a database to be the most appropriate way of structuring research data etc. Even if employees put time and effort into the database's collection or presentation, the employer is the party making the investment if this work takes place during working hours. In connection with externally funded projects, it is possible that rights are afforded to the institution and the external funders jointly.

Regulations to the Norwegian Copyright Act Section 4: Reproductions and extractions for research purposes

Section 4 of the Regulations to the Copyright Act permits the Ministry of Culture and Church Affairs to authorise educational and research institutions to produce copies of intellectual property for research purposes, including in other formats than the original. Section 4 second paragraph also explicitly states that the National Library of Norway may produce copies of intellectual property held in its collections, including in other forms than the original, as source material for language corpora. The National Library of Norway's exclusive right to produce such source material for language corpora for research purposes could result in inexpedient case processing and delays for research institutions that need access to the material.

The new Digital Single Market Directive[22] adopted by the EU in 2019 addresses this issue in Article 3. The Government's position paper[23] states that an exception is to be made from this exclusive right for reproduction and extraction of works by research organisations and cultural heritage institutions, cf. Article 3. The conditions that apply are that the institution must have lawful access to the works or material and that the purpose is text and data mining for the purpose of scientific research. A general exception for text and data mining is provided for in *Article 4*. The exception applies in cases where the rights holders have not expressly forbidden text and data mining and the condition for lawful access to the material is met. In the Government's position paper, these considerations

[21] https://www.duo.uio.no/bitstream/handle/10852/66933/525.pdf?sequence=1&isAllowed=y (In Norwegian only)

[22] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC

[23] https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2016/des/dsm-direktivet/id2556742/ (In Norwegian only)

are deemed at present to be mainly covered by Section 4 of the Regulations to the Copyright Act, which permits institutions to produce source material for research purposes. However, a significant limitation applies in that all institutions, with the exception of the National Library of Norway, must apply for permission to do so. The Digital Single Market Directive is currently under consideration by the EEA/EFTA states, including Norway. The Committee is of the opinion that if the Directive is implemented, removal of the requirement for permission from the ministry in Section 4 first paragraph of the Regulations should be considered for cases covered by the Directive's Article 3 on text and data mining for the purpose of scientific research. If relevant, the Directive's condition for lawful access to the works or material could be included in the wording of the Regulations.

Exclusive right to inventions and publication of patent applications (patent rights)

Patent law can give an inventor (or another party to whom the rights have been transferred) the exclusive right to exploit a technical solution. In cases where research data form the basis for the development of an invention, an exclusive right can be obtained to exploit the invention, but not to the data. In order for patent rights to be obtained, information about the invention must not be known before the patent application is filed. This could mean that research data are kept secret until a patent on the invention has been applied for. However, it is a requirement that information about the invention be made publicly available by publication of the patent application. The research data contained in a patent application will thus become publicly available and can be used in further research and development as long as the patent rights are not infringed on by exploitation of the financial value of the patented invention. In other words, patent rights do not present an obstacle to the invention being used as a source of knowledge, also for purposes of research and experimentation, including commercial research and development.

Employers' right to inventions and prohibition against sharing information

In situations where an employer is entitled to take over the rights to an invention pursuant to the Employees' Inventions Act, this will limit the employee's right to share information about the invention to the extent that such sharing could impair the possibilities of patenting or enable a third party to exploit the invention, cf. Section 6 second paragraph. Lecturers and academic staff at universities and university colleges nonetheless have a right to publish pursuant the Section 6 third paragraph, and this right takes precedence over the employer's right to an invention. In cases where a patent is applied for, the application, including any research data it may contain, will be published in accordance with ordinary procedures. Once a patent application has been filed, no limitations apply to the inventor's sharing of information.

6.3.2 Rules that facilitate, regulate or require sharing of research data

This section gives an account of the regulatory framework that, in principle, facilitates and requires the publication of research. The Freedom of Information Act confers right of access to documents held by public sector bodies, including documents containing research data. The Research Ethics Act and the Universities and University Colleges Act set requirements for transparency and publication of research, both by the institutions and by the researchers themselves. The Health Research Act also sets strict requirements for the publication of results in its field. However, the provisions on publication and public disclosure must be considered in light of the restrictions and exceptions that follow from the individual acts, but also restrictions on data sharing imposed by other legislation and regulations. Subchapter 6.4 deals with rules that primarily restrict the sharing of data.

The Freedom of Information Act

The Freedom of Information Act confers right of access to documents held by public sector bodies. Such documents could be research data, and the Freedom of Information Act can also grant right of access to these. In other words, the Act requires research data to be shared under certain circumstances.

The Freedom of Information Act only applies to the legal persons listed in Section 2 of the Act. This includes the state, the county authorities and the municipal authorities (letter a), other legal persons that make individual decisions (letter b), independent legal persons in which the state or municipal authority has an equity share that gives it more than half of the votes in the supreme body of that legal person (letter c), and independent legal persons in which the state or municipal authority directly or indirectly has the right to elect more than half of the voting members in the supreme body of that legal person (letter d). In cases where public sector bodies that fall within the scope of the Freedom of Information Act Section 2 commission or purchase services, such as reports, from entities not covered by Section 2, the rights to the service and any underlying material should be regulated in an agreement to ensure transparency and access on the part of the public sector body.

The Ministry of Justice and Public Security wrote the following in an interpretation statement: 'As regards bodies that fall within the scope of the Freedom of Information Act, the Act shall apply to the body's areas of activity – public administration activities, service provision, business activities, internal administration etc.' If the public sector body's activities include research, the Freedom of Information Act will also apply to this area of activity. The Regulations to the Freedom of Information Act list some exceptions from the Act's scope in Section 1. Exceptions are made for certain legal persons and for certain documents of independent legal persons. One of these exceptions concern documents related to cases concerning commercial utilisation of research results and cases concerning grant-supported research and commissioned research carried out by legal persons in the higher education sector, cf. Section 1 third

paragraph letter d). In this context, independent legal persons means limited liability companies, foundations etc.

The right of access is contingent on someone requesting access, cf. Section 3 ('Any person may apply to an administrative agency for access'). This provision also states that case documents etc. are public 'except as otherwise provided by statute or by regulations pursuant thereto'. The right of access can thus be limited by other statutory provisions restricting the sharing of information. Examples include provisions on duty of secrecy, data protection or protection of trade secrets. Contractual provisions that stipulate that information is to be kept confidential do not constitute a sufficient independent basis for exemption from right of access.[24] The right of access is not the same as a general order to make information public, and requests for access are considered on a case-to-case basis.[25] Requests for access may be made orally or in writing, but it must in principle relate to 'a specific case', cf. Section 28 second paragraph. This means that the request must be sufficiently specific – the person requesting access must know what he or she wants access to. When it comes to research data that have not yet been made public, their existence will not be public knowledge, and such data may therefore be less available to requests for access. It can be easier to formulate a sufficiently specific request for access in relation to the underlying data when an article or other publication exists. In principle, the right of access applies to 'case documents' (in addition to 'journals' and 'similar registers'), cf. Section 3.

The term 'document', which falls under 'case document', is defined in Section 4 of the Act. A 'document' is any logically limited amount of information stored in a medium for subsequent reading, listening, presentation, or transfer or the like. This is a broad definition and covers more than we traditionally think of as a text document. The requirement is that it must be a logically limited amount of information that can be stored and read or similar. This could be an ordinary text document, but the definition also covers sound, images, drawings, models etc. It is a technology-neutral term, and how the information is expressed is irrelevant.[26] The Ministry of Justice and Public Security's Legislation Department specified in an interpretation statement[27] that the Freedom of Information Act does not warrant a general exemption from public access to research reports, except in cases where the report contains information subject to a duty of secrecy or similar protection. Therefore, any exemptions must be based on content and considered on a case-to-case basis for each request for access. The statement goes on to say that the Freedom of Information Act does not warrant a general exemption for documents relating to research activities.

The Research Ethics Act

It follows from the provision relating to the purpose of the Act in Section 1 that the Act 'seeks to ensure that research carried out by public and private institutions is conducted in accordance with recognised ethical standards'. The Act regulates the prevention and breach of *recognised ethical standards for good scientific practice*.[28] In the event of conflict, statutory ethics requirements set out in the Health Research Act and the Personal Data Act, among others, will take precedence over the Research Ethics Act.[29] The Act also requires the appointment of National Committees for Research Ethics, cf. Section 9, and Regional Committees for Medical and Health Research Ethics (REC), cf. Section 10.

In the General guidelines for research ethics prepared by the Norwegian National Committees for Research Ethics, Section 11 assumes that as a rule, research results should be made available.[30] Failure to make research data available could thus constitute a breach of research ethics norms. Other research ethics norms that apply to the collection and publication of research data include requirements for integrity (Section 7) and a duty to observe good reference practices when using other researchers' material (Section 8).

The Research Ethics Act can require research data to be made available, to the extent that this follows from recognised ethical standards. When using other people's research data, norms concerning integrity and good reference practices must be observed. To the extent that breach of a research ethics norm can be described as serious and the researcher has acted with intent or gross negligence, a case could conclude in a statement that a researcher has acted in a manner that constitutes scientific misconduct.

Breach of research ethics norms in connection with research can also result in a statement that a scientific work should be corrected or withdrawn.

The Act relating to Universities and University Colleges

Among other things, the Act relating to Universities and University Colleges shall facilitate research and academic and artistic development work at a high international level by universities and university colleges, cf. Section 1-1 letter b). The Act should also help these institutions to 'disseminate knowledge of the institution's activities and promote an understanding of the principle of academic freedom and application of scientific and artistic methods and results', cf. Section 1-1 letter c).

[24] https://www.regjeringen.no/no/dokumenter/-23-og-29---forsking-i-offentlege-forval/id636453/ (In Norwegian only)

[25] It is worth mentioning in this context that the Ministry of Local Government and Modernisation commissioned an assessment of the possibility of a *duty* to *actively make available* public information, carried out by the law firm Kluge Advokatfirma.

[26] Ibid., note 68.

[27] https://www.regjeringen.no/no/dokumenter/-2---forsking-i-offentlege-forvaltningso/id636199/ (In Norwegian only)

[28] Lovkommentar til forskningsetikkloven, Rettsdata, by Bjørn L. Zwilgmeyer, note 1. (In Norwegian only)

[29] Ibid.

[30] https://www.forskningsetikk.no/en/guidelines/general-guidelines/

Section 1-3 lists different ways in which universities and university colleges are to promote the purpose of the Act. It is explicitly stated that the institutions must '[help] to disseminate and communicate the results of research and of academic and artistic development work' (letter d). This can be said to entail an obligation to encourage and facilitate such dissemination.[31] However, it also follows from this provision that the institutions must also '[contribute] to innovation and value creation based on results from research' (letter e). Comments to the act emphasised that it is desirable for results of research and academic development work with a commercial potential to be disseminated to enterprises capable of contributing to innovation and value creation. Such knowledge transfer can create new jobs and strengthen the Norwegian knowledge industries.[32]

The purpose of commercialising results from publicly funded research is that the results are further developed in a way that benefits society at large. There may be, or at least appear to be, a conflict between the goal of commercialisation and the goal of publication. If commercialisation is dependent on patent applications, then the data cannot be made public until the application has been filed, as it would otherwise be impossible to obtain a patent. This is not the only scenario in which commercialisation may depend on (temporarily) having exclusive access to data. How far and for how long exclusive access should be accepted in order to facilitate commercialisation must be determined through concrete assessments on a case-to-case basis, and the legislation does not provide much guidance in this process.

It also follows from the Act relating to Universities and University Colleges Section 1-5(6) that universities and university colleges 'must ensure transparency regarding the results of research'. *The researchers' right and duty to publish* are also regulated in detail by this provision. Persons appointed to a position where research is part of their duties, are *entitled to publish* their results and *must make sure* such publication takes place. However, the institution cannot instruct them to publish at any particular time or frequency. The researcher decides when and how the results are to be published.[33]

It follows from Section 1-5(5) that a researcher is entitled to choose the topic and method for his/her research and development work within the framework of the employment contract or a special agreement. It is also stated in Section 1-5(6) that '[t]he relevant research basis must be made available in line with good practice in the field'. The Act thus specifies that not only must the results be published, but the 'research basis' must also be made available. That could mean e.g. basic statistical data or physical material.[34] That it must be made available means that it must be available on request.[35] However, information subject to a duty of confidentiality must in principle be protected or anonymised.[36]

The board of the institution can consent to 'deferred publication' of research results where indicated by legitimate considerations. Examples of such legitimate considerations are protection of patent rights or competition interests or consideration for ongoing research.[37] However, it is explicitly stated that no 'permanent restrictions' on the right to publish results can be agreed or stipulated beyond what follows from the law.

The Health Research Act

The purpose of the Health Research Act is to promote good and ethically sound medical and health research, cf. the Health Research Act Section 1. The Act applies to all medical and health research on human beings, human biological material or personal health data, cf. Section 2.

Section 8 of the Act states that '[c]ommercial exploitation of research participants, human biological material and personal health data in general is prohibited.' This provision is fairly vague, and the more detailed demarcation is left to the field of practice.[38] This provision is not intended to prevent research for commercial purposes, for example drug trials.[39] The sale and commercialisation of research results and processed material also fall outside the scope of this provision.[40]

Research projects are required to be approved by the Regional Committee for Medical and Health Research Ethics (REC) on the basis of an application, cf. Sections 9 and 10, cf. the Research Ethics Act Section 10 first paragraph. Section 12 states that when a research project is concluded, the project manager must submit a final report to REC. The report must present the findings objectively and methodically, ensuring that both positive and negative findings are presented. The purpose of this provision is to contribute to control and transparency.[41] Society should be able to keep track of what medical and health research is conducted, and it must be ensured that results are not withheld, for example if they are unfavourable for the party funding the research.[42]

---

31 Lovkommentar til universitets- og høyskoleloven. Rettsdata, by Jan Fridthjof Bernt, note 25. (In Norwegian only)
32 Ibid., note 26.
33 Ibid., note 67.
34 Ibid., note 68.
35 Ibid., note 69.
36 Ibid.
37 Ibid., note 72.
38 Lovkommentar til helseforskningsloven. Rettsdata, by Sigmund Simonsen, note 40. (In Norwegian only)
39 Ibid.
40 Ibid.
41 Proposition No 74 to the Odelsting (2006–2007) Section 11.3.4.2.
42 Ibid.

The Health Research Act also contains rules on research biobanks that are established in connection with the collection, storage and use of human biological material. These rules are found in Chapter 6 of the Act. Section 31 states that the person or body responsible for the research must grant other researchers access to human biological material in the enterprise's research biobanks, unless the person or body responsible for the research needs the material or other exceptional grounds exist.

Chapter 7 sets out rules on the processing of personal health data, which must, among other things, comply with the principles described in the GDPR Article 5 and have expressly indicated objectives, cf. Section 32. It follows from Section 34 that personal health data may be processed, compiled and shared in keeping with the objective of the research project, any consents pursuant to Section 33 and in accordance with the research protocol. However, REC can deny such compilation or sharing etc. of data if it is deemed to be medically or ethically unsatisfactory, cf. Section 34 second paragraph. It is emphasised that compilation and sharing of personal health data can take place to a data controller or person or body responsible for the research that has the right to process such information pursuant to the GDPR Articles 6 and 9.

Chapter 8 sets out rules for *transparency and right of access to the research*. It follows from Section 39 that the person or body responsible for the research and the project manager must ensure transparency in relation to the research. Section 41 regulates public right of access. Anyone who contacts REC shall be informed about which research projects a particular person or body responsible for the research or project manager is (or has been) involved in, as well as the objective of the project. The exceptions in the Personal Data Act Sections 16 and 17 apply correspondingly here, cf. Section 42. Finally, it follows from Section 44 that REC is to keep a systematic register of ongoing and completed research projects based on information in applications and final reports, and that these registers must be public. However, Section 45 provides for the possibility of applying for deferred publication in cases where this is necessary to protect legitimate interests linked to patents or competition, or in the interests of ongoing research. Also, REC may decide that sensitive information about a project shall not be recorded in the register or that access may not be demanded for a defined period, if publication could harm significant private or public interests.

6.4 Restrictions on sharing of data

In this subchapter, we will take a closer look at legislation that imposes restrictions on sharing of data for reasons related to the *contents* of the data. This includes restrictions put in place to protect the data subject, including the data protection legislation and GDPR, restrictions intended to protect other interests, such as trade secrets, and restrictions in place for reasons of national security. If research data are shared in violation of these rules, the party sharing the data will have broken the law.

Restrictions on the sharing of data due to rights to data, data collections or inventions based on data are discussed in Subchapter 6.3.1. The prohibition on the sharing of information in the Employees' Inventions Act, as explained above, applies to information about an invention and not to individual data points as such.

6.4.1 Restrictions established to protect the data subject

The Personal Data Act

The Norwegian Personal Data Act implements the EU's General Data Protection Directive (GDPR) and regulates processing of personal data. The Act also applies to processing of personal data in connection with scientific research. When research data include personal data, the Personal Data Act will stipulate requirements concerning the processing of such information. The Personal Data Act may impose restrictions and requirements on the publication of research results and research data that include personal data.

The objective of the GDPR is to protect natural persons in connection with the processing of personal data. The Regulation protects their right to protection of personal data, while the free movement of personal data within the EEA shall be neither restricted nor prohibited (see Article 1). In what is known as the Schrems II judgment,[43] the Court of Justice of the European Union concluded that personal data protection pursuant to the GDPR follows the data where they go. When personal data are transferred out of the EEA, they must have the same or corresponding protection as they have in the EEA.[44] 'Personal data' means any information relating to an identified or identifiable natural person ('data subject'), cf. Article 4. An identifiable natural person is defined as one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or similar. Recitals 34 and 35 of the preamble contain statements about personal data in the form of genetic data and information about health status.

Section 8 of the Personal Data Act states that personal data can be processed pursuant to the GDPR Article 6(1) letter e) if necessary 'for scientific or historical research purposes or

[43] https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:62018CJ0311
[44] https://edpb.europa.eu/system/files/2021-06/edpb_recommendations_202001vo.2.0_supplementarymeasurestransferstools_en.pdf

statistical purposes'. Recital 159 of the preamble to the GDPR states that '[f]or the purposes of this Regulation, the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research'.[45] In such cases, the processing shall be covered by the necessary safeguard pursuant to the GDPR Article 89(1). This includes safeguards for the rights and freedoms of the data subject. It also follows from recital 159 of the preamble to the GDPR that '[t]o meet the specificities of processing personal data for scientific research purposes, specific conditions should apply in particular as regards the publication or otherwise disclosure of personal data in the context of scientific research purposes. If the result of scientific research in particular in the health context gives reason for further measures in the interest of the data subject, the general rules of this Regulation should apply in view of those measures.'

Concerning research, recital 33 of the preamble to the Regulation reads as follows: 'It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research.' If the data subject withdraws her consent, the withdrawal shall not affect the lawfulness of processing based on consent before its withdrawal, cf. Article 7(3).

Section 9 of the Personal Data Act deals with the processing of special categories of personal data without consent for scientific research purposes. Such processing shall be covered by the necessary safeguards pursuant to the GDPR Article 89(1). This means implementing the technical and organisational measures necessary to ensure compliance with the principle that no more personal data than is necessary to fulfil the purpose of the processing shall be made available. Before a researcher may process personal data without consent, the data controller, for example a research institution, must seek the advice of the data protection officer to determine whether the processing meets the statutory requirements. This is not necessary if the data controller has carried out a data protection impact assessment pursuant to the GDPR Article 35.

It is worth mentioning that the Norwegian Data Protection Authority provides further information on its website about the processing of personal data in connection with research. The Data Protection Authority emphasises that the data controller must assess and document that the planned personal data processing meets the statutory requirement. This requirement applies to all research where personal data are used. No exceptions are made even if the research project has been approved by REC, and REC's decisions do not provide a legal basis for the processing of personal data. The Data Protection Authority also underlines that in many cases, exemption from a duty of secrecy will be required to process personal data, and that REC will consider whether such an exemption can be granted for patient records. The European Data Protection Board (EDPB) has adopted guidelines on the processing of data concerning health for the purpose of research and on the use of location data and contact tracing tools in the context of the COVID-19 outbreak.

Duty of secrecy rules

The Public Administration Act Section 13 a–g is one source of duty of secrecy rules. Section 13 imposes a duty of secrecy on public sector employees when it comes to information about (1) an individual's personal affairs, and (2) technical devices and procedures, as well as operational or business matters which for competition reasons are important to keep secret in the interests of the person whom the information concerns. Section 13 e) imposes a duty of secrecy on researchers in relation to certain types of information, including information subject to a duty of secrecy and information received from private sources upon pledge of secrecy in connection with research. The duty of secrecy pursuant to Section 13 of the Public Administration Act shall not prevent information from being made known if those to whom the duty of secrecy is owed consent thereto, cf. the Public Administration Act Section 13 a(1). Nor shall this duty of secrecy prevent information from being used when the need for protection must be deemed satisfied by the information being presented in the form of statistics or by otherwise eliminating identificatory characteristics, cf. the Public Administration Act Section 13 a(2).

6.4.2 Restrictions intended to protect other interests

The Trade Secrets Act

The Norwegian Trade Secrets Act is intended to protect trade secret holders against unlawful acquisition, use or disclosure of a trade secret, cf. the Trade Secrets Act Section 1. The term 'trade secret' is defined in Section 2 of the Act and comprises information that is (a) 'secret' in the sense that it is not generally known or readily accessible, (b) has commercial value because it is secret, and (c) that the holder has taken reasonable steps to keep secret. According to the preparatory works to the Act, information has commercial value if infringement is likely to harm the interests of the trade secret holder in that it undermines that person's scientific and technical potential, business or financial interests, strategic positions or ability to compete.[46] If a trade secret is infringed on in the process of obtaining or publishing research results, this could be a violation of the Trade Secrets Act and result in sanctions being imposed pursuant to Sections 5–10 of the Act. If the research results are commercialised and exclusive access to research data constitutes a competitive advantage that is important to

[45] Norsk Lovkommentar's comments to the Personal Data Act, note 8-7.
[46] Proposition 5 to the Storting (Bill and Resolution) (2019–2020), Section 5.1.3.

the commercialisation, these data may constitute trade secrets subject to protection under the Trade Secrets Act.

The Security Act

The Norwegian Security Act is intended to help protect national security interests and prevent activities that present a threat to security, cf. the Security Act Section 1-1. The Act applies to governmental, county and municipal bodies, cf. Section 1-2. National security interests include general political security interests related to, inter alia, the activities of the highest state bodies, defence, security and contingency preparedness, relations with other states, economic stability, fundamental societal functions and the basic security of the population, cf. Section 1-5. In connection with research that involves information that could be protected under the Security Act, the provisions of the Act may restrict the obtaining and publication of research data and research results.

6.5 About ownership of research data – can data be owned?

Discussions about sharing of research data often focus on who 'owns' the data; see also Report No 22 to the Storting (2020–2021) Data som ressurs – Datadrevet økonomi og innovasjon ('Data as a resource: data-driven economy and innovation' – in Norwegian only). It is difficult to give a clear answer to this question. Legally speaking, an 'owner' is a person who has right of ownership of a physical object, for example a car. The right of ownership is exclusive, and other persons cannot use the car or have it at their disposal without the owner's consent. Data, on the other hand, are information rather than a physical object. Information can be used by a large number of people at the same time, and this will indeed often make it more valuable. The rules of law are therefore based on the view that data or information should, as far as possible, be freely available, and exclusive rights to data are only envisaged where indicated by particular needs. Examples include the rules on copyright, protection of databases and trade secrets, which allow for a time-limited exclusive right to utilise information subject to strict conditions. An overriding purpose of these rules is to encourage innovation.

However, the term 'ownership' of data signals an expectation of a right to determine the use of the data, as would be the case with the car in the example given above. As regards research data, this is linked to the researcher's sense of 'ownership' of his or her research, the institution's perception of employers' rights to the product of employees' work, and also data subjects' perception of a right to control the further use of personal data. These expectations are only partially met by their legal rights. Issues between the researcher and the institution, as employer, must be resolved in accordance with the provisions of labour law and intellectual property law. If the data or databases in question are subject to copyright, the copyright arises as a personal right of the researcher, while the database rights and rights to trade secrets are afforded to the institution. For data not protected by any rules, the prevailing view is that the institution has the same right to determine the use of the data as it has in relation to physical research material and other products of its employees' work. As regards the expectations of the data subjects, these are addressed by the Personal Data Act and any other relevant statutes.

The term 'right of ownership' signals a form of control, but also a responsibility to manage what you have the right to control in accordance with the applicable rules in force at all times. In relation to right of ownership of data, this will entail an expectation that the party determining the use of the data will also have a responsibility to manage them appropriately. When political guidelines apply to the management of data, for example in the form of greater sharing, the need to define a responsible subject, i.e. an 'owner', for the data arises. It is possible that several persons could determine the use, and have responsibility, perhaps particularly where research data are concerned, but it is important that this is clarified in each case. In an employment relationship between an institution and a researcher, the right to determine the use and corresponding responsibility for managing research data will usually go to the institution pursuant to the provisions of labour law and intellectual property law. Making research data available entails determining what, where, how and when to share data. The best practical solution to the responsibility for making research data available is to divide responsibility between the institution and the researcher. The institution must ensure that it has clear internal guidelines and governing documents, as well as a competent support system to assist researchers with the practical aspects of making data available in accordance with the applicable procedures, regulations and ethical assessments.

6.6 The Committee's assessment of the legal framework

The sharing of research data takes place within the complex regulatory framework that is outlined above and described in more detail in Appendices 1 and 2. There are several other Norwegian acts that the Committee has not detailed in the above, but that nevertheless apply to certain disciplines. They include the Environmental Information Act,[47] the Nature Diversity Act,[48] the Nature Inspectorate Act[49] and the Archives Act,[50] of which the latter applies to all public sector bodies, with some exceptions.

The Committee considers the intention and direction of the legislation to be clear: *Data shall to the least possible extent be*

[47] https://www.regjeringen.no/en/dokumenter/environmental-information-act/id173247/
[48] https://www.regjeringen.no/en/dokumenter/nature-diversity-act/id570549/
[49] https://lovdata.no/dokument/NL/lov/21/06/1996-38 (In Norwegian only)
[50] https://lovdata.no/dokument/NL/lov/04/12/1992-126 (In Norwegian only)

subject to exclusive rights, and transparency shall be facilitated to the greatest possible extent. In the Committee's view, the fragmentary nature of the regulatory framework represents the greatest challenge for the researchers and institutions that are going to share research data. It entails finding answers to many different questions before being able to request access to data or ensure that data can be shared in a lawful manner. In brief, the person who is to decide whether and how a dataset can be shared must consider whether the data are subject to exclusive rights (such as copyright or database rights), who, if any, hold such rights, whether the nature of the data entails restrictions on sharing, and, if so, whether adequate consent has been obtained. All of these assessments require knowledge of the applicable legislation.

The legal expertise of the person making the assessments must be sufficiently broad to cover several areas of law, and he or she must have sufficient in-depth knowledge to be able to make the assessments necessary to apply the rules correctly. It is hardly feasible for all individual researchers to develop such expertise, and the institution must therefore provide it. In the Committee's opinion, the fragmentary regulatory framework combined with the mismatch between expectations or perceptions of rights to data and the rules that apply, gives rise to a particular need for institutions to ensure that they possess sufficient expertise to share data and that they put in place a system that makes it easy for researchers to share data in a secure and lawful manner. These systems must be adapted to deal with the fact that different disciplines produce very different types of data that may be subject to quite different rules concerning whether, and how, data can be shared.

Although a simple and comprehensive regulatory framework regulating the sharing and re-use of research data would be convenient, it is challenging to establish. Due to the heterogeneity of the data, regulations for research data will inevitably be fragmentary compared with those for other types of data. A comprehensive regulatory framework would also require regulations to be coordinated across sectors. The Committee is therefore of the opinion that it is probably not expedient to regulate research data through separate legislation at present.

The Committee wishes to emphasise that the current regulation facilitates fairly extensive sharing of research data, even though the rules of law impose certain restrictions for the protection of other interests, such as personal data protection, trade secrets and security. Personal data are also subject to restrictions on ethical grounds. For rules that target particular types of data, such as the data protection rules, an assessment of the nature of the data ('Are these data personal data?') will be crucial to whether or not the rules are applicable. This is a limited and, in principle, not particularly complex question of application of law. When it comes to rules that protect interests relating to the data that are not linked to their nature, however, such as the rules on trade secrets, the assessments conducted to determine whether or not the rules are applicable can be somewhat more complex.

The most important element of uncertainty appears to be linked to whether the different parties involved have an 'ownership' of the data. As explained above, there is a discrepancy between the expectation of 'ownership' and the assumption on which the legislation is based, namely that data, in the sense of information, should in principle not be subject to exclusive rights. This introduces a certain level of uncertainty in the relationship between the researcher and the institution as regards who is entitled to and, if relevant, responsible for making research data available. The implication of this for the relationship between the researcher and data subjects, for example informants in a research project, is that it must be clarified whether sufficient consent has been obtained to share any personal data.

Other involved parties, be it other institutions, funders, or private or public sector partners in a research project, may hold different expectations of control over further use of the data. An expectation associated with the word 'ownership' is not necessarily linked to an exclusive right of use, but to the right to oppose further sharing of 'our' data, or a right to some form of gain should the data form a basis for financial profit further down the line. Such expectations can give rise to uncertainty among some researchers in relation to whether data can be shared, whether restrictions are necessary and whose consent is required to share data. Uncertainty associated with such expectations can thus indirectly prevent the sharing of research data. This uncertainty cannot be resolved by means of licences, as data sharing under a licence is dependent on the party sharing the data having sufficient rights to share them on the conditions set out in the licence. In the Committee's opinion, there is a need for continuous assessment of whether the rights to and responsibility for publicly funded research data should be enshrined in law in order to clarify expectations regarding 'ownership'.

7. LICENCES AND DEDICATIONS TO PUBLIC DOMAIN FOR SHARING AND RE-USE OF RESEARCH DATA

The purpose of this chapter is to provide an overview of licences and dedications to public domain for sharing and use of research data. In the sense in which the term is used by the Committee in this chapter, 'research data' do not include research-related digital objects such as source code, software, models and methodology protocols etc. We have made this delimitation because the Committee's work was limited in time and scope, but the report will deal briefly with the matter of licensing of research-related digital objects under Subchapter 7.3.6.

7.1 About licences and their function

When research data are made available to others, the use of such data is usually linked to a licence. A licence means a permission. The licence that applies to the data is an agreement between the party making the data available and the party making use of the data. The licence regulates any terms and conditions for use of the data, for example that they are only to be used for research purposes or that the party that produced the dataset must be credited.[51] A licence is required if the data to be shared are subject to exclusive statutory rights, such as copyright, or if the use of the data is limited by previous contracts, for example limitations relating to consent. Licences are also currently used as a means of ensuring that the use of the data lives up to research ethical 'best practice'. If the data to be shared are neither protected by law nor subject to statutory or other restrictions on sharing (e.g. contracts or consent), no licence is required for use of the data to be lawful. However, if the party making the data available wishes to impose its own restrictions on the use of the data, for example that they can only be used for research purposes or that their origin must be attributed to the source, a licence will be required. Even in cases where a party does not wish to impose restrictions on the use of data, a licence is nevertheless orderly, practical and recommended in order to make it clear that the researcher and institution waive any rights to the data.

A licence is a contract that forms the basis for and describes the content of rights and obligations that exist between the party sharing the data and the party wishing to use them. Restrictions on the use of research data that follow from rules of law intended to safeguard other interests, such as the Personal Data Act, the Security Act, the Patents Act and the Personal Health Data Filing System Act, cannot be changed or regulated through a licence. The same is mostly true of third party rights to the data – these rights must be respected and incorporated into a licence. This means, for example, that if the data were collected on the basis of a limited consent, this limitation must be respected and integrated into the licence if the data are to be shared further.

Strictly speaking, a licence with no restrictions at all on use is not a licence at all, but a *Dedication to Public Domain* (data fall into the public domain). There are two main types of dedications to public domain: Public Domain Mark (PDM) and CC0. PDM is used to document that no known copyright applies to the data, while CC0 is used to show that the researcher relinquishes all copyright to the data.

A licence thus serves several purposes: It may be necessary to make use of the data lawful, it may stipulate terms and conditions for the use, and it can document that rights are waived.

Depending on whether a dedication to public domain or some form of licence is used, the degree of openness can vary from data in the public domain (e.g. PDM or CC0), where no restrictions on use apply, to licences with terms and conditions relating to attribution, processing and adaption, share-alike and commercialisation (e.g. CC 4.0-BY). We will discuss this in more detail in Subchapter 7.3.

7.2 Conditions and considerations concerning sharing and re-use of research data

Publicly funded research data should ideally be shared subject to as few restrictions on access and purpose as possible. However, sharing and re-use must take place within a complex legal framework, and research ethics norms with a bearing on which data can be shared, and when, must also be considered. In the following, we will describe different conditions and considerations on which choices of licence or dedication to public domain should be based.

7.2.1 Licensor's responsibilities – clarifying rights and restrictions on sharing

In this case, a licensor is a researcher or an institution where the research data were produced by employees in the course of their employment relationship. The licensor is responsible for deciding which licence or dedication to public domain to use and for granting it. Before a licence is granted, it is important that the licensor has clarified whether, and if so, to what extent, the research data are subject to legal rights such as copyright and database protection, and whether any third party rights exist. A licensor does not necessarily have to be the holder of the rights to the research data in cases where rights are conferred by law or agreement, but the licensor must in such case obtain permission for licensing from the rights holder.

In addition, the party making the data available must ensure that the data are not subject to legal restrictions on sharing, for example pursuant to the Personal Data Act or for security considerations. It should also be considered whether sharing the data will disrupt the institution or researcher's own utilisation of them, for example if the data form part of an

[51] Licence conditions for attribution are often stricter than the recognised research ethical standards for good reference practice.

invention that is to be patented or if they are for other reasons deemed to constitute trade secrets.

The distinction between source data and data generated from a research project is significant in relation to rights to data and the use of licenses. OpenAire provides a useful guide to using someone else's research data. Researchers often use source data for further research, in which case they must comply with licences and agreements that regulate the use of such data. In many cases, research will entail a combination of the use of source data and data generated by the projects. It may complicate the consideration of further sharing if licences that apply to source data do not correspond to or are not compatible with licences one wishes to use for derivative data. This is discussed in more detail in Subchapter 7.3.2.

In addition to discussing which data to share, a licensor must decide when to share the data. This is a complex question to which there is no simple answer. Research data that form the basis for a scientific publication, for example a peer-reviewed scientific article, should as a rule be shared no later than at the time of publication. 'Research data' here refers to the minimum of data needed to validate and expand on the conclusions in the publication. Applying an embargo, meaning deferred sharing, to research data on which published scientific work is based should therefore be avoided, as it would harm the integrity of the research. Several scientific journals require research data and research-relevant objects on which a scientific publication is based to be shared at the time of publication.

When applying for a patent, it is a requirement that the invention must not be known. Therefore, data on which a patent application is based cannot be made publicly available until the application has been filed. The main purpose of patent rights is to create incentives for further research and innovation. This is why patent rights confer an exclusive right to utilise the invention, but the information about the invention contained in the patent application must be made publicly available, including research data included in the application. Patenting, primarily in connection with the commercialisation of results, will therefore often result in the data not being shared until a later date. The Act relating to Universities and University Colleges Section 1-5 deals with academic freedom and responsibility, and specifies the relationship between the institution and persons appointed to a position where research or academic or artistic development work is part of their duties. Persons appointed to a position where research or development work is part of their duties are *entitled to publish* their results and *must make sure* such publication takes place. The institutions cannot instruct them to publish their results at a specific time or frequency,[52] but the Act states that '[n]o permanent restrictions in the right to publish results can be agreed or stipulated beyond what follows from statute or pursuant to statute.'[53] Consideration for applications for patent rights will fall under the scope of this provision, and deferred publication will in such cases be in accordance with the law.

It is not always clear what constitutes legitimate and non-legitimate reasons for applying an embargo to research data. Uncertainty relating to expectations of 'ownership', for example on the part of a researcher or an informant who has contributed personal data to a research project, cannot be resolved by means of licensing. In its recommendations concerning the legal framework for sharing of research data, the Committee has therefore recommended that the authorities should consider enshrining rights to and responsibility for publicly funded research data in law. In addition to legal assessments, it must be considered whether there are ethical issues related to the sharing of such data. Ethical assessments are inherent parts of a research project, from the planning stage to its conclusion. The EU project RESPECT has drawn up a Code of Practice for socio-economic research. The Code of Practice is based on three underlying principles: upholding scientific standards, compliance with the law, and avoidance of social and personal harm. The EU has also drawn up the European Code of Conduct for Research Integrity (2017), which is based on the principles of reliability, honesty, respect and accountability. The national guidelines for research ethics also contain overriding principles for data sharing. These principles are to guide the execution of research projects, including the choice of licence.

7.2.2 Licensee's responsibilities – use and utilisation of research data in accordance with research ethics norms and the terms and conditions set out in licences

In the context of re-use of research data, a licensee is the person granted permission to utilise research data under the terms of the licence. This means that the licensee must understand the underlying terms and conditions, as well as whether they are compatible with the envisaged use, in order to utilise the data. In some cases, the licensed material can be subject to more than one licence and different terms and conditions, and it is important for the licensee to be aware of this. A database, for example, can comprise several smaller datasets licensed under separate licences, each with its own set of terms and conditions.

Some licences have three layers: the human-readable common deed, the machine-readable code and the legal code. The legal code is the primary licence level that describes the legal basis for utilising the licensed material, and legal expertise is required to understand the content. It is also necessary to note whether the licence is intended for international use, meaning that the terms and conditions are not based on the legislation of specific countries or jurisdictions. The final version of the

[52] Lovkommentar til universitets- og høyskoleloven. Rettsdata, by Jan Fridthjof Bernt, note 67. (In Norwegian only)
[53] https://lovdata.no/NLE/lov/2005-04-01-15/§1-5

CC licences (4.0)[54] are examples of licences intended for international use. A ported licence, on the other hand, is a licence where the terms and conditions are adapted to e.g. the copyright law of a specific country or jurisdiction. The use of such licences can have consequences for the intended use and utilisation of the research data if a licence is granted in one jurisdiction subject to certain legal terms and conditions, while the user falls under another jurisdiction where the same conditions may not apply.

In addition to what follows from the terms and conditions of the licence, research ethics norms also define requirements for use of data in research. The Research Ethics Act, discussed in more detail in Subchapter 6.3.2, applies to research and researchers in Norway, and it does not distinguish between research carried out by public and private institutions. The obligation to comply with recognised ethical standards lies with both the researcher and the institution. Honesty and good reference practice are key aspects of research integrity. Although the licence does not explicitly set e.g. attribution as a condition, the licensee must, when using licensed material for research purposes, comply with research ethics norms and discipline-specific standards for good reference practice; see Subchapter 7.3.5.

It is important to be aware that even though research data come with a licence or a dedication to public domain, this does not speak to the quality of the data and does not mean that the data are reliable, valid or re-usable. Good data quality is a multidimensional concept that may include how re-usable data are, how reliable and valid they are, whether open standard formats have been used for metadata and machine-readable data, etc. Data quality can for example be addressed by using a FAIR certified repository, as European Open Science Cloud is planning for, for data curation and documentation of data provenance.

7.3 About terms and conditions and permissions in licences and dedications to public domain

The term 'open knowledge' is defined by the Open Knowledge Foundation: 'Knowledge is open if anyone is free to access, use, modify, and share it – subject, at most, to measures that preserve provenance and openness.'. Further, they use the term 'open work' when describing open research processes. The use of open licences or dedications to public domain is one of the principles for characterising research as open.

7.3.1 Open standard licences

Licences that stipulate terms and conditions that are in accordance with open knowledge as defined in the paragraph above, are called conformant licences. The Open Knowledge Foundation has drawn up a list of such licences. However, there is no consensus about which types of licences can be called conformant licences. The Open Access movement, which primarily focuses on open access to scientific publications, has allowed for the use of licences that impose restrictions on derivatives, known as No Derivative or ND licences, but is restrictive when it comes to recommending this type. This would for example mean articles subject to such licences cannot be translated from English into other languages, which would make the article inaccessible for persons who do not master English.[55] The use of this type of licence for research data can also greatly restrict re-use of the data. The Committee will discuss this in more detail in Chapter 8 Consequences of licence choices.

Both national and international governing documents, policies and guidelines recommend the use of conformant licences as the primary choice for publicly funded research data, as they allow for the widest possible access, re-use and distribution. When a licensor produces research data that is desirable to share, the licensor should explicitly state whether the licensee can remix, change, adapt and add data, and on what conditions. There are licences that are in common use internationally and are considered good practice for licensing of research data. Open data, in the strictest interpretation of the term, means that the data are shared with no limitations on who can use data (restrictions on access) or what the data can be used for (restrictions of purpose). A broader interpretation of the term is that data can be characterised as open even if they are subject to restrictions, but only if the restrictions are necessary based on legislation or other characteristics of the data. Some Creative Commons 4.0 licences (CC licences) are defined as conformant licences, but in the next subchapter, we will problematise why these licenses could nonetheless restrict further use of the data.

7.3.2 Open Creative Commons 4.0 licences (CC licences)

Some conformant licences, such as Creative Commons 4.0 licences, were designed with regards to regulating the use of copyrighted material such as music, images, literary texts and scientific articles. These licences are also frequently used to license non-copyrighted research data. Two of the CC licences, CC-BY-4.0 and CC-BY-SA-4.0, are listed on the Open Knowledge Foundation's list of licenses that are conformant with the principles laid out in its open knowledge definition. The CC licences are the natural choice for research data that contain artistic or literary works that meet the requirements for copyright. Research data that contain pure facts or objective information about the world, however, are not subject to copyright or similar rights.

All CC licences require attribution, and they are therefore popular with researchers. The attribution condition, including

[54] CC licences refers to licences from Creative Commons 4.0, six in total, which all contain attribution conditions. CC0, which is described in the report as a public domain dedication tool rather than a licence, is not included. The distinction between licence and dedication tool without restrictions on use is consistently applied throughout the present report.

[55] https://creativecommons.org/2020/04/21/academic-publications-under-no-derivatives-licenses-is-misguided/

identification of the creator(s) of the licensed material, shall be satisfied in any reasonable manner, as described in the licence text,[56] Section 3 letter a) (2). In Section 1 letter g), however, it is stated that this is limited to licensed material subject to copyright and similar rights. The CC licences are the natural choice for research data that contain artistic or literary works that meet the requirements for copyright. The use of CC licences for data without copyright protection as a way of obtaining references and citations contributes to confusion about what is protected by law and what the licensor wants to achieve by the licence. Such use should therefore be avoided.

Some CC licences also have share-alike or copyleft conditions, known as ShareAlike or copyleft licences. They allow the licensee to remix, change and add data as well as distribute them, but only if you share them under the same license as the original licensed data. Licences with this condition are also referred to as viral licences,[57] as copyleft licences require users of the data to license derivative works, also called derivative data, using a copyleft license with compatible terms and conditions when sharing the data. This contributes to the continuous spread of the licence. This is problematic for reasons of legal interoperability and re-use of licensed datasets, for example if you want to combine two datasets that are both subject to a share-alike condition, but otherwise have different terms and conditions. In such cases, it will not be possible for the licensee to share the derivative data subject to the same conditions as the original datasets, as they do not have the same terms and conditions. We will return to this issue in Chapter 8 of the report, which deals with the consequences of licence choices.

7.3.3 CC0 and Public Domain Mark – data in the public domain

Copyright and database protection arise automatically if the conditions of the Copyright Act are met. The rules on rights to and protection of databases, on the other hand, are less aligned internationally than copyright rules for traditional intellectual property, which in a research context is typically books and articles. The rules on database protection are aligned within the EEA, but they differ, in some cases significantly, from those that apply in other jurisdictions, including the USA. An overview of the diversity of copyright protection in different European jurisdictions has been prepared by CESSDA.

When a rights holder uses CC0 or other declarations to waive rights, this means that the rights holder relinquishes the right to exercise copyright or database rights. The data are thereby made available in the public domain with no conditions attached. The data can then be freely re-used, adapted and commercialised, and as CC0 is worded, with no requirement for the licensee to credit the originator(s) of the data. As regards data collections protected as intellectual property, it follows from Norwegian law that there are some rights that the author/originator cannot waive. This applies to the rights referred to as the author/originator's 'moral rights', i.e. the right to attribution and respect, cf. the Copyright Act Section 5. Although Norwegian law does not permit authors/originators to waive their moral rights, a researcher can issue a CC0 for research data. This signals that they will not demand attribution or take action if the data are used in ways that are not in accordance with the Copyright Act, and thus the right of respect, cf. the statement of purpose Section 2.[58]

When data are not protected by copyright or database rights or by other statutory provisions, a Public Domain Mark can be used. This mark shows that the material is free of known restrictions under copyright law and therefore falls within the public domain.

7.3.4 Legal interoperability and research data

In the report EOSC Interoperability Framework, legal interoperability is about facilitating cooperation between organisations operating under different legal frameworks, policies and strategies. Legal interoperability for research data is thus a matter of the international research community's ability to share, access and re-use research data from different sources that fall within different jurisdictions, policies and strategies. RDA-CODATA's Legal Interoperability Interest Group has listed six overriding principles for legal interoperability of research data in its implementation guidelines: 1) facilitate the lawful access to and re-use of research data, 2) determine the rights to and responsibilities for the data, 3) balance the legal interests, 4) state the rights transparently and clearly, 5) promote the alignment of rights in research data, and 6) provide proper attribution and credit for research data.

Compatible licences are different licences whose terms and conditions are aligned with each other, and are used to legally share combined or compiled datasets. These should be used whenever possible to license research data in order to promote principles of legal interoperability for research data. This entails using licences with as few restrictions as possible on access and purpose to avoid unintended restrictions on derivatives from datasets in future. When using other people's licensed datasets, it must be considered whether the intended use is in accordance with the terms and conditions of the licence, which licence must be used to share derived data, and what impact the licence will have on future derivative data. When combining two or more datasets licensed with different terms and conditions, the most restrictive set must be complied with when sharing derivatives of the dataset in accordance with the principle of the lowest common multiple. In extreme cases, it may even become impossible to combine datasets at all because of conflicting terms and conditions for use of the data under the respective licences. For example, datasets with a share-alike condition cannot be

[56] https://creativecommons.org/licenses/by/4.0/legalcode
[57] https://en.wikipedia.org/wiki/Viral_license
[58] https://creativecommons.org/publicdomain/zero/1.0/legalcode

combined with other share-alike datasets if the other terms and conditions are not compatible.

It is difficult to determine whether licences are compatible with other types of licences without sufficient legal understanding and expertise. Creative Commons has drawn up a list of licences that are compatible with CC BY-SA and CC BY-NC-SA. OER Africa, an initiative for the development and sharing of open educational resources, and OpenMinTeD, an e-infrastructure for text and data mining of scientific content, have each developed a compatibility matrix for Creative Commons licences.

7.3.5 Good reference practice for research data

Good reference practice when using other people's research data is enshrined in the Research Ethics Act in the due care obligation that applies to researchers (Section 4) and the requirement for institutions to ensure that research that takes place at the institution is carried out in accordance with recognised standards of research ethics. This requirement means that the institution is responsible for providing the necessary training in the recognised standards of research ethics and ensuring that everyone who takes part in or conducts research are familiar with these standards. The general guidelines for research ethics specify that 'Researchers must adhere to good reference practices, which fulfil requirements for verifiability and form the basis for further research.' In practice, researchers often use attribution licences to ensure that they are credited. CC licences, which all contain an attribution condition, are problematic and confusing when used on data that do not meet the conditions for copyright or database protection. In such cases, research data users can in principle disregard the attribution condition without breaking the law, since, according to the licence, attribution only applies to data that meet the conditions set out in the Norwegian Copyright Act.

The Dryad Digital Repository advocates using the CC0 dedication to public domain for research data, even if it does not contain an attribution condition. CC0 was designed to minimise both intended and unintended legal and technical impediments to the re-use of research results. CC0 cannot change the data's legal status by removing their statutory protection, as facts and information in and of themselves are not eligible for copyright in most jurisdictions. What CC0 does, however, is allow for the possibility that one may have copyright or related rights in jurisdictions one is not aware of, and CC0 provides a way for researchers to waive all their copyright and related rights in their works to the fullest extent allowed by law. Using CC0 for research data does not release the licensee from the responsibility under research ethics to comply with good reference practice.

Dryad believes the best way of addressing this principle is through normative and discipline-specific guidelines and good source citation practice in research, rather than through licence terms and conditions that must be enforced through legal means. It is good practice when making a dataset available to clearly refer to norms and guidelines, known as 'community norms', for source citation. The Dataverse project[59] provides one example of such community norms. The project has developed a set of normative guidelines for data citation as well as an automatic standardisation process for citation of datasets for all data made available in a Dataverse repository. Disciplines and research communities should strive to develop normative guidelines for good reference practice. This will highlight expectations of good reference practice, both for the parties that make datasets available and for those who use the data.

In the Committee's view, there should be alignment between the terms and conditions of licences and the terms and conditions that should apply to the use of data. However, no licence or dedication to the public domain exists that is adapted to research data not protected by law and which at the same time complies with research ethics norms and scientific standards for citation, attribution and accountability in research. In cases where a future user of the research data is obliged to observe research ethics guidelines by virtue of belonging to an academic community, it is possible that re-use of data with a dedication to public domain will not in practice constitute an obstacle to good reference practice. However, when data are re-used for purposes other than scientific research, for example commercialisation by private parties, the user is not obliged to observe research ethics guidelines and good reference practice. Any expectations of attribution and citation when a researcher wants to share their data, will not be in accordance with the terms and conditions of a dedication to public domain. The Committee therefore sees a need to develop a dedication to public domain that addresses these expectations.

7.3.6 Licensing of research-related digital objects

The Committee has not focused on licensing of research-related digital objects, such as source code, algorithms, models, methodology protocols and electronic laboratory notebooks in its work. Nevertheless, such objects make up much of the research results in many research projects. Good data management practices can in many cases be transferred to e.g. source code, but can differ in some areas from making research data available. Open Science Toolbox makes reference to useful tools developed for publication and version control of source code, and provides an overview of the licences most commonly used when publishing code developed as part of research projects. Making research-related digital objects available under open licences safeguards the integrity of research results by ensuring the verifiability and validity of analyses, models and methods. The Open Source Initiative has defined ten principles that source code must comply with in order to be considered open source and has also drawn up a list of licences that comply with these principles.

[59] https://dataverse.org/

8. CONSEQUENCES OF LICENCE CHOICES

The choice of licence when sharing and making research data available can have intended and unintended consequences for both the licensor and the licensee. For this reason, it is important to have a legal understanding of the licence contract and its terms and conditions for use of the research data. If a licence is not used when sharing research data, it could lead to uncertainty about the legal status of the data and what rights apply to all or parts of the material shared, as well as about any restrictions on purpose that may apply. Such a situation will limit re-use, among other things because data users will be reluctant to take legal risks or not wish to use the data unlawfully. In the following, the Committee will discuss some potential consequences of the choice of licence.

8.1 More closed than necessary

Inadequate legal knowledge and understanding of the legal framework within which data operate could contribute to research data being more restricted by licensing terms and conditions than necessary, often unintentionally. Uncertainty on the part of the licensor and data user about which rights and restrictions on sharing apply can thus reduce sharing and re-use of research data.

Processed or adapted data originating from a combination of two or more datasets with different degrees of licence restrictions will in principle have to be licensed under the licence with the most restrictive terms and conditions, in accordance with the principle of the lowest common multiple. This means that some of the data will be subject to restrictions on access and purpose

LICENSING IN PRACTICE – INTERNATIONAL DATA PRODUCTS

The Norwegian Polar Institute hosts the collaborative project Quantarctica, which is a curated collection of thematic map data about Antarctica. Quantarctica is distributed as a package of data layers and open-source GIS software. Since the package can be installed on an ordinary computer and used offline, it has been extensively used in connection with research and fieldwork in Antarctica. Quantarctica comprises more than 250 data layers. They originate from institutes and research groups worldwide, cover a range of different research disciplines and are selected by an international editorial board before undergoing cartographic adjustment and being entered in Quantarctica.

The wide range of international sources has presented a number of challenges in terms of rights management. Datasets received from the contributors can be subject to a variety of different licences and terms of use, ranging from standard licences such as CC BY, via institution-specific re-use agreements, to ‘all rights reserved’. Some contributors have not set any conditions, but only expressed an informal wish to be cited when the data are re-used. Data from public institutions in the USA may be made publicly available subject to a provision for free use 'within the USA', but with no information about which, if any, conditions apply to use outside the USA. Yet other data may have been shared subject to a condition of non-commercial use.

At the same time, Quantarctica's users naturally wish to deal with terms of use that are uniform, easy to understand and easy to comply with. Licences and clauses that set restrictive conditions for re-use, for example restricting it to non-commercial use or ‘share-alike’, will complicate matters for those wishing to use Quantarctica to produce their own products, such as illustrations for articles or printed thematic maps. A complete absence of terms and conditions can also give rise to problems. Once a dataset has been shared or made available without a licence, formal terms of use or a dedication to public domain, it will in some countries automatically be considered protected, and re-use will therefore be unlawful.

The most recent version of Quantarctica was published under one common CC BY 4.0 licence. During the work, we have been in contact with all contributors with special terms of re-use and entered into agreements to ensure that the contributions can be published under the same licence. In a small number of cases, it has been impossible to achieve such agreements. For this reason, Quantarctica requests that the user consults each layer’s dataset description before re-using the data.

In order to avoid these complications, the editorial board has decided that the next version of Quantarctica will only accept data published under an internationally recognised, open standard licence compatible with CC BY 4.0 or a dedication to public domain. Since most of Quantarctica’s contributors are also users, it is rarely difficult for them to understand this position.

that are not, strictly speaking, necessary. In cases where the datasets are subject to share-alike licences, it could also mean that the datasets cannot be combined at all if other terms and conditions in the licences are not compatible.

In order to minimise such challenges when combining datasets, it is important for licensors to be aware of this issue and strive to use the most compatible licence available for the data to be licensed. The fewer restrictive terms and conditions a licence contains, the more compatible it will be in combination with other licences. In cases where some of the data possess characteristics that make it necessary to license them on restrictive conditions, for example if they are subject to copyright, differentiated licences or a dedication to the public domain for parts of the data could be a solution. This is not optimal, but it is preferable to imposing unnecessary restrictions on data that do not in principle require protection.

The requirements for open data sharing and protection of data containing personal data may appear to conflict with each other. OpenAire , CESSDA and ELIXIR all provide comprehensive descriptions of how to process research data that contain sensitive data. The main message is that there are ways of balancing the requirement for openness with the requirement to protect data that contain sensitive information. Anonymising data is one such measures, but it might mean that the research data lose some of their value. Anonymisation means to process the data in such a way that they cannot be traced back to individuals, and the process is irreversible. Pseudonymisation is a less invasive and final measure that will nevertheless provide sufficient protection to allow the data to be used. Encryption, the use of a key or additional information for the re-identification data, stored securely and separately from the data, is one example of pseudonymisation. The GDPR distinguishes between anonymisation and pseudonymisation of data containing personal data. Anonymisation means that the data are no longer considered personal data, while pseudonymised data still are. Data platforms with restricted access and a dedicated analysis infrastructure to safeguard data protection and good information security is another way of sharing sensitive research data, though not completely open. This allows for research data containing personal data to be re-used in a controlled and secure manner.

If anonymising data containing personal data would diminish their value and render them useless, informed consent could be a solution. This involves data subjects, or informants, consenting to the sharing of data. The purpose of the research project must be presented in such a way that it is clear to the informer how the data are to be processed, what they will be used for, and how they are to be made available. The GDPR and Norwegian law allow for the use of broad consent, in accordance with research ethics standards, for the purpose of scientific research. The consent must nevertheless be limited to certain areas of scientific research. Since it can be difficult to predict what data will be used for if shared openly, informed consent can only be used to share data containing personal data if the data subject has been explicitly informed of the potential use and consented to it. CESSDA's resource page on informed consent describes examples of 'granular consent'. Another possibility, outlined by OpenAire in its guide to dealing with data containing personal data, is to transfer and share data in individual cases based on a special assessment of legality. Research data that contain personal data cannot therefore be completely openly shared. The data must be processed and made available in accordance with the GDPR and relevant Norwegian legislation, including the Personal Data Act, as explained in the section Legal framework for sharing of research data.

8.2 Attribution stacking

Attribution stacking is a technical, legal and ethical problem with no obvious solution. The Committee wishes to point out various issues relating to attribution stacking in the following. It may be useful in this context to distinguish between attribution, which is often used to acknowledge the work done by crediting the people who obtained the data, and citation, often used to refer to the data themselves by means of a link or an identifier, for example a digital object identifier (DOI). Citation helps to document the origin of the data, promotes re-use and can also be used as a key performance indicator to measure the impact of the data.

Attribution stacking can occur when people combine many datasets from different sources or re-use research data based on many sources. Attribution licences set attribution as a condition for the re-use of research data. This requires users to identify the author/originator of the data, cite and refer to the data themselves, make reference to the licence and licence text, and indicate any changes as well as keep indications of previous changes. These conditions are often stricter than the recognised ethical standards for good reference practice. A legal argument to avoid attribution stacking would be to recommend the use of dedications to public domain with no attribution condition. Any expectations of attribution and citation when a researcher wishes to share their data will not be in accordance with the terms and conditions of a dedication to public domain. However, as discussed in Subchapter 7.3.5, good reference practice when using other people's research data is a statutory requirement as well as a principle of research ethics and an obligation in research, and it should therefore in principle be superfluous to define this explicitly through an agreement on use.

Unclear and differing attribution and citation requirements can hinder a streamlined data sharing process, and are particularly demanding in cases involving a wish to combine several datasets automatically. This can give rise to problems relating to machine-readability and compliance with terms of use of the data.

Attribution stacking is not just a problem and a challenge for individual researchers. Research communities within different scientific field should discuss the use of common standards and conventions that are aligned with international standards in

their field. Standards should balance the challenges of attribution stacking against the research ethics principles of integrity, accountability and good reference practice. This is particularly important in disciplines that routinely make use of many different datasets in research and depend on an automated process to combine and utilise different data sources.

8.3 Open data sharing = negation of responsibility?

One concern many feel about using open licences and dedications to public domain is the feeling of losing control over user access and what the data are used for. A transition from the more traditional sharing based on restricted access, for example by contacting the person(s) who produced the research data personally, to sharing research data as openly as possible can present a challenge to the ethical responsibility researchers have for 'their' data and the expectation of 'ownership'. It is therefore important for the research support system to possess sufficient legal expertise to assist researchers in clarifying legal issues so that data can be shared in a lawful and responsible manner in an international context. The support system will also need discipline-specific knowledge to be able to address requirements related to research ethics. The framework for ethically sound and open sharing of research data should be developed within the respective research communities with a view to open sharing in an international context.

Another concern associated with open sharing is degradation or dilution of the data. As the Committee has pointed out above, open sharing of data should comply with good reference and citation practices that enable users to identify the original source data.

8.4 Barriers to commercialisation

Research data should be a driver of value creation and innovation of benefit to society. The use of licences containing

LICENSING IN PRACTICE – LANGUAGE CORPORA

Language corpora are scientific annotations of text or speech (including sign language), and each piece of input can be independent intellectual property, subject to copyright. In addition, an utterance could concern persons and thus be restricted under data protection rules. Studies of translation and the development of machine translation requires corpora where each sentence is aligned with its translation. Professional translations of novels provide a good qualitative basis for such corpora, but negotiations with authors and publishers are time-consuming and their outcome uncertain.

When a Nordic research network compiled a corpus based on Jostein Gaarder's novel Sophie's World, licensing negotiations took place with the rights holders for each language. The author himself and the publisher of the Norwegian, Danish, Swedish, Icelandic, Estonian, English, German and Georgian versions all agreed to grant a licence for research purposes (attribution, datalinguistic research, no redistribution). The exception was the translator of the Finnish version, who refused. This illustrates how important it is to enter into a licensing agreement with the rights holders before conducting scientific processing of data.

Another project involved analysing thousands of Norwegian books to support lexicographic work on the Norsk Akademi dictionary. It was not possible to negotiate for every single book. The University of Bergen therefore applied to the Ministry of Culture for permission to reproduce the works for research purposes pursuant to the Regulations to the Copyright Act Section 1-4, Reproduction for research purposes. After such permission was granted, the National Library of Norway was able to provide a large number of digital books for analysis, and the analyses were made available under an academic licence (CLARIN ACA, ID, AFFIL=EDU, BY, NORED).

Another potential challenge arises when spontaneous speech and personal texts are collected from informants as a basis for language databases. Spontaneous speech will always be considered personal data, and personal texts will be considered personal data if they contain metadata that provides information about e.g. someone's age, sex and residence, or if their content may identify the informant or other people. In such cases, the metadata and, if relevant, the data, must be anonymised, and the researcher must obtain the informant's consent to use the data collected. Often, the researcher/collector is more restrictive than is necessary and expedient. There are many examples of conditions set out in consent forms that restrict the use of data to research activities, often even to the specific research project in question. This makes it impossible to verify and reproduce the research, and it also means that valuable data, which could have been included in language corpora in anonymised form and provide input to the development of language technology products and services, are not made available.

conditions for non-commercial use imposes restrictions on re-use of data for commercial purposes. In some cases, such as when research data form the basis for a patent, it may be necessary to place the data under a licence that prevents using the data for commercial purposes. In other cases, it could be an expectation of 'ownership' and a desire to prevent others from making money from 'my' data that leads a researcher to stipulate such conditions without legitimate grounds.

In addition to the obvious negative effect on value creation and innovation in general, conditions restricting commercial purposes could have unintentional consequences. One example of this is reduced utilisation of valuable language data if the consent obtained from informants limits the use of the data to research purposes. The wording of the consent was perhaps chosen by the researchers to prevent large commercial players from profiting from information provided by private individuals, or for the purpose of strengthening confidence in the project among informants who may be sceptical of commercialisation. However, closing the possibility of commercial utilisation of these data will also make them inaccessible to small technology businesses developing language technology tools – tools that could greatly benefit other private individuals, and perhaps even the informants themselves.

USEFUL RESOURCE PAGES

Open science
http://openscience.prototyp.io/

Interoperability
https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en/format-PDF/source-190308283
https://www.codata.org/uploads/Legal%20Interoperability%20Principles%20and%20Implementation%20Guidelines_Final2.pdf
https://www.ga4gh.org/genomic-data-toolkit/regulatory-ethics-toolkit/

Research Data Management tools
https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide
https://rdmkit.elixir-europe.org/index.html

Data management plan – services
https://www.nsd.no/en/create-a-data-management-plan
https://documentation.sigma2.no/services/easydmp-user-documentation.html
https://argos.openaire.eu/splash/

Attribution and citation
https://uc3.cdlib.org/2016/09/19/cc-by-and-data-not-always-a-good-fit/
https://www.library.ucdavis.edu/news/cc0-and-data-citation/
https://blog.datadryad.org/2011/10/05/why-does-dryad-use-cc0/
https://dataverse.org/best-practices/data-citation

Licence selector
https://ufal.github.io/public-license-selector/
https://chooser-beta.creativecommons.org/
https://choosealicense.com/
https://opensource.org/licenses/
https://docs.data.world/en/59261-59714-2--Common-license-types-for-datasets.html
https://help.data.world/hc/en-us/articles/115006114287-Common-license-types-for-datasets

Overview of compatible licences
https://openminted.github.io/releases/license-matrix/
https://www.oerafrica.org/FTPFolder/Website%20Materials/UnderstandingOER/licwiz/english.html

It is not always clear what constitutes commercial and non-commercial purposes, and there are grey areas. As such, researchers may then choose licences that restrict the use of the data for such purposes, just to be on the safe side. This could present an unnecessary obstacle to value creation and innovation, and restrictions on sharing of research data that exclude commercial purposes should not be imposed unless there are legal or ethical grounds for doing so.

The Committee will discuss value creation and innovation based on research data in more detail in Chapter 15.

9. PREMISES FOR MAKING RESEARCH DATA AVAILABLE

In the Committee's view, the use of licences and dedications to public domain for research data must be based on certain premises. In this chapter, we will discuss what we consider to be the most important premises.

9.1 Clear division of responsibility between individual researchers and the institution

The division of responsibility between institutions and researchers as regards transparency relating to research results is clearly set out in the Act relating to Universities and University Colleges (Section 1-5(6)). Individual researchers are entitled to publish their results and are obliged to make sure such publication takes place. Universities and university colleges have an obligation and a responsibility to ensure transparency regarding the results of research or academic or artistic development work. In the Committee's opinion, this division of responsibility should apply to all institutions and researchers, regardless of sector, whose activities include research. The division of responsibility should be clearly set out in institutional guidelines for data management.

9.2 Legal expertise in the research support system

The legal landscape surrounding openness of research results and licensing of research data is a complex one. The individual researchers cannot be expected to have sufficient legal expertise to facilitate lawful access and re-use of research data and choose the right type of licence or dedication to public domain. The institutions should have this expertise available to researchers who are going to share data by facilitating research support systems that simplify data management. This entails, among other things, specialist assistance in legal matters, such as clarification of rights to and statutory protection of research data. The institutions should also take steps to ensure that research-oriented study programmes provide sufficient knowledge of rights and responsibilities associated with research data. Research support systems that make data management easier for researchers will require resources, and the need will only become greater in the time ahead. As one of its recommendations for a legal framework for sharing of research data, the Committee has recommended that authorities provide financial resources that enable the institutions to provide good support services to researchers and develop such expertise.

9.3 Clarification of rights to and restrictions on sharing of research data

Before research data can be shared, any rights to the data and statutory or other restrictions on their sharing must be clarified. Many Norwegian institutions have data management guidelines or an IPR policy in place, but they differ in terms of both clarity and level of detail. Institutions should have clear guidelines in place for determining who may hold rights to data produced at the institution and to what extent such rights can be controlled by others, for example the institution's employees. It must also be clear who should be recognised and credited in connection with re-use of data in cases where the rights holder is an institution. If the research data to be shared comprise a combination of several datasets from different sources, the status of the data as regards rights should be accurate and transparent. The status should state 1) who, if any, holds rights to parts of the dataset, and 2) who the rights holder is for the dataset as a whole. Clear and transparent clarification of rights to research data is in line with one of the overriding principles for legal interoperability of research data as drawn up by RDA-CODATA. One of the Committee's recommendations for a legal framework for sharing of research data is to consider on a continuous basis whether there is a need to enshrine the rights to and responsibility for research data in law in order to clarify expectations regarding 'ownership'.

10. THE COMMITTEE'S SUMMARY OF USE OF LICENCES FOR RESEARCH DATA

When research data are to be made openly available to others, terms and conditions should be applied that reflect the permitted use while also addressing research ethics standards for good reference practice. Ideally speaking, it should be possible to do this by issuing a licence or a dedication to public domain, but at present, none exist that are particularly suitable for research data. Many researchers use CC licences, which are based on copyright, to ensure that they are credited and cited when the data are used. If research data are not subject to copyright, but a CC licence nevertheless is used, this will be problematic and confusing for people wishing to use the research data. Attribution conditions are also stricter than good reference practice dictates. The Committee therefore recommends that an international standard licence or dedication to public domain for open sharing of research data should be devised that reflects the permitted use and addresses the research ethics principles of integrity, accountability and good reference practice. Norway is well positioned to take the initiative in such a process.

Before a licence or dedication to public domain is granted, several conditions and assessments must be in place, and the process requires legal expertise as well as knowledge of discipline-specific research ethics guidelines. Any rights and restrictions on sharing must be clarified, and it must be made clear between researchers and institutions whose responsibility it is to make research data available and set terms and conditions for their legal use. The institution should ensure that researchers have clear internal guidelines and governing

documents for making research data available, as well as a support system with legal and research ethics expertise to provide assistance in making assessments regarding legal and ethically sound sharing. The researcher should be responsible for the practical aspects of data sharing and arrange for data to be shared in line with established practice in the research field. The consequences of not carrying out necessary assessments and measures before sharing research data could be that more restrictions than necessary are imposed, which could hinder value creation and innovation of benefit to society.

11. THE COMMITTEE'S RECOMMENDATIONS

In this chapter, the Committee presents its recommendations regarding the legal framework and its 'licence etiquette rules' that describe recommended use of licences and dedications to public domain when publishing research data.

11.1 The Committee's recommendations concerning the legal framework for sharing of research data

Below, the Committee presents its recommendations based on the review of the legal framework for sharing of research data in Chapter 6. The recommendations are addressed to the Government and its ministries, and the research institutions, respectively.

THE GOVERNMENT AND MINISTRIES

The Committee recommends:

1. continuous assessment of whether the rights to and responsibility for publicly funded research data should be enshrined in law, for example in the Act relating to Universities and University Colleges, in order to clarify expectations regarding 'ownership';

2. considering removing the requirement for permission from the Ministry in the Regulations to the Copyright Act Section 4 first paragraph in cases that fall within the scope of the Digital Single Market Directive's Article 3 on text and data mining for the purposes of scientific research;

3. clarifying the scope of the publication requirement in the Open Data Directive Article 10(2) '(...) have already made them publicly available through an institutional or subject-based repository' when research data are stored only on the institution's own servers/shared drives or directories or similar;

4. that the scope of right of access pursuant to the Freedom of Information Act should be clarified for research data that are not published or publicly known;

5. that an international standard licence or dedication to public domain for open sharing of data should be devised that supports research ethics principles and academic standards for citation and integrity in research. This should be an international process involving relevant actors, and Norway is well positioned to take the initiative and play a leading role in such a process.

THE RESEARCH INSTITUTIONS

The Committee recommends:

1. that the institutions clarify the division of rights to research data between the researcher and the institution, particularly as regards the sharing of research data;

2. that the institutions draw up clear guidelines for whether, how, where and when research data are to be archived and shared in line with national, regional (European) and international guidelines;

3. that the institutions ensure that the research support function possesses the legal expertise required to facilitate lawful and expedient sharing, and that expertise be developed for relevant disciplines;

4. that the institutions draw up clear procedures for obtaining sufficient consent from data subjects, for example informants in research projects, where required.

11.2 The Committee's recommendations regarding use of licences – the 'Licence etiquette rules'

Based on the review of the legal framework and discussions on the use of licences presented in this part of the report, the Committee has prepared a set of overriding recommendations on licence use. We have called these recommendations the 'Licence etiquette rules', and they are intended to provide guidance on how to use licences in a way that will be in line with the ambition of more sharing and re-use of research data.

LICENCE ETIQUETTE RULES

USE OF LICENCES AND DEDICATIONS TO PUBLIC DOMAIN FOR PUBLICATION OF RESEARCH DATA

1. A licence or dedication to public domain should contain terms and conditions that reflect the permitted further use of the research data.

2. Research data should have international standard licences that:
   a. have as few restrictions on access and purpose as possible
   b. promote the principles of legal interoperability for research data to the greatest possible extent
   c. are both human-readable and machine-readable

3. Metadata should always have a dedication to public domain, typically a CC0 or PDM.

4. Standard licences intended for intellectual property should only be used for data that meet the copyright or database protection requirement.

5. In cases where different legal terms and conditions apply to different parts of the dataset, differentiated licences should be used in accordance with Recommendation 2 in order to avoid imposing a restrictive licence on all of the data because some data require it.

6. The institutions should have an overriding responsibility for ensuring that a licence or dedication to public domain as mentioned in Recommendation 1 is issued for research data produced by its employees in the course of their work.

Research ethics norms and discipline-specific standards must be complied with at all times when using other people's data.

Part 3

Non-legal barriers to sharing and re-use of research data

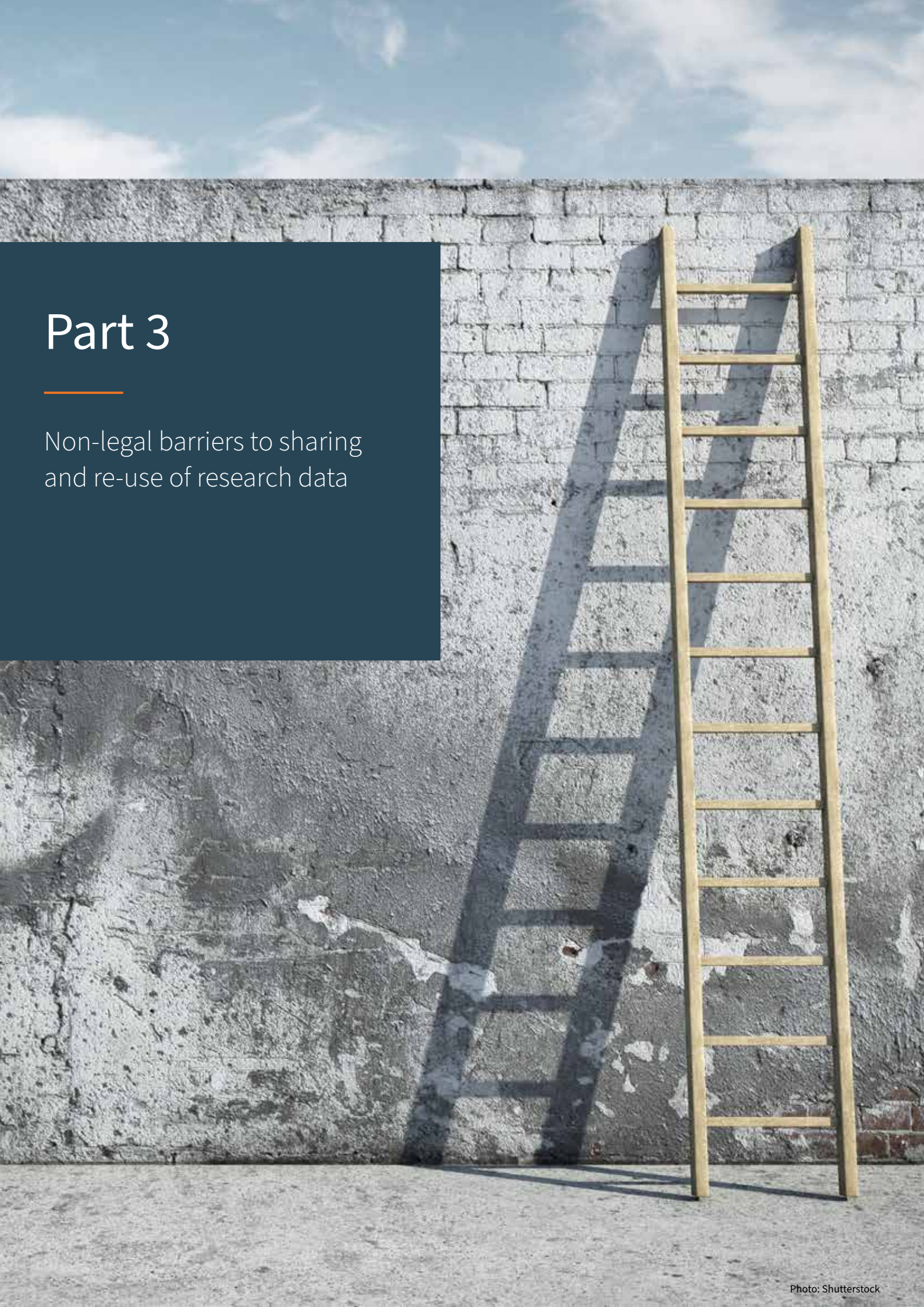


Photo: Shutterstock

In the preceding parts of this report, the Committee has reviewed the legal framework for sharing and re-use of research data, provided an overview of different aspects of the licensing of research data and provided its overriding recommendations regarding the use of licences. We believe that this will make it easier for researchers and institutions to make well-considered and expedient choices when licensing their research data. At the same time, the Committee is of the clear view that many other factors also play a role in this process, and that these factors will have at least as much of an impact on whether we succeed in achieving more sharing and re-use in future.

In this part of the report, we will therefore revisit the political, structural and cultural factors described in the memo that the Committee submitted to the Ministry of Education and Research in October 2020. These factors are not directly linked to legal issues, but are significant when researchers prepare to share their data for re-use. The Committee believes that finding good solutions to reduce the non-legal barriers to sharing, in combination with improving Norwegian researchers' overview of the legal framework, will help the Norwegian research system to develop a strong culture for sharing.

We received a great deal of useful and relevant input and comments to the challenges discussed in the memo during the consultation period, and this input has now been incorporated into the present part of the report. We are grateful that so many have taken an interest and believe that by using an open process, we have succeeded in identifying not only the many challenges facing us, but also good and well-considered ideas on how to resolve these challenges.

The Committee concludes the report with a set of recommendations on how these challenges may be approached and resolved. These recommendations are based on the Committee's many good discussions and input received in the course of its work, and they are addressed to the Government and ministries, funders and the policy implementation system, and the research institutions, respectively. It is important to the Committee to point out that these recommendations must be considered a steppingstone in an area characterised by continuous development, and that they are in no way intended as definitive answers. This is not an ending, but rather the beginning of a process of development towards a more open research sector that is prepared for the future.

12. COORDINATED EFFORTS TO INCREASE SHARING AND RE-USE OF RESEARCH DATA

During the past two or three years, there have been several Norwegian initiatives that share the objective of increasing sharing and re-use of data in order to achieve more knowledge, better services and value creation across disciplines and sectors. It is a generally accepted view that this will contribute to more efficient solutions to the challenges facing society and place Norway in a good position to take part in the international data economy. It is positive that investments are being made in this field, and the Committee considers this a sensible prioritisation in the present situation. However, it is challenging that the many initiatives are not sufficiently well-coordinated from the top. This leads to partly overlapping, partly contradictory conclusions, and could make it difficult for those in charge of data sharing to navigate the complex landscape of good advice and recommendations.

12.1 The need for data policy alignment

The Committee's view is that there is a great need for data policy alignment. It is important for those who are actually involved in sharing data for re-use to have clear guidelines in place. If there are contradictory requirements and expectations from different quarters, then there is a risk that the party with data to share will feel powerless and may choose not to comply with any of the requirements. Another possible outcome is that people may choose which requirements to comply with based on self-interest and not on the intentions behind the recommendations and requirements.

Data policy alignment is required across sectors to enable sharing between different administrative areas and knowledge systems. Different administrative areas could potentially emphasise different aspects of data sharing because they are unaware of the needs of other areas. The Committee's view is that the Norwegian authorities should do more to coordinate policies across sectors by facilitating good arenas for cooperation between developers of policies and guidelines from different areas of the public administration.

Data policy alignment is also required across different levels within each administrative area. For the knowledge sector, this means that policies and guidelines from the ministries, the policy implementation system, service providers, infrastructures and institutions must be adapted to each other and work together towards a common goal. Large institutions will also need internal alignment of guidelines across departments, administrative units and disciplines. For example, researchers may receive conflicting advice and requirements from their department's management and the institution's central support services.

12.2 International orientation

Research is international by nature. Many Norwegian researchers have closer relationships with international partners and funders than Norwegian ones. Many international communities of researchers have shared data across borders for years, and Norwegian researchers that apply for funding from the EU will face stricter open science requirements in their projects. Open science and sharing of research results are on the agenda of many global organisations and associations (including OECD, UNESCO,

ISC, WMO), and it will be essential going forward that Norwegian actors are internationally oriented in their approach to data policy development. The EU is a strong driving force for open science at the European level. This makes alignment between Norwegian and international bodies as important as alignment between sectors at the national level.

At a more operational level, there are many international organisations, initiatives and projects that develop standards, solutions and infrastructures for increased sharing and re-use of research data. Important examples include ESFRI, EOSC, RDA, OpenAire, FAIRsFAIR and CODATA. There are also a number of discipline-specific cooperation forums that it is important for Norwegian groups to participate in. These forums are perhaps where the most important discussions take place, and the presence of Norwegian research communities is therefore crucial if we are to play a role in developments.

The Committee urges the ministries to prioritise and facilitate Norwegian participation in international forums where open science and sharing and re-use of research data are discussed. This will allow us to learn from others and improve, while also contributing to steering international developments in a direction that is compatible with Norway's situation and needs.

13. INCENTIVE MECHANISMS FOR INCREASED SHARING AND RE-USE OF DATA

Publication of research results is an important aspect of academic activity and one of the yardsticks by which researchers are measured, both in their employment relationship and when competing with other researchers for funding. Scientific publication is often a crucial factor in building a career in academic research, for example to defend a doctoral thesis or achieve tenure. Moreover, the Norwegian Publication Indicator, which forms part of the basis for funding of research institutions, primarily confers merit for published articles, and not for the data on which the articles were based. This serves to further reinforce the focus on scientific publication. In other words, the combination of circumstances that arise when the datasets form a valuable basis for other profitable activities, but do not give any credit in themselves, can have a counterproductive effect on data sharing by causing the datasets to be withheld for longer than necessary. Requirements and expectations of sharing of research data could thus contribute to conflicts of values between researchers, institutions and society's investments in publicly funded research data as a common good.

In light of the above, it appears that the importance attached to the publication of articles is not adapted to the new research standard. Both funders and other potential users of research results now hold clear expectations of open sharing of data and other types of material and resources. The Committee therefore sees a need to review the current incentive system and consider whether it should be modified or adjusted to make open sharing the natural choice for researchers.

The Committee also believes that the principles enshrined in the Declaration on Research Assessment (DORA), which has been signed by the Research Council and many Norwegian research institutions, should be better reflected in assessments relating to awarding of external funds for research as well as in connection with appointments and promotions. DORA also explicitly states that datasets and software should also be taken into consideration in such cases. In addition to impact, the dataset's scientific productivity (degree of re-use and contribution to new research results) can also be used as a criterion for assessing its quality. It is important that the criteria used are clear enough to provide incentives for making data available, but also flexible enough to facilitate interdisciplinary research and other research where data are not a natural output. The NOR-CAM report recently published by Universities Norway (UHR) presents a new framework for assessment in academic careers based on a more flexible and comprehensive assessment. This framework entails more openness, broader assessments, and a more comprehensive approach than traditional assessments, which tend towards a one-sided focus on indicators. The Committee believes that this framework could also be useful in different types of evaluations of research groups, research disciplines and research projects.

External funding to realise research projects is a strong incentive for research groups. Criteria that research funders set for their assessment of projects can therefore become a powerful instrument that can bring about changes in practice. The Horizon Europe programme has now introduced open science practices as part of its assessment basis when awarding funding. The programme guide to the framework programme refers to both mandatory and recommended open science practices and how the applicants' descriptions of their own practice will have a bearing on the projects' marks. It will be natural for Norwegian funders to follow the EU's lead in this respect, but the Committee wishes to emphasise that, in addition to such incentives, a change of focus is also required. There is a need to shift attention away from the disadvantages and obstacles to sharing of data and emphasise researchers' ethical responsibility to share their findings with the world, as well as the benefits they themselves will reap by gaining access to many more data sources in a more open research society. Another positive effect of better data publication practices is that they will encourage the publication of complete, quality-assured, and well-documented data products.

The feedback the Committee received on its proposal to change the incentive system revealed concern among the institutions and in the research community. They are concerned that new ways of assessing and rewarding research and researchers will give rise to more reporting and more red tape. The Committee understands this concern and wishes to emphasise that a future system must ensure that this does not happen, and that efficient, automatic ways of sharing and finding information must be developed. This should be realistic in a more open research sector.

The Committee would also like to underline the importance of designing future incentive mechanisms in such a way as to ensure equal opportunities for all disciplines. Many disciplines will rarely generate or collect data that are suited for sharing and re-use, and such disciplines must not be disadvantaged by systems that favour data-intensive fields.

14. INVESTMENT IN FAIR RESEARCH DATA

More sharing of publicly funded research data is an ambition but sharing in itself will not necessarily bring about more re-use. Data that can only be accessed will be useful for purposes of documentation and for research verifiability, integrity and efficiency, but the data will be of limited value to society if they cannot also be used to explore new ideas, create new insights or develop new products. To be possible to actually re-use data, they must be made FAIR, which will require extra resources. This must be considered a long-term investment, and high-quality research data in particular can be a good investment for the future. Investments in FAIR research data at several levels will be required if Norway is to succeed.

14.1 Investment in infrastructures for FAIR data

In recent years, investments in generic e-infrastructure, such as data networks with high capacity and resources for data-intensive computing and storage, have grown significantly. Such e-infrastructure is crucial to advanced and resource-intensive computing, simulations and analyses. Large investments have also been made in area-specific data infrastructures. Their purpose is to collect data, add value and make them FAIR by processing, storing, curating and structuring them for meta-analysis and re-use. Specialised personnel at the research infrastructures also contribute knowledge about licensing and systems for authorised access to data and licence approval. However, such infrastructures are not in place in all areas, and established infrastructures are not always utilised as well as they should be. Continued and increased investment is therefore necessary to establish, develop and promote research infrastructures that ensure researchers access to the right tools and the user support they need for sharing and re-use. National and international infrastructures that are well-coordinated and aligned with discipline communities and institutions will give better access to and enable appropriate licensing for secure sharing and re-use of research data.

Knowledge-based data infrastructures must be continuously developed to meet the needs of society and the research communities, and this will require regular investments and upgrades. There are also great expenses associated with operating such infrastructure. Long-term plans for data infrastructure funding based on different sources of income can help to make the infrastructures more robust. There are currently divergent opinions as to how operations should be funded, and some disciplines have a long-standing tradition of access to data free of cost. Some infrastructures have developed their own funding models, such as user fees, while others have major difficulties in their operations because the willingness to pay is low among their users.

The Committee believes that the need for further investments in infrastructure for sharing, adaptation and management of FAIR data (generic as well as area-specific data infrastructures) will be important going forward if we are to achieve the goal of increased sharing and re-use of research data. Research institutions, relevant public administration bodies and business sector players should join forces and define the infrastructure needs that must be covered within the different scientific and administrative areas. These needs should be described at the institutional, national and international levels. A committee was recently appointed to provide a background document to the Ministry of Education and Research concerning investment in infrastructures for FAIR research data and public administration data of particular relevance to research. This document will describe which needs must be covered in the time ahead and should, among other things, result in coordinated applications to calls for proposals from e.g. the Research Council of Norway and Horizon Europe. Such applications should clearly demonstrate how national and international infrastructures can contribute to more sharing and re-use within and across disciplines.

The issue of sustainable funding models for operating data infrastructures must also be put even more firmly on the agenda. In order for infrastructures to make a positive contribution to more sharing and re-use of research data, good funding mechanisms must be established. Such mechanisms should reduce the challenges that infrastructures face in covering development and operating costs. This will be particularly important for costly generic research infrastructures intended to serve the research sector as a whole.

14.2 Investment in FAIR data expertise

Making data FAIR requires resources not only in terms of investment in expedient infrastructures and systems for sharing and re-use, but also in the form of human resources. Research data are often complex and preparing them for use by others will require both time and skills. One cannot expect the researchers that produce data possess this expertise, which is why expert support services for the researchers are necessary to provide guidance and advice. There will also be a growing need for professionals with expertise at the intersection between law, digitalisation and technology, open science and scientific methodology. At the same time, we see a growing need for expertise in artificial intelligence, machine learning and deep learning using high-performance computing processes when generic e-infrastructures are to provide assistance in research projects with such needs.

Open science and data publication should become more prominent in researcher training programmes so they become an integral part of Norwegian researcher's expertise. At the same time, the higher education sector must train more specialists in this field, as such expertise will be in high demand in the future and will be a critical factor in efforts to achieve open science ambitions. It is important that the authorities that demand a more open research sector also help to enable institutions to provide good support services to researchers. It can be difficult to

achieve good dimensioning between research and support functions when important expertise is subject to rationalisation requirements, for example through the de-bureaucratisation and efficiency reform.

14.3 Priority data for FAIR sharing

Data that are interoperable with other data and can be re-used by others will form an important basis for successful data-driven knowledge development and economy. The high cost of making research data FAIR means that an assessment of which data are to be made fully FAIR will be required, and it may become necessary to set criteria to define which data are deemed most important to make FAIR in an initial phase. Such a priority will also lead to more efficient and targeted efforts to make as much data as possible available as quickly as possible. The report from the expert committee on data sharing in business and industry (2020) (in Norwegian only) points out that it is unrealistic and inexpedient to demand that all datasets be made available at a level suitable for re-use for all possible areas of application. The report therefore proposes establishing mechanisms to bring together actors in areas where access to and utilisation of data will be of great public interest.

On the other hand, there is an argument to be made for the view that we cannot afford not to make data FAIR. According to a report commissioned by PWC by the European Commission in 2018, this works out as a total loss of EUR 10.2 billion a year for the EU countries. This calculation has only included the financial loss and not taken into account other losses, such as the possibility of quality-assured access to large quantities of machine-readable data.

Despite the great potential value of investing in FAIR data, there are of course financial limitations. The Committee therefore proposes to initially prioritise research data of particularly high research value and value to society modelled on Horizon Europe's approach. This will make it easier to allocate the appropriate resources to the right area of investment in order to achieve the objective of more re-use. It will also be sensible to direct efforts to areas where much is already in place and where minor investments could yield great benefits. This does not mean, however, that other publicly funded research data should not be made available. It must be a goal to ensure that as much data as possible are made FAIR, but it may be necessary to take a more long-term perspective in our ambitions for some areas. Despite it not being possible for all research data to be fully compliant with the FAIR principles, it should be a minimum requirement for all research data to be shared with sufficient information to allow the data to be found, verified and quality assured. It will be necessary to make a plan for how the more heterogeneous data, also known as the long tail of research data, can become a valuable contribution to a data-driven society and economy.

15. VALUE CREATION AND INNOVATION BASED ON RESEARCH DATA

Data sharing constitutes an important underlying premise for data-driven value creation and innovation. This is one of the reasons for the clear guidelines that the Government has issued to directorates, the municipal sector and other public sector bodies over the past 18 months. This is in line with the recommendations of the Digital21 strategy, is a key aspect of the work of the expert committee on data sharing in business and industry, and is repeated in Report No 22 to the Storting (2020–2021) Data som ressurs – Datadrevet økonomi og innovasjon: Sharing of and access to data will be important going forward.

Transparency does not mean full openness about everything in all contexts, however. For example, it is obvious that there must be transparency about the evaluation process an application for research funds is subject to, but nobody expects the actual content of the application to be shared. When an enterprise has been granted funding for a research and development project, there must be transparency regarding this funding – how much funding the enterprise will receive and what the funds are to be spent on. However, it is not necessarily the case that all the results from the completed project must also be openly shared. Transparency regarding (fully and partly) publicly funded collaborative projects between industry and research is an important principle. However, transparency to realise the social mission of the publicly funded research must be weighed against consideration for the competitiveness of the private parties involved.

The public sector funds research and development for several purposes using different policy instruments, and this is also reflected in the funding percentage and evaluation criteria. Research funds go to universities and university colleges to generate knowledge that benefits society and facilitate research-based education, social innovation and innovation in public administration. Such projects are evaluated on criteria including scientific quality, project management and the research group, feasibility, dissemination and communication, and international collaboration. These are projects with universities and research institutes in the driving seat, and they typically receive 100 or 80 per cent funding. The requirement for transparency and sharing of results must be explicit.

In addition, public funds are used for a different purpose: to stimulate value creation and innovation and to strengthen competitiveness in Norwegian business and industry. This is vital to a sustainable welfare society. To fulfil this part of the social mission, the public sector contributes financial risk relief through the Research Council and other funders. Typically, enterprises are awarded 25–50 per cent of the project costs to encourage them to use research and development in their innovation work. The enterprises are in the driving seat in such projects. They take most of the risk, and it is up to them to ensure that results are developed into commercial activities and new value creation. The Committee points out that clear guidelines for sharing of data are also needed for this type of projects so that the parties involved

know what is expected of them. There is a potential for conflict here, as it might be in the enterprise's interest to keep things secret and in society's interest that data from research that is partly publicly funded is shared.

It is a political ambition that Norwegian (and foreign) enterprises should choose to interact closely with the best (Norwegian) research communities. This contributes to keeping the communities relevant and ensuring that their expertise benefits Norwegian enterprises. There is the question of whether a requirement to share results from projects where Norwegian enterprises use the expertise of Norwegian researchers could result in the enterprises not wishing to use these research communities. If that is the case, then that would impair innovation work both in the enterprises and in the Norwegian research communities. A requirement for full or partial sharing of research results from projects where the enterprises themselves provide the majority of the funding should not be a policy unique to Norway. If Norway's policy differs significantly from that of other countries, it is conceivable that Norwegian enterprises will instead choose to seek collaboration with foreign research communities and through funding mechanism that do not stipulate such requirements.

It is a statutory requirement that some of the activities of universities and university colleges must contribute to innovation and value creation based on results from research and academic and artistic development work (Act relating to Universities and University Colleges Section 1-3 letter e). At the same time, there are strong guidelines that instruct the higher education sector to ensure transparency regarding the results of research (Act relating to Universities and University Colleges Section 1-5(6)). These two objectives will often go hand in hand because transparency in research will help to disseminate knowledge and form a basis for further research activity, while it will contribute to innovation and value creation in that business, industry and public sector enterprises can develop products and services based on the results. When there is a commercialisation aspect associated with innovation, however, there is a potential for challenges at the intersection between the expectation of transparency regarding results and expectations of contributing to innovation and value creation in the higher education sector. In cases where institutions cooperate with business and industry, the issue of to what extent data should be withheld to protect trade secrets could arise. It could be claimed based on other public interests that the publication of data is a reasonable condition for public funding as it gives something concrete back to society. This issue partly overlaps with the issue discussed above concerning collaborative projects that are partly publicly funded through the Research Council. These are complicated issues that should be clarified in more detail. The Committee had neither the time nor the expertise required to study these topics in depth, and one solution would be to assign this task to others.

16. THE COMMITTEE'S RECOMMENDATIONS

In the following, the Committee sets out its recommendations relating to the non-legal barriers to data sharing discussed in the present chapter. The recommendations are addressed to different research sector actors, and they are divided accordingly into sections addressed to the Government and ministries, funders and the policy implementation system, and the research institutions, respectively.

All the recommendations below must be viewed in light of the fact that research data are a heterogeneous category, as described in Part 1 of this report. The Committee's overall view is that the degree to which research data are suitable for sharing as pure data, differs between different types of research data, and several of the recommendations depend on a differentiation between shareable and non-shareable data. In some cases, data cannot be shared for legal reasons, as described in Part 2, while in other cases, the format or other characteristics of the data make it impossible or inexpedient to share them. Examples of the latter include qualitative assessments and analyses that are only presented in scientific articles and monographs, or new interpretations of e.g. law texts, literary texts or other artistic material.

THE GOVERNMENT AND MINISTRIES

1. It should be an overriding research policy goal to introduce policy instruments that stimulate sharing and re-use of data from publicly funded research. The Government should ensure that strategies, regulations and guidelines that deal with and/or regulate the sharing and re-use of data is aligned across ministries, across sectors, and across all levels of the knowledge sector to ensure that they all work together towards a common goal.

2. The current incentive system for the higher education sector should be adjusted or replaced by new mechanisms that incorporate the broad range of results from research, and that also encourage the production and sharing of scientific research data.

3. If the Government/ministries define focus areas in research, this must also entail investments in the establishment, maintenance and further development of good research data infrastructures that support research and allow research data to be made available in accordance with the FAIR principles.

4. The Government should facilitate Norwegian participation in international work that seeks to promote open science and sharing of research data in accordance with the FAIR principles. Norway should take an active role in the EOSC to ensure that Norwegian interests are safeguarded in the development of the European infrastructure for sharing and re-use of research data.

5. Norway should continue to increase its investments in data infrastructures to ensure that Norwegian research communities have access to the services and tools they need to make data FAIR. Work on funding models for the establishment and operation of such infrastructures must continue.

6. The skills needed in open science must be highlighted and incorporated into the follow-up of the Norwegian Strategy for Skills Policy and the white paper on *lifelong learning for reorganisation and competitiveness (Lærekraftig utvikling — Livslang læring for omstilling og konkurranseevne – in Norwegian only).*

7. The institutions must be given the resources they need to be able to raise the level of expertise on open science and provide the necessary support services to researchers.

8. The appointment of a group/committee should be considered comprising representatives from business and industry, the institute sector, the higher education sector, the policy implementation system and the general public, tasked with proposing solutions for sharing of research data produced in collaborative projects partly funded by industry.

FUNDERS AND THE POLICY IMPLEMENTATION SYSTEM

1. Open science practices, including making datasets available in accordance with the FAIR principles and good licensing agreements, should confer merit and be included in the overall assessment of projects in connection with funding allocation (in line with the Declaration on Research Assessment (DORA)). The weight attached to data sharing as an assessment criterion must be adapted to the policy instrument used, and the assessment should differentiate between data-intensive fields of research and fields that produce less data.

2. The Research Council should follow up the data management in projects more closely than is currently the case and offer guidance and consider taking action if the projects do not share and manage data in line with the projects' intentions. Incentives and follow-up of major allocations, such as Norwegian Centres of Excellence (SFF), Centres for Research-based Innovation (SFI) and research schools, should also be linked to good data management.

3. Funders should be clear about when the results of publicly funded research projects are to be shared, and it should be a rule that they are to be made available as soon as possible in their complete and most re-usable form.

4. The Research Council and other funders should consider new calls for proposals and other policy instruments to encourage more use of existing data in research and innovation. Own allocations of funds can also be directed towards collection, development, maintenance and sharing of important datasets in order to stimulate such activities.

5. In publicly funded collaborative projects that involve employers, business and industry, sharing of data should be regulated through agreements. Research results should be shared as openly as possible. What can be openly shared from such projects and the conditions for sharing may depend on the proportion of public funding and considerations of competitiveness. The conditions should be publicly known before the project is initiated.

6. The public policy implementation system should be better coordinated and merge similar and overlapping services that currently exist.

THE RESEARCH INSTITUTIONS

1. The institutions should develop long-term strategies for sharing and management of data. These should describe how the institution is to proceed to ensure that data are shared in re-usable form as soon as possible after they have been collected, including the institution's infrastructure and expertise. Such strategies must be in line with national requirements and guidelines.

2. The institutions should introduce comprehensive career assessments where publication of re-usable datasets and other research results and best practice in open science also confer merit. NOR-CAM is a potentially useful tool in this context.

3. The institutions must facilitate a gradual increase in data management and FAIR data skills throughout the research system, for example by making data sharing as a compulsory part of doctoral degree programmes, postdoctoral programmes and research management courses. (EOSC's Digital skills for FAIR and open science is a skills framework that can be used for such facilitation.) Among other things, this means that sufficient funds should be allocated to technical support/research support required to achieve good management of research data.

4. Open science in general, and data management in particular, including licensing and citation, should be included in all bachelor's and master's degree programmes, as should research ethics and integrity in research.

5. The institutions should ensure that guidelines for intellectual property rights and commercialisation are aligned with guidelines for open science and making data available.

6. Institutions and research communities that apply for external funding must highlight the costs associated with making data FAIR and enter the information in their project applications.

7. The institutions should consider having their own committees/data sharing officer to provide advice and recommendations on data sharing, for example on legal matters and choice of licence. Institutions that are not able to provide these services themselves should arrange for their researchers to have access to such services provided by other parties.

8. Research communities should be active participants in national and international debates on data sharing with a view to arriving at data sharing systems that are beneficial to their specific field of research.

Appendix 1

Legal framework for sharing of research data (in Norwegian only)

Photo: Shutterstock

Skrevet av Kristina Stenvik,
stipendiat ved Institutt for privatrett,
Det juridiske fakultet, UiO

1. INNLEDNING

Formålet med dette kapittelet er å gi en oversiktlig og kortfattet redegjørelse for relevante nasjonale rettsregler som er aktuelle for deling av forskningsdata, både når det gjelder rettsregler som gir grunnlag for deling, og rettsregler som setter eventuelle begrensninger for deling. Oppdraget er avgrenset i tid og omfang, og det vil ikke tas sikte på å gi noen uttømmende eller detaljert redegjørelse.

Det finnes ikke et enhetlig lovverk for forskningsdata, dette er regulert gjennom fragmentariske deler av en rekke ulike lover. I denne oversikten vil regelverket deles inn tre kategorier: (1) immaterialrettslig regelverk som gir rettigheter til data, samling av data, eller oppfinnelser som bygger på data, (2) regelverk som tilrettelegger for, regulerer eller krever deling av data, og (3) regelverk som setter begrensninger på deling av data. Sistnevnte omfatter for eksempel personopplysningsloven og regler om taushetsplikt.

Inngåtte avtaler og lisenser kan også gi rettigheter til, legge begrensninger på, eller stille krav om deling av forskningsdata. Det rettslige regelverket setter rammene for avtaler og lisenser på forskningsområdet. I den grad lovbestemmelsene er fravikelige, kan man avtale seg bort fra reglene. En del av regelverket kan man ikke avtale seg bort fra, herunder reglene om innsyn i offentlighetsloven, reglene om forskningsetikk mv., og reglene om personopplysninger, nasjonal sikkerhet og taushetsplikt (men regelverkene kan selv tenkes å oppstille unntak, for eksempel ved samtykke). På immaterialrettsområdet er det en viss frihet til å inngå avtale som avviker fra lovens reguleringer, og det er også i utgangspunktet anledning til å overdra rettighetene. Denne friheten oppstår imidlertid etter at rettighetene har oppstått - bestemmelser som oppstiller krav for å oppnå rettigheter etter åndsverkloven og patentloven er absolutte, herunder nyhetskravet i patentretten. Og selv om rettighetene kan overdras, kan man i begrenset grad frasi seg de ideelle rettighetene etter åndsverkloven § 5, herunder retten til å navngis som opphavsmann. Arbeidstakeroppfinnelsesloven som regulerer arbeidsgivers rett til arbeidstakeres oppfinnelser er som hovedregel fravikelig, med noen unntak.

2. IMMATERIALRETT OG FORSKNINGSDATA – RETTIGHETER TIL DATA, SAMLING AV DATA, ELLER OPPFINNELSER SOM BYGGER PÅ DATA

2.1 Innledning

Rettslig kan ingen «eie» informasjon eller opplysninger. Man kan imidlertid ha rettigheter til samlinger av data etter åndsverkloven, i form av opphavsrett eller databasevern (se punkt 1.2.2.2). Man kan også i visse tilfeller ha opphavsrett til data som utgjør eller inneholder åndsverk, og man kan ha opphavsrett til presentasjonen av forskningsresultatet, for eksempel en artikkel (se punkt 1.2.2.1). Dersom forskningsdata leder frem til utvikling av en oppfinnelse, man kan få rettigheter til denne oppfinnelsen etter patentloven, og opplysninger som inngår i patentsøknaden vil bli offentlig tilgjengelige (se punkt 1.2.3.1). Patentrettigheter forutsetter imidlertid at oppfinnelsen ikke er offentliggjort forut for inngivelse av patentsøknad. I tilfeller der oppfinnelser blir til i et arbeidsforhold, kan arbeidsgiver under visse omstendigheter kreve å få rettighetene til oppfinnelsen overdratt til seg, og arbeidstaker har da en rettslig forpliktelse til å ikke publisere informasjon om oppfinnelsen (se punkt 1.2.3.2). Lærere og vitenskapelig personale ved universiteter og høyskoler har imidlertid en særskilt publiseringsrett som går foran arbeidsgivers rett til oppfinnelsen, på visse vilkår (se punkt 1.2.3.3).

2.2 Opphavsrett og databasevern

2.2.1 Forskningsdata som utgjør kunstneriske eller litterære verk (opphavsrett)

Opphavsretten gir vern for kunstneriske og litterære verk som er «uttrykk for original og skapende åndsinnsats», jf. åndsverkloven § 2 annet ledd. Dette omfatter blant annet faglitterære tekster og muntlige foredrag, jf. åvl. § 2 annet ledd bokstav a og b. Slike tekster eller foredrag kan inneholde forskningsdata. Men selv om man kan få rettigheter til teksten eller foredraget, er det imidlertid bare selve den konkrete og individuelle utformingen man får opphavsrettigheter til, og ikke det faktiske innholdet. Man får ikke opphavsrett til ideer, fakta eller teorier. Ulike publikasjoner som inneholder forskningsdata vil altså kunne være underlagt opphavsrett, men omfatter ikke forskningsdata som ligger til grunn for innholdet som sådan. Opphavsretten kan dermed gi opphavspersonen enerett til en artikkel eller annen publikasjon, men forskningsdata som sådan kan fritt benyttes av andre uhindret av opphavsretten. Et illustrerende eksempel på grensedragningen er avgjørelsen fra Oslo tingrett i tvisten mellom DN og Retriever (18-107547TVI-OTIR/06, 7. desember 2020). Selv om DNs artikler er beskyttet av opphavsrett, omfatter ikke vernet innholdet og fakta i artiklene, men snarere presentasjonen, utvalget og sammensetningen. Retrievers sammendrag av artiklene utgjør ikke en krenkelse av opphavsretten, ettersom det dreier seg om gjengivelse av meningsinnholdet, og ikke gjengivelse av sammenhengende tekst.

I den grad forskningsdata inneholder kunstneriske eller litterære verk som er underlagt opphavsrett, vil dette selvsagt stille seg annerledes. I utvalgets rapport fra oktober 2020 nevnes fagområder som studerer kunstneriske uttrykk, og det gis et eksempel med en musikkvitenskapelig database. I slike tilfeller kan videre deling, gjenbruk og videreutvikling av forskningsdataene bare skje så langt det ikke krenker de opprinnelige rettighetene. Forskningsdata som inneholder rene fakta eller informasjon er imidlertid ikke underlagt opphavsrett. Det samme gjelder vitenskapelige teorier eller ideer som bygger på forskningsdata.

Under visse omstendigheter kan man også ha opphavsrett til en database, og ikke bare databasevern (nedenfor under punkt

1.2.2.2). Dette forutsetter at vilkårene for åndsverk er oppfylt ved utformingen av databasen – nemlig at utvalget og innsamlingen er uttrykk for individuell skapende åndsinnsats. Det er sammenstillingen som eventuelt kan ha vern, og ikke de enkelte data. Vitenskapelige databaser vil nok sjelden kvalifisere som åndsverk, ettersom utformingen og sammenstillingen typisk ikke vil bero på kreative valg, og man må som regel basere seg på databasevernet.

Man kan også ha opphavsrett til en sammenstilling av flere selvstendige åndsverk, et såkalt samleverk, for eksempel en artikkelsamling, jf. § 7. En slik rett gjør ingen innskrenkninger i opphavsretten til de enkelte verk.

Opphavsretten oppstår umiddelbart (krever ikke registrering e.l.) og gir en enerett til å råde over åndsverket ved å fremstille et eksemplar av det (for eksempel kopiere verket), eller gjøre verket tilgjengelig for allmennheten (for eksempel publisering på nett), jf. § 3. Eneretten gjelder verket i sin opprinnelig form, men omfatter også bearbeidelser av verket, jf. 3 tredje ledd. Bearbeidelser kan være opphavsrettslig vernet, men utnyttelse krever da samtykke fra opphaveren til originalverket, jf. § 6. Dersom bruk av et eksisterende verk resultater i et nytt og selvstendige verk foreligger det imidlertid ikke en krenkelse ved utnyttelse av det nye verket, jf. § 6 annet ledd.

Opphavsretten varer i 70 år etter opphaverens levetid, jf. § 11. Opphavsretten oppstår hos opphavspersonen, den som har utøvet den individuelle skapende innsatsen, i motsetning til databasevernet som oppstår hos den som har gjort investeringen. Opphavsrett vil derfor typisk oppstå på forskerens hånd, mens databaserettigheter kan oppstå hos institusjonen. Flere personer kan i fellesskap ha utøvet en slik innsats at opphavsrett oppstår. Hverken veiledning, laboratoriearbeid, analyse av data eller finansiering kan imidlertid gi opphavsrett. Det understrekes også at det kun er fysiske personer som kan være opphavere. Dersom institusjonen skal få opphavsrettigheter, må opphavsretten overdras gjennom avtale.

Eneretten som oppstår omfatter som nevnt bare den konkrete utformingen (artikkelen, presentasjonen av forskningsresultatet mv.), og ikke innholdet (forskningsresultatet som sådan). Opphavsretten verner ikke ideer, faktiske opplysninger, synspunkter, hypoteser, matematiske formler, eller tolkningen av en kilde. Derimot kan man ha vern for det konkrete språklige uttrykk, illustrasjoner og fotografier mv.

Opphaveren får også visse ideelle rettigheter i tillegg til de økonomiske, jf. § 5. Mest sentrale er retten til å bli navngitt ved offentliggjøring av verket, og vern mot endring eller tilgjengeliggjøring av verket som er «krenkende» for opphaveren. Selv om opphaveren har gitt noen rett til å endre verket gjennom avtale, må slike endringer ikke krenke opphaverens ideelle rettigheter etter § 5.

Kapittel 3 inneholder bestemmelser om avgrensning av opphavsretten. Det er blant annet en rett til å sitere fra et offentliggjort verk i samsvar med god skikk og i en viss utstrekning, jf. § 29. Dessuten er det etter § 37 en rett til å gjengi offentliggjort kunstverk og offentliggjort fotografisk verk «i tilslutning til teksten i kritisk eller vitenskapelig fremstilling» som ikke er av allmennopplysende karakter. Dette må skje i god skikk og i den utstrekning formålet betinger. Bestemmelsen gir imidlertid ikke rett til gjengivelse i digital form.

2.2.2 Vern for samling av forskningsdata (databasevern)

Etter åndsverkloven § 24 kan den som fremstiller en «database» få enerett til å råde over denne databasen, dersom innsamling, kontroll eller presentasjon av innholdet innebærer en «vesentlig investering». Eneretten er tidsbegrenset og varer i 15 år etter fremstilling eller offentliggjøring – etter utløpet av denne tidsperioden kan databasen fritt utnyttes av enhver.

En «database» er definert i databasedirektivet som en samling av selvstendige verk, data eller annet materiale, som er strukturert systematisk eller metodisk, og som er tilgjengelig individuelt ved bruk av elektroniske eller andre midler[1]. Databasevernet gir altså vern for en samling av data som er strukturert systematisk eller metodisk, og ikke for rådata eller enkeltdata som sådan. Det er selve samlingen som er vernet.

Vilkåret for vern av databaser er at innsamlingen, kontrollen eller presentasjonen av innholdet utgjør en «vesentlig investering». Det er verdt å merke seg at investeringene skal knytte seg til innsamlingen, kontrollen eller presentasjonen, altså arbeidet med struktureringen og sammensetningen av databasen. Investeringer i forbindelse med produksjon av data gir ikke vern, slik at man for eksempel lettere vil kunne få vern for en database som inneholder «eksternt» innhold, der investeringer er nedlagt i selve innsamlingen.

Rettighetshaver til databaser er den som gjør den «vesentlige investeringen». Det vil typisk være en institusjon. Rettighetene vil her oppstå hos institusjonen selv om det for eksempel var en ansatt som kom på at en database var den beste måten å strukturere forskningsdata på. Og selv om det er ansatte som legger ned tid og arbeid i innsamlingen eller presentasjon av databasen, vil det være arbeidsgiver som gjør investeringen, forutsatt at arbeidet foregår i arbeidstiden. Ved eksternt finansierte prosjekter, kan det tenkes at rettighetene oppstår i sameie mellom institusjonen og den eksterne finansiør.

Databasevernet innebærer en enerett til å råde over «hele eller vesentlige deler av databasens innhold ved uttrekk fra eller gjenbruk av databasen», jf. § 24 første ledd. Eneretten beskytter verdien av samlingen som sådan, og ikke de enkelte data som inngår i samlingen. Også gjenbruk i forskning er omfattet av eneretten, det spiller ingen rolle hva man benytter databasen til

[1] Databasedirektivet 96/9/EC, Artikkel 1 nr. 2.

så lenge det dreier seg om uttrekk eller gjenbruk i lovens forstand. Rettighetshaveren kan også motsette seg «gjentatt og systematisk uttrekk eller gjenbruk av *uvesentlige deler* av databasen», forutsatt at dette er «handlinger som skader den normale bruken av databasen eller urimelig tilsidesetter fremstillerens legitime interesser», jf. § 24 annet ledd. Dette gir tilsynelatende et ganske bredt vern for bruk av også små deler av en database, altså små mengder av data, men det er verdt å merke seg at det kun er «gjentatt og systematisk» uttrekk eller gjenbruk som er omfattet, og bare så langt det «skader den normale bruken av databasen», eller er urimelig tilsidesetter rettighetshavers legitime interesser.

I den konkrete vurderingen av om bruk av en database krenker databasevernet, må eneretten til databasen som en eiendomsrett veies opp mot andre grunnleggende rettigheter og friheter, herunder ytringsfriheten[2]. Denne balanseringen av rettigheter ligger allerede til grunn for åndsverklovens bestemmelser, men det følger av EU-domstolens praksis at det skal foretas en reell avveining mellom opphavsretten og ytringsfriheten[3]. Høyesterett legger til grunn at bruk av en database som i utgangspunktet faller innenfor forbudet i loven, kan være lovlig dersom bruken er begrunnet i ytringsfrihet eller andre grunnleggende friheter[4].

2.3 Forskningsdata som ligger til grunn for patenterbare oppfinnelser (patentrett)

2.3.1 Enerett til oppfinnelser og offentliggjøring av patentsøknader

Patentretten gir enerett til oppfinnelser, jf. patentloven § 1 første ledd. En oppfinnelse kan defineres som en 'praktisk løsning på et teknisk problem'[5]. Eksempler på patenterbare oppfinnelser kan være et legemiddel mot kreft, en installasjon på en oljeplattform eller en metode for å fjerne lus på fisk. Patentrettigheter kan kun oppnås ved registrering hos Patentstyret, gjennom inngivelse av en patentsøknad, og gir oppfinneren en tidsbestemt enerett som i utgangspunktet varer i 70 år. Rettighetene er også nasjonale, dvs. at man har rettigheter i det eller de landene man har registrert patent. Eneretten gir innehaveren av patentet rett til å hindre andre i å utnytte oppfinnelsen.

For at man skal kunne oppnå patentrettigheter må det altså dreie seg om en oppfinnelse. Loven avgrenser oppfinnelsesbegrepet mot «oppdagelser, vitenskapelige teorier og matematiske metoder», samt «fremleggelse av informasjon», jf. patl. § 1(2). Oppdagelser, teorier eller fremleggelse av informasjon som sådan løser ikke noe problem, og er dermed ingen teknisk løsning. I den grad man benytter oppdagelser eller informasjon til å løse et problem vil det derimot kunne utgjøre en oppfinnelse. Poenget er at data som sådan ikke kan patenteres, men benyttelsen av data til å løse et problem kan utgjøre en teknisk løsning som kan patenteres. Selv om løsningen som sådan kan patenteres, innebærer ikke det at man får enerett til eventuelle data som ligger til grunn for løsningen. Data som inngår i patentsøknaden vil tvert imot offentliggjøres, og være tilgjengelig for andre til å benytte seg av.

Hovedformålet med patentretten er å skape incentiver til videre forskning og utvikling, gjennom at oppfinneren får en belønning i form av patentrettigheter. «I bytte» mot denne eneretten, offentliggjøres patentsøknaden. Patentsøknaden definerer oppfinnelsen, og beskriver hvordan den fungerer. Denne beskrivelsen skal være så grundig at en fagperson på det aktuelle tekniske området skal kunne utøve oppfinnelsen. I den grad forskningsdata ligger til grunn for oppfinnelsen, må patentsøknaden inneholde slike data i den grad det er nødvendig for å gi en tilstrekkelig presis beskrivelse av oppfinnelsen. Denne offentliggjøringen av informasjon skal bidra til å bringe den tekniske utviklingen videre, gjennom at andre kan benytte seg av informasjonen til å utvikle nye tekniske løsninger. På den måten kan systemet med offentliggjøring av patentsøknader bidra til at forskningsdata deles. Dersom man ønsker å få enerett til kommersiell utnyttelse av en oppfinnelse, må man tåle at informasjonen om oppfinnelsen offentliggjøres, og kan benyttes av andre.

Patentretten hindrer andre i å «utnytte oppfinnelsen», dvs. den tekniske løsningen som er beskrevet i patentsøknaden (det er oppfinnelsen slik den er definert i patentkravene som er gjenstand for eneretten). Kriteriet om at oppfinnelsen må «utnyttes» innebærer at det bare er virksomhet som utnytter oppfinnelsens økonomiske verdi som rammes[6]. Patl. § 3 første ledd lister opp ulike måter for utnyttelse av en oppfinnelse som andre enn patenthaveren ikke kan gjøre uten samtykke. Eksempler er tilvirkning eller anvendelse av et produkt som er beskyttet ved patentet, for eksempel et legemiddel, eller anvendelse av en fremgangsmåte som er beskyttet ved patentet. Eneretten omfatter imidlertid bare utnyttelse i nærings- eller driftsøyemed, og privat bruk er ikke omfattet, jf. § 3 tredje ledd nr. 1. Utnyttelse ved eksperiment som angår selve oppfinnelsen er dessuten unntatt etter § 3 tredje ledd nr. 3. Utnyttelse av oppfinnelsen som kunnskapskilde, i motsetning til utnyttelse av dens økonomiske verdi, er altså tillatt. Bruk av oppfinnelsen til forskning og eksperimentelle formål faller utenfor eneretten, og dette gjelder også kommersiell forskning og utvikling[7]. For å omfattes av unntaket må eksperimentet angå «selve oppfinnelsen». Dette omfatter undersøkelser for å klarlegge hvordan oppfinnelsen fungerer eller for å finne nye

[2] Se HR-2019-1725-A, der Høyesterett tok stilling til spørsmålet om et annet nettsted som hadde publisert høyesterettsavgjørelser hentet fra denne databasen hadde krenket databasevernet til stiftelsen lovdata.
[3] Sak C-469/17 Funke Medien
[4] HR-2019-1725-A
[5] Stenvik, Patentrett, 2020, s. 13.
[6] Stenvik, 2020, s. 322.
[7] Jf. Rt. 2009 s. 1665 (SINTEF) avsnitt 45 og 46.

bruksområder, men det omfatter ikke utnyttelse av patenterte stoffer o.l. i forbindelse med andre eksperimenter[8].

2.3.2 Arbeidsgivers rett til oppfinnelser og forbud mot informasjonsdeling

Arbeidstakeroppfinnelsesloven regulerer arbeidstakeres rettigheter til oppfinnelse i arbeidsforhold. Loven er i utgangspunktet fravikelig, dvs. at den bare kommer til anvendelse når annet ikke er eller må anses avtalt, jf. § 2. Lovens bestemmelser i § 7 første ledd, § 9 og § 10 er imidlertid ufravikelige og man ikke avtale seg bort fra disse. § 7 første ledd oppstiller retten til rimelig godtgjøring for oppfinnelsen. § 9 bestemmer at en avtale mellom arbeidsgiver og arbeidstaker om å begrense arbeidstakerens rett til å forføye over en oppfinnelse som gjøres *senere enn ett år etter tjenestens opphør*, ikke er bindende for arbeidstakeren. § 10 bestemmer at fastsatt godtgjøring etter § 7 kan endres dersom forholdene har endret seg vesentlig.

Som utgangspunkt har en arbeidstaker samme rettigheter til en oppfinnelse som andre oppfinnere, jf. arbeidstakeroppfinnelsesloven § 3, men loven bestemmer at arbeidsgiver har rett til å overta rettighetene til en oppfinnelse som utvikles i forbindelse med et arbeidsforhold på visse vilkår. Lovens § 4 oppstiller hovedreglene om dette. I tilfeller der arbeidstaker har forsknings- eller oppfinnervirksomhet som sine vesentligste arbeidsoppgaver, og har nådd frem til oppfinnelsen gjennom utføring av disse oppgavene, eller oppfinnelsen er et resultat av en nærmere angitt oppgave som arbeidstakeren er pålagt, kan arbeidsgiver kreve retten til oppfinnelsen overført til seg. Det forutsetter imidlertid at utnyttelsen av oppfinnelsen faller innenfor bedriftens virksomhetsområde. Arbeidsgiver kan kreve en utnyttelsesrett i tilfeller der oppfinnelsen er blitt til i annen forbindelse med tjenesten, så lenge den faller innenfor bedriftens virksomhetsområde.

Det har vært diskutert hvorvidt utnyttelsen skal regnes for å falle «innenfor bedriftens virksomhetsområde» der bedriftens virksomhet hovedsakelig går ut på lisensiering eller videre overdragelse av oppfinnelse, som for eksempel for forskningsinstitutter. Det er argumentert for at det i slike tilfeller er naturlig å legge til grunn det som er institusjonens forskningsområde, altså at en oppfinnelse som ligger innenfor forskningsområdet skal anses for å falle innenfor «bedriftens virksomhetsområde»[9]. For at arbeidsgiver skal ha rett til å overta rettighetene til oppfinnelsen, kreves det skriftlig underrettelse til arbeidstakeren innen 4 måneder etter arbeidstaker har gitt arbeidsgiver melding om oppfinnelsen, jf. § 6 jf. § 5.

I tilfeller der arbeidsgiver har rett til å overta rettighetene til en oppfinnelse etter arbeidstakeroppfinnelsesloven legger dette begrensninger på arbeidstakers deling av informasjon om oppfinnelsen, i den grad dette kan forringe mulighetene for patentering eller muliggjør at andre kan utnytte oppfinnelsen, jf. § 6 annet ledd. Dersom informasjon om oppfinnelsen offentliggjøres før inngivelse av patentsøknad, vil dette være til hinder for å oppnå patentrettigheter, ettersom det er et krav for patent at oppfinnelsen er «ny», dvs. at informasjon om den tekniske løsningen ikke allerede er allment tilgjengelig. Denne begrensningen i å dele data vil også omfatte forskningsdata som ligger til grunn for oppfinnelsen.

I tilfeller der patentsøknad inngis, vil denne offentliggjøres på vanlig måte, inkludert eventuelle forskningsdata som søknaden inneholder. Etter at patentsøknad er inngitt er det dessuten ingen begrensninger på oppfinnerens deling av informasjon, ettersom det kun er handlinger som «forringer mulighetene for patentering» som er forbudt etter § 6 annet ledd.

2.3.3 Publiseringsretten for lærere og vitenskapelig ansatte

Etter § 6 tredje ledd er det en spesialregulering av retten til å publisere oppfinnelsen for lærere og vitenskapelig personale ved universiteter og høyskoler. Disse har rett til å publisere oppfinnelsen til tross for begrensningene i § 6 annet ledd, forutsatt at de har varslet arbeidsgiveren i den meldingen de skal gi om oppfinnelsen etter § 5, og at tredjeparts rett ikke er til hinder. Dersom disse vilkårene for publisering er oppfylt, har arbeidsgiver ikke rett til å overta oppfinnelsen etter § 4. Men, dersom oppfinneren ikke innen ett år etter varselet etter § 5 har «tatt skritt til publisering», kan arbeidsgiver likevel overta oppfinnelsen. Dersom arbeidstaker har forbeholdt seg retten til å publisere oppfinnelsen i meldingen som er gitt etter § 5, kan arbeidstaker ikke søke patent på oppfinnelsen uten skriftlig samtykke fra arbeidsgiveren. Publisering av oppfinnelsen vil gjøre at man ikke kan oppnå patentrettigheter, dersom patentsøknaden skal kunne føre til meddelt patent krever det at oppfinnelsen ikke er publisert forut for inngivelsen av patentsøknaden, på grunn av kravet til nyhet. Formålet med spesialreguleringen i § 6 tredje ledd er at arbeidstakeren skal kunne velge spredning av kunnskap fremfor kommersiell utnyttelse[10]. Dersom man velger publisering skal kunnskapen fritt kunne utnyttes av alle, og arbeidstakeren kan da ikke gis rett til å søke patent.

2.3.4 Oppsummering

Patentretten kan gi en oppfinner (eller andre som rettighetene er overdratt til) enerett til å utnytte en teknisk løsning. Der forskningsdata gir grunnlag for å utvikle en oppfinnelse, kan man altså få enerett til utnyttelse av oppfinnelsen, men man får ikke enerett til dataene som sådan. For å oppnå patentrettigheter er det imidlertid et krav at informasjonen om oppfinnelsen offentliggjøres, gjennom offentliggjøringen av patentsøknaden. Forskningsdata som inngår i en patentsøknad vil dermed bli offentlig tilgjengelig, og kan benyttes i videre forskning og utvikling, så lenge man ikke krenker patentretten gjennom å utnytte den patentbeskyttede oppfinnelsens økonomiske verdi. Patentretten er ikke til hinder for bruk av oppfinnelsen som kunnskapskilde, herunder til forskning og eksperimentelle

[8] Stenvik, 2020, s. 322.
[9] Stenvik, Norsk lovkommentar, note (8). https://min.rettsdata.no/Dokument/gL19700417z2D21?directHit=1&dq=arbeidstakeroppfinnelsesloven&noteid=gN19700417z2D21z2E8
[10] Ibid., note (21).

formål, inkludert kommersiell forskning og utvikling. Patentretten setter i utgangspunktet altså ingen hindringer for deling av forskningsdata, men kan tvert imot innebære at forskningsdata må deles i den grad de inngår i en patentsøknad. I utgangspunktet innebærer patentretten derfor ingen begrensning for deling av forskningsdata, men kan snarere bidra til informasjonsspredning. I situasjoner der arbeidsgiver har rett til å overta rettighetene til en oppfinnelse etter arbeidstakeroppfinnelsesloven, kan arbeidstakeroppfinnelsesloven sette begrensinger på oppfinnerens rett til å dele informasjon om oppfinnelsen forut for inngivelse av en patentsøknad. Lærere og vitenskapelig personale ved universiteter og høyskoler har imidlertid en publiseringsrett etter lovens § 6 tredje ledd, som går foran arbeidsgivers rett til oppfinnelsen. I tilfeller der patentsøknad inngis, vil denne offentliggjøres på vanlig måte, inkludert eventuelle forskningsdata som søknaden inneholder. Etter at patentsøknad er inngitt er det dessuten ingen begrensninger på oppfinnerens deling av informasjon.

3. REGLER SOM TILRETTELEGGER FOR, REGULERER ELLER KREVER DELING AV FORSKNINGSDATA

3.1 Innledning

Formålet her i punkt 1.3 er å redegjøre for de regelverk som i utgangspunktet tilrettelegger for og krever offentliggjøring av forskning. Vi har blant annet offentlighetsloven som gir rett til å kreve innsyn i offentlige organers dokumenter, herunder dokumenter som inneholder forskningsdata. Videre har vi forskningsetikkloven og universitets- og høyskoleloven som stiller visse krav til åpenhet og tilgjengeliggjøring av forskning, både av institusjonene og av forskerne selv. Helseforskningsloven stiller dessuten strenge krav til blant annet offentliggjøring av resultater på sitt område. Bestemmelser om tilgjengeliggjøring og offentlighet må imidlertid sees i lys av de begrensninger og unntak som følger av den enkelte lov, men også de begrensninger som legges på deling av data gjennom annet regelverk. Regler som hovedsakelig legger begrensninger på deling av data er tema nedenfor i punkt 1.4.

3.2 Offentlighetsloven

Offentlighetsloven gir rett til innsyn i dokumenter som gjelder et offentlig organs virksomhet. Slike dokumenter kan tenkes å inneholde forskningsdata, og offentlighetsloven kan gi rett til innsyn også i disse. Offentlighetsloven krever således deling av forskningsdata under visse omstendigheter.

Offentlighetsloven gjelder bare for de rettssubjekter som er listet opp i offl. § 2. Dette er blant annet stat, fylkeskommune og kommune (bokstav a), andre rettssubjekter som gir enkeltvedtak (bokstav b), selvstendige rettssubjekter der stat eller kommune har en eierandel som gir mer enn halvparten av stemmene i det øverste organet i rettssubjektet (bokstav c), og selvstendige rettssubjekt der stat eller kommune direkte eller indirekte har rett til å velge mer enn halvparten av medlemmene med stemmerett i det øverste organet i rettssubjektet (bokstav d).

Forskrift til offentlighetsloven §1 lister opp visse unntak fra virkeområdet til offentlighetsloven. Det er gjort unntak for visse rettssubjekter, og for enkelte dokumenter hos selvstendige rettssubjekter. Dette gjelder blant annet dokumenter knyttet til sak om kommersiell utnyttelse av forskningsresultat og sak om bidrags- eller oppdragsforskning i rettssubjekter innenfor UH-sektoren, jf. §1 tredje ledd bokstav d.

Retten til innsyn etter offentlighetsloven er betinget av at noen fremsetter et krav om innsyn, jf. offl. § 3 («alle kan krevje innsyn»). Innsyn kan kreves skriftlig eller muntlig, men det kravet må i utgangspunktet gjelde «ei bestemt sak», jf. § 28 annet ledd. Dette innebærer at kravet om innsyn må være tilstrekkelig spesifisert – den som krever innsyn må vite hva han eller hun ønsker innsyn i. Forskningsdata som enda ikke er offentliggjort, kan derfor være vanskelig tilgjengelig for et innsynskrav, mens der det foreligger en artikkel eller annen publikasjon, kan det være enklere å formulere et tilstrekkelig presist innsynskrav knyttet til de underliggende data. Innsynskravet kan i stedet for en bestemt sak også gjelder «i rimeleg utstrekning saker av ei bestemt art», jf. § 28 annet ledd. Dette kravet tar bare sikte på den arbeidsbyrden forvaltningsorganet blir påført hvis det skal oppfylle kravet[11]. Dette innebærer at det for eksempel kan nektes innsyn der organet må finne frem til sakene manuelt og antall saker er betydelig, eller der utlevering vil kreve en betydelig arbeidsinnsats i form av en gjennomgang av opplysningene for å sortere ut taushetsbelagt informasjon eller annen informasjon som er unntatt fra innsyn[12].

Innsynsretten gjelder i utgangspunktet for «saksdokument» (i tillegg til «journalar» og «liknande register»), jf. § 3. Både «dokument» og «saksdokument» er definert i lovens § 4. Et «dokument» er en logisk avgrenset informasjonsmengde som er lagret på et medium for senere lesing, lytting, fremvisning, overføring eller lignende. Dette er en vid definisjon, og omfatter mer enn det man gjerne tenker på som et tradisjonelt tekstdokument. Kravet er at det er en logisk avgrenset informasjonsmengde som kan lagres og leses e.l. Dette kan være et alminnelig tekstdokument, men omfatter også lyd, bilde, tegninger, modeller osv. Begrepet er teknologinøytralt, og det har ikke noe å si hvordan informasjonen kommer til uttrykk[13]. Definisjonen av «saksdokument» innebærer at det må dreie seg om dokumenter som er kommet inn til eller lagt fram for et organ, eller som organet selv har opprettet. I tillegg må det gjeldet organets *ansvarsområde eller virksomhet*. Dette kravet skal ikke tolkes strengt, det er bare tale om et krav om en innholdsmessig sammenheng med virksomheten i organet,

[11] Norsk Lovkommentar til Offentleglova, Rettsdata, ved Jan Fridthjof Bernt, note 400.
[12] Ibid.
[13] Ibid., note 68.

og det er tilstrekkelig at dokumentet knytter seg til ansvarsområdene til organet eller til virksomheten mer generelt[14]. I § 4 tredje ledd er det listet opp visse dokumenter som ikke skal regnes som saksdokument, blant annet aviser og tidsskrifter som organet mottar, og dokumenter som en medarbeider mottar i annen egenskap enn som ansatt i organet.

I tillegg til en rett til dokumentinnsyn, gir offentlighetsloven også en rett til innsyn i «ei samanstilling av opplysningar som er elektronisk lagra i databasane til organet», jf. § 9. Gjennom en slik sammenstilling oppstår det et nytt «dokument» i lovens forstand[15]. Retten etter § 9 forutsetter at sammenstillingen kan gjøres «med enkle framgangsmåtar». Dette kriteriet knytter seg til den arbeidsbyrden som påføres organet ved å lage sammenstillingen – dersom sammenstillingen krever tidkrevende manuelle operasjoner, vil vilkåret ikke være oppfylt[16].

Hovedregelen om innsyn følger av offentlighetsloven § 3. Der heter det at saksdokumenter mv. er åpne for innsyn dersom «ikkje anna følgjer av lov eller forskrift». Retten til innsyn kan altså begrenses av andre lovbestemmelser som setter begrensninger på deling av informasjon. Dette kan være for eksempel regler om taushetsplikt, personvern eller regler om vern av forretningshemmeligheter. Offentlighetsloven spesifiserer i tillegg visse unntak fra innsynsretten i lovens kapittel 3. Etter § 13 er opplysninger som er underlagt taushetsplikt i lov eller i medhold av lov, unntatt fra innsyn.

Et eksempel på en slik bestemmelse er forvaltningsloven § 13 første ledd som pålegger taushetsplikt om (1) «noens personlige forhold», og (2) «tekniske innretninger og fremgangsmåter samt drifts- eller forretningsforhold som det vil være av konkurransemessig betydning å hemmeligholde av hensyn til den som opplysningen angår». Når det gjelder alternativ 1 er det gitt en oppramsing i § 13 annet ledd med forhold som ikke regnes som personlig – en slags nedre grense. For alternativ 2 kreves det for det første at det dreier seg om informasjon om *tekniske innretninger, fremgangsmåter, driftsforhold eller forretningsforhold, og for det andre at det er av konkurransemessig betydning å hemmeligholde opplysningene.* Dette forutsetter i utgangspunktet at det er fare for økonomisk tap ved offentliggjøring av opplysningene[17]. Det kan tenkes at forskningsdata kan bli underlagt taushetsplikt i medhold av denne bestemmelsen, i tilfeller der det er av konkurransemessig betydning å hemmeligholde dataene, for eksempel der offentliggjøring av dataene vil kunne være til hinder for en senere patentsøknad. Meddelelse av patent krever at oppfinnelsen er «ny», og det vil være nyhetshindrende dersom informasjon om oppfinnelsen som er tilstrekkelig til å kunne utøve den blir offentliggjort forut for inngivelse av patentsøknad (se ovenfor under punkt 1.2.3).

Det kan understrekes at der informasjon er underlagt taushetsplikt etter loven, er brudd på taushetsplikten straffbart, jf. strl. § 209.

Man kan i denne forbindelse også merke seg fvl. § 13 d og § 13 e. § 13 d sier at departementet kan bestemme at et forvaltningsorgan kan eller skal gi opplysninger til bruk for forskning *uten hinder av organets taushetsplikt etter § 13,* dersom det finnes rimelig og ikke medfører uforholdsmessig ulempe for andre interesser. § 13 e pålegger forskere taushetsplikt om visse opplysninger, blant annet opplysninger undergitt taushetsplikt og opplysninger som er mottatt fra private under taushetsløfte i forbindelse med forskningsarbeidet.

Et annet eksempel på informasjon som er unntatt fra innsyn er opplysninger om forskningsideer og forskningsprosjekt i saker som gjelder økonomisk støtte eller rådgivning fra det offentlige i forbindelse med forskningsprosjekt, jf. offl. § 26 (4). Slike opplysninger vil i noen tilfeller være underlagt taushetsplikt som forretningshemmeligheter etter fvl. § 13 første ledd nr. 2, men denne bestemmelsen gjelder bare taushetsplikt for opplysninger der det kan være fare for økonomisk tap ved offentliggjøring, hvilket kan være vanskelig å dokumentere for forskningsprosjekter utenfor det kommersielle området, eller der det ikke er nærliggende at forskningen vil ha kommersiell verdi på et tidlig tidspunkt[18].

Offentlighetsloven kan altså gi rett til innsyn i offentlige organers dokumenter, også der disse inneholder forskningsdata. Dette forutsetter imidlertid at det er fremsatt et tilstrekkelig spesifisert krav om innsyn, og at informasjonen ikke er underlagt taushetsplikt eller for øvrig unntatt fra innsyn.

3.3 Forskningsetikkloven

Det følger av forskningsetikklovens formålsbestemmelse i § 1 at loven «skal bidra til at forskning i offentlig og privat regi skjer i henhold til anerkjente forskningsetiske normer». Loven regulerer forebygging og brudd på *anerkjente forskningsetiske normer for god vitenskapelig praksis*[19]. Lovfestede etiske krav som blant annet finnes i helseforskningsloven og personopplysningsloven vil ved motstrid gå foran forskningsetikkloven[20]. Forskningsetikkloven § 4 bestemmer at forskere har en aktsomhetsplikt, og den pålegger forskningsinstitusjonene et ansvar for å sikre at forskningen ved institusjonen skjer i henhold til anerkjente forskningsetiske normer, jf. § 5. Forskningsinstitusjonene har selv ansvar for å behandle saker om mulige brudd på anerkjente forskningsetiske normer av institusjonens forskere, og skal blant

[14] Ibid., note 73.
15 Ibid., note 133.
[16] Ibid., note 137.
[17] Ibid., note 386.
[18] Ibid., note 386.
[19] Lovkommentar til forskningsetikkloven, Rettsdata, ved Bjørn L. Zwilgmeyer, note 1.
[20] Ibid.

annet ha et redelighetsutvalg, jf. § 6. Det er i tillegg oppnevnt et granskningsutvalg, som er klageinstans for uttalelser om at en forsker har opptrådt vitenskapelig uredelig, og som skal veilede forskningsinstitusjoner om behandling av saker, jf. § 7. Loven bestemmer også at det oppnevnes nasjonale forskningsetiske komiteer, jf. § 9, og regionale komiteer for medisinsk og helsefaglig forskningsetikk (REK), jf. § 10.

I uttalelser fra forskningsinstitusjoner, redelighetsutvalg eller granskningsutvalget i saker om mulig brudd på anerkjente forskningsetiske normer skal det alltid tas stilling til a) om forskeren har opptrådt vitenskapelig uredelig, b) om det foreligger systemfeil ved institusjonen, og c) om det aktuelle vitenskapelige arbeidet bør korrigeres eller trekkes tilbake, jf. § 8 første ledd. En slik uttalelse er å anse som en tilråding – det er opp til institusjonen om tilrådingen følges opp[21]. Sanksjonene vil være gjenstand for alminnelig domstolsprøving eller prøving hos Sivilombudsmannen (for eksempel SOM-2012-1158 om underkjenning av doktorgradsavhandling)[22]. Vitenskapelig uredelighet, jf. 8 første ledd bokstav a, er i annet ledd definert som «forfalskning, fabrikkering, plagiering og andre alvorlige brudd på anerkjente forskningsetiske normer», i forbindelse med forskning. Det kreves også at det må foreligge forsett eller grov uaktsomhet hos forskeren. Et eksempel på slike anerkjente forskningsetiske normer er De nasjonale forskningsetiske komiteenes retningslinjer[23]. I de Generelle forskningsetiske retningslinjer, som er utarbeidet av De nasjonale forskningsetiske komiteene, er det lagt til grunn i punkt 11 at forskningsresultater som hovedregel skal tilgjengeliggjøres[24]. Manglende tilgjengeliggjøring kan således tenkes å utgjøre brudd på forskningsetiske normer. Øvrige forskningsetiske normer i forbindelse med innsamling og publisering av forskningsdata, omfatter krav til redelighet (punkt 7) og plikt til å følge god henvisningsskikk ved bruk av andres materiale (punkt 8).

Forskningsetikkloven kan altså kreve tilgjengeliggjøring av forskningsdata i den grad det følger av anerkjente forskningsetiske normer, og dersom man benytter andres forskningsdata må man følge normer om redelighet og god henvisningsskikk. I den grad et brudd på en forskningsetisk norm kan sies å være alvorlig, og forskeren har opptrådt forsettlig eller grovt uaktsomt, kan saken tenkes å resultere i en uttalelse om at forskeren har opptrådt uredelig. Brudd på forskningsetiske normer i forbindelse med forskningen kan også resultere i en uttalelse om at et vitenskapelig arbeid bør korrigeres eller trekkes tilbake.

3.4 Universitets- og høyskoleloven

Universitets- og høyskoleloven skal blant annet tilrettelegge for at universiteter og høyskoler utfører forskning og faglig og kunstnerisk utviklingsarbeid på høyt internasjonalt nivå, jf. universitets- og høyskoleloven § 1-1 bokstav b. Loven skal også tilrettelegge for at disse «formidler kunnskap om virksomheten og utbrer forståelse for prinsippet om faglig frihet og anvendelse av vitenskapelige og kunstneriske metoder og resultater», jf. § 1-1 bokstav c. I § 1-3 er det listet opp ulike måter hvorigjennom universiteter og høyskoler skal fremme lovens formål. Her er det eksplisitt uttalt at institusjonene skal «bidra til å spre og formidle resultater fra forskning…» (bokstav d). Dette kan sies å innebære en forpliktelse til å oppmuntre og legge forholdene til rette for slik formidling[25]. Det følger imidlertid også av bestemmelsen at institusjonene skal «bidra til innovasjon og verdiskapning basert på resultater fra forskning» (bokstav e). I Norsk Lovkommentar er det fremhevet at det er ønskelig at resultater av forskning og faglig utviklingsarbeid som har et kommersielt potensial skal formidles til virksomheter som kan bidra til innovasjon og verdiskapning – slik kunnskapsoverføring kan skape nye arbeidsplasser og styrke norsk kunnskapsindustri[26]. En forutsetning for slik kommersialisering er at forskningens frihet og integritet ivaretas, og at forskningsresultatene offentliggjøres innen rimelig tid[27]. Ønsket om kommersialisering og kravet til offentliggjøring kan imidlertid tenkes å komme i konflikt dersom kommersialisering forutsetter inngivelse av patentsøknader, hvilket legger begrensninger på offentliggjøring av forskningsresultatene forut for inngivelse av slik søknad (jf. ovenfor om patentrettigheter under punkt 1.2.3).

Det følger videre av § 1-5 (6) at universiteter og høyskoler «skal sørge for åpenhet om resultater fra forskning…». I denne bestemmelsen er også *forskernes rett og plikt til offentliggjøring* utførlig regulert. Den som er ansatt i stilling hvor forskning inngår i arbeidsoppgavene, har *rett til å offentliggjøre* sine resultater og *skal sørge for* at slik offentliggjøring skjer. Institusjonen kan imidlertid ikke pålegge offentliggjøring på et bestemt tidspunkt eller med en bestemt hyppighet, og det er uansett forskeren selv som har råderetten over resultatene og offentliggjøring av disse[28]. Det følger av § 1-5 (5) at forskeren har rett til å velge emne og metode for sin forskning og sitt utviklingsarbeid innenfor de rammer som følger av ansettelsesforholdet eller særskilt avtale. Videre heter det i § 1-5 (6) at «det relevante forskningsgrunnlaget skal stilles til rådighet i overensstemmelse med god skikk på vedkommende fagområde». Her er det altså spesifisert at ikke bare skal resultatene offentliggjøres, men også «forskningsgrunnlaget» skal stilles til rådighet. Dette kan være statistiske grunndata eller fysisk

[21] Ibid., note 22.
[22] Ibid.
[23] https://www.forskningsetikk.no/ressurser/fbib/lover-retningslinjer/forskningsetikkloven/
[24] https://www.forskningsetikk.no/retningslinjer/generelle/
[25] Lovkommentar til universitets- og høyskoleloven, Rettsdata, ved Jan Fridthjof Bernt, note 25.
[26] I bid., note 26.
[27] Ibid.
[28] Ibid., note 67.

materiale[29]. At det skal stilles til rådighet innebærer at det skal gjøres tilgjengelig ved forespørsel[30]. Opplysninger som er underlagt taushetsplikt må imidlertid som utgangspunkt skjermes eller anonymiseres[31]. Styret ved institusjonen kan samtykke til «utsatt offentliggjøring» av forskningsresultatene dersom det foreligger legitime hensyn. Slike legitime hensyn kan være beskyttelsen av patentrettslige eller konkurransemessige interesser, eller hensynet til løpende forskningsarbeid[32]. Det er imidlertid uttrykkelig fastslått at det ikke kan avtales eller fastsettes «varige begrensninger» i retten til offentliggjøring utover det som følger av lov. Situasjoner der retten til offentliggjøring er begrenset ved lov kan være regler om taushetsplikt eller vern av forsvarshemmeligheter[33]. Det følger dessuten av lovens § 7-6 at bestemmelsene om taushetsplikt i fvl. §§ 13 til 13e gjelder for universiteter og høyskoler.

3.5 Helseforskningsloven

Helseforskningsloven har til formål å fremme god og etisk forsvarlig medisinsk og helsefaglig forskning, jf. helseforskningsloven § 1. Loven gjelder for medisinsk og helsefaglig forskning på mennesker, humant biologisk materiale eller helseopplysninger, jf. § 2.

Lovens § 8 bestemmer at «kommersiell utnyttelse av forskningsdeltakere, humant biologisk materiale eller helseopplysninger som sådant er forbudt». Bestemmelsen er vag, og den nærmere grensedragningen er overlatt til praksis[34]. Bestemmelsen er ikke ment å forhindre forskning i næringsformål, for eksempel legemiddelutprøving[35]. Omsetning og kommersialisering av forskningsresultater og bearbeidet materiale faller også utenfor bestemmelsen[36].

Det stilles krav til at den regionale komiteen for medisinsk og helsefaglig forskningsetikk (REK) godkjenner forskningsprosjektet etter søknad, jf. §§ 9 og 10, jf. forskningsetikkloven § 10 første ledd. I § 12 er det bestemt at når forskningsprosjektet avsluttes skal det sendes en sluttmelding til REK. I denne skal resultatene presenteres på en objektiv og etterrettelig måte, som sikrer at positive og negative funn fremgår. Formålet med bestemmelsen er å bidra til kontroll og åpenhet rundt forskningen[37]. Samfunnet skal ha mulighet til å følge med på hva slags medisinsk og helsefaglig forskning som foregår, og det må sikres at resultater ikke underslås for eksempel dersom de er til ugunst for den som finansierer forskningen[38]. Et sammendrag på én til to sider anses som hensiktsmessig[39]. Det er utarbeidet et eget skjema for sluttmeldingen av REK. Sluttmeldingen må også sees i sammenheng med at REK har plikt til å føre et systematisk og offentlig tilgjengelig register over innmeldte og avsluttede forskningsprosjekter, jf. § 44 (se nedenfor).

Helseforskningsloven har også regler i kapittel 6 om forskningsbiobanker som opprettes i forbindelse med innsamling, oppbevaring og bruk av humant biologisk materiale. I § 31 er det bestemt at den forskningsansvarlige skal gi andre forskere tilgang til humant biologisk materiale i virksomhetens forskningsbiobanker, med mindre den ansvarlig selv har behov for materialet eller det foreligger andre særlige grunner.

I kapittel 7 er det regler om behandling av helseopplysninger, som blant annet skal være i samsvar med prinsippene i personvernforordningen artikkel 5 og ha uttrykkelig angitte formål, jf. § 32. Det følger av § 34 at helseopplysninger kan sammenstilles, tilgjengeliggjøres og behandles i tråd med forskningsprosjektets formål, eventuelle samtykker, forhåndsgodkjenningen etter § 33 og i samsvar med forskningsprotokollen. REK kan imidlertid nekte slik sammenstilling, tilgjengeliggjøring mv. dersom denne anses som medisinsk eller etisk uforsvarlig, § 34 annet ledd. Det er understreket at sammenstilling og tilgjengeliggjøring av helseopplysninger kan skje til dataansvarlige eller forskningsansvarlige som har adgang til å behandle personopplysningene etter personvernforordningen artikkel 6 og 9. REK kan også bestemme at helseopplysninger kan eller skal gis fra helsepersonell til bruk i forskning, uten hinder av taushetsplikt, etter § 35. I § 36 er det regler om retting eller sletting av helseopplysninger, som den registrerte kan kreve etter personvernforordningen artikkel 16 og 17. I § 38 er det gitt forbud mot lagring av unødvendige helseopplysninger.

Kapittel 8 har regler om *åpenhet og innsyn i forskningen*. Det følger av § 39 at den forskningsansvarlige og prosjektlederen skal sørge for åpenhet rundt forskningen. Allmennhetens rett til innsyn er regulert i § 41. Enhver skal ved henvendelse til REK få vite hvilke forskningsprosjekter etter helseforskningsloven som en forskningsansvarlig eller prosjektleder er eller har vært involvert i, samt formålet med prosjektet. Unntakene i personopplysningsloven §§ 16 og 17 gjelder tilsvarende her, jf. § 42. Endelig følger det av § 44 at REK skal føre en systematisk fortegnelse over innmeldte og avsluttede forskningsprosjekter basert på opplysninger i søknad og sluttmelding, og at registrene skal være offentlig tilgjengelig. Man har imidlertid mulighet til å søke om utsatt offentliggjøring etter § 45, der dette er nødvendig for å beskytte legitime patentrettslige eller konkurransemessige interesser, eller av hensyn til et løpende forskningsarbeid. REK kan dessuten bestemme at sensitive

[29] Ibid., note 68.
[30] Ibid., note 69.
[31] Ibid
[32] Ibid., note 72.
[33] Ibid., note 75.
[34] Lovkommentar til helseforskningsloven, Rettsdata, ved Sigmund Simonsen, note 40.
[35] Ibid.
[36] Ibid.
[37] Ot.prp. nr. 74 (2006-2007) pkt. 11.3.4.2.
[38] Ibid.
[39] Ibid.

opplysninger om et prosjekt ikke skal inngå i registeret eller ikke kunne kreves innsyn i, for et klart avgrenset tidsrom, dersom offentlighet kan skade vesentlige private eller offentlige interesser.

4. BEGRENSNINGER PÅ DELING AV DATA

4.1 Innledning

Her i punkt 1.4 skal vi se nærmere på lovverk som legger begrensninger på deling av data på grunn av *innholdet* i dataene. Dette omfatter begrensninger som skal beskytte datasubjektet, herunder personvernregelverket og GDPR, begrensninger som skal beskytte andre interesser, for eksempel forretningshemmeligheter, og begrensninger av hensyn til nasjonal sikkerhet.
Begrensninger på deling av data som skyldes *rettigheter* til data, samling av data, eller oppfinnelser som bygger på data er redegjort for ovenfor i punkt 1.2. Forbudet mot informasjonsdeling etter arbeidstakeroppfinnelsesloven er også redegjort for ovenfor – dette knytter seg som nevnt til informasjon om en oppfinnelse, og ikke enkeltdata som sådan.

4.2 Begrensninger som skal beskytte datasubjektet

4.2.1 Personopplysningsloven

Personopplysningsloven gjennomfører personvernforordningen, og regulerer behandlingen av personopplysninger. Personopplysningsloven gjelder behandling av personopplysninger også i forbindelse med vitenskapelig forskning, og som det vil redegjøres for i det følgende er det enkelte særlige omstendigheter som gjør seg gjeldende. Der personopplysninger inngår i forskningsdata, vil personopplysningsloven altså stille krav til behandlingen av slike opplysninger. Ved offentliggjøring av forskningsresultater og forskningsdata hvor personopplysninger inngår, vil personopplysningsloven kunne sette begrensninger og krav.

Personvernforordningen har til formål å verne fysiske personer i forbindelse med behandling av personopplysninger, og sikrer deres rett til vern av personopplysninger, samtidig som fri utveksling av personopplysninger i EØS ikke skal begrenses eller forbys (se artikkel 1). Personopplysninger er enhver opplysning om en identifisert eller identifiserbar fysisk person («den registrerte»), jf. artikkel 4. En identifiserbar fysisk person er definert som en person som direkte eller indirekte kan identifiseres, særlig ved hjelp av en identifikator. En slik identifikator kan være navn, identifikasjonsnummer, lokaliseringsopplysninger, nettidentifikator e.l. I fortalepunkt 34 og 35 er det uttalelser om personopplysninger i form av genetiske opplysninger, og opplysninger om helsetilstand.

Behandling av personopplysninger må oppfylle et av vilkårene i artikkel 6 nr. 1, herunder (a) at det foreligger (frivillig og informert) samtykke fra den registrerte, (b) at behandlingen er nødvendig for å oppfylle en avtale som den registrerte er part i, eller (e) at behandlingen er nødvendig for å utføre en oppgave i allmennhetens interesse, eller (f) behandlingen er nødvendig for formål knyttet til de berettigede interessene som forfølges av den behandlingsansvarlige eller en tredjepart.

Om forskning heter det i forordningen fortalepunkt 33 at det for formål knyttet til vitenskapelig forskning ofte ikke er «mulig fullt ut å identifisere formålet med behandlingen av personopplysninger på tidspunktet for innsamlingen av opplysningene. De registrerte bør derfor kunne gi sitt samtykke til visse områder innen vitenskapelig forskning når dette er i samsvar med anerkjente etiske standarder for vitenskapelig forskning». Personopplysningsloven § 8 bestemmer at personopplysninger kan behandles på grunnlag av personvernforordningen artikkel 6 nr. 1 bokstav e), dersom det er nødvendig for «formål knyttet til vitenskapelig eller historisk forskning eller statistiske formål» (se nærmere nedenfor).

I forordningens kapittel III (artikkel 12 til 23) gis den registrerte visse rettigheter, herunder til informasjon, innsyn og sletting, og i kapittel IV (artikkel 24 til 43) er det gitt regler om behandlingsansvarlig og databehandler. Ved behandling av personopplysninger stilles det krav til en behandlingsansvarlig som skal gjennomføre egnede tekniske og organisatoriske tiltak for å sikre og påvise at behandlingen av personopplysninger er i samsvar med forordningen (artikkel 24). Og dersom en behandling skal utføres på vegne av en behandlingsansvarlig, stilles det krav til at den behandlingsansvarlige bruker databehandlere som gir tilstrekkelige garantier for at de vil gjennomføre tilsvarende tiltak (artikkel 28). Slike tiltak kan omfatte pseudonymisering og kryptering av personopplysninger, og evne til å sikre vedvarende konfidensialitet i behandlingssystemene (artikkel 32).

For behandling av personopplysninger utelukkende «med henblikk på akademiske ytringer» gjelder bare visse bestemmelser i personvernforordningen og personopplysningsloven, jf. personopplysningsloven § 3. Dette gjelder personvernforordningen artikkel 24, 26, 28, 29, 32 og 40 til 43, jf. personvernforordningen kapittel VI og VIII og personopplysningsloven kapittel 6 og 7. Artiklene det vises til i personvernforordningen dreier seg blant annet om reglene om behandlingsansvarlig og databehandler, sikkerhet ved behandlingen, og atferdsnormer og sertifisering. Dette er materielle bestemmelser som gjelder uansett. Kapitlene det vises til er prosessuelle regler om tilsyn med etterlevelsen av reglene og sanksjoner ved brudd på de materielle reglene. De prosessuelle reglene gjelder bare i samme utstrekning som de materielle bestemmelsene, og Datatilsynet kan bare føre tilsyn med og sanksjonere brudd på de materielle bestemmelsene som er listet opp i § 3[40]. Når det

[40] Norsk Lovkommentar til personopplysningsloven, Rettsdata, ved Thomas Olsen, note 3-10.

gjelder innholdet i begrepet «med henblikk på akademiske ytringer» tilsier ordlyden at behandlingen ikke nødvendigvis må resultere i en ytring, så lenge den skjer med henblikk på en ytring[41]. I forarbeidene til personopplysningsloven er det uttalelser om forholdet mellom akademiske ytringer og forskning, og om rekkevidden av unntaksregelen i lys av kravene til ytrings- og informasjonsfriheten:

«det må trekkes en grense mot behandling av personopplysninger i forbindelse med forskning, selv om resultatene av forskningen eventuelt skal publiseres, og behandlingen derfor kan sies å foranledige en akademisk ytring. At forordningen skal gjelde for forskning er slått fast i fortalepunkt 159 og 160, jf. også 156, og adgangen til å gjøre unntak fra forordningens regler for behandling av personopplysninger i forbindelse med forskning er regulert i forordningen artikkel 89. Ettersom det ikke skal gjøres unntak fra personopplysningsvernet i større grad enn det som er nødvendig av hensyn til ytrings- og informasjonsfriheten, jf. artikkel 85 nr. 2 og fortalepunkt 153, vil rekkevidden av unntaksregelen måtte bero på hvilke krav reglene om ytrings- og informasjonsfrihet stiller, jf. Grunnloven § 100 og EMK artikkel 10.» [42]

Personvernforordningen artikkel 89 regulerer adgangen til å gjøre unntak fra forordningens regler for behandling av personopplysninger i forbindelse med forskning. Her er det blant annet bestemt at når personopplysninger behandles for formål knyttet til vitenskapelig forskning, kan det fastsettes unntak fra rettighetene i artikkel 15, 16, 18 og 21 på visse vilkår. Disse bestemmelsene gjelder den registrertes rett til innsyn, rett til retting, rett til begrensning av behandling og rett til å protestere.

Personopplysningsloven § 8 bestemmer som nevnt at personopplysninger kan behandles på grunnlag av personvernforordningen artikkel 6 nr. 1 bokstav e, dersom det er nødvendig for «formål knyttet til vitenskapelig eller historisk forskning eller statistiske formål». Vitenskapelig forskning skal tolkes vidt, og det er uttalt i personvernforordningen fortalepunkt 159 at «I denne forordning bør behandling av personopplysninger i forbindelse med formål knyttet til vitenskapelig forskning tolkes vidt og f.eks. omfatte teknologisk utvikling og demonstrasjon, grunnleggende forskning, anvendt forskning og privatfinansiert forskning»[43]. Behandlingen skal i slike tilfeller være omfattet av nødvendige garantier i samsvar med personvernforordningen artikkel 89 nr. 1, dvs. nødvendige garantier for å ivareta den registrertes rettigheter og friheter. Det følger også av fortalepunkt 159 at «for å ta hensyn til de særlige forholdene som gjelder ved behandling av personopplysninger i forbindelse med vitenskapelig forskning, bør særlige vilkår få anvendelse, særlig med hensyn til publisering eller annen utlevering av personopplysninger i forbindelse med nevnte formål. Dersom resultatet av den vitenskapelige forskningen, særlig på området helse, berettiger ytterligere tiltak i den registrertes interesse, bør de allmenne bestemmelsene i denne forordning få anvendelse med henblikk på nevnte tiltak.»

I personopplysningsloven § 9 sies det noe om behandling av særlige kategorier av personopplysninger uten samtykke for formål knyttet til vitenskapelig forskning. I § 17 er det gjort unntak fra den registrertes rett til innsyn mv. ved behandling av personopplysninger for vitenskapelige formål.

Det kan nevnes at Datatilsynet gir nærmere informasjon på sine sider om behandling av personopplysninger i forbindelse med forskning: https://www.datatilsynet.no/regelverk-og-verktoy/sporsmal-svar/forskning/jeg-skal-bruke-personopplysninger-i-forskningen-min.-hva-skal-jeg-gjore/. Datatilsynet understreker at det er den behandlingsansvarlige (for eksempel en forskningsinstitusjon) som må vurdere og dokumentere at den planlagte behandlingen av personopplysninger oppfyller kravene i loven. Kravet omfatter all forskning hvor det brukes personopplysninger. Det gjøres ikke unntak fra kravet selv om forskningsprosjektet må ha etisk godkjenning fra REK, og REKs vedtak gir ikke rettslig grunnlag for behandling av person-opplysninger. Datatilsynet fremhever også at det i mange tilfeller kreves dispensasjon fra taushetsplikt for behandling av personopplysninger, og at REK vurderer om det kan gis dispensasjon fra taushetsplikt i pasientjournaler. Det kan også nevnes at Personvernrådet (EDPB) har vedtatt retningslinjer om behandling av helsedata til forskningsformål og bruk av lokasjonsdata og kontaktsporingsverktøy i forbindelse med COVID-19 pandemien: https://www.datatilsynet.no/regelverk-og-verktoy/internasjonalt/retningslinjer-og-uttalelser-fra-personvernradet/nye-retningslinjer-fra-personvernradet-om-helsedata-og-lokasjonsdata/v

4.2.2 Regler om taushetsplikt

Når det gjelder regler om taushetsplikt har man blant annet regler om taushetsplikt i forvaltningsloven §§ 13 flg. (se om dette ovenfor under punkt 1.3.2). Forvaltningsloven § 13 pålegger offentlig ansatte taushetsplikt om opplysninger som gjelder (1) «noens personlige forhold», og (2) «tekniske innretninger og fremgangsmåter samt drifts- eller forretningsforhold som det vil være av konkurransemessig betydning å hemmeligholde av hensyn til den som opplysningen angår». Etter § 13 d kan departementet imidlertid bestemme at et forvaltningsorgan kan eller skal gi opplysninger til bruk for forskning uten hinder av organets taushetsplikt etter § 13, dersom det finnes rimelig og ikke medfører uforholdsmessig ulempe for andre interesser. § 13 e pålegger forskere taushetsplikt om visse opplysninger, blant annet opplysninger undergitt taushetsplikt og opplysninger som er mottatt fra private under taushetsløfte i forbindelse med forskningsarbeidet.

Taushetsplikten etter fvl. § 13 er ikke til hinder for at opplysninger gjøres kjent dersom den som har krav på taushet samtykker, jf.

[41] Ibid., note 3-7.
[42] Prop. 56 LS (2017-2018) s. 100-101.
[43] Norsk Lovkommentar til personopplysningsloven, note 8-7.

fvl. § 13 a nr. 1. Taushetsplikten er heller ikke til hinder for at opplysningene brukes når behovet for beskyttelse ivaretas ved at de gis i statistisk form eller ved at individualiserende kjennetegn utelates på annen måte, jf. § 13 a nr. 2.

Brudd på lovpålagt taushetsplikt er som nevnt straffbart, jf. strl. § 209. Dette omfatter naturligvis også brudd på taushetsplikt i forbindelse med forskning, så langt det ikke finnes et rettslig grunnlag for å unnta taushetsplikt. Det finnes en rekke andre lover som pålegger taushetsplikt, herunder helsepersonelloven § 21, pasientjournalloven § 15, helseregisterloven § 17, spesialisthelsetjenesteloven § 6-1 og helse- og omsorgstjenesteloven § 12-1. Også sikkerhetsloven har regler om taushetsplikt, se nærmere nedenfor.

4.3 Begrensninger som skal beskytte andre interesser

4.3.1 Lov om vern av forretningshemmeligheter

Lov om vern av forretningshemmeligheter skal sikre innehavere av forretningshemmeligheter vern mot urettmessig tilegnelse, bruk og formidling av hemmeligheten, jf. forretningshemmelighetsloven § 1. «Forretningshemmeligheter» er definert i lovens § 2, og omfatter opplysninger som (a) er «hemmelige», dvs. at de ikke er allment kjent eller lett tilgjengelig, (b) har «kommersiell verdi fordi de er hemmelige», og (c) innehaveren har «truffet rimelige tiltak for å holde hemmelige». Når det gjelder kravet til at opplysningene «har kommersiell verdi fordi de er hemmelige» omfatter dette både faktisk og potensiell verdi[44]. Etter forarbeidene har opplysninger kommersiell verdi der inngrep vil kunne skade interessene til innehaveren av forretningshemmeligheten «på en måte som undergraver personens vitenskapelige og tekniske potensiale, forretningsmessige eller finansielle interesser, strategiske plasseringer eller evne til å konkurrere»[45].

Lovens § 3 oppstiller et forbud mot å gjøre «inngrep i en forretningshemmelighet». Slikt inngrep kan bestå i å «oppnå kunnskap om eller rådighet over en forretningshemmelighet» gjennom å urettmessig skaffe seg adgang eller gjennom annen adferd «i strid med god forretningsskikk» (§ 3 første ledd). Inngrep kan også bestå i å urettmessig «bruke eller formidle en forretningshemmelighet» man har fått kunnskap om eller rådighet over (a) i strid med § 3 første ledd, (b) i anledning tjenesteforhold e.l., eller (c) i medhold av lovbestemmelse (§ 3 andre ledd). Inngrep i forretningshemmelighet kan resultere i forbud etter § 5, pålegg om korrigerende og forebyggende tiltak etter § 6, fortsatt bruk mot rimelig vederlag etter § 7, vederlag og erstatning etter § 8, og straff etter § 9 og § 10.

Dersom man i innhenting av forskningsdata eller i publisering av forskningsresultater gjør inngrep i en forretningshemmelighet kan dette være i strid med forretningshemmelighetsloven, og således resultere i sanksjoner etter lovens §§ 5 til 10.

4.3.2 Sikkerhetsloven

Sikkerhetsloven skal bidra til å trygge nasjonale sikkerhetsinteresser og forebygge sikkerhetstruende virksomhet, jf. sikkerhetsloven § 1-1. Loven gjelder for statlige, fylkeskommunale og kommunale organer, jf. § 1-2. Nasjonale sikkerhetsinteresser omfatter overordnede sikkerhetspolitiske interesser knyttet til blant annet de øverste statsorganers virksomhet, forsvar, sikkerhet og beredskap, forholdet til andre stater, økonomisk stabilitet, og samfunnets grunnleggende funksjonalitet, samt befolknings grunnleggende sikkerhet, jf. § 1-5. Departementet er pålagt ansvar og myndighet for forebyggende sikkerhetsarbeid etter § 2-1, og sikkerhetsmyndighetens ansvar for forebyggende sikkerhetsarbeid er regulert i § 2-2. Sikkerhetsmyndigheten fører tilsyn med virksomheter som er omfattet av loven, jf. § 3-1. Sikkerhetsloven inneholder reguleringer av informasjonssikkerhet og informasjonssystemsikkerhet (kapittel 5 og 6), herunder regler om beskyttelse av skjermingsverdig informasjon (§ 5-2), sikkerhetsgradert informasjon (§ 5-3) og taushetsplikt (§ 5-4). Loven regulerer også objekt- og infrastruktursikkerhet (kapittel 7), personellsikkerhet (kapittel 8), sikkerhetsgraderte anskaffelser (kapittel 9) og eierskapskontroll (kapittel 10). Brudd på sikkerhetsloven kan medføre tvangsmulkt (§ 11-2), overtredelsesgebyr (§ 11-3) og straff (§ 11-4). I forbindelse med forskning som omfatter informasjon som kan være vernet etter sikkerhetsloven, kan bestemmelsene i sikkerhetsloven sette begrensninger i tilknytning til innhenting og publisering av forskningsdata og forskningsresultater.

[44] Norsk Lovkommentar til forretningshemmelighetsloven, Rettsdata, ved Tore Lunde, note 14.
[45] Prop. 5 LS (2019-2020), punkt 5.1.3.

Appendix 2

The Open Data Directive in relation to publicly funded research data (in Norwegian only)

Photo: Shutterstock

Skrevet av:
Heather Broomfield, seniorrådgiver i Digitaliseringsdirektoratet og stipendiat ved institutt for offentlig rett (UiO) og Espen Dennis Kristoffersen, utredningsleder i Kommunal- og moderniseringsdepartementet

INNLEDNING

Dokumentet omhandler åpne data-direktivet (EU) 2019/1024[1] i relasjon til forskningsdata. Det er også sektorspesifikke initiativ knyttet til forskningsdata, men det er utenfor formålet med dokumentet. Dokumentet omhandler heller ikke EU-kommisjonens forslag til datastyringsforordning (COM (2020) 767 final), som muligens vil ha relevans i relasjon til forskningsdata. Dette fordi forslaget fortsatt er under bearbeidelse. Åpne data-direktivet er en omarbeidelse av det tidligere viderebruksdirektivet (2003/98/EF) med endringsdirektiv (2013/37/EU). Det innebærer at åpne data-direktivet fortsatt inneholder mye fra de tidligere direktivene, men også at det er en del ny regulering. Det viktigste med åpne data-direktivet i denne sammenheng er at det utvider virkeområdet til, på visse vilkår, å omfatte offentlig finansierte forskningsdata.

KORT INTRODUKSJON TIL ÅPNE DATA-DIREKTIVET

For sammenhengens skyld kan det nevnes at selv om EU-initiativet for deling av offentlig sektorinformasjon (PSI) kan spores tilbake til 1970-tallet (Bates, 2012), så kom den første lovgivningen først ved viderebruksdirektivet i 2003, endret i 2013. Åpne data-direktivet fra 2019, som erstatter begge de to foregående direktivene, var begrunnet i ønsket om fullt ut å kunne utnytte potensialet i informasjon fra offentlig sektor. Hovedfokuset i nytt direktiv er på:

- levering av sanntidstilgang til dynamiske data gjennom tilstrekkelige tekniske midler,
- økt utlevering av verdifulle offentlige data til viderebruk, herunder fra offentlige foretak, forskningsorganisasjoner og forskningsfinansierende organisasjoner,
- håndtering av framveksten av nye former for avtaler om enerett,
- bruk av unntak fra prinsippet om å pålegge gebyrer som tilsvarer marginalkostnaden, og
- forholdet til beslektede rettsakter som f.eks. personvernforordningen[2] og databasedirektivet[3].

Direktivet er et minimumsdirektiv og oppfordrer medlemslandene til å gjøre mer enn det som følger av direktivet mht. å gjøre informasjon (data) tilgjengelig for viderebruk. Sentralt er å fjerne hindringer og innføre ensartede regler om sentrale spørsmål som bl.a. berører vilkår for viderebruk, likebehandling og konkurranseregler, som igjen omhandler alt fra formater, gebyrprinsipper, standardlisenser og enerettsavtaler til særregler om forskningsdata.

Åpne data-direktivet innførte ingen formell definisjon av hva åpne data er, men i fortalens punkt 16 fremgår det at (uoffisiell oversettelse4):

"Åpne data som begrep forstås vanligvis som data i et åpent format som kan brukes fritt, brukes på nytt og deles av hvem som helst for et hvilket som helst formål."

Informasjon som ikke er å forstå som åpne data er bl.a. data der det foreligger tredjepartsrettigheter og andre juridiske begrensinger for deling. Dette kan eksempelvis være knyttet til immaterielle rettigheter eller persondata. Et dokument som i utgangspunktet ikke er å betrakte som åpne data, kan bli klassifisert som åpne data, eksempelvis gjennom anonymisering av persondata. Direktivet gjør det likevel klart at det ikke påligger en plikt for de som sitter med offentlige data til å fremstille eller tilpasse dokumenter eller fremlegge utdrag av dokumenter ut fra brukerbehov, med mindre det kun innebærer enkle operasjoner, jf. art. 5 nr. 3. Dette gjelder også offentlige finansierte forskningsdata.

UTVIDELSE AV DIREKTIVETS VIRKEOMRÅDE TIL Å OMFATTE OFFENTLIG FINANSIERT FORSKNING

Deling og viderebruk av forskningsdata er tidligere ikke regulert i EU-lovgivningen. Begrunnelsen bak utvidelsen av direktivets virkeområde i artikkel 1 nr. 1 bokstav c) til å omfatte offentlig finansierte forskningsdata i samsvar med vilkårene i artikkel 10 nr. 2, er at de allerede er betalt for av fellesskapet og at dataene derfor bør gjøres tilgjengelig i størst mulig grad for å bidra til ny forskning og nye innovasjoner. Åpen tilgangspolitikk sikter særlig mot å gi forskere og allmenheten tilgang til forskningsdata så tidlig som mulig i formidlingsprosessen, samt å forenkle bruk og viderebruk, jf. fortalens punkt 27.

Samtidig ligger det i de ytterligere vilkårene som må være oppfylt for at offentlig finansierte forskningsdata skal omfattes av direktivet også en erkjennelse om at slike data skiller seg noe fra f.eks. administrative data, jf. art. 10 nr. 2, bl.a. ved at offentlig finansierte forskningsdata først omfattes dersom de allerede er offentliggjort gjennom et sentralt datalager (engelsk: repository) i institusjonen eller et emnebasert datalager.

Art. 10 er den sentrale bestemmelsen for å fastslå om forskningsdata er omfattet eller ikke. Art. 10 nr. 1 omhandler krav til landenes politikk og tiltak for tilgang til offentlig finansierte forskningsdata, mens art. 10 nr. 2 mer konkret omhandler når slike forskningsdata er omfattet av reguleringen og kan viderebrukes.

[1] Directive on Open Data and Public Sector Information ("the Open Data Directive")
[2] Europaparlaments- og rådsforordning (EU) 2016/679 av 27. april 2016 om vern av fysiske personer i forbindelse med behandling av personopplysninger og om fri utveksling av slike opplysninger samt om oppheving av direktiv 95/46/EF (generell personvernforordning) (EUT L 119 av 4.5.2016, s. 1).
[3] Europaparlaments- og rådsdirektiv 96/9/EF av 11. mars 1996 om rettslig vern av databaser (EFT L 77 av 27.3.1996, s. 20).
[4] Gjengivelse av Åpne data-direktivet på norsk i dette dokumentet baserer seg på en uoffisiell oversettelse foretatt av Utenriksdepartementet

Definisjon av forskningsdata
Forskningsdata er definert i art. 2 nr. 9:

"«forskningsdata» dokumenter i digital form, unntatt vitenskapelige publikasjoner, som er innsamlet eller produsert ved vitenskapelig forskningsvirksomhet, og som brukes som dokumentasjon i forskningsprosessen, eller som er allment akseptert i forskningsmiljøet som nødvendig for å validere forskningsfunn og -resultater,"

Av fortalens punkt 27 følger det at forskningsdata omfatter statistikk, forsøksresultater, målinger, observasjoner fra feltarbeid, undersøkelsesresultater, intervjuopptak eller bilder, men også metadata, spesifikasjoner og andre digitale objekter. Vitenskapelige artikler redegjør for og kommenterer resultater av vitenskapelig forskning, og er således noe annet enn data som fremkommer ved forskningsvirksomheten, og derfor unntatt. Andre publikasjoner enn vitenskapelige publikasjoner vil også være unntatt i den grad de kommenterer resultater av forskningen.

Artikkel 10 nr. 1 – nasjonal politikk og tiltak for offentlig finansierte forskningsdata
Det følger av artikkel 10 nr. 1 at:

"Medlemsstatene skal støtte tilgangen til forskningsdata ved å vedta nasjonal politikk og relevante tiltak med sikte på å gjøre data fra offentlig finansiert forskning åpent tilgjengelige (i tråd med retningslinjer for åpen tilgang), etter prinsippet om «åpenhet som standard» og i samsvar med FAIR-prinsippene. I denne sammenheng bør betenkeligheter med hensyn til immaterialrettigheter, vern av personopplysninger og fortrolighet, sikkerhet og legitime forretningsinteresser tas hensyn til, i samsvar med prinsippet «så åpent som mulig, så lukket som nødvendig». Disse retningslinjene for åpen tilgang skal være rettet mot forskningsutøvende organisasjoner og forskningsfinansierende organisasjoner."

Direktivet forplikter altså landene til å vedta nasjonal politikk, eksempelvis retningslinjer, og relevante tiltak som støtter opp om tilgangen til offentlig finansierte forskningsdata[5]. Formålet er å gjøre offentlig finansiert forskning åpent tilgjengelig, basert på retningslinjene for åpen tilgang. Tilnærmingen er basert på prinsippet om «åpenhet som standardinnstilling», jf. også fortalens punkt 16. Eventuelle begrensninger på tilgang til dataene bør således begrunnes, eksempelvis i offentlig sikkerhet eller i vern av personopplysninger. Dette i motsetning til at dataene i utgangspunktet er lukket og at en deretter vurderer åpenhet. FAIR-prinsippene er et akronym for "findable, accessible, interoperable og reusable", og er oversatt til søkbare, tilgjengelige, interoperable og gjenbrukbare. Det siktes her til at en bør forsøke å sikre at "planlegging av databehandling blir standard vitenskapelig praksis", jf. fortalens punkt 27.

Retningslinjene for åpen tilgang retter seg for øvrig ikke mot enhver forskningsorganisasjon, men mot forskningsutøvende- og forskningsfinansierende organisasjoner.

Med åpen tilgang siktes det til nettbasert og vederlagsfri tilgang uten andre begrensninger på bruk og viderebruk enn retten til å kunne kreve anerkjennelse av forfatterskapet, altså en kreditering av forfatteren(e), jf. fortalens punkt 27. Dette innebærer ikke at artikkel 8 om standardlisenser ikke gjelder for offentlig finansierte forskningsdata. Også for slike forskningsdata vil det være mulig å stille objektive og rimelige vilkår så lenge de er begrunnet i mål av allmenn interesse og ikke medfører forskjellsbehandling, se nærmere direktivets kap. 4.

Landene har forskjellige utgangspunkter for hvor langt de er kommet mht. åpen tilgangspolitikk og det å implementere prinsippet om åpenhet som standard (open by default) i samsvar med FAIR-prinsippene. Bestemmelsen angir en retning for politikken og tiltakene som gjennomføres, jf. fortalens punkt 16 om at medlemsstatene oppfordres til å fremme generering av data basert på de nevnte prinsippene. Den konkrete tilnærmingen vil det imidlertid være opp til landene å bestemme, forutsatt at de følger den generelle veiledningen som følger av direktivet.

Det følger for øvrig ikke av direktivet et tidspunkt for når landene skal ha nådd formålet om åpen tilgang og prinsippet om innebygd åpenhet som standardinnstilling. En må likevel forvente at EU vil følge opp med innstramninger gjennom nytt regelverk dersom overgangen går for sent og/eller det blir "stort strekk i laget".

Av artikkel 10 nr. 1 andre punktum følger det at eventuelle betenkeligheter med hensyn til immaterialrettigheter, vern av personopplysninger og fortrolighet, sikkerhet og legitime forretningsinteresser skal hensyntas i samsvar med prinsippet "så åpent som mulig, så lukket som nødvendig". Setningen er kun en klargjøring av det som allerede vil følge av artikkel 1 nr. 2, nr. 4 og nr. 5, og setningen ble tatt inn på forespørsel fra flere av medlemslandene og enkelte forskningsorganisasjoner.

Det bør derfor understrekes at direktivet for eksempel ikke berører det vern av personopplysninger som bl.a. følger av personvernforordningen (GDPR)6, og som innebærer at eventuell videre-bruk må være i tråd med formålsbegrensningene som er fastsatt der (artikkel 5 nr. 1 bokstav b) og artikkel 6). Artikkel 10 nr. 1 andre punktum skal således ikke forstås som en innskrenkning av vernet, og en kan ikke gå utover den fleksibiliteten som personvernforordningen og andre relevante rettsakter ev. gir.

[5] Norge har i motsetning til mange andre land allerede en nasjonal politikk på området. Det er imidlertid utenfor formålet med dokumentet å vurdere om norsk politikk er i tråd med direktivet eller ikke. Forslaget til forordning er til behandling i EU og videre prosess vil kunne medføre endringer. https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-tilgjengeliggjoring-og-deling-av-forskningsdata/id2582412/
[6] Europaparlaments- og rådsforordning (EU) 2016/679 av 27. april 2016 om vern av fysiske personer i forbindelse med behandling av personopplysninger

Artikkel 10 nr. 2 – viderebruk av offentlig finansiert forskning

Artikkel 10 nr. 2 omhandler mer konkret tilfeller der offentlig finansierte forskningsdata er omfattet og kan viderebrukes i tråd med direktivet:

"Uten at det berører artikkel 1 nr. 2 bokstav c) skal forskningsdata kunne viderebrukes for kommersielle eller ikke-kommersielle formål i samsvar med kapittel III og IV, i den grad de er offentlig finansiert og forskere, forskningsutøvende organisasjoner eller forskningsfinansierende organisasjoner allerede har offentliggjort dem gjennom et sentralt datalager i institusjonen eller et emnebasert datalager. I den forbindelse skal det tas hensyn til rettslige forretningsinteresser, kunnskapsoverføring og allerede eksisterende immaterialrettigheter."

Et vilkår for viderebruk er at dataene allerede må være offentliggjort i institusjonelle- eller emnebaserte datalagre av forskere, forskningsutøvende organisasjoner eller forskningsfinansierende organisasjoner for å være omfattet av direktivet. Før en slik offentliggjøring er ikke forskningsdataene omfattet av direktivet.

Direktivet er tydelig på at viderebruken både omfatter kommersiell og ikke-kommersiell bruk. Det er således viderebrukeren som har råderett over hva de offentlig finansierte forskningsdataene brukes til når tilgang først er gitt, så lenge det er i samsvar med kapittel III om vilkår for viderebruk og kapittel IV om likebehandling og konkurranseregler – se her kap. 4 om andre bestemmelser av betydning for forskningsdata. Dette følger for øvrig også av artikkel 8 nr. 1 om adgangen til å sette vilkår.

Direktivet er et minimumsdirektiv og landene er dermed gitt en rett til å utvide anvendelsen av direktivet til også å omfatte forskningsdata som gjøres offentlig tilgjengelig gjennom annen datainfrastruktur enn datalagre, eksempelvis gjennom åpent tilgjengelige publikasjoner, eller som vedlegg til en artikkel eller et dokument i en datajournal, jf. fortalens punkt 28.

Direktivet gir ingen brukerveiledning mht. når forskningen er å anse som offentlig finansiert. Det kan forstås som at landene er gitt en viss fleksibilitet til å innfortolke et minstenivå på finansieringen, men dette trekker i så fall i motsatt retning av målet om mer harmonisering av landenes regler og praksis, som er et sentralt aspekt ved direktivet. På bakgrunn av de ytterligere kvalifiserende vilkårene som må oppfylles før offentlige finansierte forskningsdata omfattes av direktivet, for eksempel det at dataene allerede må være offentliggjort, trekker også i retning av at landene har et begrenset skjønn til å innfortolke et "nivå på finanseringen". Offentlige bidrag utover det "ubetydelige" bør derfor som utgangspunkt antas å være omfattet av offentlig finansierte forskningsdata.

Det kan i denne sammenheng nevnes at "***Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata***" refererer til offentlig finansierte forskningsdata som:

"data som er samlet inn eller frembrakt til bruk for eller som et resultat av offentlig finansiert forskning, og

data som utgjør grunnlaget for publikasjoner som er et resultat av offentlig finansiert forskning, uavhengig av hvilken kilde dataene kommer fra."[7]

Det er ikke et krav at forskningsutøvende organisasjoner og forskningsfinansierende organisasjoner må være offentlige. Slike organisasjoner kan være organisert både som offentlige organer og offentlige foretak, og det er presisert i fortalens punkt 28 at direktivet får anvendelse på slike hybridorganisasjoner, men da bare i deres egenskap som forskningsutøvende organisasjoner og på deres forskningsdata.

Det følger videre av fortalens punkt 28 at visse forpliktelser som følger av direktivet, bør utvides til å omfatte forskningsdata som kommer fra vitenskapelig forskningsaktivitet subsidiert av offentlig finansiering eller samfinansiert av offentlig og private sektor. Dette er ikke et krav, og det er derfor opp til landene å beslutte en slik utvidelse.

Det bør legges til grunn at direktivet mht. subsidiering ikke sikter til en generell skattemessig avskrivning som det Skattefunnordningen representerer.

Direktivet sier ingenting om hvilke forpliktelser som ev. også bør gjelde for data der forskningen er samfinansiert av offentlige og private aktører. Det vil her være opp til landene å beslutte dette på nasjonalt nivå. Slike forpliktelser vil uansett være begrenset til de forpliktelser som følger av direktivet, som eksempelvis innebærer at forskningsdata ikke vil omfattes av direktivet med mindre de allerede er offentliggjort. Forskere og de som har rettigheter til forskningen har således fortsatt kontroll på når forskningen ev. skal omfattes av direktivets regler, og også her er vitenskapelige artikler unntatt da dette ikke er å anse som data som fremkommer ved forskningsvirksomheten. Det kan for øvrig nevnes at mange tidsskrifter krever at forskningsdataene som ligger til grunn for publikasjon skal offentliggjøres, og at dataene dermed også blir tilgjengelige. Et annet overordnet hensyn ved direktivet er å ikke etablere regler som vil motvirke selskaper fra å investere i offentlig-private partnerskap. Medlemsstatene bør altså finne en balanse mellom behovet for at data skal være åpne og de økonomiske realitetene.

[7] Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata (regjeringen.no)

ANDRE BESTEMMELSER ENN ARTIKKEL 10

Vilkår for viderebruk – kapittel III
Artikkel 5 omhandler formater som dataene (dokumentene) skal gjøres tilgjengelig i, og omfatter også språkversjoner. Disse bestemmelsene gjelder i utgangspunktet også for forskningsdata. Nytt i bestemmelsen er særlig reglene om dynamiske data og datasett med høy verdi. Hvilke datasett som omfattes av datasett med høy verdi er foreløpig ikke avklart gjennom en gjennomføringsrettsakt (ventes andre halvår 2021), men det er tvilsomt om dette vil være særlig relevant for forskningsdata da en her sikter til særlig verdifulle og etablerte datasett som allerede er tilgjengelig og i bruk, eksempelvis en del geografiske data og meteorologiske data. Heller ikke reguleringen av dynamiske datasett har forskningsdata spesifikt i sikte. Det betyr ikke at dynamiske forskningsdata ikke vil kunne omfattes. Om forskningen har bidratt til at slike datastrømmer tas i bruk i offentlig sektor etter at forskningen er avsluttet, så vil viderebruk normalt vurderes etter direktivets "standardregler" og ikke etter særreglene for forskningsdata.

Artikkel 6 omhandler gebyrer, og det følger av nr. 6 bokstav b) at offentlig finansierte forskningsdata i samsvar med vilkårene i artikkel 10 skal være vederlagsfrie, jf. artikkel 1 nr. 1 bokstav c). Her må det imidlertid legges til at det er først når vilkårene etter artikkel 10 er oppfylt at slike forskningsdata i det hele tatt er omfattet av direktivet. Det innebærer at slike forskningsdata allerede er offentliggjort, og da enten av relevante forskere, forskningsutøvende organisasjoner eller av forskningsfinansierende organisasjoner gjennom et sentralt datalager i institusjonen eller et emnebasert datalager.

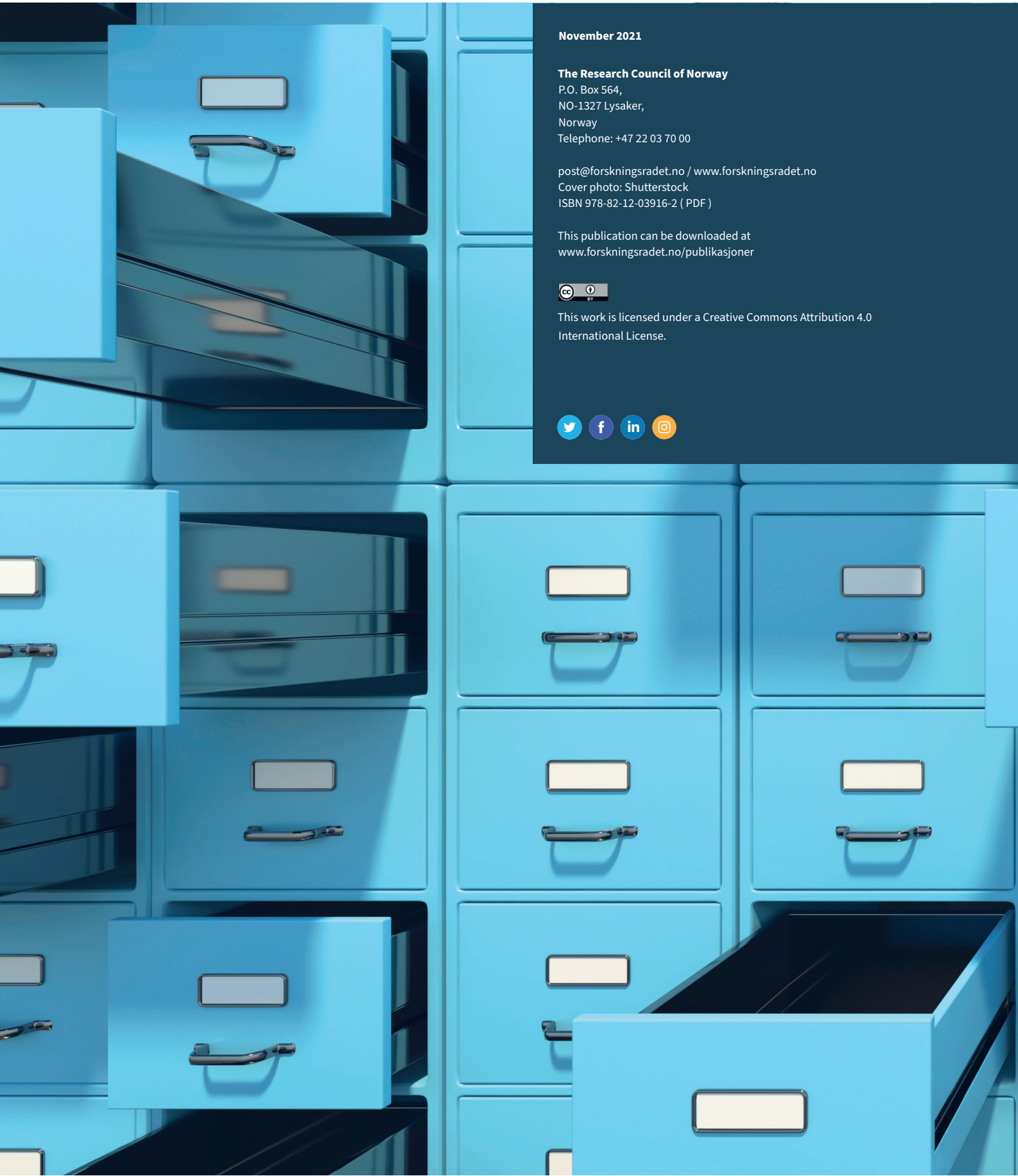
Artikkel 8 om standardlisenser, definert i artikkel 2 nr. 5, stadfester prinsippet om at viderebruk ikke skal være underlagt vilkår. Det kan likevel settes vilkår dersom de er *"objektive, rimelige, ikke medfører forskjellsbehandling og er begrunnet med hensynet til et mål av allmenn interesse"*. Utgangspunktet er likevel at det skal stilles så få vilkår som mulig. Innen de angitte rammene kan det stilles vilkår for viderebruk mht. ansvar, vern av personopplysninger og riktig bruk av dokumenter, garanti for at de ikke endres og kildeangivelse, jf. fortalens punkt 44. Direktivet oppfordrer landene til å sørge for at standardlisenser er tilgjengelige, og oppmuntre til at de blir brukt. Offentlige organer er imidlertid ikke forpliktet til å bruke dem, med mindre landene gjør dem obligatoriske. Her vil det for forskningsdata, som for andre data, være viktig å etablere klare anbefalinger om hvilke lisenser som bør brukes til ulike datatyper slik at man bl.a. unngår utfordringer knyttet til interoperabilitet.

Artikkel 9 om praktiske ordninger oppfordrer landene til å fortsette arbeidet med bedre tilrettelegging for viderebruk, og videre at landene skal jobbe sammen med EU-kommisjonen om å lette tilgangen til datasett og å etablere et felles tilgangspunkt. Bestemmelsen skiller ikke mellom typer av datasett og omfatter derfor også offentlige finansierte forskningsdata. Selv om bestemmelsen i seg selv ikke avgrenses mot forskningsdata, så er det ikke slike data bestemmelsen først og fremst sikter til.

Likebehandling og konkurranseregler – kapittel IV
Som nevnt over gjelder direktivets kapittel IV også forskningsdata. Dette er i stor grad bestemmelser som også gjaldt under de tidligere viderebruksdirektivene, som åpne data-direktivet erstatter. Det nye er at praktiske ordninger, som har som mål eller som med rimelighet kan forventes å føre til begrenset tilgjengelighet for viderebruk, nå likestilles med enerettsavtaler. Dette gjelder også for forskningsdata regulert i artikkel 10.

Særlig om forholdet til databasedirektivet
Det følger av artikkel 1 nr. 6 at der databaser faller inn under direktivets virkeområde, så skal ikke offentlige organer utøve sin rett som databaseprodusent til å hindre eller begrense viderebruk utover de grensene som er fastsatt i direktivet. Dette er en klargjøring av rettstilstanden, og ikke en nyvinning. Innføringen er et resultat av at det i EU-området har vært innrapportert tilfeller hvor offentlige organer nettopp har hindret eller begrenset en slik viderebruk med hjemmel i databasedirektivets artikkel 7 nr. 1.

Utdanningsinstitusjoner, forskningsutøvende- og forskningsfinansierende organisasjoner er for øvrig ikke forpliktet til å følge saksbehandlingsreglene om forespørsel om tilgang til data i artikkel 4. De er imidlertid fortsatt forpliktet til å ikke utøve sine eventuelle rettigheter som databaseprodusent til å hindre eller begrense viderebruk utover direktivets grenser.



November 2021

The Research Council of Norway
P.O. Box 564,
NO-1327 Lysaker,
Norway
Telephone: +47 22 03 70 00

post@forskningsradet.no / www.forskningsradet.no
Cover photo: Shutterstock
ISBN 978-82-12-03916-2 (PDF)

This publication can be downloaded at
www.forskningsradet.no/publikasjoner