

# DARE TO SHARE

OPEN ACCESS AND DATA SHARING IN SCIENCE



The Advisory council for science, technology and innovation (AWTI) advises the Dutch government and parliament on policy in the areas of scientific research, technological development and innovation. The AWTI provides advice when it is applied for or when the council deems it necessary. It enjoys an independent position towards the Ministers and their Departments, as well as towards other parties involved. The advisory council consists of a maximum of 10 members, each originating from different sectors of society, such as research institutes and trade and industry. The members do not represent any special interests.

The Advisory council on science, technology and innovation consists of the following members:

prof. dr. U. Rosenthal (voorzitter)

prof. dr. ing. D.H.A. Blank

mrs. ing. T.E. Bodewes (until September 1, 2015)

mrs. prof. dr. R. Cools

mrs. prof. dr. V.A. Frissen

prof.dr. ir. T.H.J.J. van der Hagen

prof. dr. E.M. Meijer

dr. ir. A.J.H.M. Peels

prof.dr. ir. M.F.H. Schuurmans

prof. dr. L.L.G. Soete

mrs. dr. D.J.M. Corbey (secretaris)

The office is located in The Hague, the Netherlands:

Javastraat 42

2585 AP The Hague

The Netherlands

t. +31(0)70 31 10 920

e. [secretariaat@awti.nl](mailto:secretariaat@awti.nl)

w. [www.awti.nl](http://www.awti.nl)

# Dare to share

Open access and data sharing in science

January 2016

Photography      Shutterstock.com

Design            2D3D Design, The Hague

January 2016

All publications may be downloaded free of charge from [www.awti.nl](http://www.awti.nl).

## Copyright

All rights reserved. Provided the correct source reference is used, this publication or parts thereof may be reproduced, stored on a retrieval system or made public without requiring the prior written permission of AWTI. A correct source reference must as a minimum include a clear statement of the name of the organisation and the name and year of the publication.

# Contents

<b>Summary</b>	<b>5</b>
<b>Part 1: Advice</b>	
<b>1. Introduction: <i>open science</i></b>	<b>11</b>
1.1 Background	11
1.2 Request for advice	13
1.3 Urgency	13
1.4 Scope and focus	15
1.5 Approach and structure of this report	16
<b>2. Findings</b>	<b>19</b>
2.1 Open science	19
2.2 Access to publications: <i>open access?</i>	19
2.3 Sharing research data	23
2.4 Role of government	27
<b>3. Recommendations</b>	<b>29</b>
<b>Part 2: Analysis</b>	
<b>4. Open access</b>	<b>35</b>
4.1 Between dream and reality...	35
4.2 Scholarly publishing: functions and stakeholders	41
4.3 Public authorities and open access	50
4.4 The impact of open access	57
4.5 Conclusion	65
<b>5. Sharing research data</b>	<b>69</b>
5.1 Retain or share?	69
5.2 Research data: functions and actors	72
5.3 Policy on sharing research data	83
5.4 The impact of (more) sharing of research data	87
5.5 Conclusion	97
<b>Appendix 1 Request for advice</b>	<b>103</b>
<b>Appendix 2 Open access and open data in different countries</b>	<b>105</b>
<b>Appendix 3 Interviewees</b>	<b>121</b>
<b>Appendix 4 References</b>	<b>125</b>





## Summary

This report is about *open science*, and more specifically about access to scholarly publications (*open access*) and research data (*open research data*). What impact is this likely to have for the world of science itself, for society and for business? What level of openness is publicly desirable and what does this imply for government policy?

AWTI is a strong advocate of making science more open and accessible, and shares the belief that this will strengthen the position of science in society. Ensuring that the transition towards open science is successful will require a broad and powerful strategy which goes beyond merely improving access.

### **Open access**

Traditionally, the 'reader' has paid for access to scholarly publications. Under a system of open access, everyone can access such articles freely. The publication costs are then borne by other parties (e.g. the authors). The impact of open access on science, at least in the Netherlands, is however likely to be limited as most researchers already have good access to scholarly publications. Open access could potentially have a larger impact on society and business, which currently have poor access. However, this impact can only be achieved if open access is embedded into a broader strategy of 'translating' scientific results and presenting them in a way that is clear for lay people, and at the same time ensuring that the 'social impact' of research is appreciated more within the world of science. Only then will open access contribute to better utilisation of scientific knowledge.

A 'bottom-up' development towards open access has been under way for some time, though it appears that the growth of open access has come to a standstill at present. For the majority of researchers, the reputation of a journal is more important than whether or not a publication offers open access. We found that the organisations funding research, in particular, have the power to enforce open access. Moreover, if governments were to take the lead here, this would help give direction to stakeholders and speed up the process. Open access to scholarly publications in itself does *not* reduce the perceived high costs of scholarly publishing; the added value of open access lies mainly in making scientific knowledge more widely available.

### **Sharing research data**

The sharing of research data is another story, with wide differences across the various scientific disciplines. In some disciplines storing and sharing research data is the norm already, but in most disciplines such storing and sharing rarely or never takes place. The quality and efficiency of science can be improved through good research data management and sharing research data with third parties. Society and business will

undoubtedly benefit from access to research data, though the extent will depend greatly on the nature of the data. A number of conditions are crucial for the successful storage and sharing of research data, such as a good storage infrastructure, a set of standards and clarity about legal aspects (such as privacy and data ownership). Sharing research data also only makes sense if the data are findable, accessible, interoperable and reusable.

It is not advisable simply to open up all research data to everyone unconditionally. There may be good reasons for not opening certain research data, such as privacy or commercial interests. Moreover, support for allowing access to data will increase if researchers and their institutions are given (some) control over who they share their research data with and for what purpose. Storing and sharing research data on a wider scale will initially cost more; the anticipated benefits are promising but uncertain, and it therefore makes sense to set priorities as to which data might be made available.

### **Recommendations**

These conclusions lead to a number of recommendations to the Dutch Minister and State Secretary of Education, Culture and Science and the Minister of Economic Affairs. They are summarised briefly below and explained in more detail in the body of the report.

#### **Recommendation 1:**

##### **Embrace the importance of openness and formulate a broader strategy for better utilisation of knowledge**

Acknowledge and embrace the importance of openness as a core value in science, and especially for publicly funded research.

- a) Do not view open access as an independent goal of science, but as a link in the process of achieving better utilisation of knowledge, within and, above all, beyond the field of science. Only as part of such a broader strategy which increases the accessibility of scientific knowledge does open access genuinely add value.
- b) Make a substantial and concerted effort to let research data be shared more widely, and aim to optimise access to research data.

#### **Recommendation 2:**

##### **Strive for effective implementation of the principles of open science in the Netherlands**

Formulate a masterplan for the Netherlands aimed at the better utilisation of scientific knowledge, with attention for the role of open access and the sharing of research data. Set aside sufficient resources for this and ensure that the importance of the 'impact on society and business' is properly anchored in research practice.



- A. Ensure that the masterplan for open access:
  - a) contains clear milestones, preconditions and, where necessary, infrastructure. Allow for differences in elaboration and actual implementation in different scientific disciplines. Use this process to showcase the results of Dutch science;
  - b) makes scientific results more accessible to society by making them findable and presenting them clearly and in a way that can be understood by a wider public;
  - c) provides for better access to scientific knowledge for business.
  
- B. Ensure that the masterplan for sharing research data:
  - a) contains clear milestones and preconditions and allows differences in elaboration and actual implementation in different scientific disciplines;
  - b) provides for general facilities for data storage and sharing;
  - c) promotes good data management by embedding data management in science training and also by rewarding good data management;
  - d) contains a clear assessment framework for deciding which type of research data should be stored at public expense;
  - e) as a minimum, provides for access to the research data that underlie a scholarly article, as long as there are no compelling reasons for not doing so;
  - f) for the rest, encourages the sharing of research data in general where this is possible, receives support and serves a useful purpose;
  - g) supports the sharing of data, preferably by rewarding it rather than through imposition;
  - h) properly addresses legal issues.

### **Recommendation 3:**

#### **Collaborate on open science at EU level**

As far as possible, ensure that Dutch policy on open access to scholarly publications and research data is in line with EU policy. Aim for an accelerated transition to open access (to publications) and research data sharing, and advocate a common strategy to achieve this. The Dutch government can then place the following items on the agenda at EU level (including during the Dutch EU presidency):

- a) the economic, social and scientific importance of sharing research data and making scientific knowledge accessible;
- b) the importance of establishing conditions necessary for open science, such as maximum embargo periods, transparency on the costs of publication, and a degree of standardisation;
- c) actively monitoring developments and progress in relation to open access and research data sharing in Member States with a view to convergence and accelerating the process;

- d) genuine accessibility of the results and research data for those projects that are (co-)funded by the EU, by ensuring open access but also by providing accessible summaries and introductions for lay readers;
- e) the creation of a European Science Cloud to serve both as a hub and as a means of publicising European scientific achievements.



**Part 1: Advice**



# Introduction: *open science*

## 1.1 Background

Globalisation and digitalisation are changing the way in which science is organised and functions. Researchers are working in new ways and are increasingly collaborating internationally; the relationship between scientists and society is changing; and digitalisation is changing the way in which research is carried out in practice. A movement is under way to make the scientific process more transparent and to improve the dissemination of the results of science to society as a whole, including business. This development is often summarised using the term *open science*.

Openness is a key value in scientific culture, and the developments referred to above offer new opportunities for increasing that openness. Openness can benefit science, but undoubtedly can also benefit society. In response to the Ebola outbreak in West Africa, for example, several scientific publishers made relevant articles freely accessible to all, in a bid to help fight the epidemic. Similarly, opting for open collaboration between public research institutes and the private sector in the first phase of the development of new drugs, with open sharing of research data and results, as in the Structural Genomics Consortium,<sup>1</sup> can speed up the launch of new drugs considerably. In addition, there are other, very different databases that can also be accessed, such as the Meertens Institute databases on first names, surnames, dialects, and pilgrimages (in the Netherlands).<sup>2</sup> Such sources are likely to be of interest to researchers, and also to the public at large.

However, there are always limits to scientific openness. There is some knowledge that people prefer not to share, for example because of its strategic military importance, or due to commercial interests, for example where a manufacturer first wishes to apply for a patent or develop a product. Personal privacy may also be an issue. And sometimes a researcher or institute may wish to defer publishing an article or research data until further research has been carried out. Lack of time and resources can also be a barrier to making research data accessible – something that often requires rather more than a mere mouse-click.

Science has to find a way through this maze of considerations in relation to openness. There is also a broader context, a striking feature of which is the ever-increasing computing power and communication capabilities and the sheer amount of data they produce (*e-science*).<sup>3</sup> Added to this are the ever more complex challenges facing society

---

<sup>1</sup> <http://www.thesgc.org/>

<sup>2</sup> See: <http://www.meertens.knaw.nl>

<sup>3</sup> See also: AWT (2011).

and the shifts in global economic and scientific relations. There is also a continuing debate about how science should be run and what kind of scientific culture we want.

This context begs the question of how scientific knowledge that is generated throughout the world can be used to best effect for the benefit of science, society and business. Good accessibility of such knowledge is generally expected to be a crucial part of this. Accessibility can improve transparency, speed up scientific processes, strengthen the links between disciplines and create new opportunities for research, for example analysis of sets of research data, possibly in combination with other (types of) data. It can also improve the links between science on the one hand and society and industry on the other. That is good for the economy, good for innovation and good for public support for science. As taxpayers, society and businesses fund a considerable share of scientific research; they have everything to gain from optimum use of the knowledge that follows from that research, and will take into account both the benefits and the costs of doing so.

In talking about creating (better) access to scientific knowledge, we are actually talking about *open content* or, more specifically, access to scientific information such as publications (journal articles and books) and research data generated or collected by researchers. The European Commission has high expectations of the increased accessibility of scientific knowledge as a driver of European economic development.<sup>4</sup> The Netherlands, too, has high expectations and, against the backdrop of the 2025 Vision for Science (*Wetenschapsvisie 2025*)<sup>5</sup> and the establishment of a National Science Agenda (*Wetenschapsagenda*),<sup>6</sup> is actively looking at ways to improve the functioning of science and to make the best possible use of scientific knowledge. These topics are one of the core focus areas for the Netherlands during its presidency of the European Union in the first half of 2016. The Dutch government has expressed its desire to encourage a move towards improved ('open') access to scholarly publications and research data.<sup>7</sup> A common approach within Europe would be helpful.

In recent years, great strides have been made in improving the access to and use of scientific knowledge. More scholarly publications can, for example, be freely accessed online, and in a number of disciplines research data are shared with other researchers or the general public. However, this is not always an automatic process: there are many disciplines which share virtually no research data, and a high proportion of scholarly publications are still locked behind a paywall. The scientific community (still) generally regards publishing in a journal with high scientific impact as more important than whether or not a publication is freely accessible. It is also notable that Europe by no means leads

---

<sup>4</sup> Presentation by EU Commissioner Moedas (2015), Brussels, ERA Conference 22 June 2015.

<sup>5</sup> Ministry of Education, Culture and Science (2014); see also Ministry of Education, Culture and Science (2015b).

<sup>6</sup> See: <http://www.wetenschapsagenda.nl/>

<sup>7</sup> See the letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 23 January 2015: TK 2014-2015, 31 288, no. 414, p. 2.

the way globally in making scientific output freely accessible. Brazil, to name but one example, is much further ahead. Moreover, too little scientific knowledge still reaches the public and business sector. This is acknowledged by stakeholders,<sup>8</sup> and the Dutch government has formulated the ambition of strengthening the links between Dutch science, society and industry, and maximising the impact of science.<sup>9</sup>

## 1.2 Request for advice

All in all, there is sufficient reason to investigate whether and how increasing access to scientific knowledge can contribute to the best possible use of that knowledge. That is AWTI's aim in this report. The emphasis is on access to scholarly publications and research data generated by others. If such access is free for all, it is referred to as 'open access' (to publications) and 'open research data', respectively.

This will enable AWTI to answer the central question in the request for advice:<sup>10</sup>  
*What social impact can be expected from the development of open science, and in particular open access to scholarly publications and open research data?*

We are explicitly concerned with the social impact in a broad sense, in other words the impact on science itself, on society and on business. This also is evident from the (possible) constituent questions mentioned in the request for advice:

- ▶ What degree of open science is socially desirable? What are the pros and cons of open access and open research data for society as a whole and for science and industry in particular? Where do the benefits accrue and who will bear the costs?
- ▶ What should the role of the Dutch government be?
- ▶ What can be learned from other countries concerning the social impact of open access and open research data?
- ▶ What role do key third countries play?

This report aims to support the Dutch government in defining its strategy on open science, partly in the light of the Dutch EU presidency in the first half year of 2016.

## 1.3 Urgency

The theme of (open) access to scientific knowledge is high on the political agenda concerning science. For the Netherlands, it is a core focus of the State Secretary for

---

<sup>8</sup> This was evident for example from discussions about trust in science: De Jonge (2014).

<sup>9</sup> Ministry of Education, Culture and Science (2014), p. 11.

<sup>10</sup> See Appendix 1 for the complete request for advice.

Education, Culture and Science, Sander Dekker, and within the European Union open science is one of the three priorities of the responsible EU Commissioner, Carlos Moedas.<sup>11</sup>

A number of developments have made this theme extremely urgent.

First, digitalisation has created the technical conditions to enable publications (e.g. articles and books), research data and other scientific knowledge to be shared on a large scale. Thus, increased access to scientific content is now achievable.

In addition, the costs of granting full access to scholarly publications are rising year on year. Not only because of the sharp rise in the global output of scholarly articles, but also due to the strong market position of the major publishers of scholarly journals. Moreover, as research budgets are under pressure in most European countries due to the economic crisis, it is becoming more and more difficult for European researchers and research institutes to access relevant scholarly literature. This is becoming a real problem.

European science still scores above average in terms of the scientific impact of its articles. But to achieve this, researchers often have to publish in journals to which there is no open access, thus reducing their social impact because society and business cannot access those articles freely. This is at odds with the view that society should be able to access the scientific knowledge for which we have all paid – a view that has gained considerable traction in politics and society in recent years.

This view not only extends to accessibility of publications, but also to (access to) research data obtained with public funding. Sharing such data is expected to lead to more transparent and more efficient science and to less duplication, or to the creation of new interdisciplinary links. This is part of a wider series of developments surrounding big data. In addition, researchers in several disciplines are investigating problems of unprecedented magnitude and complexity, such as studying the climate, human cognition, or the cell. Traditional empirical methods are inadequate for studying such complex phenomena and the use of computer and database technology to process and interpret the data is indispensable.<sup>12</sup> This also holds for creating effective and usable access to documents such as old manuscripts or historical financial data. There are several new opportunities here, and the European Union is keen to be ahead of the curve.<sup>13</sup> That demands action now!

---

<sup>11</sup> Presentation by EU Commissioner Moedas (2015), Brussels, ERA Conference, 22 June 2015.

<sup>12</sup> AWT (2011), p. 5.

<sup>13</sup> The 'connected digital single market' is one of the priorities of the present European Commission; see e.g. the European Commission Communication dated 6 May 2015 on A Digital Single Market Strategy for Europe COM(2015) 192 final.



At the same time, the volume of research data is increasing steadily due to the research activities being carried out worldwide. Substantial financial and other investments are needed to document this data, render it suitable for reuse, store it and make it accessible. This demands an effective strategy in which priorities are set now for the storage and sharing of research data.

## 1.4 Scope and focus

As a European Commission consultation also showed, (open) access to scholarly publications and to research data are seen as particularly important aspects of *open science*<sup>14</sup> and as central elements in better use of scientific knowledge. Therefore these two topics take a central position in this report. However, it should however be noted that, in reality, *open science* is a much broader concept. For the European Commission, open science is “about the way research is carried out, disseminated, deployed and transformed by digital tools, networks and media”.<sup>15</sup>

In the first instance, therefore, this report is concerned with access to research results in the form of publications, including journal articles and books. If these are freely accessible, we use the term *open access*. The scope for open access has been the subject of heated debate for a number of years, focusing mainly on the questions of who should have access, when and under what conditions, to information (in the form of scholarly publications) that is already being shared, albeit in the restricted circles of subscribers or people who are willing to pay for access to a specific publication.

The second topic is the accessibility of research data for (re)use by third parties, for example for carrying out (other) research. Where these research data are freely accessible, the term *open research data* is used. The discussion on this topic is more recent and still has to crystallise. Although there are scientific disciplines where research data sharing forms part of the culture, (more) open access to research data would mark a radical change for many other disciplines, where *not* sharing data is the norm. In particular in the latter cases, research data are viewed as a key asset of an individual researcher and/or their research institute.

In this report we prefer to approach this topic from the perspective of ‘sharing research data’ rather than ‘open research data’, given the major step between not sharing research data at all (the practice in some scientific disciplines) and full open access to the data for reuse (the situation possibly evoked by the term ‘open research data’). This

---

<sup>14</sup> European Commission (2015a).

<sup>15</sup> Website of the European Commission: <https://ec.europa.eu/digital-agenda/en/open-science> (most recently accessed on 16 October 2015).

leaves scope for choices on subjects such as who the data is shared with or what kinds of reuse are permitted. Thus, it more accurately reflects the range of options available in practice when promoting the storage and sharing of research data.

Another term that can be interpreted in different ways is 'research data'. In this report, we define research data in principle as all data collected (used) for research. This includes all data *generated for the purposes of research*: measurement data (produced) by researchers themselves in and for research. But in principle it also includes existing datasets that are *used* for research. In this perspective, 'research data' then constitute a subset of (big) data in general.

This report discusses the topics mentioned explicitly against the broader backdrop of the interaction between science, society and business. As the Vision for Science 2025 rightly states, scientific knowledge must be shared and acquires societal value if it is deployed in real-world solutions by society and/or business.<sup>16</sup>

## 1.5 Approach and structure of this report

The first part of this report briefly sets out the key findings (Chapter 2). This is followed by recommendations for policy (Chapter 3). Together with Chapter 1, this constitutes the 'advisory report section'.

In the second part, we describe how we arrived at our findings and recommendations. A detailed analysis is presented of the themes of open access (Chapter 4) and research data sharing (Chapter 5). Chapters 2 and 3 are based entirely on this analysis.<sup>17</sup>

The analysis of open access and research data sharing is based on a literature review and interviews with experts and stakeholders.<sup>18</sup> As far as possible, we adopted the same approach for both topics; this is described in Chapters 4 and 5. We first outline the 'functions' associated with each topic which need to be safeguarded. We describe how those functions are currently given form and what developments are taking place.

We then consider developments from the perspective of the main stakeholders: what are their interests and what scope ('power') do they have to initiate changes? What is most likely their position towards open access to scholarly publications and the sharing of research data? We devote special attention to the roles of the government as a

---

<sup>16</sup> Ministry of Education, Culture and Science (2014), pp. 9 and 39. See also; AWT (2013).

<sup>17</sup> Chapters 4 and 5 accordingly contain references to the literature, etc..

<sup>18</sup> See Appendix 3 for a list of interviewees.

policymaker and lawmaker. What is the government's position, and how much scope does it have for effectively promoting open access and research data sharing?

Finally, in order to form an opinion on the social desirability<sup>19</sup> of the degree and form of open access and research data sharing, we analyse the effects at 'system level'. We believe this is the right level because social desirability mainly involves looking at the general interest – which is also the government's perspective. The following three dimensions seem to us to be the most relevant in assessing the impact of open access or research data sharing:

1. **Impact on science.** What is the impact on the functioning of science? How does open access or research data sharing affect the *quality* and efficiency of research, and the effectiveness and speed of the dissemination of research results? Does it foster cooperation and/or cross-fertilisation between different scientific disciplines? What impact does it have on scientific integrity?
2. **Impact on business and society.** What is the impact on the use of scientific knowledge by businesses and civil-society organisations? How does open access and data sharing influence processes of innovation? What is the impact on collaboration between science on the one hand and business and society on the other?
3. **Financial impact.** What are the costs and benefits, and where do they accrue? The costs of (paid) access to scholarly publications are rising year on year. Does a move to forms of open access deliver any financial benefits? A move to open research data will lead to additional direct costs involved in storing and providing access to the data, while at the same time there may be indirect benefits. How do they compare?

This analysis at system level provides a basis for formulating a (policy) objective in relation to open access to scholarly publications and the sharing of research data. If we add the insights from the stakeholder analysis, this creates a picture of the drivers for the desired level of openness and illustrates where the main obstacles lie. This then works through into our recommendations on how the envisaged level of openness can be achieved in practice.

To support our analysis, we compiled an inventory of developments in other (relevant) countries in relation to open access and research data sharing. Some of this information was provided by sister organisations of AWTI or through the network of Neth-ER.<sup>20</sup> This information was supplemented with our own literature review for a number of countries. The results are shown in Appendix 2. The experiences from the different countries are incorporated in the analysis presented in the report and also play a role in the positioning

---

<sup>19</sup> This was one of the constituent questions in the government's request for opinion to the AWTI (see Appendix 1).

<sup>20</sup> Neth-ER is the *Netherlands house for Education and Research*, and is the representation of the Dutch knowledge field in Brussels; website: <http://www.neth-er.eu/>

of the Netherlands and its policy in the international context, for example in the European Union.

This report was prepared by a project group consisting of Council members Valerie Frissen (chair), Arno Peels and Luc Soete and staff members Paul Diederer, Isabelle van Elzaker, Bart van Gent, Hamilcar Knops (lead author) and Ruud Verschuur. Interviews were held during the preparation of this report with more than 60 experts and stakeholders; input was also received from respondents from 11 countries (see Appendix 3 for the names).

## Findings

### 2.1 Open science

In this report we look at two forms of greater openness in science, namely facilitating easier access to scholarly publications (*open access*) and making research data more accessible. These are both forms of *open content*. These developments may be of great value to science, society and business. Such openness fits in well with the traditionally open character of science, in which researchers communicate with each other about the results of their research, build on each other's work or work together. Moreover, in the case of publicly funded research, openness aids accountability. Under *certain conditions*, more openness can also contribute to the better utilisation of scientific knowledge.

Those conditions are important, because openness also brings risks, and therefore the pros and cons have to be weighed carefully. Moreover, open access to publications and the sharing of research data are in itself not enough to make science truly accessible for society and business. For that to happen, these trends need to be embedded in a broader strategy aimed at strengthening the link between science, society and business.

### 2.2 Access to publications: *open access*?

#### 2.2.1 Open access needs a breakthrough

Traditionally, readers had to pay to see the results of scientific research. Scholarly articles were generally published in journals which required a subscription or a one-off payment to read an article, while scientific books had to be purchased. Books are still an important publication medium, especially in the arts, humanities and social sciences.

A trend has begun in recent years whereby scholarly publications are -as far as possible made available free of charge (*open access*). This trend is growing: between a quarter and half of recent scholarly publications worldwide are already freely accessible online in some way or another. The EU is not leading the way here: the proportion of open access is substantially higher in countries such as Brazil and the United States, partly helped by mandatory rules imposed by research funders and/or governments. Yet despite this trend, the growth in the number of freely accessible articles worldwide has been slowing in the last few years. The trend towards open access for (scholarly) books is in any event (still) going more slowly, and is really still in its infancy.

Thus far, the move towards open access has mainly been a bottom-up process, supported by a few large research funders. In practice, there are differences between scientific disciplines in the degree of open access and the way it is organised. Broadly speaking, however, this bottom-up process has not proved strong enough to bring about a broad-based system change in favour of open access. There is little resistance to the principle of open access, but neither is it the number one priority of most researchers – that honour is reserved for publishing in journals with substantial impact in the scientific world. This type of impact is often (still) achieved by publishing in renowned, ‘traditional’ subscription-based journals. In addition, a number of the key players do not (yet) recognise the importance of a large-scale move to open access.

Achieving a genuine breakthrough will require robust leadership. We found that research funders, including public authorities, have the power to realise open access publishing for the research they fund by imposing a requirement for open access. This is evident among other things from developments in the US and Brazil. Enforcing free access does, however, cost more, at least in the short term.

### **2.2.2 Different systems in parallel**

In practice, no single open access model has emerged as dominant. There are several different models for organising open access to articles, with the ‘green’ and ‘gold’ variants being the most common. In the ‘green’ variant, the author makes an article, or a version of it, freely accessible by publishing it in a scientific journal and at the same time placing it in a publicly accessible repository, often subject to an embargo. In the ‘gold’ variant, the journal itself provides open access to the article. In this variant, someone other than the reader has to pay; at present, this is often the author, who pays for the publication of his article. This differs from the traditional model in which the user (reader) pays for scholarly journals/articles. Similar models could also be used for providing open access to books.

Both models (‘green’ and ‘gold’) occur, often alongside each other. Green open access is a robust model that can exist alongside the traditional scholarly journals. However, its disadvantage is that it requires a ‘dual’ infrastructure (hence the higher costs) and that open access is often granted only after a period of embargo. The gold model does not have these disadvantages, but it only works well in practice if it is the ‘standard’, something that requires a wide-ranging system change. This in turn only happens if there is sufficient critical mass, which is difficult to achieve in the fragmented international field of science. Moreover, there are clear risks in a model where authors have to pay for publication in scholarly journals: in such model there is a much weaker incentive to put quality first in the selection of contributions. There are also several disciplines, especially in the arts and humanities, where it seems unlikely that budgets will be adequate to pay the required authors’ fees.

It thus remains unclear what might constitute a sustainable open access model in which there is free access and scientific quality is assured. Alternative models are currently being tested, but these have yet to prove themselves as viable and lasting open access standards. They include models in which research institutes and/or funders bear the journal costs directly, or where authors become 'members' of a journal and can then publish a certain number of articles. One initiative for books is the pooling of the procurement budget of several libraries to facilitate open access publishing of as many books as possible.<sup>21</sup>

Internationally, there are clear differences in preferences and policy with regard to open access. Within the EU, some Member States operate an explicit open access policy. Most of these countries support both the 'green' and 'gold' models. In many cases, the aim is to establish a green model as a minimum standard, based on the realisation that no single country will be able to enforce a system change towards 'gold' open access on its own. The United Kingdom is one of the few countries where the government has expressed a preference for the 'gold' route. Yet in the UK, too, the reality offers a mixture of green en gold as some research funders demand gold while others prefer green. In the United States, the federal government has opted for the 'green' route. The European Union has not shown a preference for either model.

### 2.2.3 Impact of open access

**A. Impact on science.** The impact on science of (merely) creating free access to scholarly publications is unclear. Impact, especially *within* disciplines, would probably be limited. In practice, researchers – particularly researchers in the Netherlands and neighbouring countries – often already have reasonably good access to the literature produced in their discipline. This may not be the case for publications from other disciplines, access to which is often more difficult. The impact of open access is therefore likely to be greater in the latter case. The impact on science will also be greater in countries where such access is currently less well established – generally the less affluent countries (although institutions in such countries often benefit from special arrangements or lower fees).

Although greater openness is regarded by many as an important and necessary development, good quality control of publications, speedy knowledge dissemination and good findability of publications are of key importance for the functioning of science. Open access only has a positive impact on science if it contributes to these areas. This is not something that can simply be assumed to be the case. In a model in which authors pay for publication of articles, the incentives to allow the quality of articles to drive the acceptance by journals are weaker than in a system where the reader pays, as at

---

<sup>21</sup> This is the 'Knowledge Unlatched' initiative (<http://www.knowledgeunlatched.org/>), which won the IFLA/Brill Open Access Award 2014.

present. In the worst case, the (model of) payment by authors could become a perverse incentive and lead to lower quality. This point is all the more pressing as, in the ever burgeoning sea of scholarly publications, the demand for mechanisms to assess the quality or reputation of a journal or individual publication will only increase.

**B. Impact on society and business.** Open access to scholarly publications for society and business is still an area of largely untapped potential. The reason being that access to such publications beyond the field of science is extremely limited (with the possible exception of companies with a strong R&D department, which often already have subscriptions or have good connections with research institutes). On its own, however, open access will have only a limited impact on society and business. Tapping into the added value of scientific knowledge requires more than opening up access to publications. Scholarly articles are mainly intended for other researchers, and may therefore not be in a form that is readily accessible to business and society. Therefore, how much of the potential added value of open access for society could actually be realised in practice will depend greatly on the effectiveness of ancillary measures and developments aimed at improving the accessibility and usability of scientific output.

Increasing the impact of science on society and business therefore requires open access to be part of a broader strategy. First, scientific results have to be findable. Second, science needs to be made genuinely accessible by 'translating' results into a form that can be digested by a wider public. Thirdly, a broader approach is needed to ensure that the 'social impact' of research is (more) valued and rewarded within the scientific world itself, for example during evaluations. Open access can be a useful part of such a broader strategy.

**C. Costs and benefits.** Scholarly publishing costs money. There is currently no reason to assume that the total costs would fall purely because of a move to open access. The relatively sharp rise in costs stems from the global growth in the number of publications and the economically strong position of some publishers, which exploit the fact that researchers want to publish globally in journals with a strong reputation (often related to citation impact). Readers of such journals also prefer publications with a good reputation. If reputation were to become less important, or if cheaper alternatives with good reputations were to appear the present high margins of a small number of large scientific publishers could be squeezed (either by the 'pressure' of newcomers or because new entrants or business models displace the existing operators/models). Better negotiation, possibly by users acting together, could also constrain the current margins of the large scientific publishers. This would require more transparency concerning subscription and publication costs. In either case, the total costs paid by readers and/or authors could reduce. In itself, however, open access to scholarly publications is not a panacea for lowering the perceived high costs of scholarly publishing.



What will happen to the total costs depends on which open access model will emerge as the standard, though in reality there is a good chance that a mixed system will continue. The 'green' variant involves extra costs (for example for repositories). The business model that is the norm in the 'gold' model, is based on the principle of 'the author pays'; this carries the risk of volume growth (the more articles that are published, the more income for the publisher) and, if the authors' fees are not capped, these could also rise to excessive levels. Moreover, in a system where the author pays, the resources available to researchers are a decisive factor in how much (and where) they can publish. Whether that is desirable is very much the question.

There will also be some cost displacement in a model based on the principle of 'the author pays'. Countries (or institutions) which relatively publish more than they read will pay more on balance. This is likely to be detrimental for the Netherlands, which is an above-average knowledge producer. In such a situation, it will be necessary to look at what other countries do. For example, if the Netherlands or the EU incur costs in making 'their' publications freely accessible worldwide, this will partly benefit third countries not willing to incur those costs for their own researchers (even where they have the resources to do so): a free-rider problem. Dutch or EU researchers will then be paying twice: for publishing their own articles and for reading articles published in the countries that do not operate a system of open access.

In short, genuinely reducing costs requires a rapid transition to a new system. The chance of success will increase as more countries and research funders participate.

## **2.3 Sharing research data**

### **2.3.1 Sharing research data still the exception**

Traditionally, results of scientific research are communicated via all kinds of publications (conference papers, articles, books, etc.). Those results are usually based on an analysis of research data that has been generated specifically for that study. This data can take very different forms, such as measurement results, photographs, interview reports or other descriptions. Large quantities of research data are collected in scientific practice, but what is done with that data differs widely from one discipline to another. There are some disciplines where storing and then sharing research data is already the norm, and others – the majority – where that does not happen, or only exceptionally. In these cases, the research data often remain with the researcher or their institution; it is unclear whether the data will be kept and others have little or no access to them. The outside world only gains an idea of the existence of the research data through the publications about the research the data were gathered for.

Experiences in disciplines where research data is already stored and shared are increasingly giving rise to calls for this to happen in other scientific disciplines as well, especially for publicly funded research. Sharing research data can increase the efficiency of research, for example through the reuse of datasets. In addition, there is growing pressure from society for more transparency about the underlying research data, for example in areas such as food safety or climate. Individual companies are also feeling this pressure for openness.

Before research data can be shared with others, however, they have to meet a number of requirements. It is generally accepted that data must be 'FAIR':

- ▶ *Findable*;
- ▶ *Accessible*;
- ▶ *Interoperable*;
- ▶ *Reusable*.

Raw research data must therefore first be made suitable for consultation and then stored in a way that makes them accessible. The difficulty and costs involved vary depending on the discipline and the type of research. Moreover, allowing different disciplines to access and reuse each other's data places heavy demands on interoperability, in turn requiring a specialist professional approach (*data science*).

The decision on whether or not to share data is mostly left up to the researchers themselves, or sometimes to their employers or funders. These actors make different judgements in deciding whether or not to share their research data (publicly). In some branches of science it is almost essential to generate research data together with others and then to share them, for example because there is a large common infrastructure or because a group of researchers are working together on a large scientific puzzle which goes beyond the scope of a single research group.

In disciplines where research data are never or almost never shared, in particular, it is unclear for many researchers what is to be gained by allowing access to their data. At first glance, it costs extra time and trouble to prepare the data for storage and reuse – interoperability can prove to be a difficult requirement to meet – while there is also the risk that another party will publish 'their' research data sooner. Moreover, currently there is still not sufficient credit given within the academic community for the fact that someone's data have been reused by other researchers, which is another disincentive for research data sharing. There are some initiatives aimed at offering ways of publishing datasets (such as data journals) and rendering them suitable for citation, but this is by no means the norm within the wider scientific community.

### 2.3.2 Impact of research data sharing

**A. Impact on science.** In those parts of the scientific world where research data are already stored (in a usable form) and shared, the benefit of data sharing has already been proven. But sharing data on a wider scale can have a major impact in disciplines where data sharing is currently limited. Sharing research data can help raise the quality of science in various ways. Individual research becomes 'better' through improved research data management; the transparency of research will improve if the data underlying scholarly articles can be viewed; and the efficiency of science can be raised by avoiding unnecessary duplication of research. At the same time, access to research data helps the 'reproduction' of research when the various 'reproductions' are accessible. It also opens the way to a new type of research that cuts across boundaries between individual datasets and disciplines. All this does mean that researchers have to learn to trust research data collected by others, and be willing to use them in their own research.

**B. Impact beyond science.** In theory, society and business stand to benefit from the free sharing of research data. This will very probably prove to be the case for some research data, though other types of data will prove to be less interesting or usable for citizens and businesses. There is also an indirect social impact: if sharing research data helps speed up research faster and make it more effective, results or applications that are useful for society and business (such as medicines) may become available more quickly.

To promote maximum use of research data by society or businesses, the data could be presented in a way that is readily accessible, free and directly usable (*open data*). Time will then tell what impact this has. Our case study on (the use of) open government data shows that the usability, accessibility and findability of data still regularly pose an obstacle to the reuse of data.

**C. Costs and benefits.** There is a clear asymmetry in the costs and benefits of providing access to research data. Clearly, the costs will increase because researchers will have to spend more time and/or money on data management, preparing data for storage and reuse, and the storage itself. The benefits of sharing research data are much less clear for individual researchers at present. An incremental approach is needed here, using pilots and trials to try and make clear the benefits of open access to scientific information. It is not wise simply to store all research data: choices will need to be made.

The costs and benefits in relation to the useful deployment of research data by society and business also need to be weighed carefully if a systemic change is to be achieved. Which costs will be borne collectively and in which cases is an individual contribution reasonable? Particularly where research data are of use to commercial application, it is not unreasonable to ask for payment for commercial use. Businesses are willing to pay

for use of data, as we saw in our case study on weather data from the Royal Netherlands Meteorological Institute (KNMI). However, in the mean time, data from the KNMI were made available free of charge, so this revenue has been lost. On the other hand, the move to make these data available for free has not led to a noticeable increase in the (commercial) use of these data.

### 2.3.3 A route map for responsible research data sharing

There is a long way to go before the sharing of research data becomes the norm across the whole scientific research spectrum. First, researchers will need to be convinced of the value of sharing, and many conditions will have to be met.

A number of aspects are crucial to the useful and responsible sharing of research data with third parties:

1. Good data management during the research;
2. A storage infrastructure;
3. Standards in relation to storage (for example for metadata).

'Good' data management means *FAIR* data management (Findable, Accessible, Interoperable, and Reusable). Data management is something that benefits every researcher, regardless of the extent to which data are shared. Facilities (e.g. an infrastructure for data storage) have to be created or used. Research data also have to be stored in such a way that they are findable. Standards are needed to make research data genuinely open and lasting. A variety of strategies can be used for this, but there is currently a lack of sorely needed overall control and coordination.

The next step is to decide how data could be (re)used (and by whom, and for what purpose) in a way that is useful.

A number of legal questions also need to be resolved in relation to aspects such as security, privacy and intellectual property. This includes the question of ownership of research data. It is very important for researchers that control over research data, especially data obtained with public funding, remains within the scientific world – with researchers, institutions or funders. By their nature, these legal questions require a response from the government, as the legislative power.

It is not wise to provide unconditional access to all (stored) research data to everyone. There can be various reasons for not doing so (such as commercial interests or privacy considerations). Moreover, to foster support for research data sharing it is wise to give researchers and their institutions (joint) control over who they wish to share their research data with, and for what purpose. It is quite plausible that reciprocity will play a role here, with researchers or institutions granting access to each other's research data.

## 2.4 Role of government

### 2.4.1 What can the EU and the Netherlands do in relation to open access and the sharing of research data?

As a key funder of scientific research, the government is in a position to influence the preferred mode of scholarly publishing and dealing with research data. The government is also the (public) regulator and in this capacity the rulemaker of matters such as intellectual property rights. Furthermore, the government has 'system responsibility' and can exercise a degree of coordination and control. The government provides a sort of 'benchmark': by setting out a clear vision, it can give direction to the choices made by the other actors. Interaction with the field is crucial here.

Our analysis of the development of open access leads us to conclude that the bottom-up process has so far not been enough to achieve a broad-based system change in favour of open access. It is important to step up this process. To achieve this, open access needs to be embedded in a broader strategy aimed at making better use of knowledge within and beyond the world of science. This requires coordinated action on multiple fronts. The government is in a position to take on this coordinating role to some degree. The same applies for promoting the sharing of research data: this requires (extra) infrastructure, all manner (pre)conditions that need to be addressed (standards, legal questions) and the release of funds in the short term to cover the initial costs. This again requires coordination, in a way that leaves scope for individual specification and implementation in each discipline.

Within the European Union, the Horizon 2020 research programme offers such a 'benchmark'. Together with the Seventh Framework Programme which preceded it, Horizon 2020 is a key driver of open access. It also includes a pilot for providing access to research data. In practice, however, open access publishing is not yet widespread among research in the EU.

The European Union is the world's biggest producer of scholarly articles. Yet that still represents only around 30% of the world total, which means the EU will not be able to impose its will unilaterally on the worldwide system of scholarly communication, either for articles or for data. This observation is of particular relevance if the EU were to express a preference for a model that only works well if everyone switches to it, such as 'gold' open access. This applies all the more for a single Member State, such as the Netherlands.

### 2.4.2 International context

The international dimension is important. The Netherlands and the EU are part of a knowledge-intensive region where high-grade research is carried out. Allowing easier access to their own scholarly publications, and especially to research data, means that

the Netherlands and the EU will effectively be 'giving away' information to other countries, though this is not necessarily a bad thing. The cost aspect and strategic considerations do mean it is necessary to make the right judgement calls. This is less relevant for open access than for data sharing: publications are, by definition, already published, and others can already gain access to them (albeit for a price); at most, there will be a shift in the proportions paid by the different parties. The consequences are more far-reaching for (research) data: much of which is currently *not* routinely shared. The sharing of such data by the Netherlands and the EU with the rest of the world would mean that third countries would gain access to new information.

The Netherlands and the EU boast a number of centres of research excellence. That is a competitive advantage in the global scientific world. Giving access to the results of this research too readily to parties outside the Netherlands or the EU could limit or even nullify that competitive advantage. This means, in the first place, that the appropriateness of sharing research data must be carefully considered. Secondly, this can have consequences for how and with whom research data is shared. This underlines yet again the importance of adopting a gradual approach in deciding whether or not to share research data. The EU and its Member States may, for example, take into account the degree of openness displayed by third countries themselves (a form of reciprocity).

At the same time, there is a strong argument for improving access to relevant scientific results by poorer countries, for example in relation to health care. Many initiatives are already being developed in this regard, and these are broadly supported by the key stakeholders.

## Recommendations

In the foregoing chapter we presented the key findings from our detailed analysis as described in Part 2 of this report. Based on that analysis and on the underlying goals of ensuring the effective use of scientific knowledge and a science system that functions well and is closely tied into society and business, the AWTI presents the following recommendations to the Minister and State Secretary of Education, Culture and Science and the Minister of Economic Affairs (and the relevant stakeholders). These recommendations relate to open access and the sharing of research data.

### Recommendation 1

#### **Embrace the importance of openness and formulate a broader strategy for better utilisation of knowledge**

Acknowledge and embrace the importance of openness as a core value in science, and especially for publicly funded research.

- a) Do not view open access as an independent goal of science, but as a link in the process of achieving better utilisation of knowledge.

Promote open access as part of a broader strategy aimed at improving the dissemination and deployment of knowledge within and, above all, beyond the field of science. Only as part of such a broader strategy which increases the accessibility of scientific knowledge does open access genuinely add value.

- b) Make a substantial and concerted effort to let research data be shared more widely, and aim to optimise access to research data.

More openness can help science to function better, lead to better utilisation of scientific knowledge within and beyond the world of science, or ensure that science is more accountable for the (public) research it carries out. Increasing openness in science should always be seen in the light of these wider objectives. The pros and cons of more openness need to be carefully considered, as does the likelihood that the right conditions can be created. This balanced assessment also makes it possible to set priorities.

For many scientific disciplines, sharing research data more widely would be a more radical step than open access, and one which can potentially have a major positive impact on science itself. On the other hand, there are inherent costs, limitations and risks. This means striking a balance to determine what constitutes the right degree of openness. There will likely be differentiation in what is shared with whom, as well as in the types of use and reuse. This will demand substantial investments, particularly in an adequate infrastructure and in the development of good data management methods and strategies.

## **Recommendation 2**

### **Strive for effective implementation of the principles of open science in the Netherlands**

Formulate a masterplan for the Netherlands aimed at the better utilisation of scientific knowledge, with attention for the role of open access and the sharing of research data. Set aside sufficient resources for this and ensure that the importance of the 'impact on society and business' is properly anchored in research practice.

This anchoring could, for example, be achieved by taking into account the impact on society and business in evaluation of research and researchers. Ask scientific disciplines to formulate plans for improving the use of scientific knowledge and provide financial support for those plans. Reward disciplines and/or research institutes accordingly, for example by introducing a ranking system which (partly) reflects the social and economic impact of their work.

#### **A. Ensure that the masterplan for open access:**

- a) contains clear milestones, preconditions (such as quality assurance) and, where necessary, infrastructure. Allow for differences in elaboration and actual implementation in different scientific disciplines. Use this process to showcase the results of Dutch science;
- b) makes scientific results more accessible to society by making them findable and presenting them clearly and in a way that can be understood by a wider public. Arrange that research programmes present the results of its research more coherently and in an accessible way, not only by providing links to the (open access) scholarly articles, but also by adding a summary aimed at lay readers. This approach could very well be applied to the Dutch National Science Agenda and its website;
- c) provides for better access to scientific knowledge for business, for example by establishing intermediaries and ecosystems in which science and business are brought together; open access is then an important link in the transfer of knowledge.

#### **B. Ensure that the masterplan for sharing research data:**

- a) contains clear milestones and preconditions and allows differences in elaboration and actual implementation in different scientific disciplines. There is still a long way to go in storing and sharing research data in a usable form. Adopt a step-by-step approach to this;
- b) provides for general facilities for data storage and sharing: ensure that the storage infrastructure is in place – set aside the necessary resources for this – and establish other basic conditions (such as standards). Organise generic aspects generically and leave the detailed specifications as far as possible to the



individual disciplines, though bearing in mind the need for interoperability of data across disciplines;

- c) promotes good data management: encourage the embedding of data management in science training; initially set aside additional resources for this; ensure that good data management is rewarded (through grants or evaluations) and seek to make good data management a standard part of the scientific process;
- d) contains a clear assessment framework for determining which type of research data should be stored at public expense;
- e) as a minimum, provides for access to the research data that underlie a scholarly article, as long as there are no compelling reasons for not doing so. This fosters transparency in science;
- f) for the rest, encourages the sharing of research data in general where this is possible, receives support and serves a useful purpose. Allow the sharing of data to grow (first build up trust) and initially allow researchers and institutions to decide who they wish to share their research data with;
- g) supports the sharing of data, preferably by rewarding it (for example through grants, evaluations or through rankings) rather than through imposition;
- h) properly addresses legal issues, for example the protection of the efforts invested in gathering data (this is currently unprotected); clarity is also needed on the ownership of research data and the responsible use of personal and other data. It is advisable for the 'ownership' of research data obtained with public funding to remain with the institution(s), funder(s) or researcher(s) that obtained them – in this way it remains within the science system.

### **Recommendation 3**

#### **Collaborate on open science at EU level**

As far as possible, ensure that Dutch policy on open access to scholarly publications and research data is in line with EU policy. Aim for an accelerated transition to open access (to publications) and research data sharing, and advocate a common strategy to achieve this.

Take into account developments within (and differences between) EU Member States. Look for and formulate common interests, such as transparency on subscription costs, shortening embargo periods, improving the accessibility of scientific knowledge and a European infrastructure to facilitate open access and research data sharing.

The Dutch government can then place the following items on the agenda at EU level (including during the Dutch EU presidency):

- a) the economic, social and scientific importance of sharing research data and making scientific knowledge accessible;

- b) the importance of establishing conditions necessary for open science, such as maximum embargo periods, transparency on the costs of publication, and a degree of standardisation;
- c) actively monitoring developments and progress in relation to open access and research data sharing in Member States with a view to convergence and accelerating the process;
- d) genuine accessibility of the results and research data for those projects that are (co-)funded by the EU, by ensuring open access but also by providing accessible summaries and introductions for lay readers. (This includes not just Horizon 2020 or European Research Council grants, but also EU research institutes such as the Joint Research Centres);
- e) the creation of a European Science Cloud to serve both as a hub and as a means of publicising European scientific achievements.

The Hague, December 2015,

Professor Uri Rosenthal (Chairman)

Dr Dorette J.M. Corbey (Secretary)



# Part 2: Analysis



## Open access

### 4.1 Between dream and reality...

Researchers have long been focused on communicating their findings and results within the scientific community: sharing results, discussing and commenting on each other's research. This regularly leads to subsequent collaboration. Science is characterised by a large extent of openness (at least within the peer community) with the common aim of further advancing knowledge.

Knowledge and research results are shared in many ways: through 'learned societies' (i.e. associations of researchers), often focusing on a specific discipline; through presentations and discussions at conferences; through books; or through articles in scholarly journals. Peer-reviewed journal articles, in particular, have become the 'standard' form of scientific output, though in the arts and humanities, books are also still important. Peer review gives these articles a scientific stamp of quality, and the reputation of the journal in which an article is published enhances its prestige. A journal's reputation is closely related to the average scientific impact of its articles. The number and impact of articles published plays an important role in a researcher's career. It strongly influences the esteem in which he or she is held by the peers, plays a major role in the awarding of grants, etc., and is an important element in evaluating a researcher's performance.

The number of scholarly articles published has grown sharply in recent decades. Not just because there are more people active in science around the world, but also because the scientific culture described above works as an incentive for a researcher to publish as many articles as possible – preferably in journals with high citation impact. As the 'readers' of scholarly journals have traditionally paid for them (authors did not pay), this increase in volume means readers are confronted with increasing costs. In addition, the journals with the best reputations are 'must-reads', research institutes are almost obliged to subscribe to them. This means that the publishers of such journals can demand sizeable margins. Furthermore, the scientific publishing 'market' is highly concentrated, which puts the publishers involved in an even stronger position. Research institutes are becoming more and more dissatisfied with the sharply rising costs of access to scientific literature.

There has been another development in the way scholarly publications are disseminated. In the past, journals and books were published on paper and ultimately ended up on a library shelf. Today most publications, especially articles but increasingly books as well,<sup>22</sup>

---

<sup>22</sup> See e.g. the OAPEN Library of freely accessible scientific books: <http://www.oapen.org/>

are also available in digital form and are published on websites, from where they can, in principle, be downloaded. Users affiliated to an organisation which has a subscription (in the case of the traditional subscription-based journals), can download the publication for free. Non-subscribers cannot readily access these publications, although they are frequently able to purchase one-off access to a specific article.

Now that digital publication has more or less become the standard, one could look for ways to ways of providing access to scholarly publications to everyone. That would fit in well with the value attached to the openness of science, which enables researchers to study each other's research results. It would also mean that society (both citizens and the organisations of which they are members) and businesses could (more easily) access the results of research. To maximise the utilisation of these scholarly publications, it would be preferable to permit the widest possible *use* of these articles: not just access to reading them, but also being allowed to reproduce and disseminate them, for example.

This is precisely the idea behind the calls for open access to scholarly publications, which were first discussed a little over ten years ago. The open access declarations of Budapest (2002)<sup>23</sup> and Berlin (2003)<sup>24</sup> are often taken as the starting point, and over the years many research institutes and researchers have signed up to these declarations. The goal of these declarations is free access for everyone (researchers, businesses and citizens) to scholarly publications and permission to use them broadly.

What should have been a rapid revolution has, over a decade later, come to resemble a chess game in which the pieces are stuck in position and developments move slowly, step by step. Each player at the board has their own interests to defend.

Progress has undoubtedly been made towards more open access. New journals have appeared in which all articles are freely accessible for everyone. Some existing journals have also converted to open access. In such open access journals, it is generally the authors who pay, in the form of a fee, for publication, although experiments are being conducted with other models. For example journals for which research institutes or funders bear the cost.<sup>25</sup> Most publishers of subscription-based journals now allow an author to pay for the article to be made accessible to subscribers and non-subscribers. The article then becomes an open access article (though the journal itself is not open access). Authors are allowed, often under certain conditions, to deposit a version of an article in an online repository where it can be freely accessed. This is sometimes only permitted after an embargo period. There are also researchers who post articles on their

---

<sup>23</sup> <http://www.budapestopenaccessinitiative.org/read>

<sup>24</sup> <http://openaccess.mpg.de/Berlin-Declaration>

<sup>25</sup> Several alternative models were discussed at a workshop organised by the European Commission on 12 October 2015 in Brussels: <https://ec.europa.eu/digital-agenda/en/news/save-date-12-oct-ec-workshop-alternative-open-access-publishing-models> (most recently accessed on 16/10/2015).

own websites or on 'sharing sites' (such as ResearchGate), sometimes without the copyright rules being fully respected.

### Open access terminology

**Open Access** - In the ideal situation as set out in the Budapest Open Access Initiative (2002), 'open access' refers to free access to scholarly articles, whereby it is permitted to read, download, copy, disseminate, print, search for, index and refer to (parts of) articles and where there are no financial, legal or technical barriers to that access. The only exception is the right of authors to protect the integrity of their own work.

It is common practice to distinguish between two 'main routes' towards open access, the 'gold' and the 'green' route:

The gold route is a new standard under which existing journals and publishers retain their central role, but publications are freely accessible with immediate effect and a party other than the reader, usually the author, pays. Publications are directly available online via the publishers' platforms. This is a substantial change in publishers' business models, from a model based on subscription income to one where authors pay.

The green route is a hybrid system of open access and subscriptions. It is based on self-archiving. Articles are published in a traditional, non-open access journal, and the authors then make their manuscript publicly accessible by archiving it in a (publicly accessible) repository. Publishers sometimes require application of an embargo.

All these possibilities have, over the last 15 years, led to an increase in the number of scholarly publications that can be freely accessed in some way. It is difficult to determine the precise percentage of open access publications. A reasonable attempt can be made for 'gold' open access articles, as this is a well-defined model. It is more difficult to find and count all variants of self-archiving ('green' model). Does one count only articles in repositories, or also articles on authors' own websites, for example – and how does one find these? Consequently, different studies come up with different percentages of open access articles.

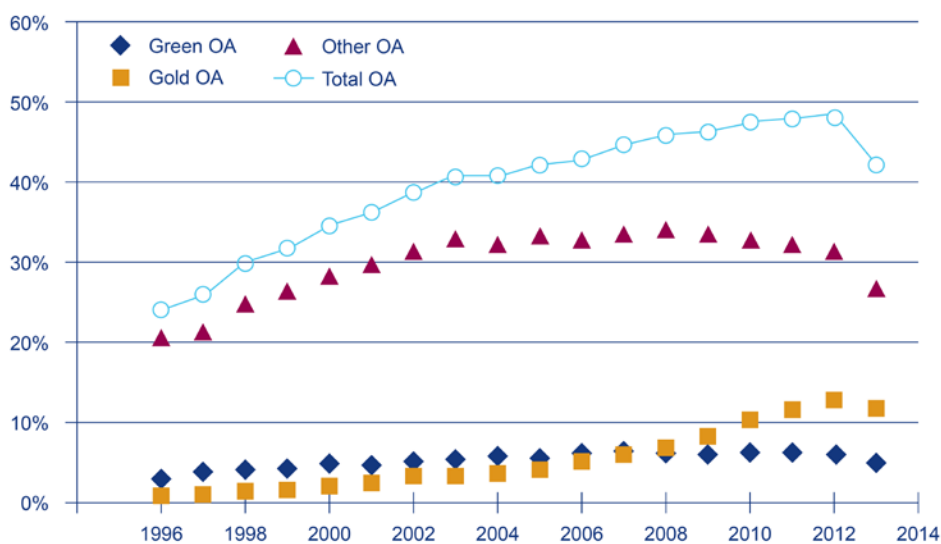
Here we present the results of two studies which can serve as a sort of upper and lower limit. In a study carried out for the British government, Elsevier (2013)<sup>26</sup> estimated the global share of 'gold open access' publications at 10.2% (plus a further 1% for articles

---

<sup>26</sup> Elsevier (2013).

made freely accessible by the publisher after a set period). Elsevier arrived at a figure of 11.4% for the global share of ‘green open access’ publications. This would mean that a total of around 23% of recent publications were freely accessible under either the gold or green model – so around a quarter of the total. These estimates are of the same order as those produced in a number of earlier studies.<sup>27</sup>

Figure 1 Percentage of freely available peer-reviewed scholarly papers, as measured in April 2014 for 1996-2013



Source: Archambault *et al.* (2014), p. 16. The percentages were computed by Archambault *et al.* using Scopus, DOAJ, ROAR, OpenDOAR, PubMedCentral and several other sources of freely downloadable articles.

Another study, carried out by Archambault *et al.* (2014) at the request of the European Commission, uses a different method. These authors produce a figure of approximately 12% for the share of ‘gold’ open access publications in 2013, comparable with Elsevier’s result. Archambault *et al.* (2014) apply an unusual definition for ‘green’ open access, as a result of which their figures comes out at 6% (2012) and 5% (2013). Thus, lower than Elsevier (2013). In addition to gold and green, Archambault *et al.* (2014) include a category of ‘other open access’, which comes out at 30.9% in 2012 and 27% in 2013. The fall between 2012 and 2013 in the ‘green’ and ‘other’ categories may be linked to the fact that self-archiving can be subject to delays, for example because of embargoes. The authors were unable to provide certainty on this, accepting that their figure could also signify a ‘genuine’ reduction. According to Archambault *et al.* (2014), almost half of all globally produced scholarly publications can be accessed freely online in some way or another. However, the category ‘other’ is so vague and wide that it is impossible to

<sup>27</sup> Björk *et al.* (2010) found a total of 20.4%, while Gargouri *et al.* (2012) arrived at a total of 21%.



guarantee that all articles in this category are genuinely easily findable. The data presented are also unclear in other respects (which are not explained by the authors), and therefore the figures in this study must be treated with caution.<sup>28</sup>

Despite these shortcomings, we can conclude that the percentage of recent scholarly articles which are open access is between a quarter and half (depending on definitions). Less is known about the figures for scholarly books; open access is more in its infancy here, and it is likely that the percentage is (even) lower.

The trend in (the percentage of) open access scholarly articles is interesting. The overview of Archambault *et al.* (2014) shows that this percentage has grown since the middle of the 1990s, but has flattened off in recent years (and actually fell in 2013). See Figure 1.

The same effect can also be seen in the Netherlands: the number of open access publications included in repositories that can be accessed via NARCIS has stopped increasing (and even appears to be falling slightly).<sup>29</sup> Figure 2 shows the number of scientific articles posted in NARCIS from 2000 to 2014.

Around 20% of the publications in NARCIS were open access in 2000, and this figure peaked in 2008 at 28%. It fluctuated around 25% in the last few years. This is probably an overestimation of the share of open access publications in the Dutch total, because although some institutions have placed all scholarly publications in their own repository, others only include the open access publications.<sup>30</sup>

The analysis by Archambault *et al.* (2014) of the degree to which peer-reviewed articles are available through open access reveals wide differences between disciplines and between countries. For example, four out of five articles in the discipline 'general science and technology' can be freely accessed online, mainly because they are published in open access journals. The percentage of open access publications in the biomedical research discipline is also high, due to the central role played by repositories in this discipline. At the other end of the spectrum is a discipline such as 'visual and performing arts', where barely one in four articles genuinely offers open access.

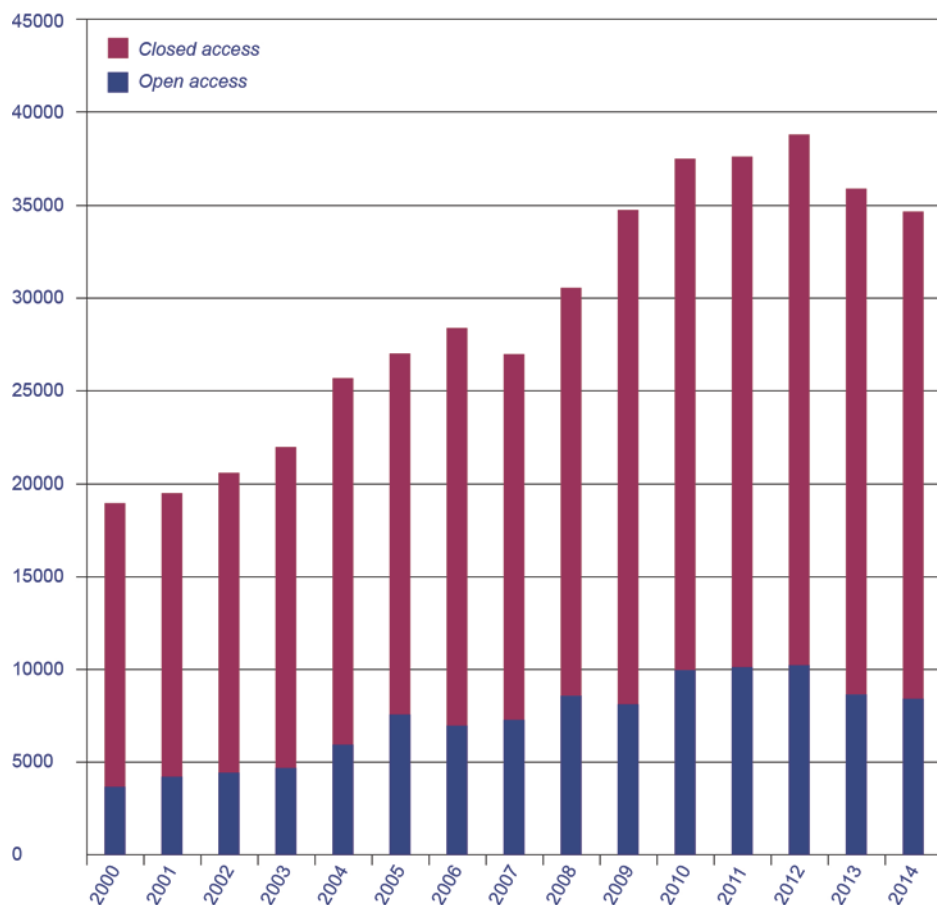
---

<sup>28</sup> Just a few examples of these unclear areas: the total percentage of open access publications is not equal to the sum of the three categories 'gold', 'green' and 'other'; in the country allocation, it is not clear how publications are allocated across countries: for example, the sum of the publications from EU Member States does not equal the total number of EU publications, and the averages per country and across the whole EU are therefore not comparable. Similar problems are evident in the overviews for each discipline.

<sup>29</sup> See the NARCIS website for the current situation: <http://www.narcis.nl/metrics/Language/en> (most recently accessed on 16/10/2015).

<sup>30</sup> A breakdown of the numbers of publications in the repositories (total) and the number of open access publications per institute may be found at <http://www.narcis.nl/reponumbers/Language/en/#show-me> (most recently accessed on 16/10/2015).

Figure 2. Number of open and closed access scholarly articles in NARCIS (by year of publication)



Source: NARCIS, see: <http://www.narcis.nl/metrics/Language/en>

Note: This chart shows the actual number of open and closed access peer reviewed articles in NARCIS, since 2000. NARCIS provides access to publications in the repositories of all Dutch universities, the Royal Netherlands Academy of Arts and Sciences (KNAW), the Netherlands Organisation for Scientific Research (NWO) and a number of research institutes. Some institutes have placed all their scholarly publications in the repository, others only the open access publications.

The open access culture differs between countries.<sup>31</sup> Brazil is the clear leader; not only are four out of ten scholarly articles published in open access journals (compared to a global average of one in ten), but ultimately two out of every three Brazilian scholarly papers are available through open access. The percentage of open access publications is also high in the US (nearly 60%), which is mainly due to the central role played by a

<sup>31</sup> The figures quoted in this paragraph come from the study by Archambault *et al.* (2014), who appear to give a (too) high estimate of the proportion of open access articles, see above; but we use them here for a 'qualitative' comparison between countries.

small number of repositories, where a version of an article that has been published elsewhere is lodged. The percentage of articles that can be freely accessed in one way or another in the EU as a whole is just above 50%. The Netherlands leads the way in the EU, with almost two out of three articles being available through some form of open access according to Archambault *et al.* (2014) (and according to their broad definition of open access). Interestingly enough, Dutch researchers publish less than average in open access journals, but evidently many papers are placed in repositories, etc. The Netherlands scores very highly in the category 'other open access', but because this is such a vague and broadly defined category, it is not entirely clear what lies behind this high figure. Other sources show a much lower open access percentage for the Netherlands; for example, only a quarter of the Dutch publications in NARCIS are open access, and even that is probably an overestimate (see above).

Like the European Union, the Dutch government is in principle a proponent of open access: everyone should have access to scientific knowledge for which they help pay (as taxpayers).<sup>32</sup> In addition, these authorities expect that a publication system built around open access will help speed up the sharing of scientific information, enhance the robustness of review processes, support quality assurance, benefit integrity in research, and help boost collaboration among researchers and between researchers and other stakeholders, such as businesses or civil-society organisations.

Some supporters of open access base their support on strongly ideological arguments, namely that science worldwide, and therefore scholarly publications, should be open. Financial arguments are also used, with open access then being seen as a way of challenging the ever-increasing costs of access to scholarly publications.

In this chapter, the AWTI examines the (potential) effects of open access and under what conditions it could make a maximum contribution to the functioning of science and to society and the economy. We draw conclusions based on this analysis, and these in turn form the basis for the recommendations for government policy presented in Chapter 3.

## 4.2 Scholarly publishing: functions and stakeholders

The culture surrounding the publishing of scientific results, and consequently also the debate about access, takes place within the context of the larger (global) 'science system'. Different kinds of stakeholders play a role, who are members of all manner of networks and bound by all kinds of rules. In addition, the world of science is linked to (the rest of) society and industry. In this section we try to analyse the position of relevant

---

<sup>32</sup> Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 15 November 2013, TK 2013-2014, 31 288, no. 354, p. 1.

stakeholders. We look at their interests and their capabilities in relation to scholarly publishing. This means addressing questions such as: 'What do they want to achieve?', 'What can they achieve and are they able to bring about change?'. We then discuss their position specifically in relation to open access. This provides insights into the drivers of open access, the bottlenecks and opportunities.

First, however, we discuss the functions that play a role in relation to the communication of scientific results:<sup>33</sup>

- ▶ Certification: providing scientific validation of results for publication (with a key role assigned to peer review);
- ▶ Dissemination: the distribution of (published) results;
- ▶ Registration: the 'official' linking of the authors to the publication;
- ▶ Archiving: filing the publication and keeping it available.

Once scientific results have been published, they can be used in all kinds of ways. They can for example be cited, something that is seen as an important form of recognition within the scientific world.

Developments in relation to open access are mainly concerned with the dissemination function. The transition from publishing on paper to digital publication has created new opportunities for disseminating. Publishing working papers or setting up preprint servers enables researchers to communicate their ideas earlier, while still making use of the traditional channels (journals) to ensure the validation and reputation of their publication. Publishers have introduced open access: complete journals or individual articles are then freely accessible to everyone. Scholarly books are increasingly offered in open access. These new forms of publishing have given rise to new business models, including a move away from the traditional 'reader pays' model toward 'the author pays' or other variants.

## Researchers

Researchers play a variety of roles in relation to scientific publishing. First, they are users (readers) of scholarly publications. In this role, they benefit from having the widest possible access to the literature that is relevant for them. In many cases, the library of the research institute where a researcher works will take it upon itself to pay for the right information (journals, books, etc.). The individual researcher then does not directly experience these costs. If a researcher wishes to consult literature outside this framework, he or she will have to pay for it, or arrange access to it some other way. When deciding which journals have priority, researchers will often be guided by the journal's reputation. That reputation will depend in part on the average (scientific) impact of the journal, its relevance for the discipline/area of research and, for example, who the editors are.

---

<sup>33</sup> See e.g. Ware & Mabe (2015), p. 16.

As well as being 'consumers', researchers are 'producers' of scientific output. In this capacity, they are primarily interested in the successful dissemination of their results and in communicating with others: members of the scientific community, but also with the broader society and the business community. Researchers consequently have an incentive to publish in journals with a wide reach, high impact and good reputation.<sup>34</sup> In itself, this preference is natural. In the current system this is reinforced by the importance of the 'weight' of publications in shaping researchers' careers and prospects. Costs can also play a role: how do different publications compare? And who will pay? A researcher who does not have an institution (or funder) to support him may well decide not to publish if he has to pay himself.

Finally, a (large) proportion of researchers play a role in quality assurance, for example as journal editors or reviewers. Most researchers now do this, partly because it helps them find their place within the networks in their discipline. Most review work is unpaid or pays little; evidently being a reviewer is so highly valued by researchers that they are still willing to do this work without being paid for it.

How much 'power' do researchers have? As the author of an article, they can decide where they offer their article for publication. They also decide in which journal they wish to place their article. This means that the leading figures in a given discipline can send out a powerful signal as to which journals matter (in their eyes).<sup>35</sup> Researchers can also exert influence on the use of journals. They influence the scientific impact by what they download, read and cite. One difficulty is that if a researcher uses the impact of a journal from the past as a yardstick for prioritisation in the present, it will take a long time for shifts to manifest themselves. It should be noted that the effect of the actions of an individual researcher will generally be small. It is what researchers do together that matters most. A critical mass is needed if researchers wish to achieve change from the bottom up. Often, such a movement will be based on a broadly supported culture within the scientific community in question.

What does the above analysis of the interests and power of researchers as stakeholders mean for open access? In general, most researchers will in principle take a neutral or positive stance on publishing their article (or book) in accessible ways. For many researchers, however, the question of 'open access or not' is not the determining factor in choosing where to publish. That is reserved for factors such as the quality and scientific

---

<sup>34</sup> See De Goede & Hessels (2014) and the draft results of a global OECD survey of researchers: OECD (2015a).

<sup>35</sup> Dutch universities hoped to be able to make use of this mechanism in the call for a boycott of Elsevier journals in order to put pressure on the negotiations on access to Elsevier publications in the Netherlands; see the article in *NRC Handelsblad* of 2 July 2015, 'Eerste stap universiteiten boycot Elsevier' ('Universities take first step in Elsevier boycott'), and also the press release by the Association of Universities in the Netherlands (VSNU), 'Taalwetenschappers publiceren wetenschappelijke artikelen voortaan in betaalbaar Open Access' ('Linguists will henceforth publish scientific articles in open access' dated 12 October 2015 (VSNU 2015).

impact of a journal.<sup>36</sup> Although there are researchers who choose open access on ideological grounds, the majority will be pragmatic with regard to open access. Any costs involved will undoubtedly play a role. If publishing in a subscription journal is free for the author, choosing an open access publication for which the author has to pay a fee is less appealing.<sup>37</sup> A level playing field between these two options only exists if these author's publication fees have already been taken care of in some way,<sup>38</sup> or if the author has access to funds for open access publication.<sup>39</sup> In certain disciplines, especially in the arts, there are major concerns that researchers will not have access to such funds.<sup>40</sup>

If there is a strong culture of sharing and open collaboration within a given scientific discipline, open access has a real chance of becoming the norm, for example because a number of 'hot shots' link their names to an open access journal, inducing other researchers to publish in it. This does occur but is still an exception.

### Research institutes

Research institutes, such as universities, have an interest in ensuring that their researchers have the best possible access to global scientific literature. A research institute wants to make as much relevant information as possible available to its researchers and students at the lowest possible cost. Research institutes are, in turn, heavily dependent for their reputation on the quantity and quality of the scientific output of their researchers. These factors go a long way towards determining their position in all manner of university rankings. Thus, universities have an interest in ensuring that their research is published via channels with high (scientific) impact.

As an employer, a research institute can of course take a directive approach towards its employees, including researchers. A university can draw up rules on which publications will be included in the evaluation of its researchers (e.g. only open access publications, or only publications which are in the university repository). It can also make available funds for certain forms of publishing (e.g. to pay the author's fees for open access publishing).

---

<sup>36</sup> De Goede and Hessels (2014) find that professors – within the Netherlands – regard publishing in journals with (scientific) impact as the most important factor compared with publishing in general. Ware and Mabe (2015, p. 71) discuss several studies on what motivates scientists to want to publish in a particular journal, and conclude that the most important factors determining that choice are the quality of the journal, its relevance and the speed of the publication process. Ware and Mabe (2015, p. 71) do however also find that the open access character of a journal is gaining in importance as a secondary factor.

<sup>37</sup> The OECD survey (OECD 2015a, pp. 37-38) showed that fewer than 20% of authors are willing to pay a fee of around \$500. In practice, the fees payable for open access are often much higher than this – often between \$1,000 and \$5,000; see Ware & Mabe (2015, pp. 93-95).

<sup>38</sup> As for example, in the deal between Dutch universities and the Springer publishing group for 2015 and 2016. The article processing charges (APCs) for publishing have been paid collectively, which means that corresponding authors involved in the deal with Springer will no longer have to pay individually for the open access publishing of individual articles. (See: <http://www.vsnunl.nl/faq-open-access-nl.html>)

<sup>39</sup> This currently occurs only rarely; see OECD (2015a), pp. 28-29.

<sup>40</sup> These concerns were expressed repeatedly during a workshop organised by the European Commission on *Alternative Open Access Publishing Models* on 12 October 2015 in Brussels; see: <https://ec.europa.eu/digital-agenda/en/news/save-date-12-oct-ec-workshop-alternative-open-access-publishing-models> (most recently accessed on 16/10/2015).

Research institutes can also play an important facilitating role, for example by providing infrastructure such as repositories or support for their researchers in the publication process.

Universities and other research institutes are increasingly coming out in favour of open access to publications and are setting up a (public) repositories themselves.<sup>41</sup> In practice, the picture is mixed. There are universities which are enthusiastically embracing open access to publications and are doing a great deal to promote it, both in their policy and in practice (e.g. the Université de Liège);<sup>42</sup> but there are also more than enough universities adopting a wait-and-see attitude.<sup>43</sup>

### Research funders

Research funders pay for research for a particular reason. They may have a particular goal such as combating and finding a cure for a particular disease in the case of health funds. But their goal may also be to finance excellent scientific research, regardless of the topic. Private companies also fund research, often with a view to creating added value for their business (e.g. a patent, a new product or a better understanding of the market). The importance of research funders lies mainly in the achievement of their goals. If a company wishes to apply for a patent, it will have no interest in rapid publication of the research results, but will want to wait until the patent has been granted. Sometimes, companies do not wish to publish at all, as, for example in the case of commercially sensitive research. In other cases, companies may benefit from openness and collaboration.<sup>44</sup> A health fund, for example, will be keen on rapid and accessible publication of research results so that the new insights can be applied and developed further as quickly as possible. A funder of excellent research will want to see that excellence confirmed by publications with scientific impact. All research funders want to see the maximum result for the money they invest in research.

Research funders do have some 'power': in principle, they are able to set the parameters for the performance of 'their' research, which means that they can impose demands regarding the method of publication. Research funders can consequently play a key role in the move towards open access. Moreover, they possess resources, which they could use to support their preferred development in other ways, for example by founding their own open access journal, such as *eLIFE*, or by maintaining a repository, or perhaps by

---

<sup>41</sup> For Dutch universities, see the VSNU website: <http://www.vsnunl.nl/openaccess>. For universities in Europe, see e.g. the websites of the European University Association (EUA): <http://www.eua.be/policy-representation/research-innovation-policy/science-2-0-open-science> and the Open Access Statement by the League of European Research Universities (LERU 2015) dated 12 October 2015.

<sup>42</sup> See e.g.: Caruso *et al.* (2013), who write on p. 13 that universities and research institutes often lead the way in the national open access initiatives.

<sup>43</sup> The overview of open access mandates in ROARMAP (Registry of Open Access Repository Mandates and Policies) in the various countries shows that by no means all universities in the Netherlands and surrounding countries have in place a (mandatory) policy on open access publishing. See: <http://roarmap.eprints.org/>

<sup>44</sup> See e.g. the Structural Genomics Consortium (SGC): <http://www.thesgc.org/>

providing a comprehensive overview of all publications stemming from the research they have funded.<sup>45</sup>

As might be expected, the position of the different types of research funders on open access varies. Some are among the strongest advocates of open access.<sup>46</sup> An example is the largest national research funder in the Netherlands, the Netherlands Organisation for Scientific Research (NWO). NWO has accordingly announced plans to tighten up the rules on publishing its publicly funded research in the future, with the aim that open access becomes the norm.<sup>47</sup>

## Publishers

Publishers naturally look first to their own business interests. Scientific publishers, such as Elsevier, Springer or Brill, have spent many years building up their position and expertise. The processes surrounding scholarly publication have evolved; new services have developed, journal reputations have been built. They know the 'market' of authors and their readers. Publishers will undoubtedly seek to maintain this position and where possible to expand it. A further factor is the concentration that has taken place in the market for scientific communication, as a result of which the three largest publishers (Elsevier, Wiley, Springer) together publish almost half of all scholarly articles that appear in journals.<sup>48</sup> Partly because of this oligopolistic market, the major publishers are able to generate high profit margins from their activities in scholarly communication. They have moreover succeeded in pricing their subscriptions in such a way that institutes are in reality forced to buy a total package from a publisher ('big deal').

For publishers, the 'traditional' business model in which the reader pays for access to publications has proved itself. If a different model is used, publishers will have to find different ways of earning money. The most obvious is that the author will have to pay to publish. In that case, the more attractive the journal for an author, the higher the fee a publisher can demand. As long as the journal rankings remain important to researchers, publishers will be in a strong position. One risk for publishers is that any constraints on the budgets available to researchers to publish their own articles, will reduce the possibility for high fees.

Some scholarly journals are published by 'learned societies': organisations of researchers and other stakeholders in a particular discipline. Membership of such a

---

<sup>45</sup> As the body that funds research in São Paulo state in Brazil, FAPESP, does, for example.

<sup>46</sup> Caruso *et al.* (2013) cite the Wellcome Trust in the United Kingdom and the National Institutes of Health (NIH) in the United States as examples. Our own research also reveals FAPESP in Brazil to be a good example. FAPESP is also the driving force behind SciELO (Scientific Electronic Library Online), the most important open access library in South America.

<sup>47</sup> See: <http://www.nwo.nl/en/policies/open+science> (most recently accessed on 16/10/2015).

<sup>48</sup> See e.g.: The Publishers Association, International Association of Scientific, Technical and Medical Publishers and The Association of Learned & Professional Society Publishers (2010).



learned society can include a subscription to one or more published journals, and for many this will be a reason for becoming a member. In addition, any income from their journals is an important source of income for these organisations, which can be used to fund their (other) activities. On the face of it, such learned societies would seem to have little financial incentive to switch to a model of (gold) open access.

Naturally, publishers have a great deal of scope to change the practice of scholarly publication, as they are the linchpins in this process. A publisher can decide whether to provide open access, can set up new open access journals, convert existing journals to open access, or allow immediate open access to individual articles in subscription-based journals. Publishers can lay down what options authors have for placing their articles in open access repositories. There are also many other ways in which publishers can innovate.

It will come as no surprise that most publishers are generally not taking a lead in the push to completely revamp the present publishing system. They will, moreover, seek to steer any transition process in directions that are beneficial for them and will want to (continue to) respond to supply and demand in the scientific world. But publishers are aware that open access is a development that has momentum. One in which publishers will participate as long as it does not harm their own position. And there are fully open access publishers, such as BioMedCentral, or publishers which are leading the field with all manner of innovations in relation to open access and scholarly publishing, such as Pensoft.

### **Libraries**

We are concerned here principally with libraries at universities and other research institutes. These libraries have traditionally played a role in providing access to global information for their researchers and students. Their importance is rooted in the fact that they make available as many relevant publications as possible. They do this by subscribing to scholarly journals, databases, etc., which means they are in the front line in the face of rising subscription and other costs. It is therefore in libraries' interests that this development is stemmed. In addition, in recent years several libraries have taken on the role of managing research institutes' repositories, which serve to make the research results produced at the institute accessible worldwide.

As libraries often decide centrally which journal subscriptions are taken out and which (other) publications are purchased, they are – in theory – in a pivotal position to determine which (types of) publications are contracted (and which ones are not). In practice, however, libraries take into account – formally or informally – researchers' wishes regarding what is relevant. It is unlikely that a library would restrict itself to 'free' open access journals, if relevant journals in a discipline are subscription-based. It is now

common practice in the Netherlands for universities to negotiate as a block with the major publishers and to agree a 'big deal' with each publisher. This has strengthened the libraries' negotiating position.

University libraries have traditionally led the way in open access. At first sight, it would seem they stand to benefit from open (and free) access – especially in times of rising subscription costs – if this means that their subscription costs are reduced or eliminated. On the other hand, libraries will have to adapt.<sup>49</sup> It is also possible that libraries will take on new roles which entail extra costs, such as maintaining a repository.

## Businesses

Private sector companies have a variety of goals, but ultimately a business has to earn money. Sometimes, companies carry out or fund research themselves. The position of companies as research funders was discussed above. Businesses are also (potential) users of scientific results, though this will apply for only a small proportion of them. A business will generally be most interested in obtaining access to relevant information. Determining what actually constitutes relevant (scientific) information will probably be the most important step in that process. Large corporations with large research departments may be able to resolve this internally (they often subscribe to scholarly journals), but smaller and medium-sized businesses (SMEs) will need to find the right 'filter'. They often resolve this by approaching an expert, such as a researcher, who briefly explains the current scientific state-of-the-art and points them to the most relevant publications.<sup>50</sup> Businesses would of course then like to be able to access this knowledge at no cost, but paying for targeted access to specific publications will not be an insurmountable obstacle for most businesses.<sup>51</sup> Not all companies may have staff who are capable of fully understanding scholarly articles. Those most likely to have such staff will mainly be specialised, for example in the biotechnology industry. For most other businesses, a simplified summary written for the interested lay reader would probably provide enough information.

What 'power' do businesses have in their role as consumers of scientific literature to bring about changes in the practice of scholarly publishing?<sup>52</sup> They are customers who purchase scientific literature and are therefore in a position to take out different or no

---

<sup>49</sup> For example, the transition from paper to digital also demanded a significant change in the working procedures of university and other libraries (this came out during our discussion with Susan Reilly from LIBER).

<sup>50</sup> This observation is based on numerous interviews conducted in compiling this report (see Appendix 3 for the complete list of interviewees), and on the findings that emerged from the project currently being carried out by AWTI on knowledge absorption capacity (that report will appear in early 2016).

<sup>51</sup> This was confirmed in a number of our interviews, including with the Confederation of Netherlands Industry and Employers (VNO-NCW) and the SME association Royal Association MKB Nederland. N.B.: We are concerned here with paying for access to a small number of articles after the most relevant have been selected; this underlines yet again the importance of helping businesses in making this selection.

<sup>52</sup> The role of businesses as funders of research has already been discussed in the section on research funders.

subscriptions) wherever appropriate. An indirect option is their choice of which institutes or researchers they are willing to work with. If that choice is based partly on the (read) publications of researchers, this could work to the advantage of researchers who publish open access articles, because their articles are then more likely to be read and studied by business users.

In their role as (potential) consumers of scholarly publications, businesses stand to benefit from open access. If scholarly journal articles are freely available, large corporations will no longer need to spend money on subscriptions, and small and medium-sized enterprises will also be able to access those publications. At present, SMEs have to pay to access this information or access it indirectly. A move towards open access would therefore be of particular benefit for this latter group in giving them wider and easier access to scientific results.<sup>53</sup>

### **Civil-society organisations and citizens**

Society helps to pay for scientific research from public funds (taxpayers). It is therefore reasonable for science to find some way to be accountable for what it does with that money. One way would be to make scientific output accessible to the public and to their organisations

Research has shown that some members of the population are genuinely interested in reading articles themselves. For example, 13% of the Dutch population aged 25 and over say they would be definitely interested in access to scholarly articles; that is a total of 1.5 million people.<sup>54</sup> This includes people who would be able to make use of scientific knowledge in their own profession, such as doctors or teachers. Civil-society organisations such as patient associations or environmental organisations, may also benefit. As with businesses, creating effective access to scientific results will usually require an intermediate step, for example a portal where research results are grouped and which provides a comprehensible explanation, or a lay summary of each article. Without this, it is uncertain how much use many citizens and organisations will be able to make of the original scholarly publications, which are generally written for scientific peers and will to non-specialist readers come across as highly technical and specialised. For example, the editor of a membership magazine for a patient association stated that they generally ask the researchers or research institute concerned to provide a summary of the relevant (scientific) publication that is digestible for lay readers.<sup>55</sup>

The 'power' of citizens and their organisations to bring about changes in the world of scholarly publishing is an indirect power. They can exert influence through the political

---

<sup>53</sup> Such a move would however mean that large corporations would lose a (potential) information lead over SMEs.

<sup>54</sup> TNS Nipo and Koninklijke Bibliotheek (2014).

<sup>55</sup> Interview with Dick de Ruiter, editor-in-chief of the Lynch Polyposis Contact magazine in August 2015.

system or by putting direct pressure on stakeholders. If public use of scholarly publications were given more weight (for example by including a ‘social impact factor’ when evaluating the work of researchers), this would provide citizens and civil-society organisations with a lever to make changes, because open access publications are highly likely to have more social impact through downloading, etc..

As potential ‘users’ of scholarly publications, citizens and civil-society organisations will likely benefit from open – and therefore easier – access to those publications. As stakeholders, they will also have views on the benefits and costs of open access. If open access, on balance, leads to higher costs (for the scientific world), these will have to be set against the benefits of improved public access to publications.

### 4.3 Public authorities and open access

The government is a relevant actor in its role as regulator. The government is also a funder of research; that role was discussed in the previous section. As a regulator, the government sets the framework for science and therefore exerts influence on the publishing of scientific results. In this section we analyse what the various governments consider important, what scope for action they have and what their position (policy) on open access ultimately is. These findings are subsequently linked to the rest of the analysis to determine what the (Dutch) government’s goal should be in relation to open access and what action it ought to be taking (see also Chapter 3).

#### UNESCO

At global level, the UN organisation UNESCO plays a particularly important role in the open access debate. UNESCO is a great advocate of open access, which it sees as contributing to better global dissemination of knowledge, strengthening innovation and promoting socio-economic development in general.<sup>56</sup> Open access is one of the three pillars of UNESCO’s policy on strengthening access to scientific knowledge. The other two are Open Software (FOSS) and Open Educational Resources (OERs).<sup>57</sup>

The tools that UNESCO has at its disposal are limited. It does not have the power to implement open access through regulation. It focuses primarily on awareness-raising, knowledge dissemination and gathering (see for example the Global Open Access Platform<sup>58</sup>), stimulating debate, supporting local and regional initiatives and providing

---

<sup>56</sup> <http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/open-access-to-scientific-information/>

<sup>57</sup> UNESCO Netherlands National Commission for UNESCO (2011)

<sup>58</sup> The Global Open Access Portal (GOAP) was founded under the auspices of UNESCO. This platform provides an overview of (the progress in) the degree of openness of scientific information throughout the world, and also analyses best practices and the problems that countries may face: <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/>

training (see for example the UNESCO OA Curriculum<sup>59</sup>). The UNESCO OA programme concentrates mainly on Africa and other developing countries where, despite major strides in ICT facilities, knowledge dissemination remains limited.

Perhaps at least as important for UNESCO is setting a good example. In 2013, UNESCO announced that from July of that year, all its publications and data would be placed in an open access repository with a multilingual user interface. The intention is to place all earlier publications in the repository as well.<sup>60</sup> Shortly afterwards, other international organisations, such as the WHO, OECD and the World Bank, followed with their own internal open access policy.

### European Union

Moving down from the global level brings us to (large) countries and regions such as the European Union. These regions can be divided into countries or, in the case of large (federal) countries such as the United States, India or Brazil, states. Control of scientific research is generally diffused across these different levels. Science is a 'world business' and the authorities that control it are highly fragmented. Moreover, in many countries the influence of the government on universities and similar organisations is constrained by the fact that universities often enjoy a certain degree of autonomy.

For this report, the European Union is the most relevant actor at this level. European unification has traditionally focused mainly on the market. However, in September 2000 the European Commission introduced the European Research Area (ERA) with a view to enabling Europe to offer attractive opportunities to researchers. The ratification of the Treaty of Lisbon in 2009 paved the way for creating a statutory framework in which to anchor the ERA. The ERA was intended to be the prime instrument in transforming Europe into a knowledge society where both scientific knowledge and researchers themselves can circulate freely.

In 2012 the European Commission followed with a Communication on 'A Reinforced European Research Area Partnership' in which the first priorities were identified.<sup>61</sup> To develop the European Research Area further, in 2014 the EU ministers with responsibility for research agreed that an ERA Roadmap should be introduced at EU level by mid-2015;<sup>62</sup> the roadmap was ratified by the Council in May 2015. It lists eight priorities,

---

<sup>59</sup> A set of online training sessions for researchers and libraries on numerous aspects related to open access: [http://www.unesco.org/new/en/communication-and-information/resources/news-and-in-focus-articles/all-news/news/unescos\\_open\\_access\\_oa\\_curriculum\\_is\\_now\\_online/#.Vcn0MrXweAg](http://www.unesco.org/new/en/communication-and-information/resources/news-and-in-focus-articles/all-news/news/unescos_open_access_oa_curriculum_is_now_online/#.Vcn0MrXweAg)

<sup>60</sup> UNESCO (2013).

<sup>61</sup> Communication from the Commission dated 17 July 2012, 'A Reinforced European Research Area Partnership for Excellence and Growth, COM(2012) 392 final.

<sup>62</sup> <http://www.neth-er.eu/nl/dossiers/onderzoek-en-innovatie/era>; Conclusions of the Council of February 2014 on progress in the European Research Area (EOR) 2013, Doc. 6945/14.

including promoting open access.<sup>63</sup> Based on these priorities, each of the Member States will draw up its own national roadmap for the implementation of the ERA.<sup>64</sup>

This attention for open access is not surprising. For some time, the European Commission has been a major proponent of open access to publicly funded research results (both publications and data). The Commission believes that this is important for the circulation of knowledge and for innovation. In the first instance, the Commission's preference manifested itself in its policy as a funder of research, about which more later. In 2012, this preference also led to a set of general policy principles set out in the Communication 'Towards better access to scientific information: Boosting the benefits of public investments in research'<sup>65</sup> and in a Recommendation on 'access to and preservation of scientific information'<sup>66</sup> The Commission stressed that: "The European Commission's vision is that information already paid for by the public purse should not be paid for again each time it is accessed or used, and that it should benefit European companies and citizens to the full. This means making publicly-funded scientific information available online, at no extra cost, to European researchers, innovative industries and citizens, while ensuring long-term preservation."<sup>67</sup>

The European Commission has also published recommendations to the Member States to encourage them to take policy steps of their own towards making publicly funded research results freely accessible.<sup>68</sup> The Commission has recommended that Member States develop policy to further open access to scholarly publications and data, storage and reuse of scientific information, and E-infrastructures. It is important that Member States first define clear objectives complemented indicators to measure progress, with implementation plans including allocation of responsibilities and financial planning. Member States are asked among other things to focus their open access policy on the functioning of the licensing system, the academic careers system, and on securing the broadest and cheapest possible access to scientific results for SMEs.

The European Commission has not opted for a directive approach, for example by creating legislation to force Member States to move towards open access. Instead, it has chosen an open dialogue with the various national, European and international stakeholders.<sup>69</sup> The Commission has largely placed responsibility with the Member

---

<sup>63</sup> ERAC (2015), Priority 5b focuses on open access.

<sup>64</sup> <http://www.neth-er.eu/nl/nieuws/Raad-Concurrentievermogen-bekrachtigt-ERA-Roadmap>

<sup>65</sup> Communication from the Commission dated 17 July 2012: 'Towards better access to scientific information: Boosting the benefits of public investments in research', COM(2012) 401 *final*.

<sup>66</sup> Recommendation of the Commission dated 17 July 2012 on access to and preservation of scientific information, (2012/417/EU), *PbEU* 2012 L 194/39.

<sup>67</sup> European Commission (2013b), p. 4.

<sup>68</sup> Recommendation of the Commission dated 17 July 2012 on access to and preservation of scientific information, (2012/417/EU), *PbEU* 2012 L 194/39.

<sup>69</sup> This choice is extensively argued in the European Commission's impact assessment (2012).

States themselves. Part of the reason for this approach is that Member States are in different phases, examples follow later.

First, we outline how the European Commission has acted in its role as a research funder in relation to open access. The first major steps were taken in 2006/2007, when the first reports and guidelines were published by the European Research Council.<sup>70</sup> It was also decided to launch an open access pilot as part of the Seventh Framework Programme (FP7), which covered 20 percent of the budget. Recipients of grants were expected to make their publications freely accessible within six or 12 months, depending on the discipline. Another key aim of FP7 was to improve the e-infrastructure for storage and access. The OpenAIRE project has played a central role here.<sup>71</sup> In the European Commission's successor to FP7, the Horizon 2020 (H2020) programme, open access publishing has become the norm, and if necessary funding will be made available for this.<sup>72</sup>

As stated, the EU Member States take slightly different positions on open access. We will discuss the position of the Netherlands first and then other relevant countries. Appendix 2 contains a detailed summary of the situation in different countries.

### **The Netherlands**

The Dutch State Secretary for Education, Culture and Science, Sander Dekker, is a strong advocate of open access to publications. His principle is that "the results of publicly and publicly/privately funded research should always be freely accessible."<sup>73</sup> He believes that open access fosters the sharing and circulation of knowledge, in turn contributing to the innovative capacity of the Netherlands. The government's initial goal is that stakeholders work together to speed up the progress towards open access, without additional government funding. Only if stakeholders do too little or fail to achieve adequate results will the State Secretary consider bringing in legislation to make open access publishing mandatory.<sup>74</sup> The target is that 60% of publications should be open access by 2018, and 100% by 2024.<sup>75</sup> Achieving this target will require a better monitoring system, because at present every new study comes up with a different open access percentage. The State Secretary initially expressed a clear preference for the 'gold' open access model, in which articles are immediately freely accessible on publication. Since then, the State Secretary appears to have come to see 'immediate

---

<sup>70</sup> [http://erc.europa.eu/sites/default/files/document/file/erc\\_scc\\_guidelines\\_open\\_access.pdf](http://erc.europa.eu/sites/default/files/document/file/erc_scc_guidelines_open_access.pdf)

<sup>71</sup> <https://www.openaire.eu/>

<sup>72</sup> European Commission (2013b).

<sup>73</sup> Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 15 November 2013, TK 2013-2014, 31 288, no. 354, p. 1.

<sup>74</sup> Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 15 November 2013, TK 2013-2014, 31 288, no. 354, p. 7.

<sup>75</sup> Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 23 January 2015: TK 2014-2015, 31 288, no. 414, p. 4.

green' open access as acceptable as well.<sup>76</sup> Although under this system the journal version of the article is still only available to subscribers, the author simultaneously places a version in a repository that is accessible to all (with no embargo).

A new clause (25fa) in the Dutch Copyright Act (*Auteurswet*) came into force on 1 July 2015. It stipulates that authors of a scholarly article funded wholly or partly with *Dutch* public resources have the (inalienable) right '(after a reasonable period of time following the initial publication to make that article freely available to the public, provided the source of the first publication is clearly stated'.<sup>77</sup> By giving researchers this right, authors are put in a stronger position vis-à-vis publishers. Yet there is a lack of clarity in the implementation of the clause. For example, the researcher and the publisher of his article have to negotiate on what constitutes a reasonable period. It is also unclear whether the version to be made accessible to the public may/must be the same as the published (journal) version. While researchers to whom this clause applies will be able to derive protection from it, a number of practical questions need to be resolved.<sup>78</sup>

## United Kingdom

Like the Netherlands, the United Kingdom is a leading player in the field of open access. The UK is the largest 'producer' of scholarly articles in the EU. There are several mandates in the United Kingdom which make open access mandatory, some prescribe the 'gold' model and others the 'green' model. In the past, the government made extra funding available for open access publishing. However, this did not lead to spectacular growth of open access. What did happen was that extra costs suddenly emerged. A detailed summary of developments in the United Kingdom can be found in Appendix 2.

## Scandinavia

The Nordic countries also take a fairly positive view of open access. The Finns have a significantly broad strategy focusing on openness in science and have developed a detailed roadmap for open access. The Danes also have a strategy in place; they have opted for the 'green' route as the 'gold' route would demand a system change that is difficult for small countries such as Denmark to enforce. One problem in Denmark is that open access is a topic that researchers are barely engaged in. Sweden has a policy of open access to publicly funded publications, with the possibility of an embargo period of six or 12 months, depending on the discipline. In Norway, the National Research Council is currently making funds available for open access publishing.

---

<sup>76</sup> In his letter to Parliament dated 23 January 2015 (TK 2014–2015, 31288, no. 414, p. 3), he refers to the decision by the NWO to tighten up the funding conditions for open access, with the NWO starting from a preference for the 'golden road' for open access, with the 'green road' as the alternative option provided no embargoes are applied.

<sup>77</sup> This clause is based on an amendment tabled by Dutch MP Taverne: Kamerstukken TK 2014–2015, 33 308, no. 11. Further details of the government position were presented to the Upper House in the Memorandum to the report (Kamerstukken EK 2014–2015, 33 308, no. E).

<sup>78</sup> See also: Visser (2015).



## Germany

After the United Kingdom, Germany is the second largest 'producer' of scholarly articles in the EU, but lags behind in the process of open access. It is a federal state, with responsibility for higher education and research delegated to the *Bundesländer*. As a consequence, the picture is highly fragmented and varied.<sup>79</sup> A relatively high proportion of scientific research in Germany is carried out by non-university institutes, such as the Max Planck institutes. The Max Planck Gesellschaft is a strong advocate of open access. Germany also has copyright legislation under which an author of a scholarly publication based on research that was at least 50% publicly funded has the right to make the manuscript freely accessible 12 months after publication.<sup>80</sup> Although there are undoubtedly a number of enthusiastic pioneers of open access in Germany, the statistics show that, on average, Germany lags behind in this area.<sup>81</sup>

## France

France occupies an intermediate position. Initially the French adopted a wait-and-see approach, but recently there has been movement. The country is strongly focused on a system of repositories, to which they are seeking to provide central access via HAL (*Hyper Article en Ligne*). This appears to show a preference for the 'green' approach, although the government has not formulated such an explicit position. France is the leader in the French-speaking scholarly publishing world, and is home to the most important publishers in that segment.

## Southern Europe

Spain stands out amongst Southern European countries, with a law stipulating that researchers must deposit a most recent digital version of a (publicly funded) scholarly article in a repository (either institutional or thematic) within 12 months of its publication.<sup>82</sup> Like France, Spain focuses heavily on repositories, though does not rule out the 'gold' route. The same applies to most other countries in Southern Europe. The general picture across these countries does not suggest that they are taking leadership in relation to open access. However, open access could be interesting for these countries as, since the recent economic crisis, their research budgets, and consequently their access to knowledge, have been under serious pressure.

## Central and Eastern Europe

Among Central and Eastern European countries, we received input from Latvia, Slovakia and the Czech Republic. The picture that emerges is that these countries are currently still working on aligning their science sectors with the top European research countries.

---

<sup>79</sup> For a detailed overview, see Appendix 2.

<sup>80</sup> §38(4) *Urheberrechtsgesetz*.

<sup>81</sup> Archambault *et al.* (2014).

<sup>82</sup> Article 37 of the *Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación*.

For some countries, therefore, it is more important to ensure that their researchers are able to publish in high-impact journals than whether or not they have open access. Also, since most scholarly publishers charge lower fees in less affluent countries, some of these countries are currently benefiting from relatively low subscription charges. There are fears that a move to gold open access would lead to the loss of these benefits if everyone then had to start paying the same authors' fees (APCs).<sup>83</sup>

### Switzerland

The Swiss government supports open access, whether green or gold. There are a number of important issues for the government, namely proper quality control and coordination between science and industry with a view to the use of research results. The Swiss government largely leaves the initiative to the knowledge institutes themselves. Notably, the Swiss science funder SNSF makes open access publishing (green or gold) mandatory and makes funding available. Since July 2014, mandatory open access publishing has also applied to books, though an embargo of up to 24 months can be stipulated.

### United States

The United States is the world leader in science, in terms of both quantity (number of publications) and quality (impact of publications). Moreover, many of the world's top research institutes are based in the US, and the country attracts top scientific talent from all over the world. A key point with regard to policy is the federal structure of the United States. The federal government (as research funder) has passed regulations on open access.<sup>84</sup> Those regulations stipulate that all research funded from the federal budget to the tune of more than \$100 million must be published with open access, in principle using a limited number of permitted repositories. In effect, therefore, the US has chosen the green route. These regulations have had a major impact; the research carried out by the National Institute of Health (NIH), for example, accounts for a budget of \$30 billion per year. PubMed Central plays a key role here, and the number of publications deposited in this repository has grown strongly since the NIH policy came into effect.<sup>85</sup> Legislation at the level of individual states is currently limited, though initiatives are being rolled out in a number of states where large amounts of research are carried out (California, New York and Illinois). Private funds (charities) are extremely important in the US. Some of these funders, too, advocate or stipulate open access, such as the Bill & Melinda Gates Foundation.<sup>86</sup> Finally, many universities have an open access policy, including a number of major establishments, such as Harvard.

---

<sup>83</sup> This point was raised during our meeting with Slobodan Radičev from Serbia on 12 October 2015.

<sup>84</sup> This is laid down in a policy memorandum from the *Office of Science and Technology Policy (OSTP)*; see: [www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research](http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research)

<sup>85</sup> OECD (2015b).

<sup>86</sup> See also ROARMAP for an overview: <http://roarmap.eprints.org/>

## Brazil

Brazil is just about the world leader in open access. The biggest funder of research, FAPESP, the research fund of the state of São Paulo, has made open access publishing mandatory. Nationally, at least 40% of scholarly articles are published using the gold open access route, a very high figure by international standards.<sup>87</sup> The government has tried to bring in legislation, but ultimately withdrew the bill.<sup>88</sup> On its website, FAPESP keeps a very clear record of which publications have appeared on different research topics (and they can be accessed immediately). Brazil is also home to SCIELO (the Scientific Electronic Library Online), the main South American repository.<sup>89</sup>

## 4.4 The impact of open access

To determine the impact of open access, we looked at the position of relevant stakeholders with regard to scholarly communication. What drives them? What do they consider important? What can they do? And what does this mean for their position on open access? All these stakeholders are small cogs in the greater system of science and its embedding in society and the business community. Can we also say something about the impact of open access at system level? What benefit can 'we' all derive from it (or not)? This analysis enables us to say something about the desirability of (the degree of) open access from the perspective of general interest. This is an important factor when the government considers taking actions. In our assessment framework, we look at the impact of open access on the quality of science, the impact of science within society and the business community, and at the magnitude and distribution of the costs and benefits.

### 4.4.1 Impact on the quality of science

#### Quality assurance of publications

In the most common variants of 'gold' and 'green' open access, scholarly journals as they now are continue to occupy a central place in the publication process. The review process will not really change in either variant of open access, with peers still commenting on a draft version of an article. In the present system, this process is highly institutionalised via the journals. They have a strong interest in ensuring that the system works well, as it contributes to the (perceived) quality of the journals (and therefore to their 'value'). This quality incentive is strong in systems where the reader pays to access articles or journals. In variants where the author has to pay to publish (often in the 'gold' route), the incentive for quality to drive the acceptance process is much weaker. The form of quality assurance now in place may therefore come under pressure, which is a point of

---

<sup>87</sup> Archambault *et al.* 2014, p. 26.

<sup>88</sup> Caruso *et al.* 2013, p. 4.

<sup>89</sup> On SciELO and its history, see Packer *et al.* (2014) and Alperin (2015).

concern. Publishers will of course have to ensure that their journals remain attractive to authors looking to publish. High quality and impact contribute to this. Therefore, things may not go wrong, but they could. In the worst case, a system based on ‘the author pays’ could become a perverse incentive. Publishers and others will naturally continue to adapt and improve the review process,<sup>90</sup> but that is a separate issue from whether or not articles are published with open access. An alternative might be a system in which ‘provisional’ results are published with immediate open access, following which comments from others provide a kind of review (which ultimately results in a certain quality opinion). This is somewhat comparable to the practice of preprints in physics or working papers in economics.

A similar analysis can be performed for the incentive to accept more or fewer articles. In the traditional model where the reader pays, it is mainly the combination of a journal’s quality and relevance which determine the (potential) subscription income. In a system in which the author rather than the reader pays (after acceptance of an article), the publisher’s income increases as more articles are accepted and published. This system puts a much weaker ‘brake’ on publishers regarding the number of articles published, and could fuel an increase in the number of scholarly publications. However, in a system where the author pays, research publication budgets could have a limiting effect on the number of publications. How these two effects may unfold in practice is difficult to predict, and will also depend on other factors. Clearly, if authors pay, their publishing budget will determine how much they can publish and in which journals. It is, at the least, questionable whether this is a desirable development.

### **Quality indication**

Scholarly publications play an important role in determining the ‘quality’ of research, researchers and research institutes, for example in evaluations and in awarding grants. Researchers and research institutes therefore pay close attention to the (scientific) reputation of the journals in which people have published, and to the impact of individual publications (e.g. in terms of citations) and to how much is published in total. We might call this the ‘scientific impact’, as these indicators focus on the importance of publications within the scientific community.

From the perspective of improving the use of scientific knowledge, it is relevant to look at the ‘social impact’ of publications as well, in other words their impact beyond the world of science. This is discussed in the section on impact on society and business.

In principle, both scientific and social impact are separate from the question of open access publishing. The reputation of a journal will always be based on its quality and

---

<sup>90</sup> The *open access journal* eLife is for example experimenting with different forms of review.

relevance. These will often depend on the selection process used by the journal (the stricter the better) and the impact of articles (the more they are read or cited, the better).

### **Impact on integrity**

Open access means that more people have access to publications, which could bring poor scientific practices to light earlier. In this sense, open access may contribute to (improving) scientific integrity. Open access to the research data underlying an article is likely to make an even larger contribution in this regard. This topic – data sharing – is discussed in the next chapter.

### **Effectiveness of knowledge dissemination**

For effective dissemination of scientific knowledge, it is vital that research results reach the right readers, and, conversely, that someone looking for literature is able to access the relevant knowledge easily. The present system of scholarly journals and discipline-specific conferences creates focus and therefore promotes the effectiveness of dissemination. In addition, the ranking of journals helps readers prioritise during a literature review. On top of that, publishers provide tools for effectively searching in their extensive portfolio of journals.

In the present system, dissemination of the results of interdisciplinary research is, however, less effective. Monodisciplinary publicity channels are often the best developed and have the most impact. Researchers do not always have access to publications in other disciplines. Open access may certainly help to make literature more accessible across disciplines.

Some of the repositories, such as the broad repositories operated by universities, lack the subject-specific focus of the journals. It is generally more difficult to find specific publications in these repositories. More generally, if publications are stored and made accessible locally (e.g. in repositories), it will be more difficult to match up the supply of and demand for scientific knowledge. Open access publications in repositories really only become effective if they can be readily found, for example if there is a central ‘gateway’ (portal) and a means of searching through the many local storage sites.<sup>91</sup>

Knowledge can of course only be effectively transferred if the publication containing it can actually be read. In a system of gold open access, everyone can do so, because that is inherent in the system. With a system of green open access, everyone can read publications archived by the researchers after a certain (embargo) period has expired,<sup>92</sup> but before then they will have to pay for access (subscription or one-off payment per

---

<sup>91</sup> Examples include the Scientific Electronic Library Online (SciELO) in South America (<http://www.scielo.br/>), the French portal HAL (<https://hal.archives-ouvertes.fr/>), and CHORUS in the US (<http://www.chorusaccess.org/>).

<sup>92</sup> For an overview of the usual embargo periods per discipline, see OECD (2015b), p. 24.

article). The present system means that people are sometimes denied access to a publication because they do not have a subscription (though in these situations researchers can often gain access to relevant publications indirectly). Within the scientific community, open access mainly benefits researchers who currently have no or only limited access to scholarly publications. This will generally be the case in less affluent countries or less well-off institutes, or in situations where people try to gain access to literature from disciplines other than their own (or that of their institution).

An important indicator of the effectiveness of knowledge transfer through a specific article is the extent to which that article is subsequently cited. Several studies have been carried out on the impact of open access on citations.<sup>93</sup> Broadly speaking, these studies conclude that the fact that a publication is available with 'open access' has a positive effect on citations by non-scientific publications (such as policy reports), while the impact of open access in terms of scientific citations depends on the specific circumstances: it can be either positive or negative.<sup>94</sup> The impact of 'open access' is difficult to measure (due to the need to control for many variables). Ware and Mabe (2015, pp. 130-131) discuss several studies and conclude that: "The effect is still unclear, but best available evidence at this point tends to suggest that open access articles in the aggregate probably do not receive more lifetime citations, but they do get them sooner due to early view and selection bias effects." The latter refers to a tendency of authors to grant open access earlier to their better-quality works. In practice, the citation score of articles published in 'gold' open access journals is on average much lower than the score of articles in non-open access journals. This is probably linked to the fact that many open access journals are young and have not yet been able to build as strong a reputation as some traditional scholarly journals.<sup>95</sup> As regards open access books, there is no evidence that open access affects sales or citation scores. There however is evidence that online use through 'book visits' and page views on Google Books does increase due to open access.<sup>96</sup>

### **Speed of knowledge dissemination**

The speed of knowledge dissemination is also a relevant indicator for the functioning of science. It is a familiar fact that scholarly publishing is traditionally a lengthy process. Time is needed for peer reviews and amendments, and only then does the production process begin. Whilst this benefits the ultimate quality, if new insights cannot be published during this period, then clearly the dissemination of those insights is delayed. (This is particularly relevant in medical science, for example). In practice, alternatives

---

<sup>93</sup> For a discussion of the most relevant studies, see Ware & Mabe (2015), pp. 130-131.

<sup>94</sup> See also: OECD (2015a), pp. 30-31. Researchers themselves generally do not believe that open access leads to more (scientific) citations, according to a study by Taylor and Francis (2014).

<sup>95</sup> See e.g. CWTS (2015) for Denmark, the Netherlands and Switzerland. Archambault *et al.* (2015) also find this result for gold open access. See Ware & Mabe (2015, pp. 130-131) for a more detailed overview of the most relevant studies.

<sup>96</sup> Ferwerda *et al.* (2013), p. 58.

have emerged such as working papers (e.g. in economics) and preprint servers (for example in mathematics and physics). These enable researchers to share their ideas before the definitive article is published. However, it also occurs in the present age of online publishing and competition that in order to be the first to make a discovery public, the whole process (review, acceptance and publication) sometimes has to be completed within a very short space of time (e.g. a matter of days).

How does open access affect the speed of dissemination? In the present system, publications in subscription-based journals are accessible to subscribers – and only subscribers – immediately after publication. No one else has access. In the gold open access model, an article is accessible to everyone immediately after publication. With green open access, it is partially accessible after publication, to subscribers, and to others only later, after an embargo period.

#### **4.4.2 Impact on society and business**

##### **Utilisation of scientific knowledge by businesses**

Having greater access to scientific publications through open access will of course not make the utilisation of knowledge by businesses worse, but the real question is what the decisive factors are for *better* utilisation. At present, larger businesses generally have access to scholarly publications because they subscribe to relevant journals. The problems with access are generally in the SME sector. Open access will undoubtedly make access easier, but the question then is what impact this will have on the utilisation of this knowledge.

In a pilot project carried out by Elsevier in the Netherlands a group of SMEs were given access to 13 million Elsevier articles. The result was that each company consulted an average of one article per week, though there were sectors where this figure was higher.<sup>97</sup> Simply providing access to a database is therefore probably not enough; there are too many articles and they are too complex. What is needed is help in searching for the right articles and having those articles presented in the right format. Ancillary actions are therefore needed. A sensible approach would be to follow the mechanisms by which companies absorb other external knowledge, a topic on which AWTI will publish a report in 2016. Open access could be embedded in such a broader strategy.

##### **Utilisation of scientific knowledge by society**

While it may be true that many researchers already have direct or indirect access to a great many scholarly publications, this applies to a much lesser extent in society at large. In reality, only open access articles are fully accessible (though an abstract will be

---

<sup>97</sup> Information taken from interviews with Messrs Kolman and Berghmans from Elsevier in July 2015.

available for most other articles). Open access thus may have an enormous impact on the ability of citizens and civil-society organisations to read about scientific results in journal articles. This general interest in scientific literature is indubitably real, as evidenced for example by a study carried out by the National Library of the Netherlands,<sup>98</sup> or the usage statistics of PubMedCentral.<sup>99</sup>

However, there is no linear progression from the public's wish to receive knowledge to more knowledge transfer and utilisation of scientific results in the public domain. Scientific literature is highly specialised, being written for other scientists in the same discipline. Society is likely to benefit more from generalist publications, such as trade journals, lay summaries<sup>100</sup> or platforms which summarise scientific knowledge and place it in context.<sup>101</sup> These 'intermediaries' can prompt readers to (want to) read the underlying scholarly publications.

### **Social impact factor**

As with the impact on business, genuinely increasing the impact of scientific endeavour on society requires more than open access alone. Without systematic attention to 'bridging the gap' between science and society, much potential value of scientific results for society will be unused or underused, even if access is open. One way of dealing with this problem is to include a 'social impact factor' when evaluating researchers, in addition to the scientific impact factor. This social impact factor should also be used by the scientific world itself (and would then play a part in researchers' careers). Although a 'social impact factor' is in principle a separate issue, it may fit in well with a world with open access. It would then for example be possible to look at the number of times an article (or book) has been downloaded, possibly differentiated by the type of reader (researcher, member of the public or business). Or other indicators of social impact could be developed.

Other ways of facilitating social impact include devoting more attention to lay summaries<sup>102</sup> or short films, as for example in the competition "Ma thèse en 180 secondes" for the French-speaking world, in which doctoral students are asked to explain their theses in 180 seconds.<sup>103</sup>

---

<sup>98</sup> TNS Nipo & Koninklijke Bibliotheek (2014): 13% of people in the Netherlands aged over 25 years are interested in reading scientific literature; 9% already do so and the remaining 4% are interested but are not yet doing so in practice. The latter group therefore constitute potential 'new' users: almost half a million people.

<sup>99</sup> According to UNESCO (2012), 40% of the daily unique users of PubMedCentral are individual citizens (this figure is cited in OECD, 2015, p. 11).

<sup>100</sup> This was for example also the central message in our interview with the editor-in-chief of Lynch Polyposis Contactblad, a Dutch membership magazine published by a patient association.

<sup>101</sup> A good example is the Cochrane Library ([www.cochranelibrary.com](http://www.cochranelibrary.com)), which includes not just publications, but also systematic reviews of different studies of comparable subjects, provided with open access summaries.

<sup>102</sup> Elsevier ('STM Digest') and *eLife* are for example developing these kinds of initiatives.

<sup>103</sup> See <http://mt180.fr>



Stronger – and also more digestible – knowledge from the scientific world about what researchers are doing and the results they are achieving may make society better informed and offers the opportunity for society to invite or challenge science to look at certain issues. This is the idea behind the Dutch National Science Agenda (*Wetenschapsagenda*).

#### 4.4.3 Costs and benefits

Changes in the way scientific output is published and paid for generate changes in costs and benefits for the various stakeholders. What happens to the distribution of those costs and benefits in a world with more open access? And can the costs and benefits be estimated at system level?

Although calls for open access are based mainly on the argument that part of the task of science is to share its results as widely as possible and that the results of *publicly* funded research should also be *publicly* accessible, the steadily rising costs of access to scholarly publications are often put forward as a key reason for placing open access on the agenda, and sometimes even as an argument for open access itself, based on the assumption that open access is a way of stemming those rising costs.<sup>104</sup>

But is open access likely to lead to lower costs? Generally, the focus in this regard is on the subscription costs paid by libraries and how these change with different variants of open access. Of the total costs involved in the ‘production’ of a scholarly article in a country such as the Netherlands, more than half goes to (the writing costs of the) researcher, a sixth goes to the review process, roughly a fifth to the publishing costs and only around 1% goes to library costs.<sup>105</sup> This puts the cost debate on open access into a different perspective.

In considering the changes in the financial landscape due to open access, there are broadly two different scenarios. In the first, payment for access shifts from the reader to the author. This is currently the dominant model in ‘gold’ open access. In the present system, with the dominance of subscription-based journals, libraries bear the costs of access. In a system where the author pays, the costs will in principle be borne by researchers. Research institute libraries will save money as they will not have to pay subscriptions, but researchers will need larger budgets in order to pay the publication fees for their publications. Therefore, there would possibly be a shift from centralised to decentralised budgets. But it is plausible that in a world of open access where – in principle – authors pay for publication, ‘big deals’ will still be negotiated between

---

<sup>104</sup> For example the European Commission’s impact assessment (2012).

<sup>105</sup> Houghton *et al.* (2009), pp. 9-10.

universities and publishers, in which all publication fees for researchers at the institutions concerned are covered. In that case, not much will change.

In addition to the question of who pays, it also matters whether publishing becomes cheaper or more expensive on balance. A shift to a system in which authors pay will, broadly speaking, lead to an increase in the costs across an entire institution or country, if a certain institution or country produces (publishes) relatively more than it consumes (reads). In an opposite situation these costs will fall, broadly speaking. That would only change if, in a system where the author pays, the total costs of all journals should also suddenly fall, which is unlikely (see below). As the Netherlands produces a relatively large number of publications, our estimation is that the Netherlands will pay more on balance.<sup>106</sup>

In the other open access scenario, in which subscription-based journals continue to exist but authors make a version of their publication available online, possibly after an embargo period, there will be double costs: the costs of the journals, which will have to be borne in some way by readers and/or authors, and the costs of maintaining and providing access to the repositories where the papers and articles are deposited. Experiences in the United Kingdom have shown that in practice research institutes incur additional transaction costs in a system of green or gold access which are of the same order as the publication charges.

How likely is it that the total costs of journals will fall in the near future, as a result of increased open access? Where publishers retain in-demand journals which have 'scarcity value' (as regards reading in the case of subscriptions or publishing in systems where the author pays), those publishers will still be able to charge subscription fees or authors' fees (well) above the production costs. This is a separate issue from open access. In a gold open access system where the author pays, publishers will (be able to) demand higher fees for top journals than for less popular journals.<sup>107</sup> This is already happening.<sup>108</sup> Moreover, the negotiating position of individual researchers is very weak, which could prompt publishers to capitalise even more on their dominant position. The present lack of transparency on the level of publication fees underlines this.

The present high margins earned by a few large scholarly publishers could be curtailed if the importance of journal reputation were to disappear or if cheaper alternatives with good reputations emerge, (either through pressure from new market entrants or because

---

<sup>106</sup> In a rough calculation, Palzenberger (2015) assumes 24,000 articles per year with a Dutch corresponding author. With an average author's fee of €2,000 (a figure also used by Schimmer *et al.* (2015) from the Max Planck Gesellschaft), this produces a total of €48 million, well above the annual amount currently spent on subscriptions (approximately €40 million).

<sup>107</sup> A worldwide survey of scientists by the OECD (2015a, p. 5) shows that they are willing to pay extra to publish in journals with more prestige.

<sup>108</sup> For an overview of the bandwidth of APCs, see Ware & Mabe (2015), pp. 93-96.

of new players or business models). Better negotiation with publishers, possibly with users acting together could also curb the margins earned by large scientific publishers. In both cases the total costs for readers and/or authors may fall. Another possibility may be that subscription charges or publication fees are regulated by law. This would be a drastic market intervention, however, and is unlikely to happen in most countries (as it will be viewed as undesirable).

Therefore, as long as the causes of the rising costs of scientific literature are not addressed (namely the growth in the number of publications and the strong position of the top journals), the total costs will not decline noticeably, with or without open access.<sup>109</sup> At most there will be a redistribution between who pays more and who pays less.

## 4.5 Conclusion

As an ideal, open access to scholarly publications sounds very attractive. It fits in nicely with the high value researchers attach to open communication in the furtherance of science. In practice, however, any system used to implement open access has advantages and disadvantages. A system of open access to publications must be evaluated on the basis of its impact on science, society and business.

There is little resistance to the principle of open access to scholarly publications, and this chapter shows that the number of scholarly publications freely accessible has grown steadily in recent years, although progress is slowing down. However, as many studies confirm, open access is not the number one priority for most researchers. That honour is reserved for publishing with impact within the scientific world. This is (still) generally achieved by publishing in renowned, traditional (subscription-based) journals. Moreover, a number of key players are (still) not very interested in a mass transition to open access.

---

<sup>109</sup> There is a series of publications by Houghton and co-authors in which they (try to) calculate what happens to the total costs of scholarly publishing in a number of open access scenarios (e.g. Houghton (2009) for Denmark, the Netherlands and the UK and Houghton *et al.* (2009) only for the Netherlands). We studied these scenarios in some detail, but regard the outcomes as insufficiently reliable. The outcomes of the model calculations depend heavily on the input parameters applied, which in many cases are the authors' own estimates, have not been validated and are not properly argued (many of the estimates have also been contested by publishers and others; see Ware & Mabe (2015, p. 119)). This is the case in particular for the cost differentials assumed by the authors between the different publication models. The findings themselves also provide food for thought: for the Netherlands, for example, the authors found (in 2009) a potential cost saving with a fully gold open access system of €133 million per year, roughly 4 times (!) the amount spent on subscriptions each year in the Netherlands. Even a system in which subscriptions remain and an extensive infrastructure of repositories is set up, the authors still find a saving of €50 million, again substantially more than the annual amount spent on subscriptions. More or less the same applies for the other countries studied.

In practice, no single overarching open access publishing model has yet emerged as dominant. There are several different models for organising open access to articles, with the 'green' and 'gold' variants discussed before being the most common.

Both models occur in practice, often alongside each other. Green open access (publishing in a subscription-based journal and putting a version of the article in a repository online) is a robust model that can exist alongside the traditional scholarly journals. Its disadvantages are that it requires a dual infrastructure (hence the higher costs) and that open access is often granted only after an embargo period. The gold model (the scholarly journal provides open access directly) does not have these disadvantages. However, it only works well if it is the 'standard', which requires a wide-ranging system change. That will only happen if there is sufficient critical mass, which is difficult to achieve in the fragmented field of international science. Moreover, there are clear risks as in this model authors generally have to pay for publication in scholarly journals. This way, the incentive to put quality first is greatly weakened. And there are disciplines, especially in the arts and humanities, where it seems unlikely that budgets will be adequate to pay the required authors' fees.

It thus remains unclear what might constitute a sustainable open access model in which freedom of access goes hand in hand with assured scientific quality. Alternative models are being tested, but have yet to prove themselves as viable and lasting open access standards. They include models in which research institutes and/or funders bear the journal costs directly, or where authors become 'members' of a journal and can then publish a certain number of articles.

Based on our analysis, we estimate that the impact on science of free access to scholarly publications by itself is unclear and, at least *within* disciplines, probably limited. In practice, researchers often already have reasonably good access to literature in their discipline. The impact may be higher for access to publications from other disciplines, and for researchers not affiliated to (large) research institutes.

Good quality control of publications, speedy knowledge dissemination and good findability of publications are key factors in the functioning of science. Open access will only have a positive impact on science if it contributes to these areas. That cannot simply be assumed to be the case. If authors pay for publication, the incentives to allow the acceptance of articles to be dictated by quality are weaker than in a system where the reader pays, as at present. In the worst case, this could become a perverse incentive. This issue is pressing because the demand for mechanisms to assess the quality of journals and publications will only increase as the amount of publications increases.

What is the (potential) impact of open access to scientific publications on society and business? The impact of simply introducing open access is likely to be limited, because the average scientific publication is hard to understand or use outside the discipline for which it was written. Increasing the impact of science on society and business therefore requires a broader strategy. First, scientific results have to be findable. Second, science needs to be made genuinely accessible by 'translating' complicated results for a wider public. Thirdly, a broader approach is needed to ensure that the 'social impact' of research is valued and rewarded more within the scientific world, for example through evaluations. By thus making scientific results more *accessible*, the currently largely untapped potential of open access to scholarly publications for society and business could be tapped into. How much of this potential could actually be realised in practice will depend greatly on the effectiveness of ancillary measures and developments.

Scholarly publishing costs money and there is no reason to assume that publishing costs will fall simply because of a transition to open access. The recent rise in costs stems from the global growth in publications and the economically dominant position of some publishers, which exploit the fact that researchers wish to publish globally in journals with a strong reputation and read such journals. If reputation were to become less important, or if cheaper alternatives were to appear which still have a good reputation, the present high margins of a small number of large scientific publishers could be squeezed (either by the 'pressure' of newcomers or because new entrants or earnings models displace the existing operators/models). Better negotiation, possibly by users acting together, could also constrain the current margins of the large scholarly publishers and reduce the cost for researchers. This would require more transparency on subscription and publication costs. In both cases, the total costs paid by readers and/or authors could reduce. In itself, however, open access to scientific publications is not a panacea for lowering the perceived high costs of scholarly publishing

What happens to the total costs also depends on precisely which open access model emerges as the standard, though in reality there is a good chance that mixed systems will continue. The 'green' variant involves double costs (for both subscriptions and repositories), and the earnings model that is the norm in the 'gold' system, based on the principle of 'the author pays', carries the risk of volume growth, and if the authors' fees are not capped, these could also rise. Moreover, in a system where the author pays, the resources available to researchers are a decisive factor in how much (and where) they can publish. Whether that is desirable is questionable.

There will also be some cost displacement in a model based on the principle of 'the author pays': countries (or institutions) which publish relatively more than they read will pay more on balance. This is likely to be detrimental for the Netherlands, which is an above-average knowledge producer. In such a situation, it will also be necessary to look

at what other countries do. For example, if the Netherlands or the EU incur costs in making 'their' publications freely accessible worldwide, this will partly benefit third countries which are not willing to incur those costs for their own researchers (even where they have the resources to do so): a free-rider problem. Dutch or EU researchers will then be paying twice: for publishing their own articles and for reading articles published in third countries which do not operate a system of open access.

How will open access develop in the future? Thus far, the move towards open access has mainly been a bottom-up process, supported by a few large research funders. However, this bottom-up process has evidently not been strong enough to bring about a system change in favour of open access. What has become apparent is that funders of research, including governments, can do something to speed up the transition towards open access, as developments in the US and Brazil have shown. Achieving a genuine breakthrough will require strong leadership.

Internationally, there are clear differences in the preferences and policy in relation to open access. Within the EU, some Member States operate an explicit open access policy. Most of them support both the 'green' and 'gold' models, often aiming in the first instance for a green route as a single country on its own is not really capable of enforcing a system change towards 'gold'. The United Kingdom is one of the few countries where the government has expressed a preference for the gold route. Yet in the UK, too, the picture is mixed in practice: some public funders demand the gold route while others opt for the green route. The federal government in the United States has opted for 'green'. The European Union has not expressed a preference for either model.

The choices made by researchers, as authors of articles, will be crucial, i.e. where and how they wish to publish. This is where open access will have to be taken up. There are 'sticks' available to force this to happen (for example the central role played by funders, with legislation as a possible backstop). But where are the 'carrots'? At present, they are almost absent.

## Sharing research data

### 5.1 Retain or share?

In the previous chapter we examined access to scholarly publications. Researchers have traditionally published their results in scientific journals (as well as in books or at conferences). These publications are essential for the furtherance of their scientific careers, and are therefore the focus of the majority of researchers.<sup>110</sup> The debate about open access is generally concerned with making scholarly articles more widely accessible.

The situation is different for the research data on which such articles are based. Traditionally, these data have been used at most in processed form in the article (for example in a table or graph). In most disciplines, however, the underlying research data are *not* published (or only in part) and remain 'private', held by the researcher and/or their research institute.

There are however disciplines where a culture of sharing data with other researchers has developed, for example because researchers work together using the same (large-scale) infrastructure, or because each contributes a small part to solving a larger scientific puzzle. Examples include research data from the CERN particle accelerator or the results from the Human Genome Project. The research based on these datasets requires collaboration by many researchers who build on each other's work. Although these collective projects differ in their approach to openness, they illustrate the importance of collecting and sharing data in joint activities.

There have been many developments in relation to research data in recent years. These data are today generally (capable of being) produced in or converted to digital format. This raises all manner of new 'questions', among which how to ensure that these digital data lasts. This is a topic that has been addressed in a broader context in the Netherlands by the 'national coalition on digital sustainability'.<sup>111</sup> Digitalisation also creates new possibilities; these data could, at least in theory, be made available to third parties via the Internet. Moreover, more and more journals are appearing which focus on publishing data (data journals), while 'normal' journals are becoming better at building in links in an article to the underlying data. The storage facilities for research datasets are also expanding, for example in repositories. In parallel with this, science is becoming

<sup>110</sup> For more on the importance of publications for a scientific career, see: AWTI (2015a), Ministerie van Onderwijs, Cultuur en Wetenschap/Ministry of Education, Culture and Science (2014), p. 53, and De Goede & Hessels (2014).

<sup>111</sup> See: [www.nccdd.nl](http://www.nccdd.nl)

increasingly data-driven.<sup>112</sup> Not only is the amount of data generated for research increasing enormously, but new ways are being found of cutting across these datasets in search of meaning and new insights.

These developments beg the question of whether, research data should be ‘published’ and made accessible to others. This sharing of research data could contribute to the integrity and quality of scientific research (verifiability and replicability), avoid unnecessary duplication of research or form the basis for comparisons of similar research projects, thus improving ‘reproducibility’. It may also prove to be a fertile breeding ground for establishing new connections. Such research data may also be of value for society or business.

Of course there are drawbacks, limitations and risks associated with sharing research data. In the first instance, it will require more work and therefore more time and money. The data could also be used ‘wrongly’ or create the danger of free-riding using research data which a researcher or institute has gone to a lot of trouble to obtain.

By analogy with open access to scholarly publications, the ideal model for access to research data could be ‘open research data’, i.e. free access to research data produced by others. A definition of this ideal is given in the text box below. With the exception of disciplines where storing and sharing research data is already the norm, this would mean a major change for most disciplines. In this report, we initially focus on *sharing* research data in some form (rather than providing complete open access to the data) in order to retain the flexibility to accommodate the required breadth and depth of the access to research data. This also better reflects the many different ways of sharing data.

**Open data** is defined by the Open Knowledge Foundation as “... *data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike* ” (Open Data Handbook)<sup>113</sup>.

Compared with open access to scholarly publications, open research data is a relatively new concept, and the debate on this topic is not as far advanced. Nonetheless, sharing and granting access to research data has been given wide support across a wide front in numerous declarations.<sup>114</sup> In practice, the picture is mixed, with a wide variation in the degree to which different disciplines store and share research data. A good overview is

---

<sup>112</sup> See AWTI (2015b) for more on the digitalisation of science and the (knowledge) economy. See also the conclusion ‘data-driven science is becoming ever more important’ during the AWTI meeting of 5 June 2015 at Science Park (AWTI, 2015c) and AWT (2011).

<sup>113</sup> Open Knowledge Foundation (2012): [http://opendatahandbook.org/guide/nl\\_BE/](http://opendatahandbook.org/guide/nl_BE/)

<sup>114</sup> For a recent example see The Hague Declaration on Knowledge Discovery in the Digital Age: <http://thehaguedeclaration.com/>



provided by the Dutch research data archive DANS (Data Archiving and Networked Services) (Dillo & Doorn (2011); see text box).

How should the sharing of research data develop further? Before arriving at conclusions and recommendations, this chapter steps back, and looks at what research data are, who plays a role in creating access to them, and what the motives are for sharing research data. It also lays down the necessary parameters for developments towards open research data: what is needed in order to fulfil the promises made, and how can we obtain the greatest possible return for society from sharing research data?

### **DANS: The Dutch data landscape in 32 interviews and a survey**<sup>115</sup>

DANS (Data Archiving and Networked Services) is an institute affiliated to the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO) which promotes sustainable access to digital research data. By providing various (online) services, including data archiving, but also training and consultancy services and participation in national and international projects and networks, DANS promotes the further opening up of access to digital research data. In doing so it works with partners such as 3TU.Datacentrum, SURFsara and the Netherlands e-Science Centre.

In 2011, DANS published a study on the Dutch 'data landscape'. This study examined how the humanities, social and natural sciences deal with research data: what differences and similarities are there, what can be improved and how? One thing is clear, different disciplines treat data in very different ways – a finding that is in line with international observations.<sup>116</sup>

Data in the natural sciences are stored on a much larger scale than in the humanities and social sciences. Sharing within those disciplines is much more accepted, facilities are better and more data are collected.

However, there are also differences, both within these groups and at the level of individual researchers. For every researcher who claims that sharing data is no more than natural, another can be found who, is unwilling to share data. Their reasons range from inadequate infrastructure, lack of standardisation and support to unwillingness to share due to scientific competition.

Although there is a growing awareness in the scientific world of the need for

<sup>115</sup> Dillo & Doorn (2011).

<sup>116</sup> For the international analysis, see: <http://exchanges.wiley.com/blog/2014/11/03/how-and-why-researchers-share-data-and-why-they-dont/>

sustainable data management, on balance it has not yet penetrated far enough. Even in disciplines where there is a willingness to store and share data, the data management is often not fit for purpose. There is little awareness of the need for sustainable storage. There is fragmentation in the allocation of tasks and responsibilities. And the data are often stored across a variety of facilities. The storage quality depends greatly on individual research groups, the research leader and culture within the discipline.

For a larger-scale transition to data sharing, more than greater awareness is needed. Good infrastructure must be established (efforts are being made to achieve this) for the long-term storage of data. Support for researchers to ensure that this storage takes up as little of their time as possible is also essential. Moreover, many researchers do not want to be the only ones to place their cards on the table: data sharing must be rewarded and must be the norm, either voluntary or mandatory, to prevent free-riding.

## 5.2 Research data: functions and actors

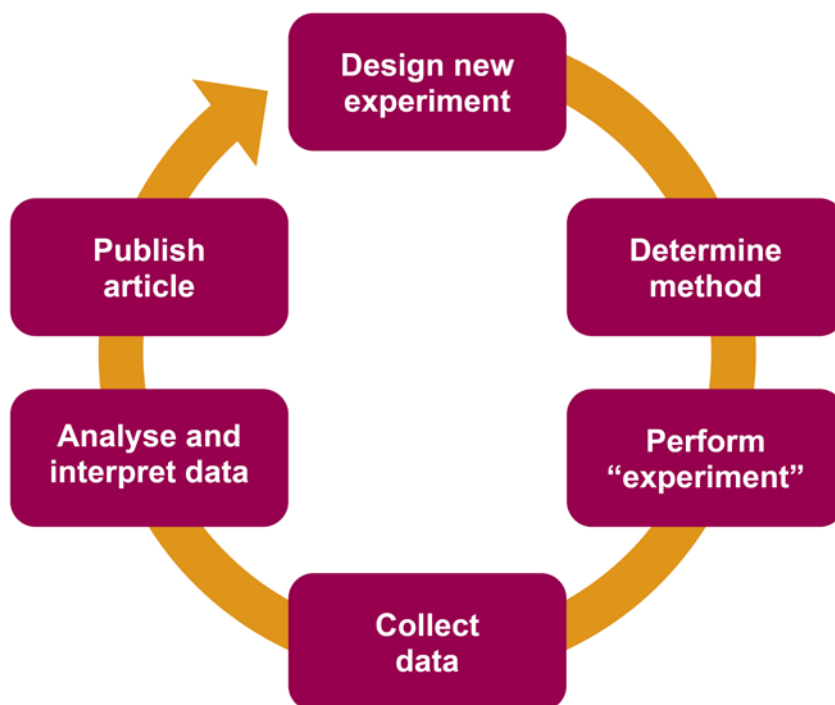
### 5.2.1 Definition: research data

What are these 'research data' that should or should not be shared? In the traditional cycle of scientific research, researchers choose a suitable research method to answer their research question. They apply this method and this generates (raw) 'research data'. If they perform a laboratory experiment, the data will consist of a set of measurement results. But if a researcher applies a completely different method, such as observing behaviour, the research data will consist of all kinds of reports. It is also possible to carry out research on an existing dataset, for example all kinds of demographic or statistical data from existing databases. The researchers will then process, analyse and interpret the (raw) data they have collected. This analysis will form the basis for the findings communicated to the scientific community in the definitive publication.

In this report, research data are regarded as all data collected (used) for the purpose of research. This includes data *generated with a view to research*: measurement data 'produced' by researchers themselves in and for their research. But it also includes existing datasets *used* for research. This might comprise databases from Statistics Netherlands (CBS), for example, or newspaper archives, or a collection of tweets covering a given period. If such datasets form the basis for a scholarly publication, it is logical, from the perspective of the envisaged transparency (and reliability) of science, to treat it in the same way as 'new' research data generated by the research project. A broad definition of research data also makes sense from the perspective of more efficient

use of scientific data, including data mining and interdisciplinary cross-fertilisation. From this perspective it is relevant if a given dataset is potentially of interest for researchers, who wish to use it as a basis for their research.

Figure 3 The research cycle<sup>117</sup>



It is important to be aware of the source of the (research) data and who is paying for it. Some research data come from public, generally accessible sources, some are generated by researchers specifically for research purposes, and some are funded by (a combination of) public or private funds. However, data (sets) from commercial sources can also be used for research.

### 5.2.2 Functions

The following functions summarised in Figure 4, are important in relation to the acquisition, processing, storage and reuse of research data.

Acquiring relevant research data ('data capture') is a crucial part of the research process. For the researchers, it is important to be able to analyse and interpret the data. If the data

<sup>117</sup> This figure focuses on the steps involved in data and publishing. Other steps are possible between the steps 'publish article' and 'think of new experiment', for example 'theory' and 'hypotheses', which naturally form the scientific context for the new experiment.

subsequently remain with the researcher and their institute and are not shared, it is unlikely that much more will be done with those data.

If someone else wishes to examine those data or reuse them at a later date, this is often not straightforward. There is a growing awareness that this is only possible if four conditions are met; the data must be 'FAIR':<sup>118</sup>

1. *Findable*;
2. *Accessible*;
3. *Interoperable (universally readable and usable)*;
4. *Reusable*.

Where this report refers to 'storing and sharing' research data, we actually mean the FAIR storing and sharing of these data.

The FAIR requirements mean that in the case of raw research data, it must be made clear what the context of the experiment was, what software was used, precisely what the measurement results measured, and so on. In short, to be suitable for reuse the raw research data must be cleaned and provided with metadata and other supplementary information or functionalities (such as open software to process the data).<sup>119</sup> This is the function of 'data curation'. To make it easy to identify the research data, 'persistent identifiers' must be added to the dataset.

Now that the research data are suitable for others to view and reuse, they need to be stored somewhere where they can be accessed (*data storage*). This storage also has to meet requirements: the data must be stored in a fixed location and in a standardised, interoperable and sustainable way – in a 'trusted digital repository'<sup>120</sup> – and then rendered findable through portals.

This then makes it possible for third parties to use the data, for example for further analysis or data mining. This is the function of 'data exploitation'.

Another option is to visualise the underlying data in some way. This is what happens with all kinds of medical scanning techniques, for example, in which the 'measurement data' are translated into a two or three-dimensional image. This is the function of 'data visualisation'.

---

<sup>118</sup> This was stated explicitly in our interviews with SURFsara and DANS, for example. On the FAIR requirements, see also: Data FAIRport (<http://datafairport.org/>) en FORCE11 (<https://www.force11.org/node/6062>).

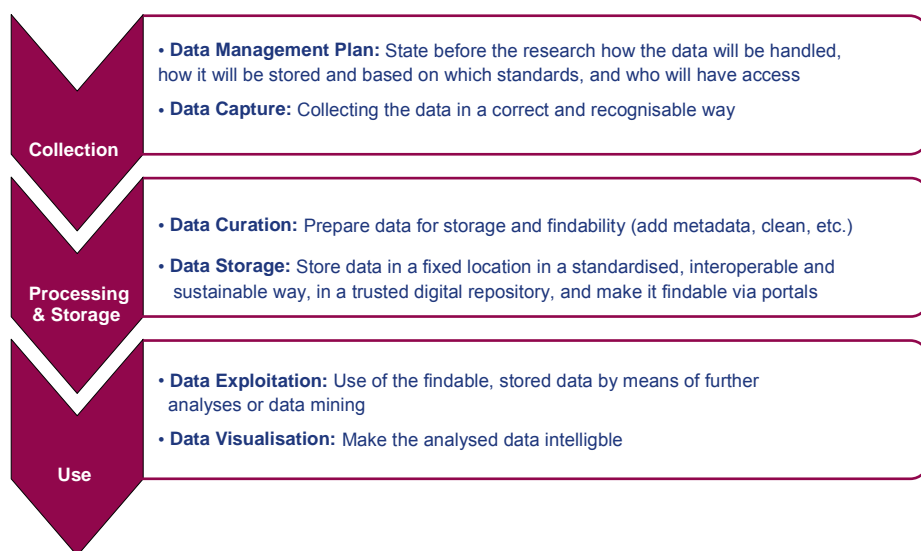
<sup>119</sup> More and more portals such as [github.com](https://github.com) are working in partnership to develop (open) software and codes. Interestingly, the software and codes can ultimately also be supplied to a repository and linked to DOIs (*Digital Object Identifiers*), making it possible not just to refer to the data itself, but also to the underlying software and codes.

<sup>120</sup> For a definition, see: [http://sedataglossary.shoutwiki.com/wiki/Trusted\\_digital\\_repository](http://sedataglossary.shoutwiki.com/wiki/Trusted_digital_repository)

Clearly, then, storing and sharing raw research data is a somewhat involved process. Storing a traditional set of research data long-term in a way that others can use or reuse it is not simply a 'click' away. Anyone wishing to make their data suitable for sharing after doing the research is very likely to discover that this involves a degree of time and effort. It is better to think carefully in advance about which research data one wishes to collect, from whom/what and in what way, and how that data can be made (and kept) readily accessible for others with whom one wishes to share the data. This is the function of 'data management'. This seems very obvious, but in practice good data management does not always receive sufficient attention.<sup>121</sup>

The following figure summarises the main functions in relation to research data.

Figure 4 Overview of the main functions in relation to research data



### 5.2.3 Stakeholders

Several stakeholders are involved with research data: collecting, processing, storing them, making them accessible and enabling them to be (re)used. They all have their own interests, preferences and capabilities, and this influences how research data sharing can and does develop. Here we try to identify the main stakeholders. Each of them is involved with one or more functions. The European RECODE research project<sup>122</sup> identifies the

<sup>121</sup> See e.g.: Dillo & Doorn (2011) for a broad overview, Bakker (2014) for psychology and Oldenburg (2015) for sociology. This was also confirmed in numerous interviews conducted for this study.

<sup>122</sup> See: <http://recodeproject.eu/>

following roles which are directly associated with the functions listed above: capture, curation, storage and exploitation and visualisation. RECODE also refers to the role of the research funder. Although this role is not directly linked to the data, it is definitely relevant because the funder sets the research parameters: for example, does the funder have sufficient money to cover any additional costs of data curation and storage?

As with open access to publications, an important consideration when making research data accessible is who has the capability to do this, who has an interest in it and why.

### **Researchers**

Researchers collect data on their own or in partnerships and also share data to a greater or lesser extent, depending on the level of collaboration within and outside the discipline. The key is that data in principle have to be made accessible by the researcher who collects them. This immediately raises questions: why would a researcher make data accessible, and what is involved in doing so?

Researchers have a sense of 'ownership' of their own data. They can use their data for more than one article, and therefore have an interest in not placing it in the public domain immediately after the first publication, let alone before they have first published. There must be clear, tangible reasons for a researcher to share their data with others, especially since that sharing requires extra efforts. Those reasons may include a mutual interest (researchers using each other's data) or a common interest (groups of researchers working on the same large puzzle), or a mandatory requirement by a research funder, or a deep conviction that research data should be shared. Some disciplines have a culture of openness and sharing (see the text box about CERN). On the other hand, there are all manner of reasons for not sharing research data, such as privacy, strategic reasons or commercial interests.<sup>123</sup>

The road to sharing research data begins with the researcher. He or she is in a position to choose whether or not to make their data available to others, for example in the form of supplementary information accompanying an article. Many publishers respond positively to this and offer this possibility. A researcher may also choose to store their data long-term in an open archive and make it accessible to others. A researcher might also publish data on separate platforms or on their own website. This enables the (raw) research data to be viewed and commented upon by others, but sometimes also to be reused for further research by other researchers.

Researchers are also potential users of research data collected by other researchers. In this capacity, they will in principle benefit from the sharing of that data. Certain types of

---

<sup>123</sup> See e.g. OECD (2015b, p. 60) for a summary.

research are only possible using existing research data, and here researchers obviously benefit from those data being available. Yet in practice, many researchers (in most disciplines) are still reticent about basing their work on other researchers' data,<sup>124</sup> feeling uncomfortable because they do not know the precise context. Ultimately, however, researchers who use data are key to determining whether research data sharing is successful, because if the stored and accessible data are not subsequently used, or only to a limited extent, there is clearly little point in storing and sharing that data for reuse. It is interesting in this context to analyse how often datasets are currently reused. DANS (2015, p. 10) presents preliminary aggregated figures on the number of annual downloads of datasets as a proportion of the total number of available datasets in the various disciplines for which DANS holds data. The figure varies between 0.5 and 2.0, so the number of downloads is of the same order of magnitude as the number of stored datasets. Of course, that tells us nothing about the distribution: it may be that the downloads are concentrated on just a small proportion of the stored data sets. Further analysis is needed to answer this. Open access to data does of course have value in verifying the integrity of scientific work, whether or not it is reused.

Finally, with increasing amounts of research data being stored, some form of 'quality indication' will become increasingly important. What value should be assigned to a given dataset? Who will do that? Will a kind of review system be introduced, comparable to that for scholarly articles? Researchers will logically play a central role in this process, creating a third role for researchers in relation to the sharing of data.

### Open Data and CERN<sup>125</sup>

The CERN particle accelerator has its own data portal where all open data generated in the experiments involving the Large Hadron Collider (LHC) are placed. The data are classified into four levels: 1) data directly related to publications; 2) simplified datasets that can be used for education and external communication; 3) data and simulations with software for use in scientific analysis; 4) raw data and software not covered under 3). All data in levels 1-3 can be accessed, sometimes after an embargo period; the raw data (level 4) will not always be accessible. These embargo periods vary depending on the experiment and are necessary among other things to give researchers the opportunity to publish.

All datasets have a persistent identifier, in this case a DOI (Digital Object Identifier). This enables the data to be cited, as agreed in the FORCE 11 Joint Declaration of Data Citation.<sup>126</sup> Data is shared using a *Creative Commons CC0 public domain*

<sup>124</sup> Based on our interview with DANS in July 2015, see also DANS (2011).

<sup>125</sup> Freely taken from: <http://opendata.cern.ch/about?ln=en>

<sup>126</sup> Data Citation Synthesis Group (2014).

*dedication* and software with an open source licence.

Some CERN researchers express reservations. They believe that the complexity of the data and the time and trouble invested in collecting and understanding them should not be underestimated, and the published data should always be seen in the complex context in which it was collected.<sup>127</sup> CERN therefore does not share its data in order to see what can be done with it, but to see how data can be presented in the best and most understandable way.

### **Research institutes (such as universities)**

Research institutes play a role by supporting their staff (scientific researchers) with facilities or through their policies as employers. They are also the owners or have a degree of control over the research data obtained. Research institutes benefit from high-quality research, and sharing research data can help improve quality as researchers have to be more rigorous in collecting it. The efficiency within an institute can also be improved if researchers have access to each other's (research) data. Like researchers, however, an institute may not always wish to share data with others outside the world of science; there must be something in it for the institute.

Research institutes have a degree of power in promoting the sharing of research data. As employers they can impose a certain policy on their researchers. They can also take on a role themselves, for example by offering storage facilities in the institute's repositories. It is legitimate to ask here whether this storage role should be left to universities, and whether it is better to store open research data centrally or locally. Local storage is closer to the researchers and gives them a degree of (physical) control over the data, but central storage makes standardisation easier. On the other hand, the reduced control, especially if data are stored abroad, makes 'giving away' data to a repository less attractive.

Although universities do support the storing and sharing of research data on paper, in practice they will also be wary of downsides such as extra costs, complications in relation to public-private partnership, and tensions concerning the task of universities in the valorisation of knowledge.

### **Research funders**

Research funders are in a powerful position because they control both the 'sticks' (setting the requirements for grants) and the 'carrots' (the grants themselves). In a sense, they are the only stakeholder that can effectively impose a requirement to share research

---

<sup>127</sup> <http://home.web.cern.ch/cern-people/opinion/2014/11/road-open-science>



data. Much research – especially in the Netherlands – is publicly funded, and this gives the government leverage through the funding. Other types of funders, such as charitable funds, can also strive for open access and data sharing.

There is of course a legitimate question as to whether research funders have an interest in storing and sharing research data. In the first instance, this is likely to entail extra costs. However, a funder may of course have good reasons for wishing to share or promote the sharing of research data, for example based on a belief that openness is an integral part of science and also helps to further it, or on a conviction that it contributes to the broader efficiency of science and that the additional costs can therefore be ‘earned back’. Or a funder may believe that publicly funded research should in principle be publicly accessible. In practice, each research funder will differ in its attitude to the sharing of research data.

A lot of research is also privately funded, at least in part. These private funders, though they are willing to share scholarly articles, are often less enthusiastic about releasing research data. In these cases, therefore, it is important to make clear agreements and where appropriate incorporate them in policy. Such agreements on aspects such as intellectual property rights, patents and commercial secrets are necessary in order not to scare off private investors from investing in R&D.

### **Data archivists**

After being collected during the research, the research data then have to be rendered suitable for storage and then stored. This is a task for data archivists. There are three central data archives in the Netherlands (3TU.Datacentrum, DANS and SURFsara), which have joined forces in the coalition Research Data Netherlands. These relatively large players, in partnership with other parties, have formulated a shared mission of promoting the sustainable archiving and reuse of research data.<sup>128</sup> These three data centres fulfil a ‘back-office function’ in the curation of research data, ensuring that it is archived in a sustainable way. The ‘front office’ role is often fulfilled by university libraries, which are closer to the researchers. Combining their strengths in this way brings specialist knowledge and skills under one roof. All affiliated archives are authenticated with the Data Seal of Approval.<sup>129</sup>

At some point, coordination is needed for the standardisation of this process of data capture, curation and storage. The government could play a role here, especially where it also funds the research and is able to provide incentives. There are standards for storing data and making it usable and reproducible, but those standards are often limited to a single research discipline or community and are rarely communicated on a wider scale.

---

<sup>128</sup> <http://www.researchdata.nl/over-rdnl/>

<sup>129</sup> See: <http://www.datasealofapproval.org>

Broader agreements are needed to ensure that these standards are not just adopted, but also broadly shared and embraced. Which party or parties ultimately fulfil this role, including in the longer term, is irrelevant, as long as someone takes responsibility for it.

### **Data storage facilities**

Data storage facilities, i.e. the sites where data are physically stored, will welcome an increase in demand for their services. They can take advantage of the exponential growth in the amount of data, though questions will need to be asked as to which data is worth storing and for how long (storing data costs money). This will make little difference to the providers of storage facilities, but will play a role in the policy of those who pay for them.

### **Publishers**

Major publishers play a role because they have the resources to develop an extensive infrastructure and the business model is attractive. Like the APCs for open access, however, there are also costs involved in storing data. Good agreements on data stewardship in the long term are important to guarantee 'perpetual' access to stored data, because publishers will not always have an interest in this themselves.<sup>130</sup> It is ultimately up to the 'market' to determine whether libraries, publishers or other parties are the best choice. Moreover, several publishers are developing new options for publishing research data.

### **Libraries**

Libraries also see a role for themselves (possibly a collective role through strategic alliances) but require funding. They have the skills but cannot compete with publishers and other private stakeholders in terms of infrastructure. Research institutes and libraries have an interest in high-quality research, something that can be fostered by the sharing of research data, but often do not have sufficient resources. They can however play an important role in encouraging 'Open Science Literacy' (Data Management Plans, etc.) among researchers in order to improve the handling of research data.

### **Businesses**

Private sector companies can play different roles. As funders of research or producers of research data, they will generally be unwilling to release their own data, or at least not immediately, because of commercial interests or a desire to retain control over what happens to their data.<sup>131</sup> This is a barrier to the sharing of data generated in research that is partly privately funded (also see the notes on research funders above: aspects such as intellectual property rights need to be properly regulated).

---

<sup>130</sup> For more on data stewardship, see Netherlands eScience Center (2013).

<sup>131</sup> This emerged among other things in the interviews with several companies.

On the other hand, a development is also under way in which companies are proactively publishing research data in order to provide transparency to their customers and society, for example in relation to food. Sometimes they are also required to share this kind of data with the authorities, for example in order to support claims about the effects and properties of their products. There are thus different kinds of drivers at work here which also result in access to research data. There are also initiatives to make unpublished research results (for example clinical trials carried out by private companies) available for secondary scientific research, as is happening in the YODA project, for example.<sup>132</sup>

Private companies could also benefit from open research data as 'users', and develop new business models based on that. Time will tell to what extent this leads to significant results. To gain an idea of what is likely to happen, it is interesting to look at what developments have been prompted by open government data; this is described in the text box below. We can conclude from this case study that the degree to which open research data is put to good use by private companies is likely to depend greatly on the nature of the company as well as the nature and usability of the data.

### **Civil-society organisations and citizens**

Civil-society organisations and citizens want maximum access to research and research results that are interesting or relevant for them. The free sharing of research data by researchers removes a financial barrier. However, many citizens and private companies will lack the resources and skills needed to interpret raw research data. Help with interpretation and a user-friendly form of presentation is needed to enable this information to be used adequately. A good example of how this can be done are the databases maintained by the Meertens Institute on its website, complete with search functions and other tools.

---

<sup>132</sup> See: [yoda.yale.edu](http://yoda.yale.edu)

### RECODE Project<sup>133</sup>

The EU co-funded RECODE Project (RECommendations for Open Access to Research Data in Europe, part of the Seventh Framework Programme) was launched in 2014 to investigate the challenges to data dissemination and preservation. The project studied various networks, initiatives, projects and research groups and took into account the diversity of disciplines, location and interests of stakeholders.

Two overarching issues emerged: 1) the need for a cohesive open data 'ecosystem'; and 2) the lack of attention for research practice and the specific role that data collection plays in it.

The stakeholders in open research data were divided into five groups: 1) funders; 2) collectors; 3) disseminators; 4) curators; and 5) users. This illustrates clearly how much still has to be done after data have been collected in order to make them genuinely accessible; the road to open data is complex and is not viewed in the same way by all stakeholders. There are also concerns about the costs of sharing research data: where does the money come from and will it be at the expense of something else?

There are also technical challenges. Data have to meet all kinds of standards: they must be heterogeneous and interoperable, accessible and findable, sustainably stored and curated, of high quality and secure. While these challenges – compared with financial, cultural and legal barriers – can be solved relatively easily, it is important not to lose sight of them. The RECODE project calls for a holistic approach in which open and interoperable standards, harmonised findability and services, persistent identifiers, a culture of good data management, user-ready data and technical solutions to security and legal problems all play a role. Once again, allowance has to be made for the diversity within the scientific community.

In addition, legal and ethical issues also need to be resolved. The main legal questions relate to intellectual property issues (copyright, trade secrets and database rights), privacy and data protection, and open access mandates. Ethical issues stem primarily from concerns about secondary users, commercial use, disproportionate distribution and unequal allocation of the costs of research data. Many solutions can be found in 'soft law'. RECODE recommends the use of Open Licensing, properly recognising and rewarding high-quality data, and only making research data accessible when it is (legally) appropriate to do so.

---

<sup>133</sup> See: RECODE (2014) and [www.recodeproject.eu](http://www.recodeproject.eu)

A number of practical issues also have to be resolved: financial and other support, continual quality assurance, evaluation and reliability of research data. Who is responsible for this? How are the different stakeholders involved? How are the benefits and limits of open research data communicated? This is a challenge for archives, libraries, universities, data centres and research funders. Clear agreements are needed on funding, information provision, evaluation and responsibility.

Although many agreements have to be made on a small-scale level, RECODE makes ten overarching recommendations:

1. Develop aligned and comprehensive policies for open access to research data.
2. Ensure appropriate funding for open access to research data.
3. Develop policies and initiatives that offer researchers rewards for open access to high-quality data.
4. Identify key stakeholders and relevant networks and foster collaborative work for a sustainable ecosystem for open access to research data.
5. Plan for the long-term, sustainable curation and preservation of open access data.
6. Develop comprehensive and collaborative technical and infrastructure solutions that afford open access to and long-term preservation of high-quality research data.
7. Develop technical and scientific quality standards for research data.
8. Require the use of harmonised open licensing frameworks.
9. Systematically address legal and ethical issues arising from open access to research data.
10. Support the transition to open research data through curriculum development and training.

### 5.3 Policy on sharing research data

As stated above, the development of open research data is proceeding more slowly than open access to scholarly publications. This also applies for government policy on sharing research data. The first thought for many governments turns to the extra funds needed to set up an e-infrastructure for data storage and the extra costs of rendering data suitable for storage, quite apart from other obstacles to sharing data (such as privacy). This often makes them initially less than enthusiastic.<sup>134</sup> These factors are all part of the reason that open research data is less well developed in terms of specific policy and measures than open access to publications.<sup>135</sup>

<sup>134</sup> Several respondents in our country comparison (see Appendix 2) explicitly raised this point as a reason that a robust policy on sharing data had not yet been developed in their country.

<sup>135</sup> This is the conclusion in OECD (2015b), for example, but also of earlier studies such as Nicol *et al.* (2013).

## UNESCO and OECD

UNESCO explicitly sees access to research data as part of its envisaged open access to scientific information.<sup>136</sup> The OECD also believes that sharing research data is a good thing and has formulated a number of principles for making research data accessible (and reusable).<sup>137</sup>

## European Union

### *The European Commission and (more) open access to research results*

In 2012 the European Commission issued guidelines to Member States, in the form of a Recommendation, on making the results of publicly funded research, including data, publicly accessible.<sup>138</sup> The Commission's aims were to optimise the impact of publicly funded scientific research and to provide a boost to the science and knowledge-based economy. Intellectual property rights, text and data mining, altmetrics, e-infrastructure and international cooperation were also addressed. This broader open science approach is part of the Digital Agenda and comes under the banners of the Digital Single Market, the European Research Area and the Innovation Union.

Like an earlier Communication on this subject,<sup>139</sup> however, a Recommendation is not binding, and Member States are free to deviate from it.<sup>140</sup> Many recommendations are however incorporated into EU policy, such as the awarding of Horizon 2020 (H2020) funding. The EU also plays a facilitating role here, for example subsidising the Open Data Institute, which works on open data skills, open data research and open data start-ups.<sup>141</sup>

The concept of a data management plan has also been incorporated in H2020 in an open research data pilot project which the Commission plans to use to evaluate experiences and provide a basis for future policy. The Commission has also used surveyed the wishes and requirements of users and collectors of data.<sup>142</sup> In addition, the results of the Seventh Framework Programme RECODE project were recently published (see text box above).<sup>143</sup>

In the H2020 open research data pilot, once funding has been allocated, in designated disciplines applicants will be required to state within six months which research data will

---

<sup>136</sup> UNESCO (2011).

<sup>137</sup> OECD (2007).

<sup>138</sup> Recommendation of the Commission dated 17 July 2012 on access to and preservation of scientific information (2012/417/EU), *PbEU* 2012 L 194/39.

<sup>139</sup> Communication from the Commission dated 17 July 2012: 'Towards better access to scientific information: Boosting the benefits of public investments in research', COM(2012) 401 final

<sup>140</sup> See also Lelieveldt & Princen (2011).

<sup>141</sup> See: <http://horizon2020projects.com/il-ict/worlds-largest-investment-in-open-data-start-ups-announced/>

<sup>142</sup> European Commission (2013a).

<sup>143</sup> RECODE (2014).

be made accessible, which will not be released, and why.<sup>144</sup> This includes data that are needed to validate the research (underlying data) as well as other, indirectly used data (including metadata). There is an opt-in system whereby recipients of funding outside the disciplines designated for the pilot are encouraged to participate in the pilot.

An opt-out arrangement is also possible in certain cases, for example if no data are collected, if results have to be protected or are confidential, where public order and safety are a concern, or where releasing the data would raise privacy issues or would jeopardise the purpose of the research. Commercial motives can thus also serve as grounds for an opt-out.

The data management plan states which data will be collected or generated, what the standards are, which data will be used and how, and by whom and how the data will be curated and stored. After completion of the research, the data can be placed in a repository chosen by the researcher. In addition to repositories for articles, OpenAIRE also provides an overview of data repositories and offers support in finding the right one.<sup>145</sup> The data base in these repositories must be accessible worldwide and must allow data mining, reuse and dissemination at no cost. This is genuinely open research data, therefore. Extensive monitoring and support systems have yet to be developed.

## The Netherlands

### *Policy in the Netherlands*

Government attention in the Netherlands initially focused on open access to scholarly publications,<sup>146</sup> but the government now seems to have recognised the importance of sharing research data as well,<sup>147</sup> though this has not yet led to any regulations. Dutch 'policy' in relation to (sharing) research data is therefore based on the rules applied by research funders and the 'policy' of research institutes themselves. In addition, the government plays a facilitating role by funding some of the necessary infrastructure.

Among research funders, the Netherlands Organisation for Scientific Research (NWO) began a data management pilot in 2015 which is comparable to the pilot in Horizon 2020<sup>148</sup> but is limited to six funding rounds.<sup>149</sup> It is intended to form the basis for a data management policy in all NWO funding instruments. The starting point is fully open

---

<sup>144</sup> European Commission (2013b).

<sup>145</sup> See: <https://www.openaire.eu/repository/ordp/select-rep>

<sup>146</sup> Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 15 November 2013, TK 2013-2014, 31 288, no. 354.

<sup>147</sup> Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament dated 23 January 2015: TK 2014-2015, 31 288, nr. 414.

<sup>148</sup> <http://www.nwo.nl/beleid/open+science/datamanagement>

<sup>149</sup> HTM (High Tech Materials) call for proposals (STW), Innovative Public-Private-Partnerships in ICT (NWO Physical Sciences division), Research Talent (NWO Social Sciences (MaGW)), Urbanising Deltas of the World of Security and the Rule of Law (NWO-WOTRO), Vici, Free Competition (NWO Earth and Life Sciences (ALW)), Projectruimte (FOM).

access to data, but privacy, public safety, property rights or commercial interests can all impose restrictions.

The costs of data management must be included in the request for funding, which also contains a section on data management. However, this section is not (yet) taken into account in assessing the funding request. The following questions have to be answered in the funding section:

1. Will data be collected or generated that is suitable for reuse? (Yes: answer questions 2, 3 and 4; No: explain why the study will not produce reusable data or will produce data that cannot be stored or that is not relevant for reuse for some other reason).
2. Where will the data be stored during the research?
3. How will the data be stored long-term after completion of the project and made available for reuse by third parties? Who will be able to access the data?
4. What facilities are likely to be needed for storing the data during and after the research? Are those facilities available?

NWO defines data as collected and unprocessed data, but also includes analysed and generated data, both digital and non-digital. The question of storage is limited to data relevant for reuse, but it is difficult to determine in advance which data this is. Data is preferably stored in a recognised data repository, or possibly in an institutional repository.

The Netherlands also supports many international partnerships, such as ELIXIR, a European intergovernmental organisation which shares biological and biometric research data, with a focus on integrating and sharing data from earlier research. Agreements are made within ELIXIR on making and keeping data accessible. The Dutch government provides €1.1 million to support this.

All in all, there is no shortage of initiatives and support in relation to the sharing of research data, but no genuine, concrete policy. Whether that is always necessary is questionable: many developments come from the bottom up, from within the scientific community. On the other hand, these developments sometimes need coordination and control or need to be enforced through policy. Policy on sharing research data also needs to address other legal issues in relation to privacy and security.

### **Other countries**

Our analysis of policy and practice in relation to open access to publications and research data in different countries shows mainly that, while there is often explicit government policy on open access to publications, this is not the case for sharing research data. This is due to access to data developing more slowly, and it is acknowledged that there are many obstacles and objections which have to be overcome, including legal issues, but also the need for storage infrastructure (and the costs that



entails). The lack of standards is also seen as a barrier. See also the country overview in Appendix 2.

## 5.4 The impact of (more) sharing of research data

Processing, combining and analysing data is becoming more and more important in the scientific world, and turning into a specialist expertise. Where in the past researchers themselves had the knowledge and skills needed to analyse their data, today they are increasingly working in partnership with data researchers who specialise in this technology. These new working methods (including data mining) offer many possibilities: much larger datasets can for example be linked and exploited, and previously unseen connections become visible. This requires open access to datasets, not just within individual disciplines, but also across disciplines; that in turn requires interdisciplinary standards and interoperability, which are currently often lacking.

Moreover, data sharing is still the exception within many disciplines, let alone sharing with other disciplines or with the outside world. In more isolated disciplines, data are currently not 'published' at all because these disciplines collaborate less with others and individual publications are more usual. Sharing data is sometimes also simply difficult because of the differing nature of the data; quantitative data are for instance easier to share than highly qualitative data. An example is ethnographic research, where it is usual to work with interview transcripts, fieldwork diaries and observation protocols. These qualitative data are something very different from the quantitative measurement data obtained with a radio telescope.

### 5.4.1 Impact on quality of science

#### **Data quality assurance**

There are several areas where we expect sharing research data to have a clear positive effect on the quality of science. First, researchers who (are required) to share their research data are forced to adopt a more rigorous and systematic approach to collecting and handling data during their research. This improves the quality of science. More standards will also be needed if storing and sharing data becomes more common. This will help professionalise research data handling. Sharing the research data that underlies a scholarly article will also undoubtedly boost transparency and therefore help with quality assurance in science, thus contributing to the 'quality control' of the article. It is plausible that testing of the data will in future form part of the review process for a publication.

## Quality indicators

If research datasets become available, a means will need to be found of establishing their quality. More technical, formal standards will apply for such datasets: data stored in a digital archive need to be well ordered, verified and documented. It is necessary to record precisely what the data represent, where they come from, how they are structured, how they were collected, and so on. This requires universal standards, procedures and protocols. Research data that are stored and shared will as a minimum need to meet these standards.

There will also be a need for indicators to mark the 'substantive' quality of research data. How interesting or useful is a particular set of research data? Which (new) indicators will be developed for this? Will there be a kind of review process for research data? Will impact be measured, for example the number of times that a set of research data has been reused or cited? Will the location of the storage be a quality indicator, with data stored in a more prestigious institute being regarded as more valuable than data stored at an institute with a lesser reputation? Time will tell, but if storing and sharing research data grows, so will the need for such indicators.

A good system of 'quality indicators' for research data can also help researchers to evaluate the research activities which generated the data. This receives little attention in the present system; researchers are only evaluated (through higher rankings, esteem, grants, etc.) on their 'end product', i.e. scholarly articles (and possibly other forms of output, such as books), but not on the work that led to that end product. Evaluating this work, too, (e.g. in citation scores and when awarding grants) will underline the importance of good experiments and proper data handling. It will also provide an additional incentive to share research data. This may help persuade researchers who currently have no interest in sharing data to do so.

## Integrity

Sharing research data could help improve scientific integrity if it makes it easier to detect 'fraud'. The Schuyt Committee set up by the Royal Netherlands Academy of Arts and Sciences (KNAW) concluded on this point that "lack of proper archiving of research data makes later verification of the data [more difficult]."<sup>150</sup> Proper archiving, storage and access to data thus make 'verification' possible. This not only boosts the transparency of science, but can also help reduce the number of mistakes, for example statistical errors.<sup>151</sup>

---

<sup>150</sup> Schuyt *et al.* (2012), p. 34.

<sup>151</sup> See e.g. Bakker (2014).

### **Speed and effectiveness of knowledge dissemination**

For this aspect, it is enough to compare the situation in which research data are not shared with one where sharing takes place to some extent. Clearly, in the latter case 'knowledge' (i.e. research data) is disseminated more widely and more quickly than in the former situation. To ensure effective dissemination of the knowledge embodied in research data, it is crucial to ensure the usability and accessibility of datasets. The easier it is for others to use data, the greater the chance that the stored 'knowledge' will be put to good use.

### **Efficiency of science**

Here again, there are a number of immediate positive effects. Current practice in many disciplines which do not share their own research data, or only on a very limited scale, is that data are left on the shelf after publication and never used again. Sharing such data would enable them to be reused, thereby increasing the impact of the original research. This would make science more efficient, because open research data not only makes it possible to 'verify' data from fellow-researchers, but also enables researchers to work together and build on the data. Data then only needs to be collected once, thus avoiding duplication. Research could then for example focus on interpreting existing data rather than collecting new data. Researchers could of course still choose to do their own research, but could aim at supplementing existing datasets in order to take science forward. In this sense, sharing research data can help the 'reproducibility' (and scaling up) of some research. Providing access to research data on a wide scale also paves the way for a new type of research which cuts across individual datasets and disciplines, opening up highly interesting scientific perspectives. Current developments in relation to big data hint at what is possible here. To make all this possible, researchers will have to learn to trust research data collected by others and be willing to genuinely use it.

Experiences in scientific disciplines where collaboration and data sharing within the relevant scientific community are already common show that this undoubtedly opens up new possibilities which would not have arisen without the sharing of data.<sup>152</sup>

Yet there are also a few points to be addressed which could have a *negative* impact on scientific effectiveness. This will apply particularly at the point when making research data freely accessible becomes mandatory (e.g. for part-publicly funded research). There may be strong (and valid) reasons for wishing to protect research data, for example because of its commercial value. This will often be the case, for example, with data

---

<sup>152</sup> Apart from collaboration in the field of high-energy physics (such as CERN) or the Human Genome Project, there are many other successful examples, such as the European Holocaust Research Infrastructure (<http://www.ehri-project.eu/>), ARIADNE (for the integration of archaeological datasets; <http://www.ariadne-infrastructure.eu/>), the Digital Collaboratory for Cultural Dendrochronology (DCCD; <http://dendro.dans.knaw.nl/>), Restructuring Global Inequality (Clio Infra; <https://www.clio-infra.eu/>) and collaboration within the Netherlands on the 1795-1971 Censuses (<http://www.volkstelling.nl/>) or on Dutch overseas trade from 1600-1795 (<https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:33898>).

collected in partnership with private stakeholders, who might in that situation decide that they no longer wish to collaborate on research with publicly funded researchers.

Imposing a duty to make data fully (and speedily) accessible can also encounter an obstacle from within the world of science. As soon as researchers lose their exclusive rights to data, they and their research institutes no longer have an incentive to invest heavily in new experiments and the necessary resources. If research data have to be released quickly, other researchers will be able to avoid having to make those investments and to free-ride using the existing data. The alternative, namely where researchers first place a series of publications on hold in order to defer sharing their data, would have the disadvantage of slowing down knowledge dissemination. If such investments are no longer made, science as a whole is ultimately worse off. However, this also means that top-quality researchers and research institutes would lose some of their competitive advantage. The same applies on an international scale: countries where very high-quality research is carried out (such as the Netherlands) would then be 'giving away' some of their advantage. This is an argument for a more differentiated approach, with scope for the different data sharing options.

#### 5.4.2 Impact on society and business

As well as within science, open access to research data can also have an impact on society and business. How can we assess this impact? There are clear success stories involving the release of research data, which can lead to exciting finds such as the discovery of Hanny's Voorwerp in the Galaxy Zoo Project.<sup>153</sup> But it can also contribute to solving global problems, such as the Ebola epidemic.<sup>154</sup> Scientific and medical data on the Ebola outbreak were for example brought together at the portal site 'eboladata.org', enabling local authorities and aid organisations to design the most effective policy for combating the epidemic. Similar projects include the Humanitarian Data Exchange portal.<sup>155</sup> Sharing research data can also have an indirect impact on society: when research and development accelerate, they help make societally useful applications, such as medicines, available earlier.

There are lots of examples of research data collections that are (potentially) of interest to the public. The Meertens Institute in Amsterdam, for example, has made numerous databases available online, including a database of first names and surnames, but also databases containing information on pilgrimages and patron saints. Making such data easily accessible by the public does however require investments (search functions, visualisations, etc.).

---

<sup>153</sup> For more information on Hanny's Voorwerp and the role of the astrophysicist and guitarist Dr. Brian May in the Galaxy Zoo Project, see: <http://tedxtalks.ted.com/video/The-discovery-of-a-Citizen-Scie>

<sup>154</sup> See Yozwiak *et al.* (2015) and Mylne *et al.* (2014) and also Plaat (2015).

<sup>155</sup> For more examples, see: <https://data.hdx.rwllabs.org/faq>

To obtain an indication of the potential effects of more accessible research data, we looked at experiences in providing access to *government data*, a development that was initiated several years ago. Here, too, a great deal was expected from (extra) use of the data by citizens and businesses. Owing to the breadth of open government data, we narrowed our search to data from the Royal Netherlands Meteorological Institute (KNMI). This case study revealed interesting facts about the use of open data by society and business (see text box). Broadly speaking, the use of KNMI data did not increase spectacularly after it was made freely available (prior to this, the data could be used in return for payment). On balance, there was a loss of income for KNMI because some of the data for which commercial organisations had paid in the past are now available for free. This case study shows that we should not expect too much of (free) open data, although there will always be examples where it is successful. That success will depend in part on the nature and suitability of the data, whether (open) government data or (open) research data. It is also likely that data with the greatest potential commercial value will be exempt from the requirement to share data freely.

### Open government data<sup>156</sup>

In 2011 the then Dutch EU Commissioner Neelie Kroes described open government data as 'a goldmine', which would stimulate innovation, lead to economic growth and increase transparency.<sup>157</sup>

#### Motives

Following the WikiLeaks affair (2008), initiatives for open government were initially motivated mainly by transparency, but more attention was also given to the possibilities hidden away in government data – even large research institutes such as TNO called for more openness.<sup>158</sup> Following similar initiatives in the United States and the United Kingdom, and as part of the Open Government Partnership and the vision on Open Government, steps were taken to promote open government data both in the Netherlands and Europe.<sup>159</sup>

#### Promises

As a result, certain government data are now open and accessible<sup>160</sup> and can be accessed via a single Dutch government portal site: [data.overheid.nl](http://data.overheid.nl). The Dutch employers' federation VNO-NCW believed that government data could unlock a market worth billions.<sup>161</sup> Open data would invite creative use of existing data, sparking

<sup>156</sup> See: RECODE (2014) en [www.recodeproject.eu](http://www.recodeproject.eu)

<sup>157</sup> [http://ec.europa.eu/archives/commission\\_2010-2014/kroes/en/blog/opendata.html](http://ec.europa.eu/archives/commission_2010-2014/kroes/en/blog/opendata.html)

<sup>158</sup> Kotterink & Huijboom (2010).

<sup>159</sup> Capgemini Consulting (2013) summarises the 'success' of open data in different countries.

<sup>160</sup> Binnenlands Bestuur (2015).

<sup>161</sup> VNO-NCW (2012).

new initiatives. This was the chief reason for the Dutch Ministry of Economic Affairs and the Ministry of the Interior and Kingdom Relations to follow the European example and introduce free data at a political level.

There have undoubtedly been some successful initiatives involving the use of government data. For example, the rainfall radar tool *Buienradar* was developed using data from KNMI, and car number plate registers are used in an app that helps in carrying out checks on second-hand vehicles.<sup>162,163</sup>

### **Impact on society**

In practice, however, no other new government data initiative has come anywhere near to emulating the (commercial) success of *Buienradar*, which was moreover actually developed before the open data era (*Buienradar* paid for the data through a data licence). To provide a further boost for successful innovation, the Open (Geo)data Breakthrough Project was launched, one of nine breakthrough projects involving ICT.<sup>164</sup> These projects are designed to prevent the promised goldmine from turning into a bubble.

One explanation for the minimal effect of open government data is that much of it requires interpretation and explanation before it can be used – an intermediary, in other words.<sup>165</sup> In addition, many developments are driven by supply and technology,<sup>166</sup> and those drivers do not change because the amount of open data increases. And a large-scale innovation stimulus based purely on making data freely available is no guarantee of success, because a great deal of data was already available, though at a cost (in the available data era, the precursor of open data, the price of a licence to access government data was equal to the cost price).

On the other hand, open data carried and (and still carries) costs. There is particular interest in time and place-specific data; this offers interesting development possibilities, but much more data would have to be made available. Time will tell whether open government data delivers a greater societal return (benefits versus costs) than the old source-holder model (with the government as the source-holder), under which users paid for a licence with conditions attached.

---

<sup>162</sup> <http://nationaleapprijns.nl/voorbeelden-van-apps/>

<sup>163</sup> NRC Handelsblad dated 15 October 2012, 'De burgerprogrammeur haalt het goud uit de overheidsdata', p. 7.

<sup>164</sup> <http://www.rijksoverheid.nl/onderwerpen/doorbraakprojecten-met-ict/economische-groei-met-doorbraakprojecten-ict> and <http://doorbraakmetopendata.nl/over>

<sup>165</sup> Binnenlands Bestuur (2012).

<sup>166</sup> See also AWTI (2015b).

As an example, KNMI previously generated annual income of €2 million from the sale of (operational) data. As much of these data are now available free of charge, that income has dwindled to around €0.5 million. At the same time, the data is managed more carefully, bringing annual cost savings of around €0.5 million. Overall, therefore, in the case of KNMI open data costs the Dutch taxpayer EUR 1 million per year.<sup>167</sup>

### **Barriers**

Even where there is a social impact in the longer term, it is a political judgement as to whether it is fairer for all taxpayers (rather than only the users) to help meet the costs of open government data and thus subsidise new initiatives.

It is also interesting to see how the 'openness offensive' began within the government. As in science, the government culture was generally a closed one, and as with researchers, civil servants had nothing to gain from openness. Yet today the government gives away (some of) its data for nothing.

Data are currently still difficult to find, difficult to use, relatively un-standardised, fragmented across different portals and requires considerable management.<sup>168</sup> Standards are essential for ensuring the usability of open data. The government recognises that it had a role here, but this too has proved more difficult than anticipated.<sup>169</sup> In addition, these technical problems, the nature of the data, the context in which the data is framed and the way the message about open data is communicated have meant there is little interest in the data. Making this project a success requires not only more time, but also clear agreements between the producers and users of the data, as well as control and coordination in order to guarantee standardisation and reliability.

### **5.4.3 Costs and benefits**

There is no doubt that storing and sharing research data initially brings extra costs. More attention and time have to be devoted to data management, data curation is necessary and storage has to be paid for. It is difficult to make a precise estimate of the extra costs, but they are often put at around 5% of the total research budget.<sup>170</sup> In practice, of course, the precise figure will depend on the nature of the data and the type of research: the easier it is to make the data suitable for storage, the less it will cost. As the estimate of 5% appears to be based mainly on examples where a lot of data is already being stored, we suspect that the costs will be higher for research data which are less easy to store.

---

<sup>167</sup> Brouwer (2014).

<sup>168</sup> Binnenlands Bestuur (2014), Sluiter (2013) and Kaasschieter (2015).

<sup>169</sup> <http://www.rijksoverheid.nl/onderwerpen/digitale-overheid/open-data-en-open-standaarden>

<sup>170</sup> From an interview with SURFSara in May 2015.

The 'costs' will be borne by researchers (who will have to spend more time on data management) and their institutes and funders. Who bears the costs of curating and storing the data will depend on the agreements made: they could be paid by individual researchers or institutes to those providing the services, or they could be facilitated at central level (as is currently the case).

On the other hand, storing and sharing research data should also generate income: from improving the research process and its efficiency, from greater system efficiency (e.g. avoiding duplication of research) and from establishing new connections (scientific innovation). The question is how this income will compare with the extra costs. A difficulty here is that most of the 'income' is currently hard to estimate in monetary terms. Sharing research data could also contribute to improving scientific discipline: if others are able to view the research data underlying a publication, researchers will have a greater incentive to go about their research with methodological rigour.

One aspect relating to the costs and benefits is that there is likely to be a learning effect on the costs side, so that the additional costs will eventually fall. On the income side, the reverse may apply: the learning effects here could lead to an increase in the benefits over time. This is an argument for adopting a cautious approach to sharing research data, for example using pilots to identify these learning effects and to enable estimates of the costs and benefits to be based on actual experiences.

Finally, it is important to be aware that data storage costs money. Although it is technically possible, it is not practically feasible or sensible to store all data. Choices have to be made. Unfortunately, it is usually only possible to say in retrospect which data are worth storing. Yet clear agreements will have to be made on this. A strategic long-term vision is needed, but also a vision that looks beyond national borders. Science is an international activity, and agreements therefore also need to be supported internationally.

#### **5.4.4 Basic conditions**

We argued above that sharing research data more widely can in principle have a number of positive effects. However, there are also a number of caveats. In order to make the 'usable' sharing of research data more the norm, certain basic conditions will have to be met. These fall into the following categories: skills, infrastructure, trust, incentives and agreements. They are briefly explained below.<sup>171</sup>

Researchers need certain skills and support to share research data successfully. Elements such as good data management could be incorporated into training programmes and curricula as part of the Open Science Literacy project.

---

<sup>171</sup> The ten principles for open research data<sup>1</sup> that were developed and recently published in the United Kingdom are also interesting (Concordat Working Group, 2015). They correspond fairly closely with the recommendations in this AWTI report.



Good infrastructure is also needed, both for users wanting to find, read and use the data, and for researchers supplying the data who need to be able to be sure that it is being properly stored.

Trust is important in providing access to and reusing data. Before sharing research data, a researcher must be able to trust that it will be used in the right way. Broadly endorsed agreements will therefore have to be made in the scientific community in relation to the sharing and reuse of data, with peers holding each other accountable. Trust in the facilities used to store research data is also important.

Conversely, potential users of stored research data must be able to rely on its quality. This is essential for reciprocity, i.e. being prepared both to share and use research data. This could be reinforced using quality marks such as the Data Seal of Approval<sup>172</sup> (quality of storage) and by means of open review platforms (quality of data).

Building this trust within the scientific community takes time. It is therefore not wise, especially in the first instance, to oblige researchers to share their data unconditionally with everyone. It is better to leave some control in the hands of the researcher and/or their institution to decide who they share their data with, for what purpose and under what conditions. The question of who 'owns' the data also arises here. For researchers, it is essential that control of research data, especially those obtained with public funding, remains within the scientific community, with researchers, institutes or funders. By their nature, these legal issues require an answer from government as the legislative power.

There are also a number of risks which should not be too easily dismissed. One is incorrect use of data: epidemiological data, for example, can be interpreted in all kinds of ways and all kinds of connections found, not all of which are correct. The national and international scientific community also has to be aware of free-riding, where researchers from elsewhere publish freely on the basis of data provided by others, possibly even before the researchers who actually collected the data. This is a real risk, as became clear in our interviews.

This latter risk is also closely linked to the 'strategic' effects if the Netherlands or the EU makes data, and especially research data, more easily accessible. This 'giving away' of research data to other countries is not by definition a bad thing, but from the perspective of costs and strategic considerations it is important to make the right judgement calls. Other countries are being given access to new information; do they provide access to their data in return? And what happens to the competitive advantage (within the global scientific community) of the locations in the Netherlands and the EU where top-quality

---

<sup>172</sup> <http://www.datasealofapproval.org>

research is carried out? Too readily providing access to the results of this research to other non-EU countries could limit or even nullify that competitive advantage.

This means first and foremost that the appropriateness of sharing data must be carefully considered, and secondly that it can have consequences for how and with whom research data is shared. This underlines yet again the importance of a gradual approach in deciding whether or not to provide access to research data. The EU and Member States could for example take into account the degree of openness applied by those third countries themselves (a form of reciprocity).

Trust alone is not enough to give genuine impetus to the wider sharing of research data. There must also be an incentive to share data, a 'carrot'. Researchers can be persuaded to release good datasets in various ways. Research Data Netherlands (RDNL) awards the annual Data Prize, for example, which rewards a researcher's hard work but also motivates other researchers to treat their data in the same way.<sup>173</sup>

However, such 'one-off' actions will not be enough to make the widespread sharing of research data a reality. Researchers want to be recognised even without a prize for the work they have invested in the data they release. Ways must therefore be sought of valuing the sharing and reuse of data. A good start has been made by linking a DOI (digital object identifier) or other persistent identifier to data so that the dataset and therefore the researcher who produced it can be cited and evaluated. The ability to publish datasets can also help improve data visibility and appreciation. More account could also be taken of data management when awarding research grants, rather than simply focusing on the end product, namely the scholarly article.

### **From data access to data mining**

There is still a long way to go in the field of text and data mining – specifically searching for statistical and other connections in a corpus of information in the form of text and/or data. The United States appears to be further advanced in this regard, as do countries in Asia and South America, according to a report by an expert group which carried out a study on standards in relation to text and data mining.<sup>174</sup> There is thus untapped potential here, because text and data mining is expected to offer opportunities for discovering new connections and links between diverse disciplines.

Researchers state that this untapped potential is due mainly to restrictions relating to copyright, database protection and data privacy. Where a fair use principle is applied

---

<sup>173</sup> <http://www.researchdata.nl/activiteiten/dataprijs/>

<sup>174</sup> European Commission(2014).

in the United States, the EU has a more complex framework for copyright. The US also has no legislation on database protection.

There are however many developments taking place from the bottom up. In practice, mining both text and data is a sensitive issue. For example, publishers have an interest in text mining, especially if carried out by themselves. They are after all in control a high proportion of articles. For example, Elsevier limits batch access to a maximum of 10,000 articles per month. The open access publisher PLOS requires authors to sign to indicate that they permit the free use of their articles, including for text mining. Data mining is an even more difficult area, simply because thus far data are not always publicly available, let alone suitable (standardised and provided with metadata) for data mining.

In seeking to allow greater use of text and data mining, a balance needs to be struck between the different valid interests. Where that balance lies is currently difficult to determine because there is little empirical evidence for the benefits. In the context of this report, the main aspect is the ability to carry out text and data mining on (publicly funded) scholarly publications and research data. Whether or not mining will be facilitated on a wide scale will be a political judgement because it will require the (partial) lifting of rules imposing limits on mining in order to protect copyright and databases.

## 5.5 Conclusion

There is growing demand in the scientific world for open access – including to data – to enable research to be validated and replicated, but also because demand for access to large datasets is growing in general and because data research is becoming a more independent discipline. Sharing research data is thus a form of ‘open content’, just like open access to publications.

Yet there is a difference between open access to publications and (open) access to research data. Researchers will not be willing to give away their research data unconditionally. Articles are written to be shared with the whole world, or at least with scientific peers, but this does not hold for data. In many disciplines, data remain with the researchers and their institutes and are often only shared where there is reciprocity and recognition of the researchers’ efforts. There are enormous differences in the way data are treated in different disciplines and even by individual researchers. Moreover, well-meaning researchers are not always in a position to share data: they do not know how to

do so, have too little (infrastructural or financial) support or are prevented from doing so by legal constraints.

The decision on whether or not to share data is generally taken by researchers themselves, or sometimes by their employers or funders. These stakeholders have different motivations for sharing or not sharing their research data, publicly or with peers. In some branches of science it is almost essential to generate research data in collaboration with others and then to share it, for example because of a large common infrastructure or because they are working together to solve a large scientific puzzle which goes beyond the scope of any one research group. In disciplines where research data are never or almost never shared, in particular, it is unclear for many researchers what is to be gained by allowing access to their data. At first sight, it costs extra time and trouble to prepare the data for storage and reuse, while there is also the risk that another party will publish 'their' research data sooner. Some researchers also do not appreciate the idea of their data being reused later by others.

The move towards open research data in publicly funded research cannot be seen in isolation from the broader debate on access to data collected using public money (see also the case study on open government data, where access was provided to data both as a matter of principle and with the promise that society would benefit from it). Context and clarification, for example including the release of software, but also the ability to process and mine data, are essential for deriving added value from open data. Time will tell whether open data – both government data and research data – genuinely delivers an adequate return for society.

In our view, the potential (positive) impact on science is greater. First, making it mandatory to store and share research data will prompt researchers to adopt a more rigorous and systematic approach to collecting and processing data. Quality standards and standardisation will also force researchers to supply their data in usable form. Good data management should in reality be a must in good scientific practice, of whether or not it is shared. Oddly enough, however, good data management is not yet generally embedded in many disciplines, or at least there is no verification of that and nothing has been published on it. Making good data management a standard component of research and development can certainly do no harm in improving the quality of science.

If the next step is a move to the proper storage and sharing of data (providing access), this would boost the efficiency of science, because public funders will get more for their money. Research will become more transparent if the data underlying a scholarly article can be consulted; the efficiency of science can be improved by avoiding unnecessary duplication of research, and it will help the reproducibility of research if the different 'reproductions' are accessible. Moreover, possibilities will emerge for a new type of

research that cuts across individual datasets and disciplines. All this does require that researchers learn to trust research data collected by others and that they are willing to use it. In addition to all of this, the wider storage and sharing of research data could also contribute to the integrity of science, because it makes verification easier.

In this sense, open data will deliver benefits to society, at least indirectly. However, it remains to be seen what the best level of 'openness' is for achieving this. Although total openness is an option, making data accessible costs money, not least because of the time and money tied up in storing it, and if knowledge is 'given away' freely to countries which share less knowledge than they receive, it is questionable whether that is wise for the country that does the giving. In some cases, therefore, it could make more sense not to share all research data entirely freely, but to adopt a 'sharing-light' mode.

Society and business could also benefit from open research data. Some data will be of direct interest and/or directly usable for citizens and businesses, other data probably less so. There are also indirect effects on society, however: if sharing research data speeds up research and makes it more effective, beneficial results or applications (such as medicines) could become available more quickly. To promote maximum use of data by society or business, research data could be made available and usable at no cost (open data). Time will then tell what effect this has in practice. Our case study on the use of open data from the Royal Netherlands Meteorological Institute (KNMI) showed that *in that* case the social impact of open data was limited.

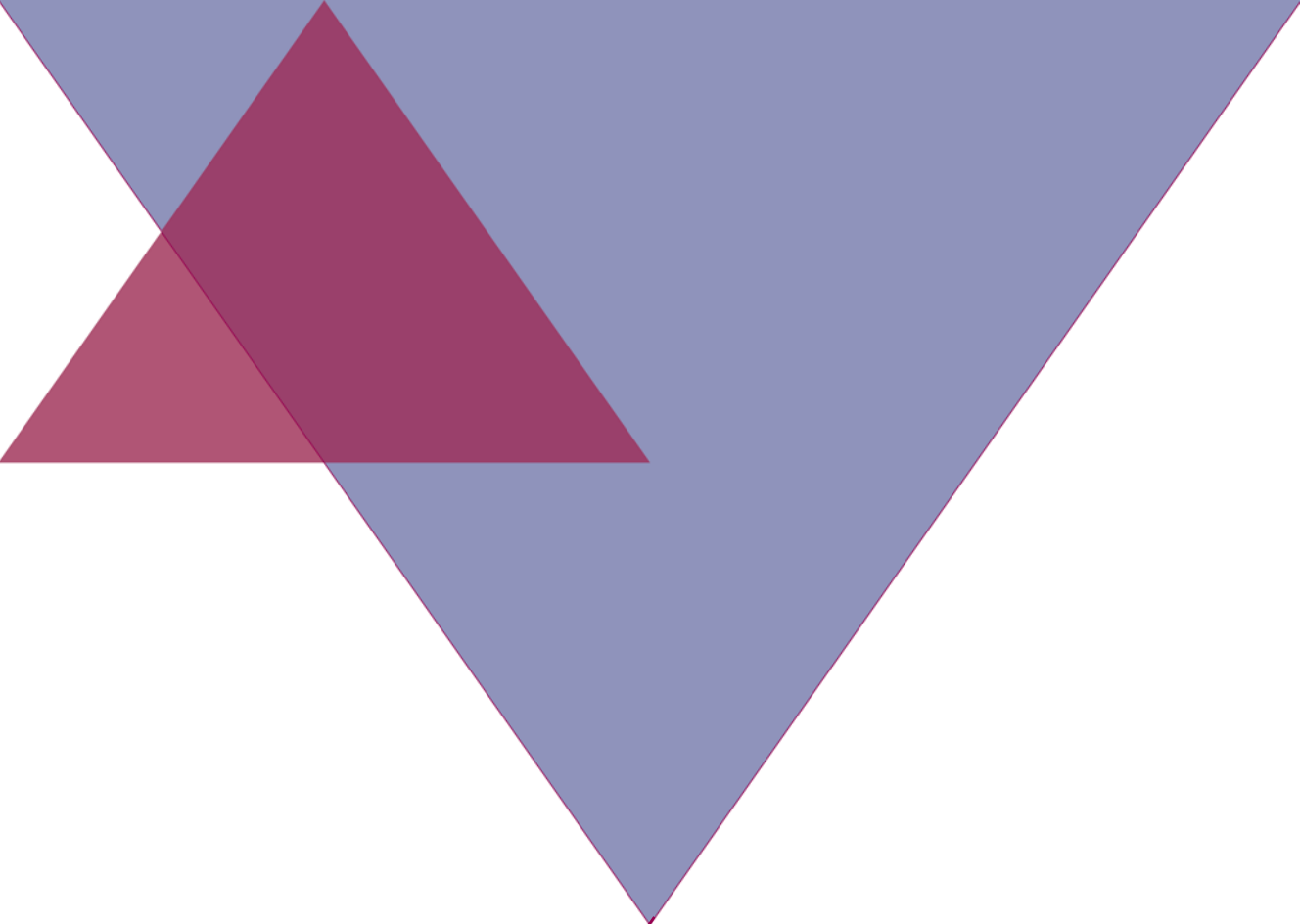
There is a clear asymmetry in the costs of providing access to and sharing research data. Logically, the costs will increase because researchers will have to spend more time and/or money on data management, preparing data for storage and reuse, and the storage itself. The benefits of sharing research data are much less clear for individual researchers at present. An incremental approach is needed here, using pilots, trials, etc., to try and clarify the benefits of open access to scientific information. It is not wise simply to store all research data: choices will need to be made.

The costs and benefits in relation to the useful deployment of research data by society and business need to be weighed carefully. Which costs are borne collectively and where is an individual contribution reasonable? Particularly where research data are of interest for a commercial application, it is not unreasonable to ask for payment for that commercial use.

All in all, there is still a long way to go before the sharing of research data has become the norm across the whole scientific research spectrum. Researchers will first need to be convinced of the value of this, and many conditions will have to be met, such as good data management, a storage infrastructure, storage standards, trust among researchers,

and the right incentives (higher rankings for sharing research data). Legal questions will also have to be resolved, for example relating to the protection of privacy, copyright and databases. The need for a cohesive approach demands some form of overarching control and coordination, a role that the government could (and should) take up.

It is not advisable simply to open up all research data to everyone unconditionally. There may be good reasons for not doing this, such as privacy or commercial interests. Moreover, support for open data access will be increased if researchers and their institutions are given (joint) control over who they share their research data with and for what purpose. This is an argument for carefully exploring the different options for sharing research data.



# Appendices





# Appendix 1 Request for advice



Ministerie van Onderwijs, Cultuur en  
Wetenschap

>Retouradres Postbus 16375 2500 BJ Den Haag

AWTI  
T.a.v. de heer prof. dr. U. Rosenthal  
Javastraat 42  
2585 AP DEN HAAG



Onderzoek en  
Wetenschapsbeleid  
IPC 4100

Rijnstraat 50  
Den Haag  
Postbus 16375  
2500 BJ Den Haag  
www.rijksoverheid.nl

Contactpersoon

R.H. Derksen  
T +31 6 31 74 91 78  
r.h.derksen@minocw.nl

Onze referentie  
775006

Datum 16 JUNI 2015  
Betreft Adviesaanvraag Open Science

Geachte heer Rosenthal,

Nederland bekleedt de eerste helft van 2016 het voorzitterschap van de Europese Unie. Bij deze gelegenheid wil het kabinet een impuls geven aan de ontwikkeling van 'open science', met name ten aanzien van 'open access' en 'open onderzoeksdata'.

Bij dezen vraag ik aan de AWTI een advies uit te brengen dat met het oog op het EU-Voorzitterschap helpt om invulling te geven aan de strategie ten aanzien van open science. Bij dit onderwerp is ook het ministerie van Economische Zaken betrokken.

De centrale vraag voor dit advies luidt als volgt: "Welke maatschappelijke effecten zijn er te verwachten door de ontwikkeling van open science, met name van open access tot wetenschappelijke publicaties en open onderzoeksdata?"

Om deze vraag te beantwoorden, kan de AWTI de volgende deelvragen adresseren:

- Welke openheid van de wetenschap is maatschappelijk gewenst?  
Wat zijn de maatschappelijke voor- en nadelen van open access en open onderzoeksdata voor de samenleving als geheel, en voor de wetenschap en voor het bedrijfsleven in het bijzonder? Bij wie komen de baten en bij wie de kosten terecht?
- Wat zou de rol van de Nederlandse overheid moeten zijn?
- Welke inzichten zijn er in andere Europese landen rondom maatschappelijke effecten van open access en open onderzoeksdata?
- Welke posities nemen voor de wetenschap belangrijke derde landen in?

Ik zie uw advies graag begin september tegemoet.

Met vriendelijke groet,  
de staatssecretaris van Onderwijs, Cultuur en Wetenschap,

  
Sander Dekker

Ministry of Education, Culture  
and Science

AWTI  
Attn. Professor U. Rosenthal  
Javastraat 42  
2585 AP THE HAGUE

Date: 11 JUNE 2015

Re.: Request for Advice on Open Science

Dear Professor Rosenthal,

The Netherlands will hold the presidency of the European Union in the first half of 2016. The Dutch government wishes to take this opportunity to stimulate the development of 'open science', particularly with regard to 'open access' and 'open research data'.

I am writing to request AWTI to prepare an advisory report which, in the light of the Dutch EU Presidency, will help in fleshing out the strategy with respect to open science. The Ministry of Economic Affairs is also involved in this project.

The central question to be addressed in this advisory report is as follows: 'What social impact can be expected from the development of open science, and in particular open access to scholarly publications and open research data?'

In order to answer this question, AWTI may wish to address the following constituent questions:

- What degree of open science is socially desirable? What are the pros and cons of open access and open research data for society as a whole and for science and industry in particular? Where do the benefits accrue and who bears the costs?
- What should the role of the Dutch government be?
- What can be learned from other European countries concerning the social impact of open access and open research data?
- What role do key third countries play?

I look forward to receiving your report in early September.

Yours sincerely,

The State Secretary for Education, Culture and Science,

[signed]

Sander Dekker

## Appendix 2 Open access and open data in different countries

### International comparison

Science is a global activity, and many disciplines operate on a global scale. Although the Netherlands makes a relatively large contribution to this global activity, in total it still accounts for not much more than 1 to 2.5%.<sup>175</sup> The European Union (EU) is by contrast a large player, producing around 30% of scholarly articles worldwide, slightly more than the US, though the publications from the US have the biggest impact.<sup>176</sup> This means that an individual country such as the Netherlands has limited scope for unilaterally changing the system of science. The EU as a whole is in a better position to push boundaries and initiate changes in the system, but the EU, too, represents a minority on the world scientific stage.

For the purpose of this report, this means it is highly relevant to identify how access to publications and research data is approached both within the EU (with a view to arriving at a common European approach) and in a number of important non-EU countries. We therefore collected further information for this report about a number of countries in order to obtain a clearer picture of the differences in approach and to bring together more specific experiences for use in our analysis.

Input was received via sister organisations of AWTI and the network of Neth-ER (the *Netherlands house for Education and Research* in Brussels) for the discussion and developments in relation to open access and open research data in different European countries. The input concerns both the government policy and the role and position of the relevant stakeholders. AWTI received information from Belgium (Flanders), Denmark, Germany, Finland, France, Latvia, Norway, Slovakia, Spain, Czech Republic, Sweden and Switzerland.

AWTI also carried out its own case studies on policy and developments in the United Kingdom, the United States and Brazil. The United Kingdom is regarded as one of the pioneers in the field of open access, particularly within the EU. It also hosts a number of important scientific publishers. The United States is the world leader in scholarly publications, in terms of both number and impact, and also has in place strong federal

---

<sup>175</sup> These figures are not easy to determine, because they depend on what is and is not included, but the Dutch Ministry of Education, Culture and Science itself states a figure of 2.5% on its website: <http://www.ocwincijfers.nl/wetenschap/inhoud/wetenschappelijke-publicaties>

<sup>176</sup> Figures from Elsevier (2013) and National Science Board (2014).

research funding regulations in relation to open access. Finally, Brazil appears to have become the standard for open access publishing, partly because of a very clear mandate imposed by the most important public funder of research.

The main findings for the various countries are set out below. The level of detail varies, depending on how full the responses from the different countries were. The study of open access and data sharing in Brazil was carried out in the context of the AWTI report on collaboration between the Netherlands and Brazil, for which reference is made to the report of the study visit to Brazil.<sup>177</sup>

### **Belgium (Flanders)**

In its policy memorandum on employment, the economy, science and innovation for the period 2014-2019 ('2014-2019. Werk, Economie, Wetenschap en Innovatie'), the Flemish government sets out a clear and broadly supported policy for open data and open access in order to promote spillover effects. The government for example wishes to encourage research institutes to publish their results in open access and to publish the research data underlying those publications as open data. Access to information on publicly funded research in Flanders, including links to publications of datasets, will be guaranteed via the research portal FRIS: [www.researchportal.be](http://www.researchportal.be). In addition, the government is following up on innovative open data projects already under way (such as the project led by iMinds and the RILOD pilot (Research Information Linked Open Data project) led by the Department of the Economy, Science and Innovation) and will investigate whether and how new initiatives can be developed.<sup>178</sup>

In addition, the federal Belgian government has signed the 'Brussels Declaration on Open Access' together with 18 Belgian research institutes and research funds.<sup>179</sup> In this declaration, researchers and research institutes commit to following the gold or green open access route. Belgium is one of the countries that mainly uses the 'green' route, and is one of the three countries with the fewest 'gold' open access journals.<sup>180</sup>

The main aim in relation to research data is to strengthen collaboration between knowledge institutes and organisations which provide digital infrastructure, as well as data archiving and curation services.

### **Denmark**

The previous Danish government began drawing up a strategy for open access. The target is that 80% of all peer-reviewed articles should be open access in 2017, and 100%

---

<sup>177</sup> Corbey (2015), *Brazilië: impressie van een land in verandering*, The Hague: AWTI, juni 2015.

<sup>178</sup> Vlaamse regering (Flemish government) (2014).

<sup>179</sup> <https://openaccessbelgium.files.wordpress.com/2012/11/signedbrussels-declaration-on-open-access.pdf>

<sup>180</sup> Archambault *et al.* (2014)

in 2022. A 'national committee' has been set up to oversee the implementation, in which representatives of stakeholders (universities, national and university libraries, science funders, private funds, etc.) participate.<sup>181</sup>

Denmark has opted for the 'green' route: the 'gold' route would require too large a system change. The specific plans for 2015 are to develop an open access indicator (to enable the open access achievements in each knowledge institute/discipline to be accurately assessed) and to set up a platform aimed at the more technological aspects.

As regards benefits to society, the Danish government thinks less in terms of businesses and more of small research institutes, teachers, general practitioners, and so on, who will benefit from open access to articles. To raise enthusiasm among most businesses and civil-society organisations for this form of knowledge utilisation, an intermediary organisation/helpdesk will be needed to help point them in the right direction and translate the information into a form they can use.

Open access is high on the political agenda, though there is a sense that it is not yet high enough on the science agenda. Better and more intensive dialogue is needed with researchers and university managements. There is a need for more awareness, more information campaigns and more training, both for open access and much more so for open research data.

## Germany

Broadly speaking, there is support both at federal level and in the *Bundesländer* for the 'green' and the 'gold' routes to open access. Policy on this is still being developed, though with the caveat that many *Bundesländer* do not yet have a strategy for open access.

The debate on open access at federal level has long been placed in a broader context. In 2009 a commission on 'The future of information infrastructures' adopted open access as one of eight themes, which also include more virtual research areas, research data, licences, and long-term archiving. The Infrastructure Council which grew out of this has not yet made any strategic choices in relation to open access or made any recommendations for the further development of open access or research data management.

The coalition agreement between the CDU, CSU and SPD parties (2013) states that an open access strategy must be developed, extending not just to open access publications, but especially to all kinds of issues relating to open data. No decision has been made on

---

<sup>181</sup> Danish Ministry of Higher Education (2014).

the preferred route, though the principle has been adopted that publicly funded research must be accessible and usable without financial, technical or legal barriers. This also applies for 'quality assured text publications as well as other digital objects, including research data'. The previous government (CDU/CSU/FPD) had already included a stipulation in the copyright legislation that the author of a scholarly article has the right to publish the accepted manuscript version no later than 12 months after publication of the article (green route).<sup>182</sup>

A number of federal states have made statements about open access:

- ▶ Baden-Württemberg: the government of Baden-Württemberg has ordered universities and higher education establishments to take measures to increase the accessibility of scientific results. The route chosen ('gold' or 'green') is left to the institutions themselves, as part of their autonomy. In the medium term, the intention is that the number of open access publications in all disciplines should rise.
- ▶ Schleswig-Holstein: the open access strategy announced in 2014 comprises three elements: a) the development of a publication server to cover the entire state; b) the formulation of an open access policy for all universities and higher education establishments; c) the setting up of a publication fund. The aim is that at least half of all publications should be open access by 2020. A key point for discussion is the development of indicators for monitoring open access. A further principle is that open access is not a stand-alone project with a clear start and end point, but is more a first step in a transformation process which also includes open access to research data and open science, i.e. all three phases of the scientific process.
- ▶ The Berlin federal state parliament decided in 2014 to develop an open access strategy in which both the 'green' and 'gold' routes will be explored. An open access network of representatives from science will be formed for this. Attention will also focus among other things on the ICT infrastructure.
- ▶ Nordrhein-Westfalen and Brandenburg are in the early stage of developing a strategy for open access.

The other federal states have not yet developed or initiated an open access strategy, though the 16 federal states do have a total of 150 repositories (mainly at institutions) which provide open access to publications.

Both universities and higher education establishments and science institutes (especially the Leibniz Gemeinschaft and the Max Planck Gesellschaft) are actively engaged in developing an open access culture. The science institutes which have joined forces in the

---

<sup>182</sup> §38 (4) Urheberrechtsgesetz reads: „Der Urheber eines wissenschaftlichen Beitrags, der im Rahmen einer mindestens zur Hälfte mit öffentlichen Mitteln geförderten Forschungstätigkeit entstanden und in einer periodisch mindestens zweimal jährlich erscheinenden Sammlung erschienen ist, hat auch dann, wenn er dem Verleger oder Herausgeber ein ausschließliches Nutzungsrecht eingeräumt hat, das Recht, den Beitrag nach 12 Monaten seit der Erstveröffentlichung in der akzeptierten Manuskriptversion öffentlich zugänglich zu machen, soweit dies keinem gewerblichen Zweck dient. Die Quelle der Erstveröffentlichung ist anzugeben. Eine zum Nachteil des Urhebers abweichende Vereinbarung ist unwirksam.“

'Allianz der Wissenschaftsorganisationen' are particularly active in the area of open access, and in February 2015 the ad hoc working group (*Arbeitsgruppe*) 'Open access Gold' published a position paper. Although the 'green' route is not ruled out, the ultimate aim is the 'gold' route.

The following are a few initial observations:

- ▶ The hybrid model can only be supported if the publisher aims for openness (i) regarding the status of the (gradual) conversion of a journal to open access; (ii) regarding a set-off in favour of the knowledge institute in order to avoid 'double-dipping'; and (iii) regarding optimum findability of open access articles.
- ▶ The provider must be transparent about how publishing fees (APCs) are calculated. It is recommended that science funders set a ceiling. [*Transparenz*]
- ▶ Contracts between providers and knowledge institutes must no longer contain confidentiality clauses. 'Big deals' are harmful to competition. [*Wettbewerb*]
- ▶ The legal and technical aspects of reuse of open access publications must be regulated through standardised licences and formats. Providers must also ensure maximum visibility of open access publications. [*Nachhaltigkeit*]
- ▶ In the transformation to open access, the different needs and principles of each discipline must be taken into account. Science, technology and medical (STM) disciplines are further advanced in the move to open access publication than the arts and humanities, where monographs are much more common. [*Pluralität*]
- ▶ The Max Planck Digital Library believes that a switch from a subscription model to an open access model can be funded in a budget-neutral way. In other words, the subscription budgets should be sufficient to pay the required APCs<sup>183</sup> (though a transitional budget will be required). [*Wirtschaftlich*]
- ▶ Still lacking, and badly needed, is an economic analysis of the emerging open access market in Germany based on the study 'Developing an effective market for APC'.
- ▶ Quality assurance is important. Publishers must safeguard existing procedures and at the same time create scope for innovative developments in the area of quality assurance, for example through open review processes.

Germany is still mainly engaged in exploring the opportunities, problems and feasibility of open access. The political will for open access is present at federal level and growing at federal state level. There is some preference – especially on the part of knowledge

---

<sup>183</sup> Ralf Schimmer and others from the Max Planck Digital Library make the following calculation. For Germany, they assume 70,000 APC-relevant articles out of a total of 103,000 scientific articles (Web of Science). With an average APC of €2,000 per article, this takes the total 'gold costs' to €140 million per year. The current total subscription fees in Germany are estimated at €200 million per year. For France, the calculation is 46,000 APC-relevant articles (out of a total of 71,000), leading to total APC costs of €92 million compared with €120 million in current subscription costs. For the UK, they arrive at 72,000 APC-relevant articles (out of a total of 109,000), leading to total APC costs of €144 million per year compared with estimated subscription costs of €218-250 million. For the Netherlands, however, we (AWTI) calculate 24,000 APC-relevant articles out of a total of 38,000 articles, giving total APC costs of €48 million compared with estimated subscription costs of €40 million per year. For the Netherlands, therefore, the cost goes up!

institutes – for the gold route, though the possibility of the green route is certainly being kept open. Of particular interest are the initial thoughts and positioning on a number of key themes by the scientific association *Allianz der deutschen Wissenschaftsorganisationen*. The focus is on aspects such as sufficient transparency, quality assurance, visibility, the level of APCs, payment of APCs where there are several authors, and legal and technical aspects.<sup>184</sup>

## Finland

Open Science is one of the top priorities in Finland's science policy. It is regarded as important for improving the quality and integrity of science and for securing a greater social and economic impact. Finland has opted for the 'gold' route, for which the Ministry of Education and Culture has set out a *roadmap*<sup>185</sup> containing an agenda for 2014-2017 on how and under what conditions the route to open science can best be paved in collaboration with the various stakeholders.

There are four goals:

1. Strengthen the functioning of science itself: openness and repetition of research strengthen the reliability and quality of research.
2. Strengthen expertise in dealing with 'openness', so that those employed in the Finnish science system know how to use the opportunities offered by openness and thus to boost Finland's competitiveness.
3. Ensure a stable basis for the research process: good, clear basic structures and services create new opportunities.
4. Strengthen the social impact of research: open science offers new opportunities for researchers, policymakers, business, civil-society organisations and citizens.

This is intended to lead to a system of open science in which scientific knowledge and data circulate freely. To achieve this, specific facilities such as storage formats, better findability and linking of scientific knowledge are needed. This will also mean a new way of working for researchers and other stakeholders, as well as much more interaction between them.

Some of the measures that need to be taken involve analysing aspects such as the present working culture in research environments or the formulation of indicators to measure the degree of openness and accessibility. Some fairly specific measures are also needed, such as setting up an accreditation system for open science or organising forums. Finally, a large number of pilots are being carried out.

---

<sup>184</sup> See the *checklist* (pp. 28-33) of the *Allianz der deutschen Wissenschaftsorganisationen* (2015).

<sup>185</sup> Ministry of Education and Culture of Finland (2014).



The Academy of Finland (the national research funding body) first of all encourages researchers to follow the 'green' open access route by placing their publications in institutional repositories or, if these are not available, on social networking sites such as [www.academia.edu](http://www.academia.edu) or [www.researchgate.net](http://www.researchgate.net). The 'gold' route is also supported. Funds are available to cover APCs. Publishing in hybrid journals is not encouraged.

Research funding is conditional on the inclusion of a publication plan and a data management plan. Long-term storage of data is also encouraged, though plans for this need to be developed further.

The national innovation funder TEKES recommends that research results be disseminated as widely as possible. The policy on open access and sharing research data within organisations is currently being brought into line with the requirements of Horizon 2020.

Finland does not have any large publishers. Most articles are published through learned societies, which charge reasonable prices for journals and have virtually no profit motive. These learned societies have their doubts about the open access business model, because they foresee that they will lose their most important and reliable source of income. The Federation of Finnish Learned Societies is exploring alternative funding models. There is a high level of willingness and enthusiasm among knowledge institutes for open access: 22 of the total of 30 rectors are committed or strongly committed to promoting openness.

## France

Three types of institutions are involved in the research process in France: universities, *grandes écoles* and research institutes. The first two also have a strong educational component. The French government and the research industry have long been advocates of open access. The government considers it important that scientific results are made public as far as possible. Since 2000, France has followed its own path here by setting up the HAL platform (*Hyper Article en Ligne*). This platform links together all institutional repositories and serves as a national open archive for all disciplines. It can be used to search for pre-final versions of articles and theses. In the years since then, policy has been aimed strongly at improving and extending the network of repositories. This policy received a boost in 2006 with the signing of a Memorandum of Understanding for a 'coordinated national approach to open archiving of scientific output' by the main research institutes, the associations (*conférences*) of universities and the *grandes écoles*. In addition, the French research funder ANR and the Inserm Institute for Medical Research require that all research funded wholly or partly by them must be placed in an open archive as soon as possible (with embargoes of 6-12 months).

France has a growing number of open access journals, but primarily in the social sciences and humanities, with far fewer in the STM (science, technology and medicine) disciplines or in economics and law. The reason is probably that the latter disciplines far more often choose to place a pre-final version in a repository and are therefore able to continue publishing simultaneously in journals with a high impact factor. It is also striking that many French open access journals do not charge a publication fee (APC). Research from 2009 shows that between 67% and 83% do not charge a fee and are dependent on other sources of income.<sup>186</sup>

The French government is especially keen on open archiving, and appears to be more focused on the 'green' route. Officially, however, politicians have not chosen a specific route. The French education minister Geneviève Fioraso spoke in 2013 in favour of open access to scientific information: *'le gouvernement français réaffirme son soutien au principe du libre accès à l'information scientifique.'* She did however express doubts about the 'gold' route, because 1) the transition costs are very high; 2) there is a risk of loss of credibility for articles paid for on a pay-to-publish basis; 3) smaller academic disciplines may be more likely to fall victim to such a change; 4) there is a danger of growing inequality in the ability to publish.<sup>187</sup> Her preference is for a 'platinum' route<sup>188</sup> in which readers and/or authors have free (online) access to scholarly publications.

## Latvia

The Latvian government takes a positive stance on open access, seeing it as a good development for the future. At present, funders do not make open access publishing mandatory. The research policy is focused more on enabling researchers to publish in journals with a high impact factor.

## Norway

The Research Council of Norway embraces the concept of open access because it is good for the integrity of the scientific process and speeds up scientific development and innovation. It also helps to give scientific institutes with little money, such as many institutes from developing countries, better access to scientific literature.

Both the 'green' and 'gold' routes should be possible, but the Council has a preference for the gold route. Funds are available in the period 2014-2019 to cover the costs of APCs. The initial funding is 9 million kroner (just over €1 million) per year, with co-funding

---

<sup>186</sup> Couperin (2010).

<sup>187</sup> Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche (2013), *discours de Geneviève Fiorasso lors de 5<sup>e</sup> journées Open Access*, 24 januari 2013.

<sup>188</sup> Platinum open access: 'this option allows open access to the published "Version of Record," without payment of an APC. Platinum OA journals are normally supported by grants or subsidies and a Board and editors who conduct editorial activities on a voluntary basis. This definition formerly applied to the term "gold open access", according to <http://editorresources.taylorandfrancisgroup.com/open-access/>

required from research institutes. After 2019, APCs must be incorporated in the indirect costs of institutes' R&D projects.<sup>189</sup> An embargo of a maximum of six months after publication applies for articles in the science, technology and medical disciplines, and 12 months for articles in the social sciences and humanities.

## Slovakia

The open access policy in Slovakia is still in its infancy. The Slovak Centre for Scientific and Technical Information (SCSTI) is the national point of reference for open access and open data. SCSTI maintains a number of information systems related to open access: a central register of (scholarly) publication activities (<http://www.creps.sk>), a central register of dissertations and theses (<http://www.crzp.sk>), a central bibliographical database (<http://scidap.cvtisr.sk>) and the Slovak information system on (current) research (<http://www.skcris.sk>).

The practical implementation of the open access philosophy in Slovakia is being developed, and is reflected mainly in the growth in the number of open access scholarly journals (currently almost 40). This development is being held back by the lack of facilitating legislation, the attitudes of researchers themselves and the lack of an adequate e-infrastructure. In March 2015, SCSTI organised a major conference on open access.<sup>190</sup>

More generally, Slovakia has joined the Open Government Partnership Initiative. The Action Plan for this initiative includes targets and actions relating to open access to research results.<sup>191</sup> Time will tell how much of this is achieved.

## Spain

Spain's policy on open access focuses mainly on the 'green' route, although both 'green' and 'gold' are permitted. The costs of open access publishing are eligible for grants within research projects carried out by public research organisations.

Spain has legislation at national level stipulating that a final digital version of a (publicly funded) scholarly article must be deposited in a repository (either institutional or thematic) within 12 months following its publication.<sup>192</sup>

Government-funded research projects have upheld these legal requirements since 2013 and are also fully geared to the requirements of the EU Horizon 2020 programme. In

---

<sup>189</sup> [http://www.forskningsradet.no/en/Open\\_access/1254008537671](http://www.forskningsradet.no/en/Open_access/1254008537671)

<sup>190</sup> The conference was entitled: *Open access policy of Slovakia in the European context* 2015, see: <http://www.oe.cvtisr.sk>

<sup>191</sup> The Action Plan initiatives for open governance in Slovak Republic 2015, see: <http://www.tretisektor.gov.sk/akny-plan-iniciativy-pre-otvorene-vladnutie-2014-%E2%80%93-2015/>

<sup>192</sup> This is stipulated in Article 37 of the Law (14/2011) of 1 June on science, technology and innovation (*Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación*).

addition to the national legislation, a number of regional authorities (Madrid, Asturias and Catalonia) have also formulated an open access policy for the research they fund. A growing number of universities and research centres are also involved in promoting open access because they manage their own institutional repositories. The government estimates that the biggest user of open access is the scientific community itself. Open access is still not widely known in the business community, and a number of cultural barriers prevent businesses from benefiting from open access.

The law stipulates that the government will facilitate centralised access to repositories and their linkage to similar national and international initiatives. In Spain, the FECYT (*Fundación Española para la Ciencia y la Tecnología*) manages the national harvester RECOLECTA (<http://recolecta.fecyt.es>). This is a platform which brings together all scientific repositories in Spain and provides services to repository managers, researchers and policymakers, just as OpenAIRE does for the European community.

The objectives of RECOLECTA are:

- a. To promote and coordinate the national infrastructure of open access digital scientific repositories in an interoperable way based on the standards adopted by the international community.
- b. To promote, support and facilitate the acceptance and implementation of open access policy by researchers at universities and research institutes.
- c. To make the results of research carried out in Spain more visible.

RECOLECTA comprises 69 institutional repositories and administers around 1,000,000 open access scientific documents produced by the Spanish research community. 85% of Spanish universities and research centres have set up an institutional repository.

In principle, Spain strongly supports the sharing of research data in the case of government-funded research and provided the sharing of that data is compatible with its (potential) commercial use or value. However, the Spanish government does have concerns about open access to research data because of (1) the need for major infrastructures for storing data (and the associated costs); (2) the need for professional data management so that the research data produced can be profitably reused; and (3) the need to find a model that does not stand in the way of public-private research.

Finally, FECYT is working to develop the right metrics for open access; it is also seeking to contribute to the establishment of the common standards, coding and metadata that are needed for the production of reliable statistics.

### **Czech Republic**

The Czech government is in favour of open access publishing and leaves it to researchers themselves to choose the method they use. Open access is largely

embedded in a broader policy framework which also includes investing in large-scale (e-)research infrastructure. Open research data has to date received little attention, though the idea of setting up a central repository has recently been mooted, primarily for institutes which do not yet have their own repository or which lack capacity.

### United States

In 2013 the White House Office of Science and Technology Policy (OSTP) announced that all (20) federal research agencies with an annual budget of more than \$100 million must make their research results publicly accessible within 12 months. This applies both for peer-reviewed articles and the underlying data.<sup>193</sup> It is a decision that builds on a development that had already been started by the National Institute of Health (NIH). The NIH invests around \$30 billion a year, half the total federal research budget, in fundamental and applied biomedical research. Following a voluntary pilot lasting three years (2004-2007), in 2008 the NIH made open access mandatory and in 2012 announced that grants would not be extended to those that were not following the open access decision. A federal bill is currently going through Congress to shorten the embargo period to six months (Fair Access to Science and Technology Research Act). A final decision on this has not yet been taken.<sup>194</sup>

The federal government has opted for the 'green' open access route and has not voiced an opinion on the 'gold' route. It must be possible to read, download and analyse articles in digital form, and to find them in the repository of the (federal) research agency or a repository approved by that agency.

A total of 128 open access mandates ('rules') have now been registered by the various knowledge institutes and funders in ROARMAP.<sup>195</sup> The list includes many of the major (research) universities (Harvard, Columbia, Princeton), as well as several government departments which have adopted an open access policy.

At state level, three states – New York, California and Illinois – have initiated legislation to adopt open access to publicly funded research.<sup>196</sup>

### United Kingdom

British higher education establishments spend just under £5 billion per year on research. The money is allocated via a dual system. The Funding Councils for England, Scotland, Wales and Northern Ireland use lump sum funding to support the research infrastructure

---

<sup>193</sup> See: [www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research](http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research), which contains a link to the OSTP *policy memorandum*.

<sup>194</sup> The bill was resubmitted to *Congress* in early 2015. Progress can be monitored at: <https://www.govtrack.us/congress/bills/114/s779>

<sup>195</sup> <http://roarmap.eprints.org/view/country/840.html>

<sup>196</sup> <http://pitt.libguides.com/openaccess/oausa>

and research priorities of research institutes, while the Research Councils award research grants for specific research projects.

In terms of the number of open access publications, the UK sits just above the average. In 2012, 5.9% of all publications were in (gold) open access journals, compared with a global average of 5.5%. The figures are further apart for hybrid journals: 2.7% in the UK compared with 0.5% globally. The same applies for depositing manuscripts in institutional repositories: 11.6% in the UK versus 5% globally.<sup>197</sup> The latter discrepancy can be explained largely by the high UK investments in infrastructure (such as repositories) and the fact that many research funders introduced policies some years ago which made it mandatory to place manuscripts of published articles in a repository from 2006 onwards. These figures also show that most articles still see the light of day via the traditional subscription system – though publishers do appear to be taking on the open access mantle. According to the Publishers Association, 75% of all journals have the option to be open access, and 96% of journals apply an embargo of 24 months or less.<sup>198</sup>

The debate around open access received a strong boost in 2011 in a roundtable organised with all stakeholders by David Willetts (Minister for Business, Innovation & Skills (BIS)), and the resultant request to the Finch Committee to publish a report with recommendations. The reason for organising this meeting was the enormous increase in subscription costs coupled with limited access to scholarly articles.

The Finch Report, which was published in 2012, recommended that “*a clear policy direction should be set towards support for publication in open access or hybrid journals, funded by APCs, as the main vehicle for the publication of research, especially when it is publicly funded.*”<sup>199</sup> The recommendations were broadly supported by Willetts and the policy that followed was aimed principally at implementing the ‘gold’ route to open access, with public money being made available to cover the costs of APCs. A system of block grants was used for this, which were awarded by the Research Councils. The Research Councils also announced that all research results funded by them would henceforth have to be published in open access or hybrid journals (gold route). In addition, all research papers would be required to include information on how and where the underlying data could be found.

In addition to the UK Research Councils (RCUK), the – private – Wellcome Trust also has an open access policy which makes publication in open access journals mandatory. Manuscripts must also be archived in the European PubMed Central. As regards open

---

<sup>197</sup> Pasteur4OA en JISC (2014).

<sup>198</sup> Research Councils UK (2015).

<sup>199</sup> Working Group on Expanding Access to Published Research Findings (2012).

data, maximum access is the norm. Both RCUK and the Wellcome Trust expect applications for research funding to include a data management plan.

The UK Funding Councils, by contrast, have opted for a model that is closer to the 'green' route. From 1 April 2016, though earlier is recommended, all research institutes must place their research results in an institutional or thematic repository. This must be done as soon as publication is assured, though an embargo is permitted. At the same time, the Funding Councils support the publication model advocated by the British government and the Research Councils, in which publishers offer open access and in which there is scope for more open access journals.

A number of key bodies (the House of Commons BIS Committee,, the Finch Working Group, the RCUK evaluation working group and several researchers) have responded over the last few years to the introduction of the gold open access route. The main observations and points of criticism are described below.

### **Costs**

The costs of introducing the gold open access route by the Research Councils (in 2013/14) was estimated at over £20 million for the first year.<sup>200</sup> In addition to £11 million in APC costs, universities would have to spend £9.2 million in preparation and administrative costs (see Figure 5). A large part of these costs could not be financed by the block grants from the Research Councils. This led among other things to a situation where research institutes held a surplus of RCUK grants, whereas the money had already been spent elsewhere. In addition to the £20 million needed to introduce open access, the costs of subscription fees were estimated at £175 million.<sup>201</sup>

### **APCs**

One point of criticism was the lack of clarity concerning the prices to be paid for APCs. There was for example a wide difference between hybrid journals and fully open access journals: an average APC of £1,150 for open access journals compared with £1,850 for articles in hybrid journals.<sup>202</sup> More openness was called for in the calculations, and also the introduction of more market forces.

The vast majority (80%) of the block grants awarded by the Research Councils went to a small number of institutes. The main beneficiaries among publishers were Elsevier (20.1%), Wiley (15.2%) and PLOS (11.2%). These same publishers reported that the amount of block grants reserved would not be enough to finance the transition to gold

---

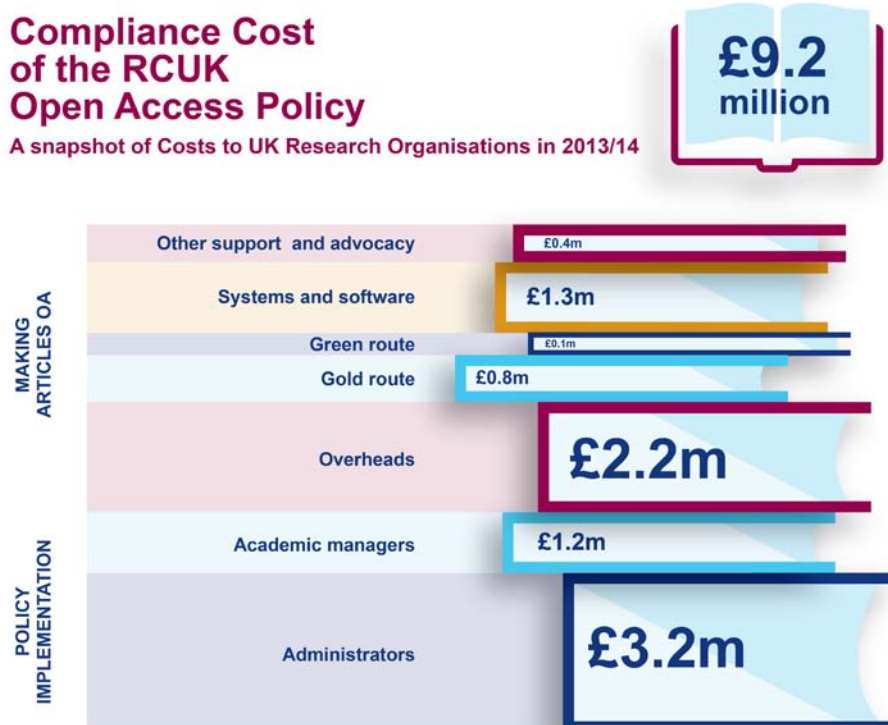
<sup>200</sup> Research Consulting (2014).

<sup>201</sup> Research Consulting (2014).

<sup>202</sup> Pinfield *et al.* (2015).

open access. Research institutes also feared that they would quickly reach a point where there was not enough money to finance all requests.

Figure 5. Administrative costs of the Research Councils UK open access policy



Source: Research Consulting (2014, p. 7)

### Perceptions of institutes and politicians

Although all institutes and disciplines support the idea of open access, there are differences in how it is perceived and accepted. The STEM disciplines (science, technology, engineering and mathematics) are much more open to the gold route and the associated CC-BY licences. For the social sciences and humanities, CC-BY licences are an unknown and therefore unloved quantity, but these disciplines also often have to deal with third-party copyrights (usage rights to image material). They also have problems with having no – or short – embargoes. The download peak and spread is much slower for these disciplines: most articles are only downloaded after two years. There are also fears that they will lose (subscription) income from their own journals, a worry that is shared by the many learned societies. Finally, they fear the spread of open access to monographs.



Most research institutes and researchers in these disciplines express a strong preference for the green route to open access.

Individual researchers are often not convinced of the usefulness of switching to open access and are quick to conclude that they are merely being saddled with extra bureaucracy.<sup>203</sup> We also found no evidence that open access had led to a major change in the publication behaviour of researchers. Authors remained loyal to their journal or publisher.

At institutional level, according to ROARMAP there are 51 self-archiving mandates, 49 of which relate to depositing research results in a repository while the other two are combined with a recommendation to publish in an open access journal. As regards open research data, 23 institutes have introduced an open access policy. Despite the preference of the British government and RCUK for the 'gold' route, few higher education establishments have adapted their open access policy in favour of this model.

The House of Commons BIS Committee voiced criticism in 2013 about the cost of APCs, and particularly the difference in costs between hybrid and 'pure' gold open access journals. The Committee also criticised the long embargoes applied by publishers and the potential consequences that open access could have for international competitiveness. Research institutes and funders had also invested heavily in the preceding years in setting up repositories, and the costs of maintaining them could not now be simply pushed on to institutes as the government was proposing. The BIS Committee advocated "*author freedom of choice between Gold and Green open access.*" The House of Lords also expressed a desire for more clarity on the embargo periods and on international developments in relation to the introduction of open access. It also recommended that the government carry out a cost-benefit analysis on open access publishing (a recommendation that was not taken up).

The UK government continued to espouse its preference for gold open access, but also acknowledged that "*decisions by researchers and the responsiveness of the publishing industry will determine whether Gold OA proves to be the prime route.*"

## Sweden

Sweden has encouraged open access since 2006. Following the recommendations of the European Commission, Sweden began drawing up guidelines for open access in 2012. The aim is to develop a uniform approach across Europe, and by 2025 to have in place a completely open policy for all scientific activities funded from the public purse. As an example, publication is only possible using the Creative Commons licensing system.

---

<sup>203</sup> Research Councils UK (2015).

The first steps were taken in 2015. All publicly funded research is freely accessible, possibly after an embargo of up to six months for articles in the science, technology and medical disciplines. An embargo of (up to) 12 months is permitted for articles in the social sciences and humanities. This embargo period will be reduced to six months in 2020, and open access will also apply for books.

The Swedish government has also produced its first proposals for open research data. These will be worked up in more detail in the period 2015-2020 on the basis of pilot projects. The main obstacles to open research data are a) privacy; b) national security; and c) commercial interests.<sup>204</sup>

### Switzerland

The Swiss science funder SNSF imposes open access publishing as a requirement, but allows either the green or gold route to be followed. SNSF makes funding available for publishing in fully open access journals. Since July 2014, open access publishing has also been mandatory for monographs and editions, though an embargo of up to 24 months may be stipulated.<sup>205</sup>

Open research data has been on the agenda since 2014. A special programme has been set up aimed at securing optimum access to, processing and safeguarding of data. The aim is to develop a joint approach with research institutes by 2020 for providing and processing scientific information.

The Swiss government also supports both the green and gold variants of open access. A number of aspects are important for the government: a) good quality control; and b) good coordination between science and the business community with a view to the use of results.<sup>206</sup> Beyond this, the Swiss government leaves the initiative to research institutes themselves.

---

<sup>204</sup> Vetenskapsrådet/Swedish Research Council (2015)

<sup>205</sup> [http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch\\_id=20143215](http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch_id=20143215)

<sup>206</sup> [http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch\\_id=20073340](http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch_id=20073340)

## Appendix 3 Interviewees

In preparing this report, interviews were held with the following persons:

- ▶ Erik van Aert NWO
- ▶ Jonas Bak Danish Agency for Science, Technology and Innovation
- ▶ Arun Balachadran Institute for Social and Economic Change, India
- ▶ Carlo Beenakker Lorentz Institute, Leiden University
- ▶ Rinze Benedictus UMC Utrecht
- ▶ Axel Berg SURFsara
- ▶ Stephane Berghmans Elsevier (RELX)
- ▶ Florent Bernard European Commission, DG RTD
- ▶ Jan van den Biesen Philips
- ▶ Magchiel Bijsterbosch SURFsara
- ▶ Margreet Bouma Dutch Ministry of Education, Culture & Science//NWO
- ▶ Judith Budde Paris Legal Publishers
- ▶ Tim Buiting Neth-ER
- ▶ Jean-Claude Burgelman European Commission, DG RTD
- ▶ João Costa Universidade Nova de Lisboa, Portugal
- ▶ Jean-François Dechamp European Commission, DG RTD
- ▶ Ron Dekker Dutch Ministry of Education, Culture & Science/NWO
- ▶ Richard Derksen Dutch Ministry of Education, Culture & Science
- ▶ Elly Dijk DANS
- ▶ Ingrid Dillo DANS
- ▶ Peter Doorn DANS
- ▶ John Doove SURFmarket
- ▶ Marc Dupuis SURFmarket
- ▶ Caroline Edwards Open Library of the Humanities, UK
- ▶ Clara Eugenia García Ministerio de Economía y Competitividad, Spain
- ▶ Thomas Grosfeld MKB-Nederland/VNO-NCW
- ▶ Lucie Guibault University of Amsterdam
- ▶ Max Haring Springer
- ▶ Wilco Hazeleger Netherlands eScience Center (NLeSC)
- ▶ Jaap van den Herik Leiden Centre of Data Science, Leiden University
- ▶ Paul IJmker Netherlands Permanent Representation to the EU
- ▶ Marijtje Jongsmá VAWO/Radboud University Nijmegen
- ▶ Margrethe Jonkman FrieslandCampina

- ▶ Thijs Jürgens Shell
- ▶ Hanne-Louise Kirkegaard Danish Agency for Science, Technology and Innovation
- ▶ Michiel Kolman Elsevier (RELX)
- ▶ Fried Kramer Neth-ER
- ▶ Arnoud Lagendijk VAWO/Radboud University Nijmegen
- ▶ Bastiaan van Loenen TU Delft/Open Data Breakthrough Team
- ▶ Dries van Loenen Dutch Ministry of Economic Affairs
- ▶ Marjan van Meerloo Dutch Ministry of Economic Affairs
- ▶ Barend van der Meulen Rathenau Institute
- ▶ Dagmar Meyer European Research Council
- ▶ Frank Miedema UMC Utrecht
- ▶ Peter Molengraaf Alliander
- ▶ Sijbolt Noorda European University Association
- ▶ Abel Packer Scientific Electronic Library Online (SciELO), Brazil
- ▶ Mark Patterson eLife, UK
- ▶ Slobodan Radičev EuroScience/University of Novi Sad, Serbia
- ▶ Susan Reilly LIBER (Association of European Research Libraries)
- ▶ M.M. Rodenburg General practitioner in Delft
- ▶ Dick de Ruiter Stichting Lynch Polyposis
- ▶ Serçe Sahin Dutch Ministry of Education, Culture and Science
- ▶ Laurents Sesink University Library, Leiden University
- ▶ Raymond Sluiter Royal Netherlands Meteorological Institute KNMI
- ▶ Daniel Spichtinger European Commission, DG-RTD
- ▶ Wim van der Stelt Springer
- ▶ Hester Tak Gunn & Twynmore
- ▶ Clifford Tatum CWTS, Leiden University
- ▶ Michiel Thijssen Brill
- ▶ Diana de Veld LUMC Magazine
- ▶ Robert van der Vooren VSNU
- ▶ Astrid van Wesenbeeck National Library of the Netherlands
- ▶ Olga Wessels Neth-ER
- ▶ Wilma van Wezenbeek University Library, TU Delft
- ▶ Ingrid Wijk University Library, Maastricht University
- ▶ David Williams International Association for Energy Economics

## European input

The following persons from European sister organisations of AWTI and organisations from the Neth-ER network provided input to help in obtaining a clear picture of open access and open data in different European countries:

- ▶ Bart Dumolyn Flemish Government, Department of Economics, Science & Innovation (Belgium)
- ▶ Monique Septon FRS-FNRS (Belgium)
- ▶ Margot Beereboom FWO (Belgium)
- ▶ Hanne-Louise Kirkegaard and Jonas Bak Uddannelses- og Forskningsministeriet/ Ministry of Higher Education and Science Danmark (Denmark)
- ▶ Rebecca Taubach Wissenschaftsrat Deutschland (Germany)
- ▶ Claudia Eggert KoWi (Germany)
- ▶ Tuomas Parkkari and Saara Vihko Research and Innovation Policy Council of Finland (Finland)
- ▶ Arnis Kokorevics Latvian Council of Science (Latvia)
- ▶ Rune Rambæk Schjølberg, Siri Lader Bruhn and Yngve Joseph Foss Norges forskningsråd/Research Council of Norway (Norway)
- ▶ Mária Žitňanská Slovak Centre of Scientific and Technical Information (Slovakia)
- ▶ Cristina González Copeiro, Andrés Martínez and Pilar Rico Castro Fundación Española para la Ciencia y la Tecnología (Spain)
- ▶ Katerina Slavikova Technology Centre ASCR (Czech Republic)
- ▶ Lisbeth Söderqvist and Sandra Olivera Vetenskapsrådet/Swedish Research Council (Sweden)
- ▶ Christian Simon Schweizerischer Wissenschafts- und Innovationsrat SWIR (Switzerland)
- ▶ Florence Balthasar Swiss Core (Switzerland)

## Conferences/workshops

The following conferences were attended:

- ▶ European Commission Open Science Visit to Elsevier (24 February 2015, Amsterdam)
- ▶ TUE Symposium Open Access 2015 (10 March 2015, Eindhoven)
- ▶ International Conference on Science 2.0 (3-4 May 2015, Hamburg, Germany)
- ▶ Conference 'Opening up to an ERA of Innovation' (22 June 2015, Brussels, Belgium)
- ▶ Workshop on 'Alternative Open Access publishing models' (12 October 2015, Brussels, Belgium)



## Appendix 4 References

- ▶ Allianz der deutschen Wissenschaftsorganisationen (2015), *Positionen zur Schaffung eines wissenschaftsadäquaten Open-Access-Publikationsmarktes*, February 2015, accessible at: <http://doi.org/10.2312/allianzoa.008>
- ▶ Alperin, J.P. (2015), *The public impact of Latin America's approach to Open Access*, doctoral thesis, Stanford University, 2015, accessible online at: <http://purl.stanford.edu/jr256tk1194>
- ▶ Archambault, E., D. Amyot, P. Deschamps, A. Nicol, F. Provencher, R. Rebout and G. Roberge (2014), *Proportion of open access papers published in peer-reviewed journals at the European and world levels 1996–2013* (D 1.8 version 11p, 22/10/2014), report for the European Commission, s.l.: Science-Metrix, 2014
- ▶ AWT (Advisory Council for Science and Technology Policy, 2011), *E-Science. De wetenschap in de 21<sup>e</sup> eeuw*, The Hague: January 2011
- ▶ AWT (Advisory Council for Science and Technology Policy, 2013), *Waarde creëren uit maatschappelijke uitdagingen*, AWT Report no. 82, The Hague: AWT, October 2013
- ▶ AWTI (Advisory Council for Science, Technology and Innovation, 2015a), *Verwevenheid van onderzoek en hoger onderwijs. Eenheid in verscheidenheid*, The Hague: AWTI, June 2015
- ▶ AWTI (Advisory Council for Science, Technology and Innovation, 2015b), *Klaar voor de toekomst? Naar een brede strategie voor ICT*, The Hague: AWTI, September 2015
- ▶ AWTI (Advisory Council for Science, Technology and Innovation, 2015c), Report of *AWTI meeting on 'Datagedreven Wetenschap' ('Data-driven Science')*, 5 June 2015, The Hague: AWTI, see: <http://www.awti.nl/documenten/verslagen/2015/06/19/verslag-awti-bijeenkomst-datagedreven-wetenschap>
- ▶ Bakker, M. (2014), *Good science, bad science: Questioning research practices in psychological research*, doctoral thesis, UvA, s.l.: 2014, accessible at: <http://dare.uva.nl/record/1/417013>
- ▶ Binnenlands Bestuur, 26 September 2012, 'Open overheidsdata: ruwe grondstof, geen oplospoeder', accessible at: <http://www.binnenlandsbestuur.nl/digitaal/opinie/ingezonden/open-overheidsdata-ruwe-grondstof-geen.8502202.lynkx>
- ▶ Binnenlands Bestuur, 25 June 2014, 'Geen progressie in open data', accessible at: <http://www.binnenlandsbestuur.nl/digitaal/nieuws/geen-progressie-in-open-data.9430257.lynkx>
- ▶ Binnenlands Bestuur, 28 May 2015, 'Plasterk: 98 procent links open data werkt na upgrade', accessible at: <http://www.binnenlandsbestuur.nl/digitaal/nieuws/plasterk-98-procent-links-open-data-werkt-na.9476561.lynkx>

- ▶ Björk, B.-C., P. Welling, M. Laakso, P. Majlender, T. Hedlund & G. Gudnason (2010), 'Open access to the scientific journal literature: Situation 2009', PLOS ONE, 5(6), e11273, doi: 10.1371/journal.pone.0011273
- ▶ Björk, B.-C. & D. Solomon (2014), *Developing an effective market for Open Access Article Processing Charges*, London: Wellcome Trust, 2014
- ▶ Blanchard, A & E. Sbuncu (2012), *Pour une meilleure visibilité de la recherche française*, (Livre Blanc), Paris: 2012
- ▶ Brouwer, F. (2014), 'Waar blijven de innovatieve weerproducten? KNMI en Open Data biedt volop kansen!', presentation accessible at: <http://bit.ly/1QiGYiZ>
- ▶ Capgemini Consulting (2013), *The Open Data Economy. Unlocking Economic Value by Opening Government and Public Data*, s.l.: 2013, at [https://www.capgemini-consulting.com/resource-file-access/resource/pdf/opendata\\_pov\\_6feb.pdf](https://www.capgemini-consulting.com/resource-file-access/resource/pdf/opendata_pov_6feb.pdf)
- ▶ Caruso, J., A. Nicol & E. Archambault (2013), *Open Access Strategies in the European Research Area*, report for the European Commission, Brussels: Science-Metrix, August 2013
- ▶ Concordat Working Group (2015), *Concordat on Open Research Data*, Version 10, 17 July 2015.
- ▶ Corbey, D. (2015), *Brazilië: impressie van een land in verandering*, The Hague: AWTI, June 2015.
- ▶ Couperin (2010), *Open Access in France – a state of the art report*, April 2010
- ▶ CWTS (2015), *Bibliometric study on Dutch Open Access published output 2000-2012/13*, Leiden: January 2015
- ▶ Data Archiving and Networked Services (DANS, 2015), *Samen data delen. Samenvatting strategienota DANS 2015-2020*, The Hague: DANS, 2015
- ▶ Data Citation Synthesis Group (2014), *Joint Declaration of Data Citation Principles*, San Diego: FORCE11, 2014; accessible at: <https://www.force11.org/datacitation>
- ▶ Dillo, I. & P. Doorn (2011), *The Dutch data landscape in 32 interviews and a survey*, The Hague: DANS, 2011
- ▶ Elsevier (2013), *International Comparative Performance of the UK Research Base – 2013. A report prepared by Elsevier for the UK's Department of Business, Innovation and Skills (BIS), 2013*; accessible at: <https://www.gov.uk/government/publications/performance-of-the-uk-research-base-international-comparison-2013>
- ▶ ERAC (European Research Area and Innovation Committee, 2015), *ERA Roadmap 2015-2020*, Brussels: 2015
- ▶ European Commission (2010), *Riding the wave – How Europe can gain from the rising tide of scientific data*, Brussels: 2010, accessible at: <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>
- ▶ European Commission (2012), *Impact assessment – accompanying the document 'Commission recommendation on access to and preservation of the scientific information in the digital age'*, SWD(2012) 222 final (17 July 2012)



- ▶ European Commission (2013a), *Report of the European Commission Public Consultation on Open Research Data*, accessible via the website on the consultation: <http://ec.europa.eu/digital-agenda/node/67533>
- ▶ European Commission (2013b), *Guidelines on Open Access to Scientific Publication and Research Data in Horizon2020*, version 1.0 (11 and 16 December 2013), Brussels: December 2013, accessible at: [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)
- ▶ European Commission (2013c), *Guidelines on Data Management in Horizon2020*, Brussels: 2013
- ▶ European Commission (2014a), *Standardisation in the area of innovation and technological development, notably in the field of text and data mining*, Report from the Expert Group, Brussels: EC-RTD, 2014, accessible at: [http://ec.europa.eu/research/innovation-union/pdf/TDM-report\\_from\\_the\\_expert\\_group-042014.pdf](http://ec.europa.eu/research/innovation-union/pdf/TDM-report_from_the_expert_group-042014.pdf)
- ▶ European Commission (2014b), *European Research Area - progress report 2014*, Brussels: 2014
- ▶ European Commission (2015), *Validation of the results of the public consultation on Science 2.0: Science in Transition*, Brussels: February 2015
- ▶ Ferwerda, E., R. Snijder & J. Adema (2013), *OAPEN-NL - A project exploring Open Access monograph publishing in the Netherlands: Final Report*, The Hague: OAPEN Foundation, October 2013
- ▶ Gargouri, Y., V. Larivière, Y. Gingras, L. Carr & S. Harnad (2012), 'Green and gold open access percentages and growth, by discipline', *ArXiv preprint*. <http://arxiv.org/abs/1206.3664>
- ▶ Goede, M. de, & L. Hessels (2014), *Feiten en cijfers. Drijfveren van onderzoekers*. The Hague: Rathenau Instituut, November 2014.
- ▶ Guibault, L. & A. Wiebe (eds.) (2013), *Safe to be open. Study on the protection of research data and recommendations for access and usage*, Göttingen: Universitätsverlag Göttingen, 2013
- ▶ Houghton, J. (2009), *Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark*, 23 June 2009
- ▶ Houghton, J., J. de Jonge & M. van Oploo (2009), *Costs and Benefits of Research Communication: The Dutch Situation*, 29 May 2009
- ▶ House of Commons Business, Innovation and Skills Committee (2013), *Open Access – fifth report of session 2013-14*, HC 99-I, London: The Stationery Office Limited, 10 September 2013
- ▶ Jonge, J. de (2014), *Verslag debatarees Vertrouwen in de wetenschap*, The Hague: Rathenau Instituut, 2014
- ▶ Kaasschieter, P. (2015), 'Open data: onvolledig, versnipperd en onleesbaar', in: *Financieele Dagblad*, 3 October 2015, p. 14

- ▶ Knowledge Exchange (2010), *The impact of open access outside European universities*, Copenhagen: 2010
- ▶ Knowledge Exchange / Van der Graaf, M. & L. Waaijers (2012), *A Surfboard for Riding the Wave. Towards a four country action programme on research data*, 2<sup>nd</sup> edition, accessible at: [www.knowledge-exchange.info/surfboard](http://www.knowledge-exchange.info/surfboard)
- ▶ Knowledge Exchange / Van den Eynden, V. & L. Bishop (2014), *Sowing the seed: incentives and motivations for sharing research data, a researcher's perspective*, s.l.: Knowledge Exchange, 2014
- ▶ Kotterink, B. & N. Huijboom (2010), 'Zet meer overheidsgegevens in een databank', in: *NRC Handelsblad*, 23 November 2010, accessible at: <http://www.nrc.nl/handelsblad/van/2010/november/23/zet-meer-overheidsgegevens-in-een-databank-11973203>
- ▶ League of European Research Universities (LERU, 2011), *The LERU roadmap towards Open Access*, Advice paper No.8, Leuven: LERU, June 2011
- ▶ League of European Research Universities (LERU, 2015), "*Christmas is over. Research funding should go to research, not to publishers!*" *Moving forwards on open access*, LERU Statement for the 2016 Dutch EU Presidency, Leuven: LERU, 12 October 2015
- ▶ Lelieveldt, H., & S. Princen (2011), *The Politics of the European Union*. Cambridge: Cambridge University Press, 2011
- ▶ Ministerie van Onderwijs, Cultuur en Wetenschap/Ministry of Education, Culture and Science (2014), *Wetenschapsvisie 2025: Keuzes voor de toekomst*, The Hague: November 2014
- ▶ Ministerie van Onderwijs, Cultuur en Wetenschap/Ministry of Education, Culture and Science (2015a), 'Non-paper on open science: open access to publications and data', The Hague: 3 March 2015
- ▶ Ministerie van Onderwijs, Cultuur en Wetenschap/Ministry of Education, Culture and Science (2015b), *Strategische Agenda Hoger Onderwijs en Onderzoek 2015-2025*, The Hague: July 2015
- ▶ Ministry of Education and Culture Finland (2014), *Open science and research leads to surprising discoveries and creative insights. Open science and research roadmap 2014–2017*
- ▶ Ministry of Higher Education Denmark (2014), *Denmark's National Strategy for Open Access*
- ▶ Moedas, C. (2015), 'Open Innovation, Open Science, Open to the World', presentation in Brussels during the conference 'A new start for Europe: Opening up to an ERA of innovation', Brussels: 22 June 2015
- ▶ Mylne, A., O.J. Brady, Z. Huang, D.M. Pigott, N. Golding, M.U.G. Kraemer & S.I. Hay (2014), 'A comprehensive database of the geographic spread of past human Ebola outbreaks', *Sci. Data* 1:140042 doi: 10.1038/sdata.2014.42

- ▶ National Science Board (2014), *Science and Engineering Indicators 2014*, Arlington VA: National Science Foundation
- ▶ Netherlands eScience Center (NLeSC, 2013), *Data-Stewardship in the Big Data Era: Taking care of big data* (2013), s.l.: NLeSC, January 2013, accessible at: [https://www.esciencecenter.nl/img/pressroom/DataStewardship\\_Final.pdf](https://www.esciencecenter.nl/img/pressroom/DataStewardship_Final.pdf)
- ▶ Nicol, A., J. Caruso & E. Archambault (2013), *Open Data Access Policies and Strategies in the European Research Area and Beyond*, report for the European Commission, Brussels: Science-Metrix, August 2013
- ▶ NRC Handelsblad, 15 October 2012, 'De burgerprogrammeur haalt het goud uit de overheidsdata', p. 7
- ▶ NRC Handelsblad, 2 July 2015, 'Eerste stap universiteiten boycot Elsevier', p. 3
- ▶ OECD (Organisation for Economic Co-operation and Development, 2007), *OECD Principles and Guidelines for Access to Research Data from Public Funding*, Paris: OECD Publishing, 2007, accessible at: <http://www.oecd.org/sti/sci-tech/38500813.pdf>
- ▶ OECD (Organisation for Economic Co-operation and Development, 2015a), Working Party of National Experts on Science and Technology Indicators, *Preliminary Findings from an OECD Pilot Survey of Scientific Authors on Access to Outputs of Scientific Research*, DSTI/EAS/STP/NESTI(2015)4, Paris: April 2015
- ▶ OECD (Organisation for Economic Co-operation and Development, 2015b), *Making Open Science a Reality*, OECD Science, Technology and Industry Policy Papers, No. 25, Paris: OECD Publishing Press, 2015
- ▶ Oldenburg, B. (2015), *Integriteit en duurzaamheid in het digitale tijdperk. Het bewaren, delen, hergebruiken en documenteren van digitale onderzoeksdata door sociologen in Nederland*, The Hague: DANS, 2015
- ▶ Open Knowledge Foundation (2012) *Open Data Handbook*, accessible online at: [http://opendatahandbook.org/guide/nl\\_BE/](http://opendatahandbook.org/guide/nl_BE/)
- ▶ Packer, A.L., N. Cop, A. Luccisano, A. Ramalho, & E. Spinak (eds.) (2014), *SciELO – 15 Years of Open Access: an analytic study of Open Access and scholarly communication*, Paris: UNESCO, 2014. Accessible at: <http://dx.doi.org/10.7476/9789230012373>
- ▶ Palzenberger, M., 'Number of Scholarly Articles per Country. Data on Web of Science listed articles and reviews 2004-2013', s.l.: Max Planck Digital Library, 28 April 2015, doi:10.17617/1.2
- ▶ Pasteur4OA & JISC (2014), *UK Open Access case study*, 17 November 2014, accessible at: <http://www.pasteur4oa.eu/sites/pasteur4oa/files/resource/UK%20Case%20Study.pdf>
- ▶ Pinfield, S., J. Salter & P.A. Bath (2015), 'The "Total Cost of Publication" in a Hybrid Open-Access Environment: Institutional Approaches to Funding Journal Article-Processing Charges in Combination With Subscriptions', *Journal of the Association for Information Science and Technology* 2015, doi:10.1002/asi.23446

- ▶ **Plaat, A.** (2015), *Data Science and Ebola*, inaugural lecture, Leiden University, Leiden: 13 April 2015
- ▶ **PRC** (Publishing Research Consortium, 2009), *Access by UK small and medium-sized enterprises to professional and academic information*, (Research report), Bristol: August 2009
- ▶ **Council of the European Union** (2015), 'Draft Council conclusions on the European Research Area Roadmap 2015-2020'
- ▶ **Raad voor het openbaar bestuur/Council for Public Administration** (Rob, 2012), *Gij zult openbaar maken. Naar een volwassen omgang met overheidsinformatie*, The Hague: Rob, September 2012
- ▶ **RDA Europe** (2014), *The Data Harvest: How sharing research data can yield knowledge, jobs and growth*, December 2014, accessible at: [https://europe.rd-alliance.org/sites/default/files/report/TheDataHarvestReport\\_%20Final.pdf](https://europe.rd-alliance.org/sites/default/files/report/TheDataHarvestReport_%20Final.pdf).
- ▶ **RECODE** (2014), *Policy Recommendations for Open access to research Data in Europe*, Brussels: 2014, accessible at: [www.recodeproject.eu](http://www.recodeproject.eu)
- ▶ **Research Consulting** (2014), *Counting the costs of Open Access. The estimated cost to UK research organisations of achieving compliance with open access mandates in 2013/2014*, November 2014
- ▶ **Research Councils UK** (2015), *Review of the implementation of the RCUK Policy on Open Access*, Swindon, UK: March 2015, accessible at: <http://www.rcuk.ac.uk/RCUK-prod/assets/documents/documents/Openaccessreport.pdf>
- ▶ **Schimmer, R., K.K. Geschuhn & A. Vogler** (2015), 'Disrupting the subscription journals' business model for the necessary large-scale transformation to open access', A Max Planck Digital Library Open Access Policy White Paper, Munich: Max Planck Digital Library, 28 April 2015, doi:10.17617/1.3
- ▶ **Schuyt, C.J.M., K.A. Algra, R.J.M. Nolte & F.R. Rosendaal** (2012), *Zorgvuldig en integer omgaan met wetenschappelijke onderzoeksgegevens*, Amsterdam: Royal Netherlands Academy of Arts and Sciences (KINAW), 2012, accessible at: <https://www.knaw.nl/nl/actueel/publicaties/zorgvuldig-en-integer-omgaan-met-wetenschappelijke-onderzoeksgegevens>
- ▶ **Sluiter, R.** (2013), 'Open data bij de KNMI', presentation on 21 May 2013, accessible at: <http://www.pbl.nl/sites/default/files/cms/WS15%20-%20Raymond%20Sluiter.pdf>
- ▶ **Taylor and Francis** (2014), *Taylor and Francis Open Access Survey, June 2014*, s.l.: Taylor & Francis/Routledge, 2014, see: <http://www.tandfonline.com/page/openaccess/opensurvey/2014>
- ▶ **The Publishers Association, International Association of Scientific, Technical and Medical Publishers and The Association of Learned & Professional Society Publishers** (2010), *Scientific Technical and Medical (STM) journal publishing in 2010*, s.l.: April 2010 (most recently accessed on 16/10/2015); accessible at: [http://www.stm-assoc.org/2010\\_04\\_16\\_STM\\_Journal\\_Publishing\\_in\\_2010.pdf](http://www.stm-assoc.org/2010_04_16_STM_Journal_Publishing_in_2010.pdf)

- ▶ TNS Nipo & Koninklijke Bibliotheek (2014), *Behoeftte aan wetenschappelijke artikelen. Verdiepingsonderzoek*, The Hague: Koninklijke Bibliotheek, 22 January 2014
- ▶ Ubaldi, B. (2013), 'Open government data: towards empirical analysis of open government data initiatives', *OECD Working Papers on Public Governance*, No. 22, Paris: OECD Publishing, 2013, <http://dx.doi.org/10.1787/5k46bj4f03s7-en>
- ▶ UNESCO (2011), *Revised Draft Strategy on UNESCO's Contribution to the Promotion of Open Access to Scientific Information and research*, Paris: UNESCO, 20 October 2011
- ▶ UNESCO (2012), *Policy Guidelines for the Development and Promotion of Open Access*, Paris: 2012.
- ▶ UNESCO (2013), *Open Access policy concerning UNESCO publications*, via [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ERI/pdf/oa\\_policy\\_rev2.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ERI/pdf/oa_policy_rev2.pdf)
- ▶ UNESCO Netherlands National Commission for UNESCO (2011), *A global perspective on Open Access*, Amsterdam: 20 January 2011
- ▶ Vetenskapsrådet/Swedish Research Council (2015), *Proposals for national guidelines for open access to scientific information*, Stockholm: Vetenskapsrådet, 2015
- ▶ Visser, D. (2015), 'The Open Access provision in Dutch copyright contract law', *Journal of Intellectual Property Law & Practice* 2015; doi:10.1093/jiplp/jpv161
- ▶ Vlaamse regering (Flemish government) (2014), *Beleidsplan 2014-2019. Werk, Economie, Wetenschap en Innovatie*
- ▶ VNO-NCW (2012), 'Kom van die miljarden af', article in *Opinieblad Forum* 2012, no. 22 (29/11/2012), accessible at: [http://www.vno-ncw.nl/SiteCollectionDocuments/Forumartikelen/Forum\\_2212\\_Overheidsdata\\_17771.pdf](http://www.vno-ncw.nl/SiteCollectionDocuments/Forumartikelen/Forum_2212_Overheidsdata_17771.pdf)
- ▶ VSNU (2015), 'Taalwetenschappers publiceren wetenschappelijke artikelen voortaan in betaalbaar open access', Press release, The Hague: VSNU, 12 October 2015, accessible at: [http://www.vsnu.nl/nl\\_NL/nieuwsbericht/nieuwsbericht/231-taalwetenschappers-publiceren-wetenschappelijke-artikelen-voortaan-in-betaalbaar-open-access.html](http://www.vsnu.nl/nl_NL/nieuwsbericht/nieuwsbericht/231-taalwetenschappers-publiceren-wetenschappelijke-artikelen-voortaan-in-betaalbaar-open-access.html)
- ▶ Ware, M. & M. Mabe (2015), *The STM Report. An overview of scientific and scholarly publishing – Celebrating the 350th anniversary of journal publishing*, fourth edition, The Hague: International Association of Scientific, Technical and Medical Publishers, March 2015
- ▶ Working Group on Expanding Access to Published Research Findings ('Finch Report', 2012) *Accessibility, sustainability, excellence: how to expand access to research publications*, June 2012
- ▶ WTI<sup>2</sup> (2014), *Wetenschappen, Technologie & Innovatie Indicatoren*, publication commissioned by the Dutch Ministry of Education, Culture and Science, Utrecht: December 2014
- ▶ Yozwiak, N.L., S.F. Schaffner & P.C. Sabeti (2015), 'Data sharing: Make outbreak research open access', in *Nature*, vol. 518, pp. 477-479

### **Netherlands Parliamentary papers**

- ▶ Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament, dated 15 November 2013: TK 2013-2014, 31 288, no. 354
- ▶ Letter from the State Secretary for Education, Culture and Science to the Dutch Parliament, dated 23 January 2015: TK 2014-2015, 31 288, no. 414
- ▶ Amendment tabled by MP Joost Taverne (section 25fa of the Dutch Copyright Act): TK 2014–2015, 33 308, no. 11
- ▶ Government information memorandum on 25fa of the Dutch Copyright Act: EK 2014-2015, 33 308, no. E

### **European ‘regulations’**

- ▶ Recommendation of the Commission dated 17 July 2012 on access to and preservation of scientific information (2012/417/EU), *PbEU* 2012 L 194/39
- ▶ Communication from the Commission dated 17 July 2012, ‘A Reinforced European Research Area Partnership for Excellence and Growth’, COM(2012) 392 *final*
- ▶ Communication from the Commission dated 17 July 2012: ‘Towards better access to scientific information: Boosting the benefits of public investments in research’, COM(2012) 401 *final*
- ▶ Communication from the Commission dated six may 2015, ‘A Digital Single Market Strategy for Europe’, COM(2015) 192 *final*