



## UvA-DARE (Digital Academic Repository)

### Introduction

Guibault, L.; Margoni, T.

#### Publication date

2013

#### Document Version

Final published version

#### Published in

Safe to be open: study on the protection of research data and recommendations for access and usage

[Link to publication](#)

#### Citation for published version (APA):

Guibault, L., & Margoni, T. (2013). Introduction. In L. Guibault, & A. Wiebe (Eds.), *Safe to be open: study on the protection of research data and recommendations for access and usage* (pp. 13-16). Universitätsverlag Göttingen.  
<http://webdoc.sub.gwdg.de/univerlag/2013/legalstudy.pdf>

#### General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## Introduction

Openness has become a common concept in a growing number of scientific and academic fields. Expressions such as Open Access (OA) or Open Content (OC) are often employed for publications of papers and research results, or are contained as conditions in tenders issued by a number of funding agencies. More recently the concept of Open Data (OD) is of growing interest in some fields, particularly those that produce large amounts of data – which are not usually protected by standard legal tools such as copyright. However, a thorough understanding of the meaning of Openness – especially its legal implications – is usually lacking.

Open Access, Public Access, Open Content, Open Data, Public Domain. All these terms are often employed to indicate that a given paper, repository or database does not fall under the traditional “closed” scheme of default copyright rules. However, the differences between all these terms are often largely ignored or misrepresented, especially when the scientist in question is not familiar with the law generally and copyright in particular – a very common situation in all scientific fields.

Public Access, for instance, is the term used by the National Institute of Health (NIH), the main US governmental funding agency for biomedical research, which is responsible for the funding of a large amount of academic research<sup>2</sup>. Since 2008 all publications that arise from NIH funds have to comply with the NIH Public Access Policy. The policy requires the final peer-reviewed paper to be deposited in PubMed Central, NIH’s digital full-text archive, upon acceptance for publication, with an indication of when, within a period of 12 months (the so-called embargo period), the paper will become accessible to the general public<sup>3</sup>. More recently, thanks to a US government directive issued by the Office of Science and Technology Policy [Public Access Directive], all federal agencies with more than \$100m in research and development expenditure are required to develop plans to make the published results of federally funded research freely available to the public within one year of publication<sup>4</sup>. Additionally, the Fair Access to Science and Technology Research Act (FASTR) was introduced in the US Parliament

---

<sup>2</sup> See <http://nih.gov> (last accessed 06/2013).

<sup>3</sup> “The Director of the National Institutes of Health shall require that all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine’s PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication: *Provided*, That the NIH shall implement the public access policy in a manner consistent with copyright law”, see Division G, Title II, Section 218 of PL 110-161 (Consolidated Appropriations Act, 2008), as confirmed by Division F, Section 217 of PL 111-8 (Omnibus Appropriations Act, 2009); for references see <http://publicaccess.nih.gov/policy.htm> (last accessed 06/2013).

<sup>4</sup> See <http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research> (last accessed 06/2013) with direct links to the Directive.

at the beginning of 2013. If passed, such a bill would back up the goals of the Directive with the more robust structure of a legislative tool. The bill is similar to the Directive with small but significant differences in terms of the number and types of agencies covered, the embargo period, and the reference to publications (both) or also other research data (Directive)<sup>5</sup>.

This is indeed a great achievement that brings a huge contribution to the dissemination of knowledge produced with public funds (i.e. basically taxpayers' money). Nonetheless, this is Public Access, not Open Access as it covers only some of the requirements of the latter<sup>6</sup>.

Indeed, the NIH Public Access Policy does not provide any explicit right or implied licence to users. This means that PubMed Central users can merely download any paper they are interested in and read it<sup>7</sup>. And that is it. In fact, following such guidelines it is not possible to reproduce the paper (make copies), nor to redistribute the paper (post it on one's own website) nor to modify the paper, outside what is allowed by fair use or other exceptions or limitations to copyright law. All these rights remain within the author's domain (more often within the publisher's). The Directive specifically calls for agencies to implement measures to prevent the unauthorised mass redistribution of scholarly publications<sup>8</sup>. In consequence, users only enjoy Public Access, but not Open Access<sup>9</sup>.

Sometimes, an exclusive right to undertake activities not covered by applicable legislation, such as data mining or bulk downloading, is also created and enforced contractually. The same NIH PubMed Central Public Access Policy prohibits the use of crawlers or systematically downloading articles that are individually available for public access on their repositories, due to alleged copyright restrictions<sup>10</sup>.

---

<sup>5</sup> The text of the bill is available at <http://doyle.house.gov/sites/dxoyale.house.gov/files/documents/2013%2002%2014%20DOYLE%20FASTR%20FINAL.pdf> (last accessed 06/2013).

<sup>6</sup> The term Open Access is discussed in more detail in Chapter 4.1.

<sup>7</sup> Interestingly the PubMed Central copyright notice prohibits bulk downloading of papers for copyright reasons: "Bulk downloading of articles from the main PMC web site, in any way, is prohibited because of copyright restrictions", available at <http://www.ncbi.nlm.nih.gov/pmc/about/copyright> (last accessed 06/2013).

<sup>8</sup> See Public Access Directive, sec. 3.

<sup>9</sup> Indeed, PubMed Central offers a specific OpenAccess subset: <http://www.ncbi.nlm.nih.gov/pmc/tools/openflist> (last accessed: 06/2013).

<sup>10</sup> "Crawlers and other automated processes may NOT be used to systematically retrieve batches of articles from the PMC web site. Bulk downloading of articles from the main PMC web site, in any way, is prohibited because of copyright restrictions. PMC has two auxiliary services that may be used for automated retrieval and downloading of a special subset of articles from the PMC archive. These two services, the PMC OAI service and the PMC FTP service, are the only services that may be used for automated downloading of articles in PMC. See the PMC Open Access Subset for information about which articles are included in this special subset, and for links to the PMC OAI and FTP services. Do not use any other automated processes for bulk downloading, even if you are only retrieving articles from the PMC Open Access Subset. Articles that are available through the PMC OAI and FTP services are still protected by copyright but are distributed under a Creative Commons or similar licence that generally allows more liberal use

The reason why a body committed to offering broader access to its funded research (although not Open Access) restricts activities nowadays so central to research (such as the mining of the data of a set of articles) beyond any legally sanctioned limits remains unclear, especially for those jurisdictions that do not know the existence of a right that protects non-original databases. Possible answers can take different angles, from lack of leadership and guidance at the policy level, to ignorance of practices in a given field, from the idea that “better to restrict access to it, one day it might be worth money”, to TTOs<sup>11</sup> that uncritically opt for a standard reservation formula employed in the past for reasons yet to be demonstrated.

On 17 July 2012 the European Commission – showing leadership and policy guidance – published its Communication to the European Parliament and the Council entitled “Towards better access to scientific information: Boosting the benefits of public investments in research”<sup>12</sup>. As the Commission observes, “discussions of the scientific dissemination system have traditionally focused on access to scientific publications – journals and monographs. However, it is becoming increasingly important to improve access to research data (experimental results, observations and computer-generated information), which forms the basis for the quantitative analysis underpinning many scientific publications”<sup>13</sup>. The Commission believes that through more complete and wider access to scientific publications and data, the pace of innovation will accelerate and researchers will collaborate so that duplication of efforts will be avoided. Moreover, open research data will allow other researchers to build on previous research results, as it will allow involvement of citizens and society in the scientific process.

In the Communication the Commission makes explicit reference to open access models of publications and dissemination of research results (either Golden or Green Road, see below Chapter 4.1), and the reference is not only to access and use but most significantly to reuse of publications as well as research data.

The Communication marks an official new step on the road to open access to publicly funded research results in science and the humanities in Europe. Scientific publications are no longer the only elements of its open access policy: research data upon which publications are based must now also be made available to the public.

---

than a traditional copyrighted work. Please refer to the licence statement in each article for specific terms of use. The licence terms are not identical for all the articles”, <http://www.ncbi.nlm.nih.gov/pmc/about/copyright> (last accessed 06/2013).

<sup>11</sup> TTO stands for Technology Transfer Office, a central asset nowadays for any public and private research enterprise, with the goal of managing and enhancing the value of investments and results in R&D.

<sup>12</sup> Brussels, 17.7.2012 COM (2012) 401 final.

<sup>13</sup> *Ibid.*, p. 3.s.

As noble as the open access goal is, however, the expansion of the open access policy to publicly funded research data raises a number of legal and policy issues that are often distinct from those concerning the publication of scientific articles and monographs. Since open access to research data – rather than publications – is a relatively new policy objective, less attention has been paid to the specific features of research data. An analysis of the legal status of such data, and on how to make it available under the correct licence terms, is therefore the subject of the following sections.