



EUROPEAN COMMISSION  
Directorate-General for Research & Innovation

# Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020

Version 2.0  
30 October 2015



These Guidelines provide context and explanation for the rules on open access applicable to beneficiaries in projects funded or co-funded under Horizon 2020.

## **1. What is open access (OA)?**

**Open access** can be defined as the practice of providing on-line access to scientific information that is free of charge to the end-user and that is re-usable. 'Scientific' refers to all academic disciplines; in the context of research and innovation, 'scientific information' can refer to (i) peer-reviewed scientific research articles (published in scholarly journals) or (ii) research data (data underlying publications, curated data and/or raw data).

**(i) Open access to scientific publications** refers to free of charge online access for any user. Legally binding definitions of 'open access' and 'access' in this context do not exist, but authoritative definitions of open access can be found in key political declarations on this subject, for instance the Budapest Declaration of 2002 (<http://www.budapestopenaccessinitiative.org/read>) or the Berlin Declaration of 3003 ([http://openaccess.mpg.de/67605/berlin\\_declaration\\_engl.pdf](http://openaccess.mpg.de/67605/berlin_declaration_engl.pdf))

These definitions describe 'access' in the context of open access as including not only basic elements such as the right to read, download and print, but also the right to copy, distribute, search, link, crawl, and mine.

There are two main routes towards open access to publications:

- A. **Self-archiving / 'green' open access** means that the published article or the final peer-reviewed manuscript is archived (deposited) by the author - or a representative - in an online repository before, alongside or after its publication. Some publishers request that open access be granted only after an embargo period has elapsed.
- B. **Open access publishing / 'gold' open access** means that an article is immediately provided in open access mode as published. In this model, the payment of publication costs is shifted away from readers paying via subscriptions. The business model most often encountered is based on one-off payments by authors. These costs (often referred to as Article Processing Charges, APCs) can usually be borne by the university or research institute to which the researcher is affiliated, or to the funding agency supporting the research. In other cases, the costs of open access publishing are covered by subsidies or other funding models.

**Misconceptions about open access to scientific publications.** In the context of research funding, open access requirements in no way imply an obligation to publish results. The decision on whether or not to publish lies entirely with the grantees. Open access becomes an issue only *if* publication is elected as a means of dissemination.

Moreover, OA does not interfere with the decision to exploit research results commercially, e.g. through patenting. Indeed, the decision on whether to publish open access must come after the more general decision on whether to publish directly or to first seek protection.

More information on this issue is available in the European IPR Helpdesk fact sheet “Publishing vs. patenting”

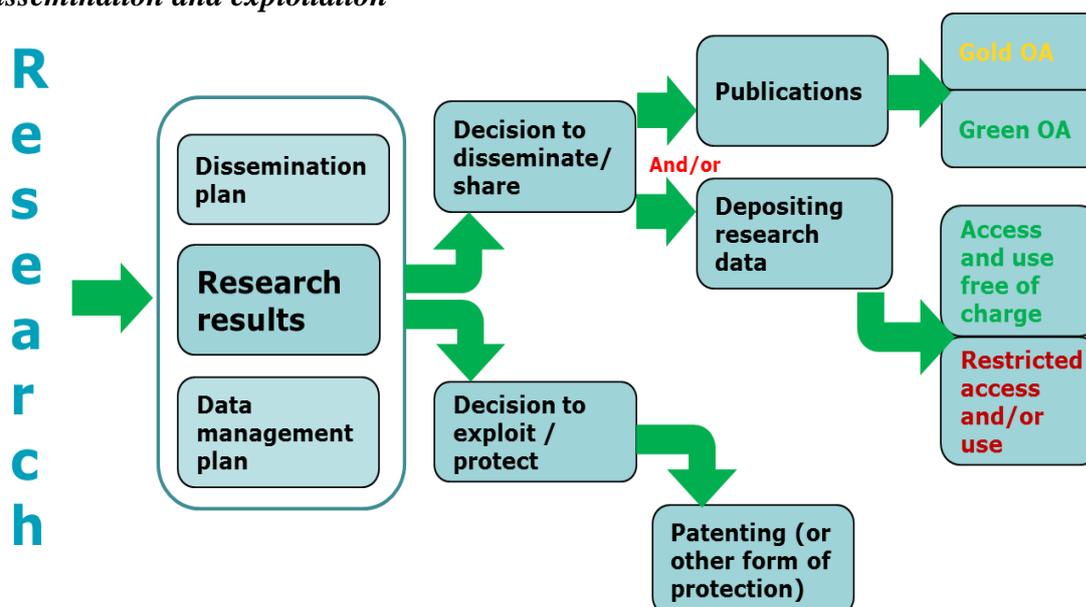
This is illustrated in the graph showing open access to scientific publication and research data in the wider context of dissemination and exploitation at the end of this section.

**(ii) Open access to research data** refers to the right to access and re-use digital research data under the terms and conditions set out in the Grant Agreement.

'Research data' refers to information, in particular facts or numbers, collected to be examined and considered and as a basis for reasoning, discussion, or calculation. In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form.

Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated free of charge for the user.

*Graph: Open access to scientific publication and research data in the wider context of dissemination and exploitation*



## 2. Why open access to publications and data in Horizon 2020?

Modern research builds on extensive scientific dialogue and advances by improving earlier work. Moreover, the Europe 2020 strategy for a smart, sustainable and inclusive economy underlines the central role of knowledge and innovation in generating growth. Fuller and wider access to scientific publications and data therefore help to:

- build on previous research results (improved quality of results);
- foster collaboration and avoid duplication of effort (greater efficiency);

- accelerate innovation (faster to market = faster growth);
- involve citizens and society (improved transparency of the scientific process).

For these reasons, the European Union (EU) strives to improve access to scientific information and to boost the benefits of public investment in the research funded under the EU Framework Programme for Research and Innovation Horizon 2020 (2014-2020).

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.

Open access as a principle in Horizon 2020 now has its legal basis in the Framework Programme itself and in its Rules for Participation. These principles are translated into specific requirements in the Model Grant Agreement under the Horizon 2020 Work Programme. The Annotated Model Grant Agreement provides explanation on the Model Grant Agreement specifically. The present Guidelines build on these documents.

### **3. Mandate on open access to publications**

The detailed legal requirements on open access to publications are contained in article 29.2 of the Model Grant Agreement.

Under Horizon 2020, each beneficiary must ensure open access to all peer-reviewed scientific publications relating to its results.

In order to comply with this requirement, beneficiaries must, at the very least, ensure that their publications, if any, can be read online, downloaded and printed. However, as any additional rights such as the right to copy, distribute, search, link, crawl, and mine increase the utility of the accessible publication, beneficiaries should make every effort to provide for as many of them as possible.

'Peer-reviewed' publications refer to publications that have been evaluated by peers, i.e. other scholars. Peer review is typically, yet not exclusively, organised by the journal or publisher to which an article or manuscript is submitted. New approaches to the organisation of peer review are expected to become more prevalent in the coming years.

The dominant type of peer-reviewed scientific publication is the journal article. In addition, however, beneficiaries are strongly encouraged to provide open access to other types of scientific publications, some of which may, in some cases, not be peer-reviewed, including monographs, books, conference proceedings and grey literature (informally published written material not controlled by scientific publishers, e.g. reports).

The open access mandate is composed of two steps: 1. depositing publications into repositories, and 2. providing open access to them. These two steps may or may not occur at the same time, depending on whether open access publishing ('gold' open access) or self-

archiving ('green' open access) is used, and, in the case of self-archiving, depending on the embargo period (if any).

**Step 1:** beneficiaries must deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications. This must be done as soon as possible and at the latest upon publication. This step must be followed even where open access publishing ('gold' open access) is chosen in order to ensure long-term preservation of the article.

The term 'machine-readable electronic copy' means that the publications must be in a format that can be used and understood by a computer. They must be stored using text file formats which are either standardised or otherwise publicly known so that anyone can develop new tools for working with these documents.

In some cases, the deposit of the final version of an article is possible before publication, for example upon acceptance of the publication by the journal. The latest acceptable time to deposit a publication is the date of publication. Where possible, the published version (in terms of layout, pagination, etc.) should be deposited.

A repository for scientific publications is an online archive. Institutional, subject-based and centralised repositories are all acceptable choices. Beneficiaries should not choose a repository which claims rights over deposited publications and precludes access. The Open Access Infrastructure for Research in Europe (OpenAIRE) is the recommended entry point for researchers to determine what repository to choose (<http://www.openaire.eu>). OpenAIRE also offers support services for researchers, such as the National Open Access Desks. Other useful listings of repositories are the Registry of Open Access Repositories (ROAR, <http://roar.eprints.org/>) and the Directory of Open Access Repositories (OpenDOAR, <http://www.openoar.org/>).

Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications, ideally into a data repository.

This requirement is based on the fact that the concept of 'publication' has rapidly evolved over the past years and in the context of the digital era. Therefore, the notion of 'publication' increasingly includes the data underpinning the publication and results presented, also referred to as 'underlying' data. This data is needed to validate the results presented in the deposited scientific publication and is therefore seen as a crucial part of the publication and an important ingredient enabling scientific best practice. Beneficiaries are also invited to grant open access to this data, but there is no obligation to do so.

**Step 2:** after depositing publications and, where possible, underlying data, beneficiaries must ensure open access to the deposited publication via the chosen repository.

Beneficiaries can choose one of two main ways to comply with this requirement:

1) Self-archiving / 'green' OA: beneficiaries can deposit the final peer-reviewed manuscript in a repository of their choice (see explanation of 'repository' above). In this case, they must

ensure open access to the publication within a maximum of six months (twelve months for publications in the social sciences and humanities).

2) Open access publishing / 'gold' OA: researchers can also publish in open access journals, or in journals that sell subscriptions and also offer the possibility of making individual articles openly accessible (hybrid journals). Monographs can also be published either via a 'pure' open access or via a hybrid business model. The author processing charges (APCs) for gold open access incurred by beneficiaries are eligible for reimbursement during the duration of the project (see article 6.2.D.3 of the Model Grant Agreement). In all cases, open access via the chosen repository must be ensured upon publication.

The costs of 'gold' open access publications incurred after the end of projects are not eligible for reimbursement via the budget of the specific action. However, a mechanism is being piloted for also dealing with open access publication charges incurred after the end of grant agreements with the Commission. This pilot is supported via the OpenAIRE2020 project: further information is available at <https://www.openaire.eu/postgrantoapilot>.

Beneficiaries must also ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication. The bibliographic metadata must be in a standard format and must include the following:

- the terms [*"European Union (EU)" and "Horizon 2020"*][*"Euratom" and Euratom research and training programme 2014-2018"*];
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable, and a persistent identifier.

The purpose of the requirement on metadata is to maximise the discoverability of publications and to ensure the acknowledgment of EU funding. Bibliographic data mining is more efficient than mining of full text versions. The inclusion of information relating to EU funding as part of the bibliographic metadata is necessary for adequate monitoring, production of statistics, and assessment of the impact of Horizon 2020. For adequate identification of the action concerned, the grant number, name and/or acronym of the action is needed (preferably all three). The publication date and embargo period enable the monitoring of the embargo periods. The persistent identifier (for example a Digital Object Identifier, DOI) identifies the publication. It allows linking to an authoritative version of the publication. For example, OpenAIRE (<http://www.openaire.eu>) will provide means to check the metadata compliance of the chosen repository.

In all cases, the Commission encourages authors to retain their copyright and grant adequate licences to publishers. Creative Commons offers useful licensing solutions in this regard (e.g. CC-BY or CC-0 licences, see <http://creativecommons.org/licenses/>). This type of licence is a good legal tool to enable open access in its broadest sense.

Where possible, it is also recommended that contributors be uniquely identifiable, and data uniquely attributable, through identifiers which are persistent, non-proprietary, open and interoperable (e.g. through leveraging existing sustainable initiatives such as ORCID for contributor identifiers and DataCite for data identifiers).

#### **4. Open Research Data Pilot**

A novelty in Horizon 2020 is the Open Research Data Pilot which aims to improve and maximise access to and re-use of research data generated by projects. The legal requirements for projects participating in this pilot are contained in the optional article 29.3 of the Model Grant Agreement. Other relevant information, such as the scope of the Pilot, is provided in the introduction to the Horizon 2020 Work Programme. The Pilot on Open Research Data will be monitored throughout Horizon 2020 with a view to further developing EC policy on open research.

**Scope of the Pilot.** The scope of the pilot is anchored in the Work Programmes. For the 2016-2017 Work Programme, the areas of Horizon 2020 participating in the Open Research Data Pilot are:

- Future and Emerging Technologies
- Research infrastructures
- Leadership in enabling and industrial technologies – Information and Communication Technologies
- Nanotechnologies, Advanced Materials, Advanced Manufacturing and Processing, and Biotechnology: ‘nanosafety’ and ‘modelling’ topics
- Societal Challenge: Food security, sustainable agriculture and forestry, marine and maritime and inland water research and the bioeconomy - selected topics as specified in the work programme
- Societal Challenge: Climate Action, Environment, Resource Efficiency and Raw materials – except raw materials
- Societal Challenge: Europe in a changing world – inclusive, innovative and reflective Societies
- Science with and for Society
- Cross-cutting activities - focus areas – part Smart and Sustainable Cities

**Voluntary participation in the Pilot on Open Research Data.** Individual projects funded under Horizon 2020 and not covered by the scope of the Pilot outlined above may participate in the pilot on a voluntary basis. The project consortia that decide to participate in the Pilot on a voluntary basis will include article 29.3 in their grant agreement and will be monitored along with and receive the same support as all other projects participating in the Pilot.

**All of Horizon 2020 is therefore covered by either through the core areas or the individual opt-in approach. The only exceptions are the instruments "cofund", and "prizes" as well as "ERC proof of concept" grants and those "ERA-Nets" that do not produce data** (in the case of ERA-Nets that do produce data, the voluntary opt-in option applies). Please note that for two stage calls the information on participation in the pilot will only be asked in stage 2 of the process. As concerns coordination and support actions, these are NOT generally excluded since a significant number of them do produce relevant data.

### **Total or partial opting out of the Pilot Action on Open Research Data.**

Projects may at any stage opt out of the Pilot for a variety of reasons, namely:

- if participation in the Pilot on Open Research Data is incompatible with the Horizon 2020 obligation to protect results if they can reasonably be expected to be commercially or industrially exploited;
- if participation in the Pilot on Open Research Data is incompatible with the need for confidentiality in connection with security issues;
- if participation in the Pilot on Open Research Data is incompatible with existing rules concerning the protection of personal data;
- if participation in the Pilot on Open Research Data would jeopardise the achievement of the main aim of the action;
- if the project will not generate / collect any research data;
- if there are other legitimate reason to not take part in the Pilot (at proposal stage – free text box provided).

The approach of the Commission can therefore be summarised as "as open as possible, as closed as needed".

Participation in the pilot is **not** part of the project evaluation. Proposals will not be evaluated more favourably because they are part of the Pilot and will not be penalised for opting out of the Pilot. During the lifetime of a project an opt out remains possible for any of the reasons above and needs to be described in the projects' Data Management Plan (DMP) For those projects participating in the pilot, the review of the data management plans is included as part of the electronic reporting process.

It should also be noted that **participating in the Pilot does not necessarily mean opening up all research data**. The focus of the Pilot is on encouraging good data management as an essential element of research best practice.

Separate guidance on research data management is also available.

The Open Research Data Pilot applies to two **types of data**:

- 1) the data, including associated metadata<sup>1</sup>, needed to validate the results presented in scientific publications as soon as possible;
- 2) other data<sup>2</sup>, including associated metadata, as specified and within the deadlines laid down in the data management plan – that is, according to the individual judgement by each project.

---

<sup>1</sup> Associated metadata refers to the metadata describing the research data deposited.

<sup>2</sup> For instance curated data not directly attributable to a publication, or raw data

**What are the requirements of the Open Research Data Pilot?** The Grant Agreements of projects participating in the Pilot include Article 29.3. Projects participating in the Pilot must comply with the following:

1) **Step 1:** participating projects are required to deposit the research data described above, preferably into a research data repository. 'Research data repositories' are online archives for research data. They can be subject-based/thematic, institutional or centralised. Useful listings of research data repositories include the Registry of Research Data Repositories ([www.re3data.org](http://www.re3data.org)) and Databib (<http://databib.org>). In addition, it is expected that the Open Access Infrastructure for Research in Europe (OpenAIRE) will become an entry point for linking publications to underlying research data.

2) **Step 2:** as far as possible, projects must then take measures to enable for third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) this research data. One straightforward and effective way of doing this is to attach Creative Commons Licence (CC-BY or CC0 tool) to the data deposited (<http://creativecommons.org/licenses/>, <http://creativecommons.org/about/cc0>).

At the same time, projects should provide **information** via the chosen repository about tools and instruments at the disposal of the beneficiaries and necessary for validating the results, for instance specialised software or software code, algorithms, analysis protocols, etc. Where possible, they should provide the tools and instruments themselves.

**Incentives / supporting measures:** Costs relating to the implementation of the pilot will be eligible. Specific technical and professional support services will also be provided (e-Infrastructures WP).

## **5. Further information and help**

Horizon 2020: <http://ec.europa.eu/programmes/horizon2020>

Participant Portal: <http://ec.europa.eu/research/participants/portal/>

Open access (Science in Society site):

<http://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=science>

Open access (Digital Agenda site): <http://ec.europa.eu/digital-agenda/en/open-access-scientific-knowledge-0>

OpenAIRE: <http://www.openaire.eu>

[IP/12/790](#) on open access in Horizon 2020:

[http://europa.eu/rapid/press-release\\_IP-12-790\\_en.htm?locale=en](http://europa.eu/rapid/press-release_IP-12-790_en.htm?locale=en)

[Open Access Policy](#)

Questions on open access: [RTD-open-access@ec.europa.eu](mailto:RTD-open-access@ec.europa.eu)