EUROPEAN COMMISSION
Directorate-General for Research & Innovation

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

Version 2.1

15 February 2016

These guidelines are intended to contextualise and clarify the rules on open access covering beneficiaries in projects funded or co-funded under Horizon 2020.

## 1. WHAT IS OPEN ACCESS?

**Open access** (OA) refers to the practice of providing online access to **scientific information** that is free of charge to the end-user and reusable. 'Scientific' refers to all academic disciplines. In the context of research and innovation, 'scientific information' can mean:

1. peer-reviewed scientific research articles (published in scholarly journals), or
2. research data (data underlying publications, curated data and/or raw data).

### 1. Peer-reviewed scientific research articles

**Open access to scientific publications** means online access, free of charge, for any user. Although there are no legally binding definitions of 'access' or 'open access' in this context, authoritative definitions of open access appear in key political declarations including:

- the 2002 Budapest Declaration
- the 2003 Berlin Declaration

Under these definitions, '**access**' includes not only basic elements - the right to read, download and print –but also the right to copy, distribute, search, link, crawl and mine.

**2 main routes to open access** are:

A. **Self-archiving / 'green' open access** – the author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online repository before, at the same time as, or after publication. Some publishers request that open access be granted only after an embargo period has elapsed.

B. **Open access publishing / 'gold' open access** - an article is immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers. The most common business model is based on one-off payments by authors. These costs, often referred to as Article Processing Charges APCs) are usually borne by the researcher's university or research institute or the agency funding 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 do not imply an obligation to publish results. Whether to publish is entirely up to the grant beneficiaries. Open access becomes an issue *only if* publication is chosen as a means of dissemination.

Moreover, open access does not affect the decision to exploit research results commercially, e.g. through patenting. The decision on whether to publish through open access must come after the more general decision on whether to publish directly or to first seek protection. This is illustrated in the chart at the end of this section, which shows open access to scientific publication and research data in the wider context of dissemination and exploitation.

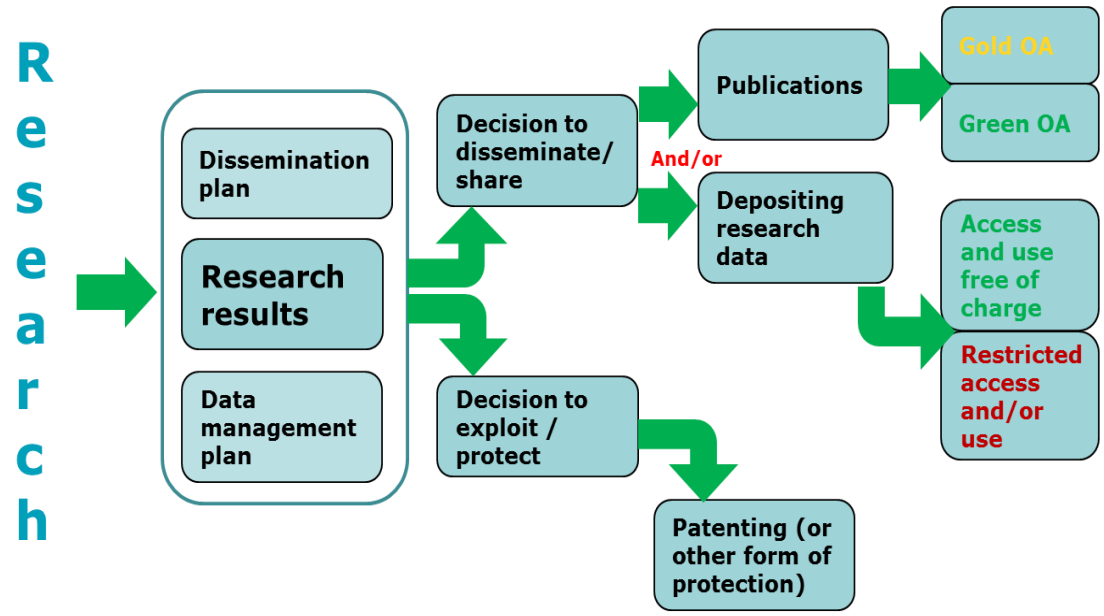
For more information, see the European IPR Helpdesk factsheet "Publishing vs. patenting".

## 2. Research data

**Open access to research data** refers to the right to access and reuse 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.

Users can normally access, mine, exploit, reproduce and disseminate openly accessible research data free of charge.




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

## 2. WHY HAVE OPEN ACCESS TO PUBLICATIONS & DATA IN HORIZON 2020?

Modern research builds on extensive scientific dialogue and advances by improving earlier work. 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 helps to:

- **build on previous research results** (improved quality of results)
- **encourage collaboration** & avoid duplication of effort (greater efficiency)
- **speed up innovation** (faster progress to market means faster growth)
- **involve citizens and society** (improved transparency of the scientific process).

This is why the European Union strives to improve access to scientific information and to boost the benefits of public investment in research funded under the 2014-2020 EU Framework Programme for Research and Innovation, Horizon 2020.

In the European Commission's view, there should be no need to pay for information funded from the public purse each time it is accessed or used. Moreover, it should benefit European businesses and the public to the full. This means making publicly-funded scientific information available online, at no extra cost, to European researchers, innovative industries and the public, while ensuring that it is preserved in the long term.

Under Horizon 2020, the legal basis for open access is laid down in the Framework Programme itself and its Rules for Participation. These principles are translated into specific requirements in the Model Grant Agreement and in the Horizon 2020 Work Programmes. The Annotated Model Grant Agreement provides specific explanations of the Model Grant Agreement. The present Guidelines build on these documents.

## 3. MANDATE ON OPEN ACCESS TO PUBLICATIONS

Article 29.2 of the Model Grant Agreement sets out detailed legal requirements on open access to scientific publications: under Horizon 2020, **each beneficiary must ensure open access to all peer-reviewed scientific publications** relating to its results.

To meet this requirement, beneficiaries must, at the very least, ensure that any scientific peer reviewed publications can be read online, downloaded and printed. Since any further rights - such as the right to copy, distribute, search, link, crawl and mine - make publications more useful, beneficiaries should make every effort to provide as many of these options as possible.

**Peer-reviewed** publications are those assessed by other scholars. Peer review is typically, though not exclusively, organised by the journal or publisher to which an article or manuscript is submitted. However, new approaches are expected to become more prevalent in years to come.

The dominant type of **scientific publication** is the journal article. Grant beneficiaries are also strongly encouraged to provide open access to other types of scientific publications including:

- monographs
- books
- conference proceedings
- grey literature (informally published written material not controlled by scientific publishers, e.g. reports)

**The open access mandate comprises 2 steps:**

1. depositing publications in repositories
2. providing open access to them

These steps are explained in more detail below. They may or may not occur simultaneously, 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 - depositing publications in repositories

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 applies even where open access publishing (**'gold' open access**) is chosen to ensure that the article is preserved in the long term.

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

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

A repository for scientific publications is an **online archive**. Institutional, subject-based and centralised repositories are all acceptable choices; repositories that claim rights over deposited publications and preclude access are not.

The Open Access Infrastructure for Research in Europe (OpenAIRE) is the recommended entry point for researchers to determine what repository to choose. It also offers support services for researchers, such as the National Open Access Desks. Other useful listings of repositories are:

- Registry of Open Access Repositories (ROAR)
- Directory of Open Access Repositories (OpenDOAR).

The beneficiary must also aim to deposit at the same time as the publication the research data needed to validate the results presented in the deposited scientific publications ('underlying data'), ideally in a data repository. The reason for this requirement is that the concept of 'publication' has changed rapidly in recent years, with the advent of the digital era. 'Publication' increasingly includes the data underpinning the publication and the results presented. 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 component of scientific best practice. Beneficiaries are also invited to grant open access to this data, but there is no obligation to do so.

**Step 2 - providing open access to them**

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 2 main ways to meet 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). They must ensure open access to the publication within at most 6 months (12 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 hybrid journals that both sell subscriptions and offer the option of making individual articles openly accessible. Monographs can also be published either on a purely open access basis or using a hybrid business model. Article processing charges (APCs) for gold open access are eligible for reimbursement during the duration of the project (as other costs defined in article 6.2.D.3 of the Model Grant Agreement). As stated, the article must also be made accessible through a repository upon publication.

The costs of 'gold' open access publications incurred once a project is completed cannot be refunded from that project's budget. However, a mechanism is being piloted to address the issue of open access publication charges incurred once a grant agreement with the Commission has expired. This pilot project, funded under the OpenAIRE2020 project, supports Open Access publications arising from FP7 projects. Detailed information and conditions to apply for reimbursements can be consulted in the EC/OpenAire FP7 Post-Grant Open Access Pilot Application Guidelines.

Beneficiaries must also provide open access, through the repository, to the bibliographic metadata that identify the deposited publication. These must be in a standard format and must include the following:

- the terms *["European Union (EU)" & "Horizon 2020"]["Euratom" & Euratom research & training programme 2014-2018"]*;
- the name of the action, acronym & grant number;

- the publication date, the length of the embargo period (if applicable) & a persistent identifier.

The purpose of the metadata requirement is to make it easier to find publications and ensure that EU funding is acknowledged. Mining bibliographic data is more efficient than mining full text versions. Information on EU funding must be included as part of bibliographic metadata so that Horizon 2020 can be properly monitored, statistics produced, and the programme's impact assessed. The project concerned should be properly identified by the grant number and the project's name and/or acronym (preferably all three).

To monitor the embargo periods, the publication date and embargo period must be provided. The persistent identifier (for example a Digital Object Identifier) identifies the publication. It enables a link to be provided to an authoritative version of the publication. For example, OpenAIRE will provide the 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 (e.g. CC BY or CC0 licences). This type of licence is a good legal tool for providing open access in its broadest sense.

Where possible, contributors should also 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 new feature of Horizon 2020 is the Open Research Data Pilot (ORD Pilot), designed to improve and maximise access to and reuse of research data generated by projects. The legal requirements for participating projects are set out in the optional article 29.3 of the Model Grant Agreement. The Pilot on Open Research Data will be monitored throughout Horizon 2020 with a view to further developing Commission policy on open research.

### Scope of the Pilot

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

- Future & Emerging Technologies
- Research infrastructures
- Leadership in enabling & industrial technologies – Information & Communication Technologies
- Nanotechnologies, Advanced Materials, Advanced Manufacturing & Processing, & Biotechnology – 'nanosafety' & 'modelling' topics
- Societal Challenge – Food security, sustainable agriculture & forestry, marine & maritime & inland water research & the bioeconomy - selected topics as specified in the work programme

- Societal Challenge – Climate Action, Environment, Resource Efficiency & Raw Materials – except raw materials
- Societal Challenge – Europe in a changing world – inclusive, innovative & reflective societies
- Science with & for Society
- Cross-cutting activities – focus areas – part Smart & 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 Open Research Data Pilot as outlined above may participate in the Pilot on a voluntary basis. Project consortia (or individual beneficiaries) that decide to take part in the Pilot voluntarily will:

- include article 29.3 in their grant agreement
- be monitored along with all other participating projects and receive the same support.

**All of Horizon 2020 is therefore potentially covered by the ORD Pilot, either through the core areas or by the individual opt-in approach. The only exceptions to this are:**

- **the "cofund" and "prizes" instruments**
- **"ERC proof of concept" grants**
- **"ERA-Nets" that do not produce data** (the opt-in option does apply to ERA-Nets that do produce data).

Coordination and support actions are included in the Open Research Data pilot, as many produce relevant data.

For 2-stage calls, information on participation in the Pilot will be requested only at stage 2.

### Opting out partially or entirely from the Pilot on Open Research Data

Projects can opt out at any stage if:

- participation is incompatible with the Horizon 2020 obligation to protect results that can reasonably be expected to be commercially or industrially exploited
- participation is incompatible with the need for confidentiality in connection with security issues
- participation is incompatible with rules on protecting personal data
- participation would mean that the project's main aim might not be achieved
- the project will not generate / collect any research data, or
- there are other legitimate reasons not to take part in the Pilot (at proposal stage – free text box provided).

⚠ Important: Participation in the pilot is **not** part of the project evaluation. In other words, proposals will not be assessed any more favourably for taking part in the Pilot, nor will they be penalized for opting out of it.

During the lifetime of a project, a complete opt-out via project amendment is possible for any of the reasons above.

Alternatively, during the lifetime of a project, a partial (e.g. for selected datasets) or even complete (i.e. for all datasets) opt out remains possible for any of the reasons above via the Data Management Plan (DMP). In this case, the project participates in the Pilot, but does not open some of/any of its data for reasons explained in its DMP.

For projects taking part in the Pilot, the review of Data Management Plans is part of the electronic reporting process.

Please see the Guidelines on Data Management in Horizon 2020 for details on DMPs, including a template.

⚠ In summary, participating in the Open Research Data Pilot does not necessarily mean opening up all research data. Rather, the focus of the Pilot is on encouraging good data management as an essential element of research best practice.

The approach of the Commission regarding the Open Research Data Pilot can be summarised as "as open as possible, as closed as necessary".

**Types of data covered by the Open Research Data Pilot**

1. **the data, including associated metadata** (i.e. metadata describing the research data deposited), needed to validate the results presented in scientific publications as soon as possible ("underlying data")
2. **other data** (for instance curated data not directly attributable to a publication, or raw data), 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.

**What are the requirements of the Open Research Data Pilot?**

The Grant Agreements of projects taking part in the Pilot include Article 29.3. Participating projects must meet the following requirements:

**Step 1** - they must deposit the research data described above, preferably in a research data repository. These are online research data archives, which may be subject-based/thematic, institutional or centralised. Useful listings of repositories include the Registry of Research Data Repositories and Databib. The Open Access Infrastructure for Research in Europe (OpenAIRE) provides additional information and support on linking publications to underlying research data. Some repositories like Zenodo (an OpenAIRE and CERN collaboration), allows researchers to deposit both publications and data, while providing tools to link them. Zenodo and some other repositories as well as many academic publishers also facilitate linking publications and underlying data through persistent identifiers and data citations.

**Step 2** – as far as possible, projects must then take measures to enable 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 Licences (CC BY or CC0 tool) to the data deposited. The EUDAT B2SHARE tool includes a built-in license wizard that facilitates the selection of adequate license for research data.

At the same time, projects should provide **information** via the chosen repository about the tools available to the beneficiaries that are needed to validate the results, e.g. specialised software or software code, algorithms and analysis protocols. Where possible, they should provide these instruments themselves.

**Periodic reporting**: check further details on how to fill in reporting tables for publications, deliverables and the process for continuous reporting in the grant management system of the Participant portal.

### Incentives / supporting measures

Costs relating to the implementation of the pilot will be eligible as part of the grant. Specific technical and professional support services are also available through OpenAIRE2020 and EUDAT2020 projects.

## 5. FURTHER INFORMATION AND HELP

- Horizon 2020 Programmes
- Participant Portal
- Open Access Policy (Science With and For Society)
- Open Access (Digital Agenda)
- Open Access Policy
- OpenAIRE
- Horizon 2020 Annotated Grant Agreement for articles 29.2 and 29.3
- Scientific data: open access to research results will boost Europe's innovation capacity (IP/12/790)
- Questions on open access by email