OPEN SCIENCE
Handbook

FOREWORD

The aim of open science is to share knowledge and to cultivate and reuse data in an effort to validate and protect the integrity of scientific research. The two most important elements of open science are open access to publications and open access to research data. This yields many benefits:

- shared knowledge that makes research more effective and more capable of addressing societal challenges

- increased visibility for publications and a broader reach for citations

- fewer issues with plagiarism; submitting a preprint to an open archive can be used to establish precedence for the research findings

- more accurate evaluations for researchers based on the content of their articles

In 2006, École Centrale de Lyon opened the HAL* collection, thereby taking a strong stand in favour of free access. Since then, the university library has developed a number of initiatives: Halathon, training programs, Open Access Week events, an escape game, exhibits, etc.

As of 2020, all École Centrale de Lyon laboratories are now listing all their publications in HAL. The HAL* collection from the École contains over 21,000 document listings, 38% of which include the full text.

As a result of the Digital Republic Act*, the National Plan for Open Science and requirements from public research funding organisations (European Commission, ANR), École Centrale de Lyon is now expanding these pro-open science efforts.

The university is encouraging all researchers to make their publications open access* and to develop an open-data management plan for all research projects in accordance with the 'as open as possible and as closed as necessary' principle.

École Centrale de Lyon has set four ambitious goals for 2022:

- to have 100% of articles published during the year listed in HAL

- to increase the share of full-text articles available for open access to 70% for articles listed during the year

- to increase the use of data management plans* for all approved/funded research projects

- to transform HAL from an École Centrale de Lyon collection* into a portal*

This handbook is the first step in the École's road map for open science. It provides the requisite information and recommendations to further these efforts at the university.
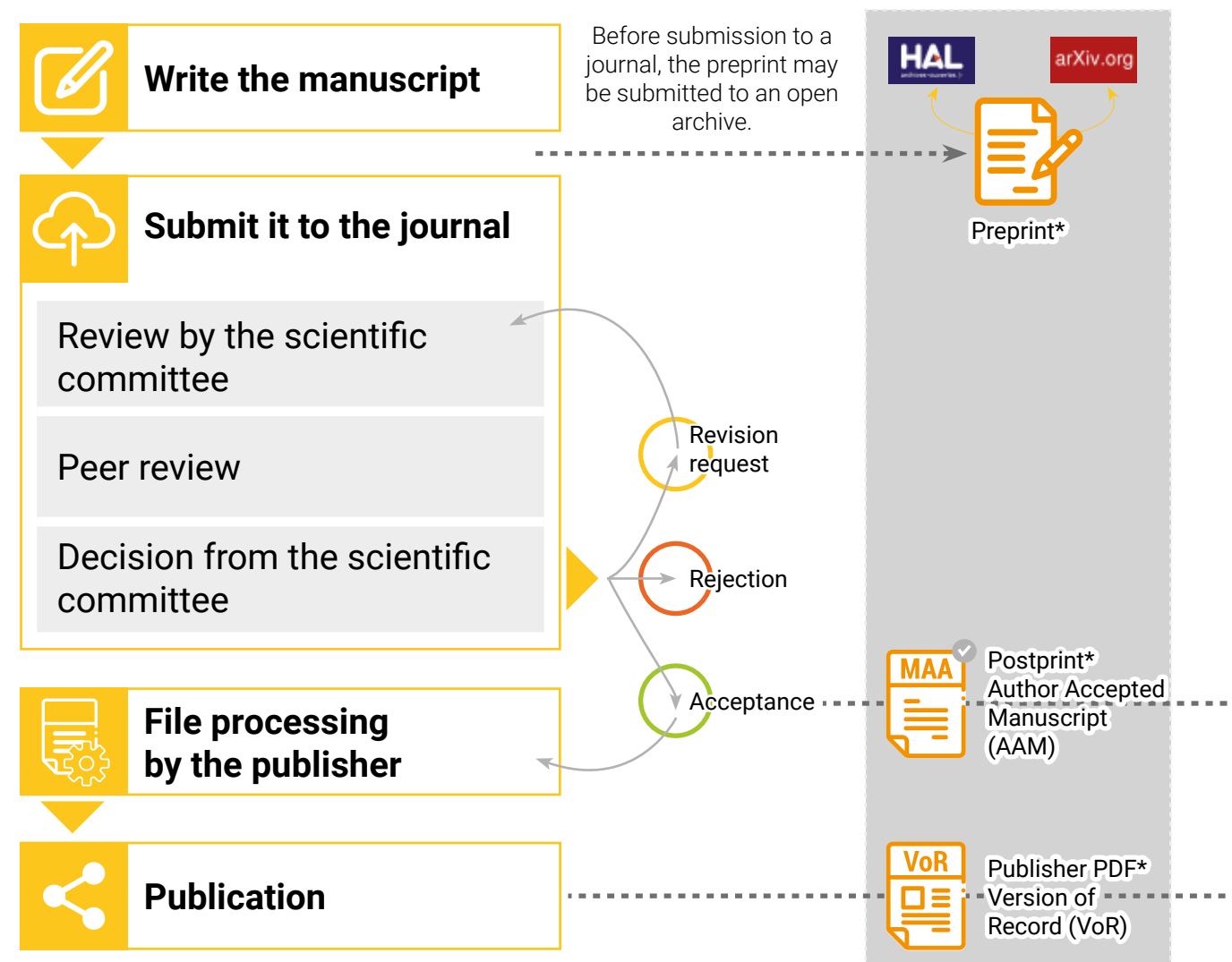
Jean-Pierre Bertoglio
Head of research
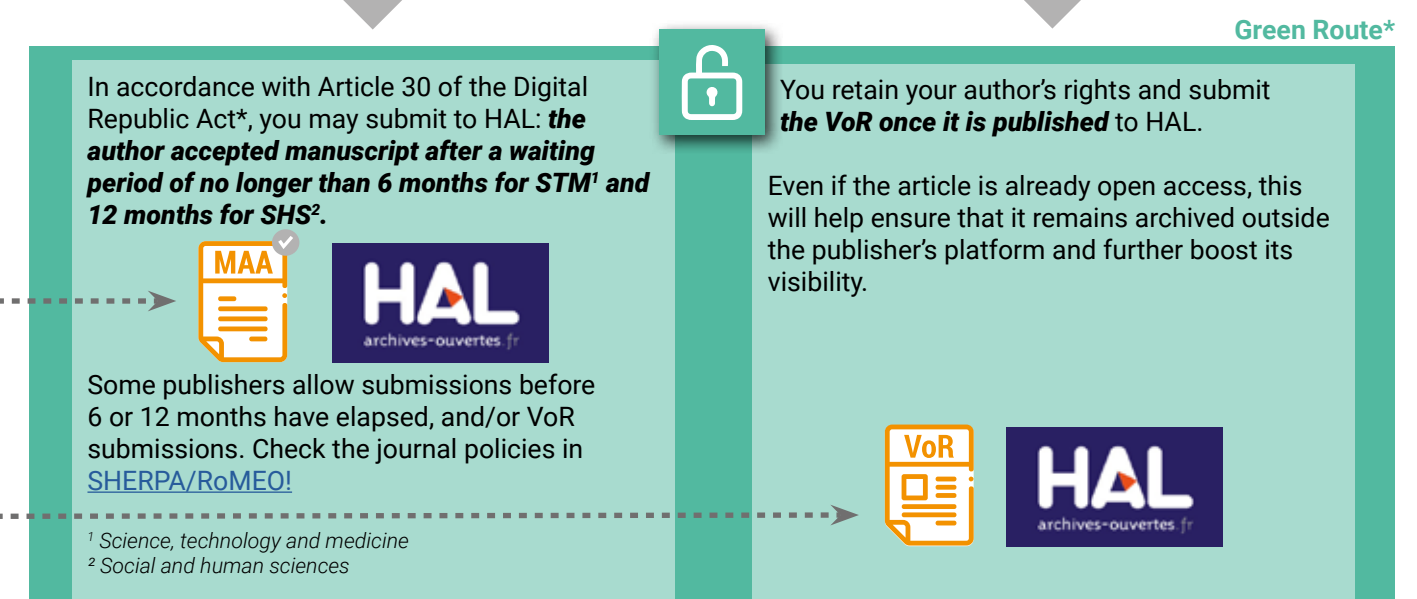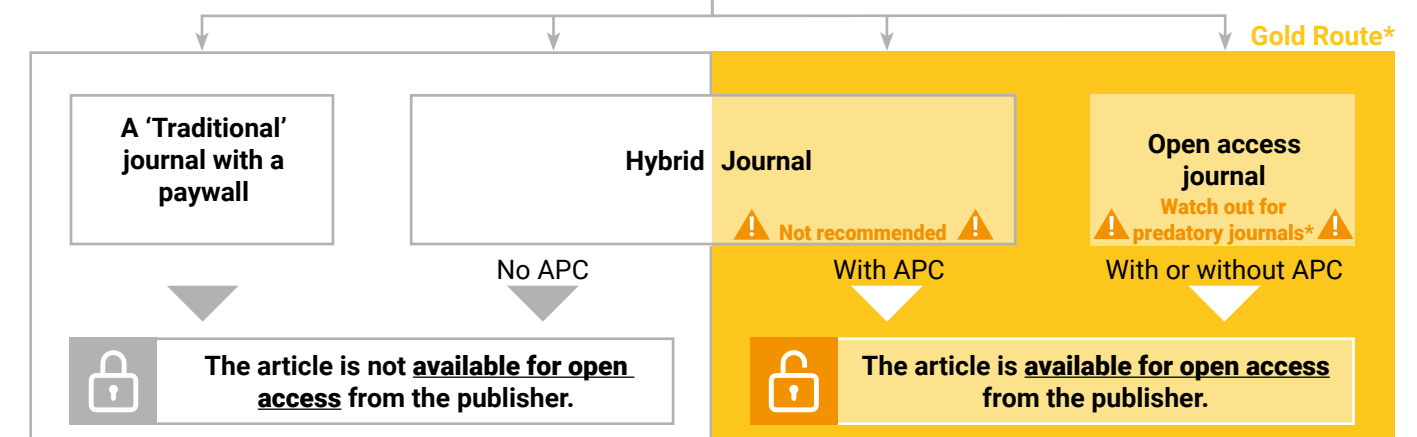École Centrale de Lyon

see glossary

PUBLICATION CYCLE

Making an article open access* can be achieved by publishing in an open access journal and/or by submitting the article to an open archive* like HAL. The 2016 Digital Republic Act* made the second option much easier. What version of the article can be open, though? And at what point in the publishing process can this happen? Under what circumstances? This two-page infographic has all the answers!

Write the manuscript

Before submission to a journal, the preprint may be submitted to an open archive.

HAL archives-ouvertes | arXiv.org

Preprint*

Submit it to the journal

Review by the scientific committee

Peer review

Decision from the scientific committee

Revision request

Rejection

Acceptance

File processing by the publisher

MAA — Postprint*
Author Accepted Manuscript (AAM)

Publication

VoR — Publisher PDF*
Version of Record (VoR)

Preprint, AAM, VoR... what does it all mean?

Preprint
The version of an article before the peer review process has taken place, which does not include any changes requested by the journal's review committee.

Postprint or Author Accepted Manuscript (AAM)
The version of the article accepted for publication that includes the changes requested by the review committee. This version, produced by the author, does not include any formatting from the publisher (page numbers, logos, copyright notices, etc.) but its content is identical to the published version (VoR).

Publisher PDF or Version of Record (VoR)
The final version of the article, formatted and published by the publisher.

*see glossary

MAKING AN ARTICLE OPEN ACCESS

The article is being published in...

Gold Route*

A 'Traditional' journal with a paywall	Hybrid Journal ⚠ Not recommended ⚠	Open access journal Watch out for ⚠ predatory journals* ⚠
	No APC	With APC	With or without APC

The article is not **available for open access** from the publisher.

The article is **available for open access** from the publisher.

Green Route*

In accordance with Article 30 of the Digital Republic Act*, you may submit to HAL: *the author accepted manuscript after a waiting period of no longer than 6 months for STM[1] and 12 months for SHS[2].*

MAA | HAL archives-ouvertes.fr

Some publishers allow submissions before 6 or 12 months have elapsed, and/or VoR submissions. Check the journal policies in SHERPA/RoMEO!

[1] Science, technology and medicine
[2] Social and human sciences

You retain your author's rights and submit *the VoR once it is published* to HAL.

Even if the article is already open access, this will help ensure that it remains archived outside the publisher's platform and further boost its visibility.

VoR | HAL archives-ouvertes.fr

Self-archiving

H2020 and ANR projects require that articles be submitted to open archives such as HAL, regardless of whether the article is published in an open-access or paywalled journal.

Open Access Journals: Handy Definitions

APC
The fees paid by an author (or their institution) to the publisher of a scientific journal for the right to publish an open access article. APC stands for 'author publication charges' or 'article processing charges', but there is good reason to be suspicious of these fees charged by certain publishers: the amount paid often has little or nothing to do with the actual publishing costs, but is instead based on the journal's prestige. So, in essence, APC often really stands for 'article prestige charges'.

Gold Open Access
A native open access journal. These are labelled 'full gold' when the journal is entirely open access (as opposed to hybrid journals). Some of these journals charge authors APCs (this is known as the author-payer model), but there are other funding models: funding from institutions or research organisations, freemium, crowdfunding, etc.

Hybrid Journals
Journals with subscription fees that charge authors APCs to publish open access articles. We do not recommend paying to publish in these journals, as they are essentially double-dipping: the author pays to publish, and the institution pays for access to articles behind the paywall.
Under Plan S (supported by the ANR and the European Commission), paying to publish in a hybrid journal is expressly forbidden – but authors may still publish in a hybrid journal as long as they do not pay any APCs and as long as they make the article freely available by submitting a version to an open archive*.

*see glossary

MANAGING AND OPENING DATA

THE DATA MANAGEMENT PLAN (DMP): A VITAL TOOL

A **DMP*** is a document that describes how the data generated by the research will be managed **during and after the project**. It is a **required** deliverable for all **H2020** and **ANR** projects.

To help you write your DMP, **templates** are available from the funding organisations and there are **contact people** you can turn to for help (see the 'Contacts' page at the end of this guide).

FIRSTLY, WHAT IS 'RESEARCH DATA'?

The OECD defines research data as 'factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings'.

EXAMPLES: experimental findings, measurements, statistical data, survey results, interview recordings, images, software source code, etc.

⚠ **NOTE**
- The term research data does not cover only raw data: it also covers processed, derived and analysed data.
- The data used to validate findings are not the only data produced by the research, nor the only data that can be made open access.

No matter what model you use, the DMP is divided into sections and must answer the following questions:

RESPONSIBILITIES	Who will manage the data? Who owns the data? What resources are being allocated to data management?
COLLECTION	What kinds of data are collected? What form does the data take? How much is produced? What equipment/software is being used?
METADATA	What metadata* and documentation* will accompany the data to make them comprehensible and reusable?
STORAGE	How will the data be stored during the project and archived afterward? How will the files be organised (folder structure, naming conventions, etc.)?
SHARING	

V0 -> V1 -> V2, ETC. The DMP* is an evolving document that must be updated throughout the project.

WHY MANAGE AND OPEN DATA?

- **To reduce the risk of data loss** and enhance security through appropriate storage and archiving systems.
- **To ensure better use of public funds** by preventing the production of redundant data and facilitating the reuse of data.
- **To promote research integrity and reproducibility** by facilitating the validation of findings.
- **To pave the way for new fields of data-based research** (data driven science, text and data mining) and promote cross-disciplinary work.
- **To promote your research** by boosting its visibility and impact: data sets may be cited using their DOI* and published in data papers*.
- **To establish ownership of data:** when data is published in a repository* it is timestamped, which helps prevent improper use.
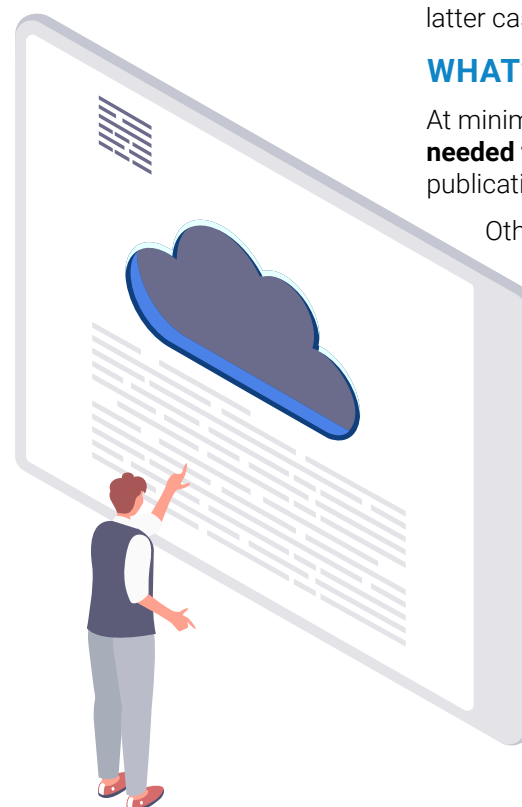
AM I REQUIRED TO OPEN MY DATA?

Opening data entails making them **freely accessible** online and making it possible to **freely reuse** them.

Recipients of funding from the **H2020** programme and the **ANR** are required to develop a **data management plan***. This plan must include how the data will be **shared** or **justify why the data will not be openly accessible**, where appropriate.

Some data may/must remain **closed** in order to comply with requirements to **protect** personal or confidential findings and data. This must be **explicitly stated in the DMP*** (see next page).

For non-H2020 or ANR projects, we still **recommend** writing a DMP* and, whenever possible, opening data.

OPENING DATA

The DMP* must state exactly which data will be made open and which data will remain closed (and, in the latter case, why that is).

WHAT?

At minimum, the data and associated metadata **needed to validate the findings** presented in the publications must be made open.

Other data may also be made open, according to the conditions listed in the DMP*.

WHEN?

The data needed to validate the findings presented in the publications must be made open **as soon as possible** (ideally, at the time of publication).

The other data will be made open according to the timeline in the DMP*.

WHERE?

If a recognised **depository*** exists for the discipline, we recommend using it.

If that is not the case, use a **general repository*** such as **Zenodo**, which is managed by CERN and recommended by the European Commission.

HOW?

▸ Add all useful **documentation*** and **metadata*** (title, date, format, etc.) so that your data may be **understood** and **reused**.

▸ Apply an **open licence** establishing **how your data may be reused** (for instance, a Creative Commons* CC-BY licence) and manage access rights if needed.

▸ Designate a **DOI*** for your data set (Zenodo does this automat.tion.

FUNDER AND EVALUATOR OBLIGATIONS

HORIZON 2020

The Horizon 2020 European research and innovation programme established the requirement that all publications resulting from projects funded by the European Commission must be submitted to an open archive*. The goal is to provide free and open online access to all scientific information discovered through EU projects so it can be reused by anyone, in accordance with the usual rules.

In addition, most Horizon 2020 research projects must also make the research data generated during the projects freely accessible and reusable online. To fulfil these requirements, a data management plan* must be developed at the start of the project and subsequently updated.

The costs associated with these requirements are eligible to be covered by EU funding and must be included in the project's preliminary budget.

If the open access publishing and/or data requirements are not met, the EU funding recipient may be subject to financial penalties in the form of repayment of some or all of the funds received.

ANR

For projects funded after the advent of the 2019 action plan from the ANR (the French national research agency - *Agence Nationale de la Recherche*) and in accordance with the funding rules, the partners in a research project funded by the ANR must agree to:

- submit the full text of scientific publications resulting from the research, development or innovation project to an open archive*, either directly to HAL or through a local institutional archive as stipulated by Article 30 of the Digital Republic Act*;

- provide, within six months of starting the project, an initial data management plan* that will be updated according to the process laid out in the Special Terms and Conditions and the Funding Rules from the ANR, which is a condition of receiving the full amount of the ANR funding.
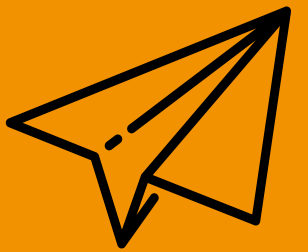
EVALUATING ORGANISATIONS

When conducting evaluations, some organisations only count publications from open archives.

For instance, when evaluating the activities of its researchers, CNRS now only counts publications that are listed in HAL and provide access to the full text of the publication.

The hope is that this will become common practice at other institutions (such as HCERES).

'THERE IS A STRONG TEMPTATION TO CLOSE OFF ACCESS TO ARTICLES AND DATA. SCIENCE BELONGS TO EVERYONE, HOWEVER, AND WE MUST SHARE IT AS WIDELY AS POSSIBLE'.

Frédérique Vidal, in his speech at the launch of the National Plan for Open Science.

*see glossary

LINKS

OPEN SCIENCE INFORMATION SOURCES

Site for the committee for open science: ouvrirlascience.fr

National Plan for Open Science ouvrirlascience.fr

Passport for open science - handbook for doctoral candidates ouvrirlascience.fr

Couperin open science site: openaccess.couperin.org

DoRANum, site for self-guided learning about research data: doranum.fr

CNRS open science roadmap science-ouverte.cnrs.fr

Plan S on the Coalition-S site: coalition-s.org

Article 30 of the Digital Republic Act:

read the law at legifrance.gouv.fr

Guidelines for applying Article 30 ouvrirlascience.fr

TOOLKIT

- **Open archives:**

HAL: hal.archives-ouvertes.fr

arXiv: arxiv.org

- **Tools for Checking Submission Policies:**

SHERPA/ROMeO: sherpa.ac.uk/romeo

FAQ for LRN at openaccess.couperin.org

- **Research Data Repositories:**

Zenodo: zenodo.org

Repository directory: re3data.org

- **Tool for Writing a DMP:**

DMP OPIDoR: dmp.opidor.fr

- **Directory of Open Access Journals (DOAJ):**

Directory of open access journals (DOAJ): doaj.org

- **Tools to Protect Yourself from Predatory Journals:**

List of predatory journals: predatoryjournals.com

Best practices for identifying reputable journals: thinkchecksubmit.org

QUIZ
OPEN SCIENCE

If I have already made a publication accessible on ResearchGate, I do not need to submit it to HAL.

FALSE. *Unlike HAL, ResearchGate is privately owned and does not maintain permanent archives: there is nothing to guarantee that publications made available there will still be accessible (or freely accessible) in the future. What's more, without oversight, there is no guarantee that publishing an article there does not violate the rights of the publishers.*

If I have signed over the exclusive rights to my article to a publisher, that means I am unable to provide free access to it on an open archive like HAL.

FALSE. *Thanks to the 2016 Digital Republic Act*, it is possible for you to submit an author accepted manuscript (see page 5 of this handbook)!*

If I open access to a preprint by submitting it to an open archive, there is a good chance that publishers will reject my article later on.

FALSE. *There may be publishers that will refuse to publish articles with open access preprint versions, but in the vast majority of cases, publishers accept this practice, recognising that submitting to a preprint server is not the same as publishing in a journal. Check the policies of journals and publishers for yourself at SHERPA/RoMEO.**

Predatory journals are hard to identify because there is nothing that distinguishes them from other open access journals.

FALSE. *There are a number of criteria you can use to identify a predatory journal: email solicitations, the lack of a bibliographic database (Scopus, Web of Science, DOAJ*, etc.), incorrect or missing information on the journal website, a vague or hasty review process and so on.*

It is sometimes possible to submit the version of my article that was published in a paywalled journal to the HAL archive.

TRUE. *It is possible if your publisher permits the submission of the published version to an open archive (this is true of the American Physical Society, for instance) – you can check the policies for journals and publishers at SHERPA/RoMEO.**

Merely listing my publications in HAL is not truly engaging in open access.

TRUE. *Listing a publication is not the same as providing access: you must submit the full text of the publication in HAL to ensure open access*.*

Projects funded by the EU H2020 programme or ANR must have a written data management plan* and publications must be submitted to an open archive* such as HAL.

TRUE. *See page 8 of this handbook.*

How much metadata (fields describing the publication: title, author, date, etc.) must I provide to submit a journal article to HAL?

☐ **4** ☑ **8** ☐ **16** ☐ **32**

The HAL interface has been streamlined in the past few years and now automatically retrieves the metadata for the article using the DOI* or the file submitted. Submitting your document takes only a few minutes!

The articles submitted to open archives are generally less cited than other articles.

FALSE. *Studies have shown that freely accessible articles are cited more overall than articles whose full text is behind a paywall.*

The Open Access Citation Advantage Service from SPARC Europe offers a list of studies on the topic up to 2015.

GLOSSARY

Article 30 of the Digital Republic Act

This article facilitates the open archive* process for scientific publications by enabling researchers to submit the author accepted manuscript no more than 6 months after publication (12 months for SHS), even when the publisher has been granted exclusive rights.

In addition to the need to get approval from co-authors, the following two conditions must be met for the law to apply:

- the article must have been produced through research for which at least 50% of the funding was from public sources (this includes authors' salaries)
- the article must have been published in a journal that publishes at least one issue each year

If one of these conditions is not met, the law does not apply, and you must check your contract or the journal's policies (by going to the SHERPA/RoMEO* database) to understand your open archive submission rights.

Creative Commons Licences

Licences that govern the rights to reuse and distribute a given work. There are 6 licences that enable authors to specify which rights are authorised (whether others have the right to modify the work, use it for commercial purposes, etc.), but in all cases, the Creative Commons licence permits the work to be shared as long as the author is credited. These licences are often attached to articles published in open access journals and may be attached to an article when it is submitted to an open archive*.

Data Documentation

All information accompanying a data set* that facilitates its comprehension and reuse by indicating, for instance, how the data were collected, what materials and methods were used, how they were processed and by whom, what variables were employed and how the folders have been structured.

Data Management Plan

A document with the purpose of summarising the description and changes to the data sets* for a research project and plan for the sharing, reuse and preservation of the data whilst complying to the fullest extent possible with the FAIR principles (data must be Findable, Accessible, Interoperable and Reusable). The DMP is divided into sections (collection, description, storage, sharing, etc.). DMP OPIDoR, a tool provided by the INIST, can be used to fill out a DMP using a pre-existing template (the models from the H2020 programme or the ANR, for instance).

Data Paper

A scientific article describing one or more data sets* produced during a given research process (collection methods, equipment used, potential for reuse, etc.) and providing access to these sets with a permanent link (DOI*) to the data repository* where they are stored.

The aim of a data paper is to inform the scientific community about the availability of these data sets and facilitate their use.

Data papers are peer reviewed and published in 'traditional' scientific journals or in data journals that specialise in these kinds of articles.

Data Repository

An online service where researchers may submit, describe, search for and publish data sets*. Repositories may be discipline-specific or general and may have an institutional, national or international scope. The re3data.org directory lists over 2,000 research data depositories. If a reputable discipline-specific repository exists for your field of research, we recommend using it. If that is not the case, you may use a general repository like Zenodo, which is recommended by the European Commission.

Data Set

A set of data files that form an intellectual unit, along with the associated documentation* and metadata*.

Open Access

Open access to scientific literature means providing free, unimpeded, digital access to scientific articles without violating copyright restrictions.

The term 'free access' is used when the content can be freely reused, adapted and redistributed.

DOAJ (Directory of Open Access Journals)

An online directory listing all full gold open access scientific journals. The DOAJ only lists peer-reviewed journals (with or without APCs) and excludes hybrid journals.

DOI (Digital Object Identifier)

A permanent ID for a digital resource. A unique DOI is assigned to every published article and is used as a unique identifier.

Embargo

The waiting period at the end of which the author may submit their article to an open archive*. The Digital Republic Act* limits this embargo to 6 months for STM (12 for SHS).

Gold Open Access

Native open access publication in peer-reviewed scientific journals. Depending on the journal and its financial model, authors may be required to pay fees known as APCs.

Green Open Access

Publishing an open access article through self-archiving (author submission) on an open archive*. This submission is free for the author and the reader, who has free access to the content.

HAL Collection/Portal

Collections and portals are subsets of the HAL platform.

A collection is made up of a set of submissions that have been selected based on criteria chosen by the individual in charge of the collection (collection manager). A collection may, for instance, include all the submissions from a certain laboratory or research team.

A portal is made up of all the submissions from an institution (research organisation, university, institute, etc.). Unlike a collection, a portal has a unique domain name and its own submission interface that can be customised by the portal administrators.

Metadata

A set of structured information used to describe an information resource. Applied to research data, it is used to describe data sets* (title, author, date, keywords, DOI*, reuse rights, etc.) and facilitate management, research and reuse, especially through data repositories*.

Open Archive

A digital platform that allows free, unimpeded access to documents produced through scientific research (publications, data, software source codes, etc.), generally submitted by their author (known as self-archiving). An open archive may cover one or more disciplines and may be institutional, national or international in scope.

HAL is an open, multidisciplinary archive with a national scope, administered by the CCSD (the CNRS centre for direct scientific communication - *Centre pour la Communication Scientifique Directe*).

Predatory Journal

A pseudo-scientific journal whose aim is to profit off of the 'author-payer' model (see open access journal). Authors, usually solicited via email, are invited to submit articles, which are automatically accepted in return for the payment of an APC, no matter how little scientific value the submitted work may contain. The Stop Predatory Journals project maintains a list of suspected predatory journals and publishers, evaluated using 10 criteria.

SHERPA/RoMEO

A platform that compiles the open access policies of scientific publishers and journals (particularly with regard to submissions to open archives*): whether authors have the right to submit the postprint* version or the VoR*, required waiting periods, etc.

In all cases, the Digital Republic Act* authorises the publishing of the postprint* version of the article after a waiting period of no more than 6 months following the publication in a journal (12 months for SHS). If your article is covered by the law, you can use SHERPA/RoMEO to check whether your publisher has less stringent terms (the ability to submit the publisher PDF* or shorter waiting periods).

CONTACTS AND REFERENCES

MICHEL SERRES LIBRARY

Responsible for research assistance and HAL contact for Centrale Lyon: .... nicolas.jardin@ec-lyon.fr

HAL LAB CONTACTS

Ampère: .... laurent.krahenbuhl@ec-lyon.fr

ICJ: .... vignes@math.univ-lyon1.fr

INL: .... sylvie.goncalves@ec-lyon.fr

LIRIS: .... webmaster-hal@liris.cnrs.fr

LMFA: .... benoit.pier@ec-lyon.fr

LTDS: .... manuel.cobian@ec-lyon.fr

COLLABORATIVE PROJECTS

National Contracts Manager: .... veronique.merat@ec-lyon.fr

Regional, Carnot, UdL Project Manager: ..... elisabeth.dalverny@ec-lyon.fr

European Affairs Manager: .... benedicte.martin@ec-lyon.fr

European Liaison for Contracts Managed by Centrale

Innovation: .... sylvie.rault@centrale-innovation.fr

École Centrale de Lyon
36 Avenue Guy de Collongue
69130 Écully
+33 (0)4 72 18 60 00

www.ec-lyon.fr

École Centrale de Lyon is part of the Écoles Centrale Group.

@CentraleLyon