couperin.org

Consortium Unifié des Etablissements Universitaires et de Recherche pour l'Accès aux Publications Numériques

Improving the chances of success of its ANR project thanks to Open Science

Guide produced by the Research Data Working Group of the GTSO-Couperin

April 2020

Version 1

DOI: 10.5281/zenodo.3749577

About the authors

Romain Féret (ORCID ID: 0000-0002-1527-1482) is in charge of research data management and Open Access at the Library of the University of Lille. Since 2017, he has regularly supported the writing of grant proposals for ANR and H2020 projects. He also follows several ANR and H2020 on-going projects. He pilots the Working Group on Research Data of the GTSO-Couperin and represents the University of Lille in OpenAIRE.

Laetitia Bracco (ORCID ID: 0000-0002-2939-9110) is data librarian at the Library of the University of Lorraine. She supports PhD students and researchers in the management of research data and she produces data visualization for the monitoring of Open Science in her institution.

Cécile Arènes (ORCID ID: 0000-0002-1839-3530) is data librarian at the library of Sorbonne Université. She is in charge of research data management and digital humanities. She assists researchers to write their data management plans and trains researchers and PhD students in open science.

Stéphanie Cheviron is data librarian in charge of research data management at the Library of the University of Strasbourg since 2015. She's also functional project manager in charge of the development of scientific databases with researchers. Since 2019, she follows ANR and H2020 projects.

Élise Lehoux (ORCID ID: 0000-0003-2929-6768) is project manager "Scientific and Technical Information and Research Data" at the University of Paris. She assists PhD students and researchers in the management of their research data. She participates in projects dedicated to Open Science carried out by the university's libraries.

Ling Li (ORCID ID: 0000-0003-0882-3266) is in charge of scientific data processing at the University Lumière Lyon 2, where she works for the Research Department and the Library. She helps researchers to improve their data management practices.

Reviewers

We would like to thank **Kenneth Maussang** (University of Montpellier, 0000-0002-8086-8461) and **Corentin Spriet** (University of Lille, 0000-0002-5805-3426) for agreeing to review this guide and suggesting useful changes. Nevertheless, the authors assume full responsibility for this guide.

Translators

This guide was translated from French into English thanks to the participation of **Laetitia Bracco** (University of Lorraine), **Stéphanie Cheviron** (University of Strasbourg), **André Dazy** (Couperin), **Marlène Delhaye** (Aix-Marseille University), **Romain Féret** (University of Lille) and **Ling Li** (University Lumière Lyon 2).

The French version of this guide is available online: https://dx.doi.org/10.5281/zenodo.3741666

About this document

Members of the Research Data Working Group of the GTSO-Couperin wrote this guide.

The **Couperin consortium** is a non-profit association. It has 242 members, mainly universities, research organizations and business schools. In addition to its missions of evaluating and organizing the purchase of electronic resources for the benefit of its members, Couperin works to improve scholarly communication.

Couperin's Open Science Working Group (GTSO-Couperin) produces tools and resources to promote Open Science in the consortium's member institutions.

How to quote this document

Féret, R., Bracco, L., Cheviron, S., Lehoux, E., Arènes, C., & Li, L. (2020, April). *Improving the chances of success of its ANR project thanks to Open Science*. Zenodo. https://dx.doi.org/10.5281/zenodo.3749577

Glossary

AAP: *Appel à projet*, Call for proposals.

AAPG: *Appel à projet générique*, Generic Call for Proposals. About ¾ of the projects funded by the ANR are funded within this call.

ANR: French National Research Agency (*Agence Nationale de la Recherche*).

CES: *Comité d'Évaluation Scientifique* or Scientific Evaluation Panels. There are 48 CES, one by research themes (35 disciplinary research themes and 13 cross-disciplinary research themes). Based on external peer-reviewers' reports, the CES decide which projects are funded by the ANR.

Consortium agreement: agreement signed between all the partners of a project.

CoSO: *Comité pour la science ouverte*[1], French National Committee for Open Science. It animates and supports the actions associated with Open Science in France.

CPP: *Comité de protection des personnes* or research ethics committee. This committee must be consulted before the beginning of any research concerning the human subject.

Data paper: paper describing one or several datasets, especially how the data can be reused. Data papers are published either in specific journals (*data journals*) or in traditional scientific journals.

DMP: Data Management Plan. Document which describes the research data collected during a project and how the data will be managed in order to facilitate data sharing at the end of the project.

DOAJ: Directory of Open Access Journals.

DOI: Digital Object Identifier. Unique identifier attributed to datasets or publications.

DORA: San Francisco Declaration on Research Assessment[2].

DPO: Data Protection Officer, ou DPD (*Délégué à la Protection des Données*).

Eligible costs/uneligible costs: expenses that the ANR reimburses or not.

FAIR: Principles to improve the Findability, Accessibility, Interoperability and Reuse of datasets.

FNS: *Fonds National Suisse*, the SNSF (Swiss National Science Foundation) is the most important public funder for research in Switzerland.

Green Open Access: is the free dissemination of publications from an Open-Access repository, by one of the authors or by a third party.

Gold Open Access: is the dissemination of publications free of charges for the reader, directly from the publisher platform, either in an Open Access journal or in a hybrid journal.

Hybrid journals: journal distributed by subscription but some articles of which may be freely accessible to the reader (Open Access), if the authors pay Open Access charges. The majority of journals from publishers such as Elsevier or Springer are hybrid journals.

H2020: Horizon 2020, the EU framework program for research and innovation (2014-2020).

JCJC: *Jeune Chercheuse Jeune Chercheur*, this call is for individual research projects coordinated by young researchers (coordinators who has defended their thesis for less than 10 years).

---

[1] https://www.ouvrirlascience.fr/the-committee-for-open-science/

[2] https://sfdora.org/

Metadata: information necessary to describe data. Metadata are usually structured according to a standard.

OpenAIRE: consortium whose main objective is to support the research work of European scientists by creating and operating an Open Access infrastructure. In particular, OpenAIRE acts as an aggregator of research outputs (publications, datasets, software) and links them to projects funded by European research funders (H2020, ANR…).

Open Access: free, permanent and unrestricted online access to scientific publications. There are two ways to disseminate publications in Open Access: green Open Access and gold Open Access.

Open Access charges: Charges paid by authors so that their articles (APC, Articles Processing Charges), books (BPC, Books Processing Charges), or book chapters (BCPC, Book Chapters Processing Charges) are disseminated in Open Access by their publisher.

Open Access journals: journals whose articles are immediately available to the public.

Open Access Repository: platform where scientific publications can be uploaded to be preserved and disseminated. Repositories can be National (such as the French National repository HAL[3]), institutional or disciplinary (i.e. arXiv[4]). On these platforms, researchers can record a publication by describing it (title, authors, abstract…). They may also upload the publication itself if they have the right to do so.

Open data: open data is data that can be freely used, re-used and redistributed by anyone.

ORCID: Open Researcher and Contributor ID. Unique identifier for researchers.

Partner scientific leader: person in charge of the achievement of the actions of one of the partner (different of the Project coordinator).

Post-print (*author accepted manuscript*): version including revisions from the peer-review process. File without or with partial editor layout.

Pre-print (*submitted manuscript*): version sent by the authors to a journal, prior to the peer review process.

Project coordinator: person who is charge of the global achievement of the project (different of the Partner scientific leader).

Publisher version (*version of record*, *final version*): article with the final layout. Version distributed by the publisher.

PRC: *Projet de Recherche Collaborative*, funding instrument for French collaborative research projects.

PRCE: *Projet de Recherche Collaborative – Entreprises*, funding instrument for collaborative research projects between public and private entities with a potential opening to the world of business.

PRCI: *Projet de Recherche Collaborative – International*, funding instrument for international collaborative research projects.

QOAM: Quality Open Access Market[5].

[3] https://hal.archives-ouvertes.fr/
[4] https://arxiv.org/
[5] https://www.qoam.eu/

Research data: any information that has been collected or created to validate research findings. Although, usually digital (pictures, measures…), research data also includes non-digital data such as biological samples and must be supported by documentation such as metadata, laboratory notebooks and protocols.

Research data repository: platforms on which datasets can be uploaded, described and preserved. Data repositories are disciplinary or general (Zenodo, Dryad…).

RSSI: Responsable de la sécurité des systèmes d'information, CISO (Chief information security officer).

SATT: *Société d'accélération du transfert de technologie*, SATTs are responsible for enhancing and accelerating the process of technology transfer from publically funded research toward industry.

TGIR: *Très Grande Infrastructure de Recherche*, very large research infrastructure.

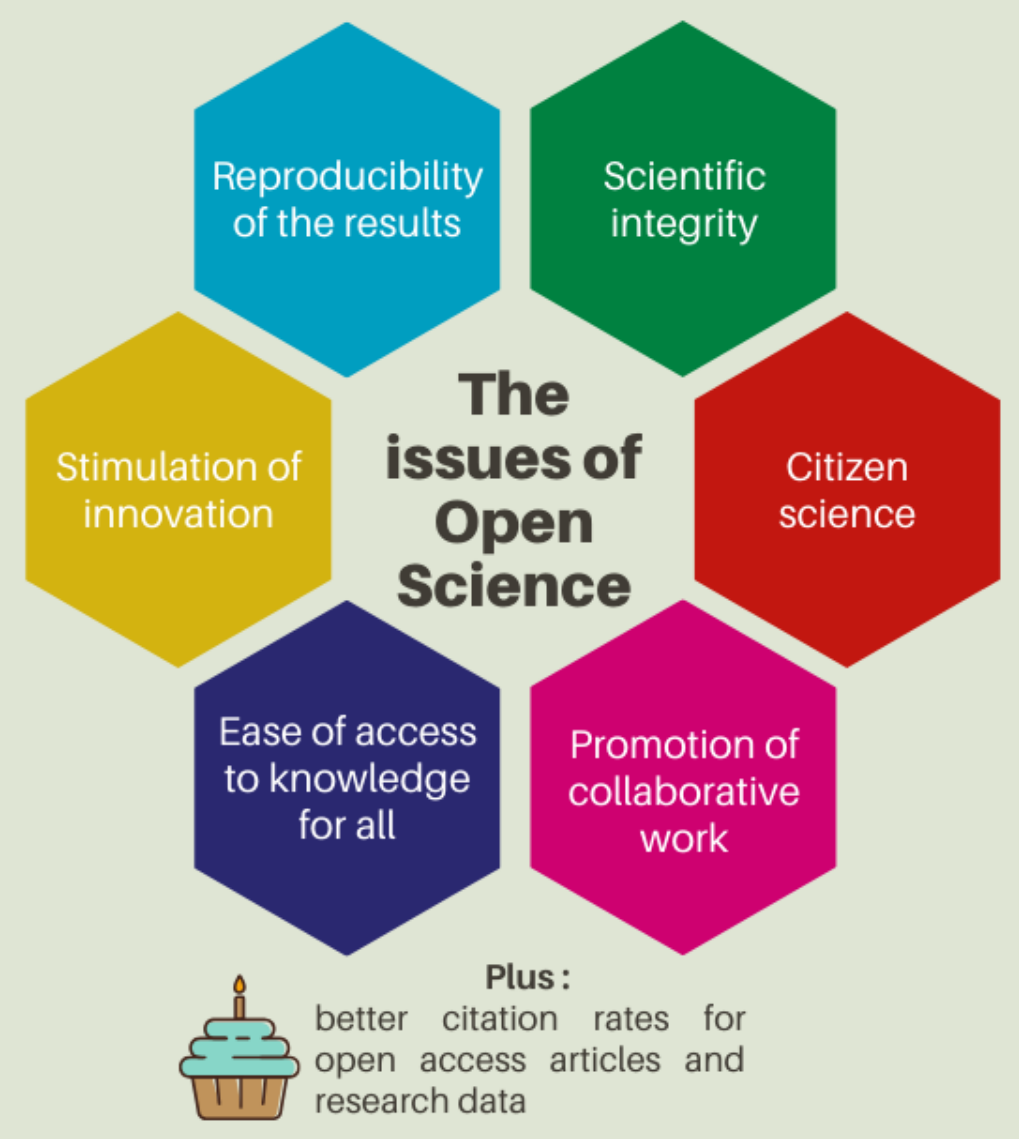
Table of contents

1. Open Science: Why? How?

1.1. French National Open Science Policy

The Open Science concept has been developed over the last ten years both to describe the **opening up of science results** and **the development of a new way of doing research through the dissemination of processes, codes, methods, and protocols**[6]. It is the result of a wide movement against the monopolistic position of major publishers.

Open Access encourages the disappearance of both economic and legal barriers to facilitate the diffusion of publications and, to a lesser extent, of research data. **The development of Open Science also responds to various issues**: **scientific** (effectiveness, integrity, and reproducibility of results), **social** (contribution to the science/society dialogue and greater transparency), and **economic** (stimulation of innovation).



[6] https://www.science-ouverte.cnrs.fr/le-mouvement-pour-la-science-ouverte/

Article 30 of the French Digital Republic Law (*loi pour une République numérique*), promulgated on the 7th of October 2016, guarantees the right for researchers to disseminate the author accepted manuscript (post-print) of their articles **6 months** after the first date of publication in Science, Technology and Medicine (**12 months** for researchers in Humanities and Social Sciences). **This law applies regardless of the contract signed with the publisher of the journal, whether the journal is French or foreign**.

Research data that have been made public by researchers or their institution may be **freely reused** if they are not protected by a specific right. A few months earlier, the loi Valter, about free access and modalities for the reuse of public data (enacted on 28 December 2015), established a principle of free reuse of public data. These two laws enshrine by default the principle of open data, which includes research data.

See : Lionel Maurel, « La réutilisation des données de la recherche après la loi pour une République numérique » [En ligne] https://hal.archives-ouvertes.fr/hal-01908766/document. (French)

The National Plan for Open Science (French), announced by the Ministry of Higher Education, Research and Innovation in July 2018, is structured along three axes. It aims **to generalise Open Access to publications**, **to structure and open up research data** and, finally, **to place France in a sustainable European and international dynamic**. This plan is declined in a series of measures.

"Open Science is not a fashion, it is not a discipline, it is a new paradigm. It therefore involves new practices and new skills. I will pay particular attention to ensuring that these skills are seen as part of the initial training of young researchers. Open Science will not be the domain of a small group of specialists; it will have to penetrate the whole world of research." Frédérique Vidal's speech (4th of July 2018)[7] (French)

1.2. ANR Open Science policy

As early as 2013, the National Research Agency (ANR, *Agence Nationale de la Recherche)* initiated an Open Science policy which intensified following the publication of the National Plan for Open Science.

See : Open Science Policy of the ANR[8] (French)

[7] https://www.enseignementsup-recherche.gouv.fr/cid132531/plan-national-pour-la-science-ouverte-discours-de-frederique-vidal.html
[8] https://anr.fr/fr/lanr-et-la-recherche/engagements-et-valeurs/la-science-ouverte/

1.2.1. ANR Open Science requirements

In **2019**, the ANR decided to enforce a new policy to promote and encourage the practical implementation of Open Science. For every funded project, the coordinator commits to:

1. **Upload the scientific publications (full text) resulting from the funded project in an open-access repository**, in accordance with the conditions stipulated by the French Digital Republic Law (*loi pour une République numérique*). Publications have to be freely disseminated at the latest 6 months after the date of publication for STM and 12 months for HSS ;
2. **Provide a Data Management Plan (DMP)** within 6 months after the beginning of the project. Coordinators and partners will have to update their Data Management Plan to provide a second version at mid-term (for projects over 30 months) and a final version at the end of the project.

The dissemination of data generated during ANR projects is not mandatory (especially if it represents a risk for natural or legal persons), but it is recommended.

The ANR's Open Science requirements are identical to those issued by the European research programme H2020, since 2014 for Open Access to publications and since 2017 for data management.

1.2.2. ANR recommendations

The ANR recommends to use its DMP template[9], which is available on DMP OPIDoR[10] where it can be completed online. However, coordinators are free to use another DMP frame if they wish. For example, some institutions offer institutional models for DMPs, such as the Paris Observatory, the INRA, the CIRAD or the University of Paris. No DMP fields are mandatory. In this case, simply indicate: not applicable.

Tip: The directory SOS-DMP[11] lists the services that support the writing of data management plans in universities and research institutions.

In addition, the ANR recommends that priority be given to publication in Open Access journals or books. It also recommends opening data whenever possible, based on the principle: as open as possible, as closed as necessary. When data are disseminated, it is preferable that they be disseminated according to the FAIR principles in order to improve their visibility and facilitate their reuse. Data may be released under certain conditions, such as after an embargo period or only on request.

See : ANR Open Science Policy in calls for projects[12] (French)

[9] https://dmp.opidor.fr/template_export/1858712127.pdf
[10] https://dmp.opidor.fr/
[11] https://openaccess.couperin.org/sos-pgd/
[12] https://anr.fr/fileadmin/documents/2019/ANR-Politique-science-ouverte-texte-AAP.pdf

1.3. How to use this guide properly

This guide is at the disposal of researchers submitting an ANR project. It is intended **to help project leaders anticipate** questions that it is sometimes too late to address once their project has been funded. It aims to **improve the viability of submitted and funded projects**, to **facilitate Open Science collaborations** between partners, and to help them **build a solid and ambitious Open Science strategy**.

This guide is not:
A collection of ready-to-use sentences or paragraphs, to be copied and pasted into project proposals.
A document to be applied in its entirety regardless of project specifics.
A document validated by the ANR that replaces or contradicts the policy and rules presented by the ANR in its official documents.

Nevertheless, this guide is based on the documentation made available by the ANR, as well as on the experience of its authors. It is also in line with the 15 Recommendations[13] (French) made in June 2019 by the CoSO Research Data College to the ANR. Several of these recommendations encourage the consideration of Open Science aspects from the project selection phase:
Recommendations n°3 and 5: mention data management in the full proposal; evaluate it during the selection process.
Recommendations n°6, 7 and 8: encourage the valorisation of data through data papers; acknowledge the ANR funding in publications and datasets published by ANR-funded projects.
Recommendations n° 12, 13 and 14: anticipate the resources needed to meet the ANR's Open Science obligations; implement them with concrete achievements.

[13] Plan de gestion de données – Recommandations à l'ANR. Zoé Ancion, Francis André, Sarah Cadorel, Romain Féret, Odile Hologne, Kenneth Maussang, Marine Moguen-Toursel, Véronique Stoll. Juin 2019. Available at: https://www.ouvrirlascience.fr/plan-de-gestion-de-donnees-recommandations-a-lanr/.

2. The full proposal, step by step

Some aspects related to Open Science can be mentioned as early as the writing of the pre-proposal for projects in which Open Science is structuring. However, the briefness of this first document rarely allows to develop these aspects. Project leaders who wish to anticipate them from stage 1 of the AAPG can refer to the corresponding sections of the full proposal (objectives, state of the art, partnership and bibliography).

2.1. Summary of the project

It is important to include the roles and responsibilities of the team members who will implement the Open Science policy in the summary table of people involved in the project. Depending on the size of the team, these roles may be filled by one person or shared among several. Answers to the following questions help identify the key elements to be included in this table.

- Research data management: Will there be a person responsible for data management within the project (data manager)? Will each partner have a "data contact" who will work in conjunction with the data manager? The data manager may be in charge of defining good practices, of coordinating the drafting of the DMP and its implementation, and of liaising with the various contacts (DPO, CISO, lawyers, IT specialists, etc.).

- Quality and documentation of research data: Who will be responsible for the quality of the data and their documentation (experimental or investigation protocols, manufacturing processes...)? This role may cover the verification of the quality of data entry, their homogeneity, the production of a codebook or data entry guide...

- Dissemination of scientific outputs: Who will manage the release of scientific publications and underlying data? Who will handle communication to a wider audience?

Summary table of persons involved in the project:

Partner	Name	First name	Current position	Role & responsibilities in the project (4 lines max)	Involvement (person.month) throughout the project's total duration
Marvel University	DANVERS	Carol	Captain	Scientific coordinator Tasks X, Y, Z	18p.month
Marvel University	OCTAVIUS OCTOPUS	Otto	PhD	Data Manager, responsible for the DMP and for data quality Task Z	12p.month
Marvel University	GREY	Jean	PhD	Writer of the codebook Task Z	3p.month
Marvel University	STRANGE	Stephen	PhD	Data Librarian, in charge of Open Access dissemination Task Z	2p.month

If the Open Science aspect was not anticipated in the pre-proposal, the main consequences on the budget should be specified now. There are costs to take into account during and after the end of the project. Indeed, the Open Science component of a project requires working time (in person.months) and may incur other types of costs.

During the project

Costs generated by data management must be taken into account, both in terms of hardware (servers, security, etc.) and human resources (data manager's workload, documentation, data cleaning and verification, etc.). Dissemination of research results and their long-term preservation may also generate costs, in particular the Open Access publication of articles, books or datasets.

After the end of the project

Data storage and preservation are also part of Open Science expenditure. These costs are eligible for reimbursement up to 5 years after the end of the project, provided they have been paid before the end of the project.

The impact of Open Science on the budget of a research project is specified in part 2.3.2 of this guide.

Tip: It is useful to get in touch with the relevant departments within your institution (grant office, lawyers, librarians) to properly analyse your project in order to identify areas of concern and avoid possible hidden or unexpected costs.

Example: A research team takes photographs of archaeological objects kept by a museum. The coordinator obtained permission from the museum to take these photographs before submitting her pre-proposal. However, she had not anticipated that she would be asked to pay fees for the re-use of the images in publications or in theses of the doctoral students involved in her project. She only became aware of this when she wrote the full proposal. It is then necessary to specify this point among the modifications made to the project in order to justify the difference between the budget of the pre-proposal and that of the detailed proposal.

2.2. Proposal's context, positioning and objectives

Part 2.2 details the main objectives of the project, highlighting in particular the potential reusability of the research results to be produced. In this regard, it should be emphasised, depending on the specificities of the project, whether the opening of data represents a specific challenge (for example, for survey data requiring complex anonymization techniques) or a particularly significant contribution to existing data. It is also important to show the added value of the collected data compared to existing data and why existing data may or may not be reused.

2.2.1. Objectives and research hypothesis

The research products to be described in this paragraph can take different forms:

- Datasets (example of a sound dataset on whale song: https://zenodo.org/record/1403351#.XoSEPnLgrIU), possibly associated with a data paper.
- Structured databases (example of the resources made available by the Ortolang[14] equipment on language processing).
- Software source code (example of the Samvera Hyrax[15] software source code).

If publications are in any case concerned by the Open Access obligation of dissemination in an open repository, research data are also a major issue of Open Science. It is therefore wise to underline in this section, where appropriate, that one of the structuring objectives of this research project is to open up the data produced in order to encourage their reuse.

[14] https://www.ortolang.fr/
[15] https://github.com/samvera/hyrax

Example: A laboratory produces reference hygrometric statistics. The creation of an open-access database, in order to enhance this work and to allow its reuse while preserving the data, can be a structuring Open Science axis of the project. In this case, it is necessary to provide the human, financial and technical resources to achieve this objective.

Tip : List examples of possible future reuse of the data produced.

2.2.2. Position of the project as it relates to the state of the art

One of the main objectives of Open Science is to encourage researchers to reuse existing data. Therefore, one should position one's project in relation to the existence or not of available data on the subject of study. If this is the case, it is firstly interesting to present these datasets and explain their interest (they are references) or the problems they raise in relation to the state of the art (these data are dated, new ones have to be produced) to possibly justify the interest of a new phase of data collection in the project.

If existing data are to be used, it should be explained how they will be reused: will the data be enriched? Will they be compared to the new data produced?

Example: two institutions have each released an open-access architecture prosopographic database on the web. A research team working on pedagogy in architecture education will reuse this data via an extraction to enrich it with the data that its members will collect in an archive collection. The resulting database will be harvestable, which means the data can be automatically retrieved by other databases. It improves the database dissemination and discoverability.

It could also occur to dispose of preliminary data, produced in previous research projects, and to be willing to build on them to pursue one's working hypothesis. This information should be mentioned in the proposal. It can then be indicated whether these data have already been disseminated and what they bring to the state of the art.

Example: a researcher has been working for several years on the constitution of an unpublished corpus of ancient statuettes. Some of the statuettes are photographed. Her proposal is based on these preliminary data to apply for funding that would allow her to finish the photographs, but also to create an Open Access database that would be a reference in its field.

2.2.3. Methodology and risk management

This section details the main stages of the project and the related scientific risks. Many Open Science issues can be addressed here.

Dividing the project into work packages or into tasks:

1. Plan within the WP on project management one or several tasks that specifically concern Open Science, with associated deliverables.

Example: Work package including Open Science tasks and deliverables

WP 1	Coordination, Management and Dissemination	Start: M1	End: M48
WP leader: Coordinator			
Contributions	Coord. (X P/M)	Partner 1 (X P/M)	Partner 2 (X P/M)
Objectives: • Administrative and financial management • Scientific management • **Data management and dissemination activities**			
Tasks			
T1.1	Administrative management, coordination (coord.)		
T1.2	Financial management (coord.)		
T1.3	Promotion of internal communication (partner 1)		
T1.4	**Writing and implementation of the DMP** (coord.)		
T1.5	**Project dissemination** (coord.)		
Deliverables	Deliverables		
D1.1	Kick-off Meeting	M1	
D1.2	Consortium agreement	M6	
D1.3	**Data Management Plan**	M6, M30, M48	
D1.4	Periodic and final reports	M12, M24, M36, M48	
D1.5	**Dissemination report**	M48	
D1.6	Workshop	M42	

Tip: identify a data correspondent per partner to facilitate the work of the person in charge of coordinating the drafting of the DMP and its implementation. Specify which partner will be in charge of this coordination.

2. Mention, if necessary, if specific actions for the dissemination of data and research results in Open Access are planned (e.g. publication of a data paper, putting a database online, dissemination of seminar reports or posters...) and indicate them as tasks and deliverables in the corresponding work packages. If a dissemination report is produced, the strategy for data dissemination and valorisation of scientific productions should be included in this report, including the target audiences.

3. Specify if the project requires access to a research infrastructure or a TGIR: will infrastructures be used to acquire data or to access existing data? Indeed, these infrastructures can offer tools and services to improve data quality and facilitate their dissemination.

Examples of TGIR: HumaNum[16] dedicated to digital humanities, Progedo[17] for data production and management in the humanities and social sciences.

Risk management:

1. Explain which types of sensitive data will be collected (personal data, sensitive data, patient data, biological data...) and which processing will be applied to them. Data protection and dissemination are not opposed in a binary way. For example, it is possible to anonymize data or to limit their reuse to certain uses.

Example: beQuali[18], a database of qualitative surveys in HSS, lists surveys that have been reversibly or irreversibly anonymized. The surveys, generally carried out in the form of interviews or observations, are then documented and made available only for reuse for academic purposes via a controlled access.

Tip : Horizon 2020 offers a guide to assess the ethical issues related to your project, which may be useful to consult : How to complete the self-ethics assessment[19].

2. Clarify whether certain scientific risks are related to data management, such as uncertainty about obtaining authorisations to collect sensitive data (or about deadlines), difficulties related to the storage or exchange of data between project partners, etc.

Tip: Identify the people who can help you manage sensitive data. If the stakes are high, it is strongly recommended that you involve them from the writing of your grant proposal.

[16] https://www.huma-num.fr/
[17] www.progedo.fr/en/
[18] https://bequali.fr/en/
[19] https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf

These support services may be: the Data Protection Officer (DPO[20]), the Committee for the Protection of Individuals (CPP[21]), the Ethics Committee or the institution's legal department. It may also be interesting to ask a librarian to seek a balance between data protection and dissemination.

To find out whether or not the research data produced as part of the project can be disseminated, the guide to analyse the legal framework in France[22] (French) about the openness of research data is a very useful resource.

2.3. Organisation and implementation of the project

Part 2.3 is an opportunity to present the experience of the team or consortium that will carry out the project, highlighting both the complementary nature of the working group and the expertise of each member. The Open Science obligations issued by the ANR require project leaders to identify the means and skills they already have and those they will need.

2.3.1. Scientific coordinator and its consortium / its team

Whether it is an individual project or a collaborative project, it will be in the interest of the scientific coordinator to highlight his or her experience and achievements in Open Science.

Scientific coordinator

Scientific expertise: In their individual presentation, the project coordinators can point out their achievements in Open Science, such as an important number of publications uploaded into open-access repository, datasets uploaded into general (Zenodo[23], Dryad[24]) or disciplinary (Protein Data Bank[25], Cambridge Structural Database[26]...) data repositories or their contribution to the creation of important databases in their field. Coordinators can, for example:

- Indicate the number of their publications in *Open Access*, thanks to the Dissemin[27] website.
- Indicate the number of datasets already disseminated, thanks to the DataCite[28] search engine.
- Specify the number of datasets disseminated in international databases.

[20] https://www.cnil.fr/en/home
[21] https://www.ars.sante.fr/comite-de-protection-des-personnes-1
[22] https://prodinra.inra.fr/ft?id=C7D38E14-877E-4883-AB91-0536DD4D63B9
[23] https://zenodo.org/
[24] https://datadryad.org/
[25] https://www.rcsb.org/
[26] https://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd/
[27] https://dissem.in/
[28] https://search.datacite.org/

Project management experience: more generally, coordinators can also indicate if they have already participated in research projects (ANR, H2020...) that were covered by Open Science obligations or if they have already contributed to the writing of a DMP in a context other than a research project (unfunded project, thesis...).

Focus: Identify the resources available within the institutions involved in the project

Just as the coordinators usually rely on a support unit for questions relating to administrative and financial aspects, it is essential that they identify the relevant services to assist them on Open Science aspects. These services can be:

- University Library: Open Access, metadata, assistance in writing a DMP, support for the repository of datasets.
- SATT or economic valorisation service: identifying data that can be subject to economic valorisation and those that can be freely disseminated.
- Data Protection Officer: identify the personal data that will need to be processed in a particular way; training in tools to manage these data responsibly (encryption of storage media, anonymization of survey or interview data).
- The Information Systems Department: data security, data storage and sharing, database hosting, custom developments, help with budgeting for IT services.

Tip: if you do not know who can help you write your data management plan, you can consult SOS-DMP[29] .This directory lists the services that support the writing of data management plans in universities and research institutions.

Other types of resources can be mentioned, such as:

- Training courses or workshops dedicated to Open Science that the consortium members have attended or are planning to attend.
- Available infrastructures and associated services: secure data storage space, data acquisition and processing equipment, etc.

Individual projects

In addition to the elements mentioned above, it is in the interest of the individual project coordinators to clarify the following points:

- Their current Open Science skills and those they intend to develop during the project.

[29] https://openaccess.couperin.org/sos-pgd/

- The distribution of roles and responsibilities within their team for the activities that fall under data management.

Collaborative research projects: overall coordination

In addition to the above points, coordinators of collaborative projects will need to ensure the quality of interactions between each partner, particularly in the area of data management. When reading the detailed proposal, it is essential to easily understand whether the partners are working on the same types of data, which implies sharing data and harmonizing practices (naming convention, organization, backup procedures...), or whether they are working on different data that only require sharing the obtained results.

This affects the solutions chosen to store and share data, as well as the degree of harmonization needed in data management practices between partners. In the detailed proposal, the following should be made clear:

- The balance between centralization and decentralization of data management by specifying the distribution of roles and responsibilities among partners. Explain the choice of whether or not to appoint explicitly a data manager.
- Where the data will be stored and how it will be shared: explain the tools used and justify their choice (robustness of the solution, security and location of the data).
- The main rules of data ownership sharing. For reused data: have they been produced outside the consortium or are they made available by one of the partners? For data acquired during the project: who will own it? This information will then be refined within the consortium agreement.

Focus: Should a data manager be appointed?

A data manager is in charge of defining good practices within the consortium and verifying their implementation by each of the partners. It does not need to be the project coordinator. It can be a researcher, engineer or technician. It is desirable for this person to have a thorough knowledge of the research practices of the project actors as well as scientific expertise to understand how the data are structured.

Indeed, the data manager role implies being involved at the operational level with all the project researchers, and more particularly with the scientific managers of each partner. It is not always necessary to recruit a dedicated person for this role. A person involved in the project may be given this role. It will then be important to position the data manager well within the team so that he/she can assume his/her role.

Example: the library of the University of Utrecht provides its researchers with data managers for 35 € per hour. These data managers are part of the projects and provide expertise in data management and data curation[30].

Collaborative research projects: partner by partner

In addition to an overall strategy, it is important for each partner to specify a number of elements concerning it. Depending on the project, each partner may, for example:

- List the main types of data it will produce.
- Identify the expertise present within its team in terms of data management.
- Allocate roles and responsibilities within its team.
- Describe the support provided by its institution: infrastructure; support by specific services (library, valorisation, DPO).

Tip: the detailed proposal should remain a concise document. You will not be able to detail all of the points listed above. Identify the specific characteristics of your project to target the main points to be mentioned in your project.

2.3.2. Implemented and requested resources to reach the objectives

The means needed to implement an Open Science strategy depend on the objectives set by the project members and the complexity of the data (data sensitivity, volume...). The means requested are dependent on the means implemented, which is why we have opted for a presentation of this section according to the typology of cost category proposed by the ANR. With regard to *Open Access* fees, commonly referred to as APC, projects are not obliged to use them. However, a focus has been made on this common practice.

In financial matters, the ANR's Financial Regulations[31] (French) are authoritative. The eligible costs sheet[32] (French) is a useful resource to find out which costs are eligible and which are not.

Staff expenses (category a[33])

These expenditures generally account for the majority of Open Science expenditures, although they are generally included in researchers' working time through routine research activities: data acquisition, data organisation and data documentation. However, the implementation of an Open Science strategy

[30] Jacques Florès. Supporting researchers with RDM costs at Utrecht university. 2019. www.dcc.ac.uk/webfm_send/2954 (slide 12).
[31] https://anr.fr/fr/rf/
[32] https://anr.fr/fileadmin/documents/2017/ANR-RF-Fiche-COUTS.pdf
[33] The category refers to the cost categories defined by the NRA in its Financial Regulations.

requires additional work, to a greater or lesser extent, depending on the degree of formalisation of the research practices of the project members.
This additional working time will be dedicated to activities such as: anonymizing data, improving data documentation (additional metadata, writing ReadMe files), systematically harmonizing file naming and organization conventions, saving files in different formats, depositing datasets in a dedicated repository, etc.

These good practices are a condition for disseminating data that can be reused by other researchers according to the FAIR principles[34]. They also have the advantage of facilitating the exchange of data between the project partners and ensuring the continuity of access and understanding of these data by the teams that produced them, even after the departure of a colleague or doctoral student.

The data curation work can be distributed among the project stakeholders or can be mainly entrusted to a data manager. In both cases, it will be necessary to estimate as accurately as possible the working time needed to manage the data according to the objectives set by the consortium.

Tools exist to facilitate the calculation of the working time to be spent on data curation, for example:

- The *Costing data management tool*[35] developed by the UK Data service
- The University of Utrecht details the costs incurred by data management[36] involved at all stages of the data life cycle

Focus: Opening interview data containing personal data

In some cases, opening the data is not very time-consuming because it does not involve modifying the data, while in other cases it is necessary to transform the data in depth. This is the case with the opening of data obtained during interviews.
To facilitate their reuse and anonymization, it may be necessary to transcribe them and then anonymize them manually or using specific software. The University of Manchester has put about ten practical sheets[37] online to inform researchers on how to process qualitative data, particularly for the transcription of interview data. To calculate this time, it is important to take into account many factors such as the number and length of interviews, the number of people interviewed at the same time, the quality of the recordings, the experience of the transcriber, etc. It takes at least 4 hours to transcribe 1 hour of interview data correctly.

[34] https://www.go-fair.org/fair-principles/
[35] https://www.ukdataservice.ac.uk/media/622368/costingtool.pdf
[36] https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management
[37] https://www.socialsciences.manchester.ac.uk/morgan-centre/research/resources/toolkits/toolkit-08/

Instruments and material costs (category b)

Within this cost category, Open Science expenses can for example cover the following needs:

- Storing and sharing data: whether the solution used is provided by the IT department or by an external service provider. This can cover the purchase of physical storage media (NAS servers, external hard drives) or access to a cloud space, as well as services to transfer large volumes of data securely. Some of these costs may be included in the costs of subcontracting (category d).
- Scan documents: if it requires the purchase or extensive use of a scanner.
- The University of Lausanne has developed a cost calculator[38] to anticipate the costs related to data storage and archiving.

Building and ground costs (category c)

This cost category has no impact on the Open Science aspects.

Outsourcing / subcontracting (category d)

In this cost category, project coordinators may include tasks that are outsourced such as:

- Access to a cloud
- Anonymization
- Legal advice
- Scanning
- Transcript

They may also include expenditure on the acquisition and maintenance of software to make data disseminable and reusable, such as software for transcription, data anonymization (ex. Amnesia[39]) or file format conversion.

It is also in this cost category that the costs of Open Access dissemination of articles, books, datasets, etc. may appear (see below the focus on Open Access charges).

General and administrative costs & other operating expenses (category e)

This category of costs does not have a major impact on the Open Science aspects, unless training on these topics is necessary and involves the payment of travel expenses for the trainers requested.

Tip: for Open Science as for the rest of the project, it is essential that the means requested are consistent with the implemented means and the objectives of the project.

[38] https://costcalc.epfl.ch/

[39] https://amnesia.openaire.eu/amnesiaInfo.html

Focus: Open Access charges (APC, BPC, BCPC, DPC)

APCs (*Article Processing Charges*) are fees to be paid for publishing an article in some Open Access journals. To find out if an *Open Access* journal requires APCs, you can consult the DOAJ (*Directory of Open Access Journals*[40]) and enter the name of the journal. Exemple for the journal *Ecology and Society*[41]: APCs are 975$ minimum per article. The QOAM (*Quality Open Access Market*[42]) database also allows you to compare the quality of the journal (contribution of the peer review process, content editorialisation, publication deadlines) and the amount of APC requested, as well as to know how much APC was actually paid by the authors. Example: the journal *Ecology and Society*[43] scores 4,4 out of 5.

APC are eligible costs only for full Open Access journals. **They are not eligible for hybrid journals.**

Hybrid journals are journals that offer the possibility of publishing Open Access articles for a fee (APC); however, other articles remain accessible only through a subscription. The business model of these journals allows them to generate income through:

- Payment of the subscription to the journal by the academic libraries
- Payment of APCs by the authors

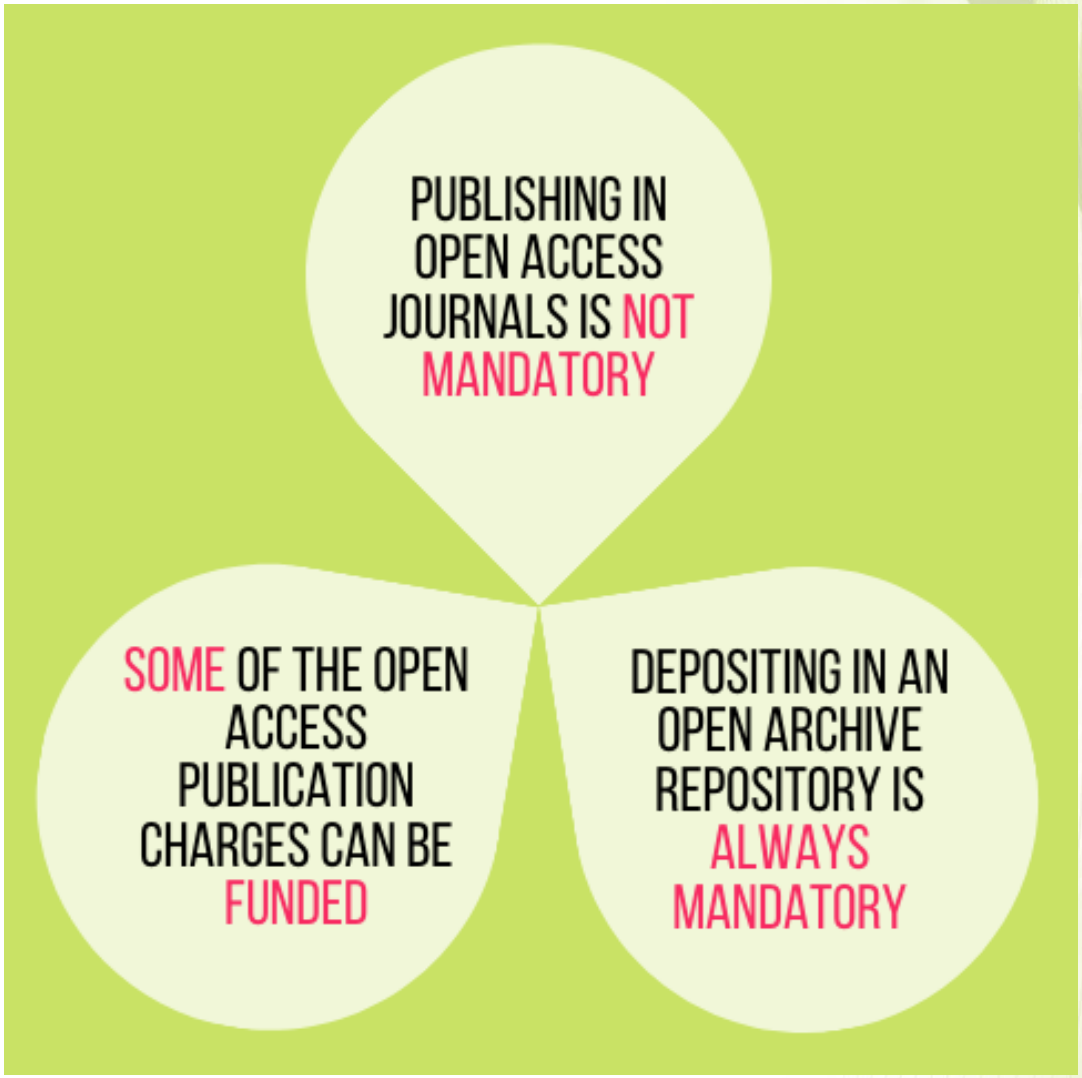
The payment of Open Access fees are eligible for other scientific productions, such as:

- Book Processing Charges (BPC) and Book Chapters Processing Charges (BCPC). For example, the publisher Berghahn Books[44] charges authors 15 000$ to publish an OA book and 2 000$ to publish a book chapter in Open Access.
- Dataset Processing Charges (DPC). For example, the research data repository Dryad[45] requires 120$ to preserve a dataset up to 50 Go and an additional 50$ per 10 additional Gb.

As with all ANR project expenditure, publication costs are eligible on condition that they have been paid before the end of the project, after the service has been provided.

[40] https://doaj.org/
[41] https://tinyurl.com/wabrydw
[42] https://www.qoam.eu/
[43] https://tinyurl.com/w8kqymh
[44] https://www.berghahnbooks.com/authors/
[45] https://datadryad.org/stash/publishing_charges

2.3.3. ANR Open Science requirements for funded projects



1) Publishing in an Open Access journal is not mandatory

ANR requires that publications resulting from the projects it funds be permanently preserved, and released in an Open Access repository, either directly in HAL[46] or through an institutional repository, such as LillOA[47] orUnivOAK[48].

The French Digital Republic Law, enacted in October 2016, authorizes researchers to submit the post-print versions of their articles in an Open Access repository, and to make them available 6 months after their publication by the publisher in TSM (12 months for researchers in Humanities and Social Sciences, Management and Law). **This law applies regardless of the contract signed with the journal's publisher, whether the journal is French or foreign**.

[46] https://hal.archives-ouvertes.fr/
[47] http://lilloa.univ-lille.fr/
[48] https://univoak.eu/

Researchers can therefore comply with the ANR's requirements without spending any APC fees.

NB: any researcher covered by French law, who is a co-author of a publication, can submit it under these conditions in an open access repository (with the agreement of its co-authors). However, this does not include publications produced exclusively by foreign partners in the case of PRCI, transnational projects and bilateral projects (France-Switzerland, France-Quebec...). In this case, it is necessary to check the policy of the journal, for example on the Sherpa/Romeo[49] website.

To find out more:

- WillO[50], a tool to check your rights and obligations to release an Open Access publication.
- The implementation guide of the Digital Republic Law[51], (French), published by the Open Science Committee.
- The FAQ on the Frech Digital Republic Law[52] (French), produced by Couperin.

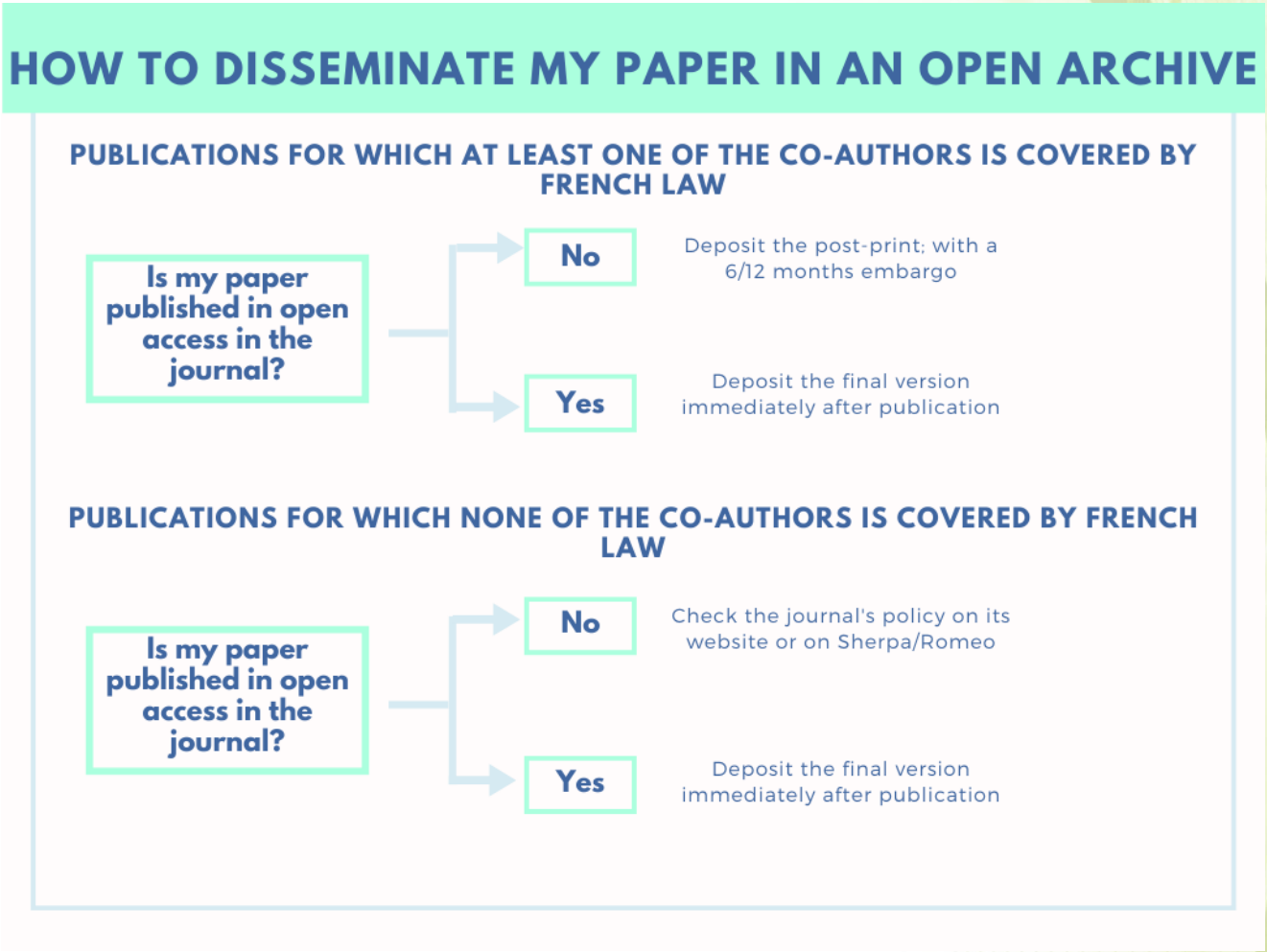
2) Submission to an Open Access repository is mandatory, regardless of the initial publication mode

Even if the article is published in an Open Access journal, the submission to an Open Access repository is mandatory. In this case, it must be done as soon as the article is published by the publisher.

Generally, when APCs have been paid for, the authors retain ownership of their article and/or the article is released under a Creative Commons[53] license. The final version of the article, known as the publisher version, can then be archived and disseminated in Open Access.

[49] https://v2.sherpa.ac.uk/romeo/
[50] https://decadoc.typeform.com/to/W2ZZMV
[51] https://www.ouvrirlascience.fr/guide-application-loi-republique-numerique-article-30-ecrits-scientifiques-version-courte/
[52] https://openaccess.couperin.org/category/faq/
[53] https://creativecommons.org/choose/

To summarise…



Tip: In order to find a balance between publishing in Open Access journals (gold road), often using APC, and the project budget, it is preferable to adopt either a self-archiving strategy only (deposit in an Open Access repository, green road) or a mixed strategy between green and gold roads.

2.4. Impact and benefits of the project

Section III of the full proposal template outlines the dissemination and communication strategy of the project. While some elements are common to all projects, the project coordinators may emphasise certain aspects to a greater or lesser extent depending on the funding instrument under which their project falls in order to stand out.

2.4.1. For every funding instruments

In terms of Open Science, it is in the interest of project coordinators to build their dissemination strategy around three main axes: the dissemination of Open Access publications, the dissemination of data and the link between Open Science and the valorisation of research.

Open Access

This section gives an opportunity to specify the dissemination strategy chosen for scientific articles: Open Access only through open archiving, or a mixed strategy that also includes the publication of a few Open Access articles in journals. The consistency of the strategy chosen by the consortia partners is essential.

It is also important to specify how other scientific productions will be disseminated, such as: books or book chapters, conference proceedings, slides presented at conferences, posters... Even if there is no ANR requirement about this, researchers would be well advised to ensure maximum dissemination of these productions which can be uploaded into Open Access repositories, acknowledging the ANR funding.

Research data

Disseminating data is now part of the ANR's expectations and it is a good way to increase the impact of a project within researchers' scientific disciplines, but also on related ones. It is also a good means to ensure that the project's research outputs are given a second life after it has ended.

It is preferable to distinguish the different ways of disseminating data by giving priority to data disseminated in international databases, the creation of a specific database or the publication of a data paper.

In a second phase, it is possible to describe how data that underlie the results presented in publications, but which can not be uploaded into disciplinary data repositories, will be archived and disseminated. General data repositories such as Zenodo or Dryad can be used to obtain a DOI for datasets and to make a reciprocal link between data and publications.

Example: researchers publish an article in the *Inorganic Chemistry*[54] journal. The corresponding dataset[55] is uploaded into Zenodo. A link between the article and the dataset is created using the DOI.

Some data will not be deemed interesting enough or can not be disseminated, according to the principle: as open as possible, as closed as necessary. It may be interesting to specify which data may not be disseminated and why, as well as what it is intended to do with them (deletion, confidential access).

[54] https://dx.doi.org/10.1021/acs.inorgchem.9b02096
[55] http://doi.org/10.5281/zenodo.3380352

From Open Science to valorisation

The societal valorisation of science is a growing challenge for research institutions and funding agencies. This notion covers very different realities depending on the projects and disciplines.

The main thing to remember is that there is a continuum between Open Science and the valorisation of research, which goes from the promotion of scientific productions to research communities and professionals from one's field (industrial sector, health workers...) to the general public, via the mediating role of, for example, school teachers or science journalists.

Presenting research results on freely accessible mainstream platforms such as The Conversation[56] comes under both Open Science and the valorisation of research, for example.

Regarding the economic valorisation of research productions (through patents for example), measures should be taken to ensure that the publication of data does not jeopardise the protection of an invention. If in doubt, it is better to wait for the patent to establish the anteriority of the invention and, in a second phase, to open the data after an embargo period. The SATT can provide information on this subject.

2.4.2. Specific approaches according to financing instruments

PRCI

Regarding the PRCI, it is interesting to show how the strategies chosen by the partners from different countries complete each other for the dissemination of publications and data. The use of European platforms or aggregators such as Zenodo[57] or OpenAIRE[58] facilitates the promotion of joint productions. Anticipation is key in this part, as each National agency may have specific Open Science requirements to which foreign partners are subject.

For example, the Swiss National Science Foundation (FNS[59]) requires a DMP as soon as the project is submitted.

PRCE

It is important to consider Open Science aspects when drafting the consortium agreement between the project partners since this may have an impact on property sharing or on the design of final products such as software. Finding the right balance between protection and valorisation of the results (patents, software, databases, computer modelling, etc.), and their dissemination can be part of a mixed strategy that should be described.

---

[56] https://theconversation.com/fr
[57] https://zenodo.org/
[58] https://www.openaire.eu/
[59] http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open_research_data/Pages/default.aspx

Example: a researcher develops a software in collaboration with a start-up during a PRCE research project. Placed under a GPL license, its basic functionalities will be free of charge, while the additional developments produced by the start-up will be subject to a fee. This is an example of commercialisation of a research product.

PRC and JCJC

The request here is similar to a dissemination plan, often required in European calls for projects. The Open Science strategy should be mentioned as an element of promotion, specifying how the dissemination of research results will be articulated towards the different actors of the academic world (researchers, students, etc.) and other circles (citizen science, press, decision-makers, etc.). Opening the results to non-academic audiences may take more time than traditional forms of academic communication. This needs to be anticipated to ensure the project's consistency.

2.5. References related to the project

The bibliography and the CV, as an annex, can also involve Open Science aspects.

Bibliography

- ANR recommends to provide a link, when existing, to an Open Access version of the articles. In addition to improve accessibility to the articles for the scientific experts and CES members, it shows the understanding of Open Science. To identify if an Open Access version of the references in the bibliography exists, the Unpaywall[60] web browser extension is useful.
- The ANR asks not to mention the journals' impact factor in the bibliography, in accordance with the San Francisco Declaration on Research Evaluation (DORA[61]). It is usually better to avoid such metrics in the full proposal. However, some CES are still receptive to this, and it is possible to mention metrics in the team or consortium members' presentation. It is preferable to favour metrics at the level of articles and individuals, such as the h-index, rather than at the journals' level (impact factor, CiteScore).

CV

- Whenever possible, it is better that the 5 major publications selected by the scientific leaders be freely accessible. If they have not been published in an Open Access journal, this is an opportunity to upload them into an Open Access repository. An Open Access mention or a link to the freely accessible version can be added.

[60] https://unpaywall.org/products/extension
[61] https://sfdora.org/read/

- If the data underlying these publications have been disseminated, they can also be highlighted with a link. The contribution these data can make should not be overlooked.
- The valorisation section can be used to present actions to promote scientific productions such as the dissemination of a database or the archiving and dissemination of an open software (in Software Heritage[62] for example).

ORCID identifier

- The ORCID[63] identifier is a 16 digit-number. It allows researchers to be uniquely identified even if they have homonyms or several author forms (e.g. name change during their career). It can be connected with other researcher identifiers such as IdHAL, ScopusID or ResearcherID.
- This number is not mandatory in the CV or in the full proposal, but it is provided for key individuals involved in the project when completing the ANR online form. It is essential that the coordinator and scientific leaders have an ORCID ID and their profile completed, at least with a list of their publications. University libraries usually offer assistance in creating and completing ORCID profiles.

[62] https://www.softwareheritage.org/?lang=fr
[63] https://orcid.org/

3. Bibliography

3.1. Open Science Public Policy and Regulation

- ANR. *Règlement financier*. 2019. https://ANR.fr/fr/rf/
- Couperin, Science ouverte France. *FAQ on the French Digital Republic Law*. 2018. https://openaccess.couperin.org/category/faq/
- LOI n° 2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public. 2015-1779 https://www.legifrance.gouv.fr/eli/loi/2015/12/28/2015-1779/jo/texte
- LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique. 2016-1321. https://www.legifrance.gouv.fr/eli/loi/2016/10/7/2016-1321/jo/texte
- MESRI. *Implementation guide of the Digital Republic Law (art. 30) – Publications*. 2018. https://www.ouvrirlascience.fr/guide-application-loi-republique-numerique-article-30-ecrits-scientifiques-version-courte
- MESRI. *Opening research data. Guide to analyse the legal framework in France.* 2017. https://prodinra.inra.fr/ft?id=C7D38E14-877E-4883-AB91-0536DD4D63B9
- MESRI. *Plan national pour la science ouverte*. 2018. https://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf

3.2. Guides and resources to implement Open Science into practices

- Commission Européenne. *How to complete your ethics self-assessment*. 2019. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf
- GOFAIR. *Fair principles*. 2016.
- https://www.go-fair.org/fair-principles/
- UK data service. *Data management costing tool and checklist*. 2015. https://www.ukdataservice.ac.uk/media/622368/costingtool.pdf
- Université de Lille. *WillO: researchers' rights and requirements to disseminate their publications in Open Access*. 2017. https://decadoc.typeform.com/to/W2ZZMV
- University of Manchester, School of Social Sciences. *Toolkit 08: Transcribing your own qualitative data*. 2010. https://www.socialsciences.manchester.ac.uk/morgan-centre/research/resources/toolkits/toolkit-08/
- Utrecht University. *Costs of Data Management*. 2018. https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management