Maple Transactions

Articles, software, data: An Open Science ethological study

TERESA GOMEZ-DIAZ*, University Gustave Eiffel-LIGM-CNRS, France
TOMAS RECIO*, University Antonio de Nebrija, Spain

Abstract.

Background. Open Science seeks to render research outputs visible, accessible, reusable. The Open Science framework is currently evolving vigorously due, among others reasons, to the UNESCO Open Science Recommendation adopted in November 2021. In this context, it is relevant to better visualize and describe the relationships that hold among the direct protagonists of this changing landscape: research teams and their research outputs, namely: articles, software and data, as their comprehension will certainly contribute to foster better Open Science practices.

Method. In this work we review and describe, through the information collected in a large number of bibliographic references, the current changing trends involving some essential, defining, characteristics and behaviors of the main components of the scientific production, namely, research teams and three kinds of research outputs they produce in many scientific areas. This comparative study is based, among others, in our recent work on the evolving concepts of research software, research data in the context of Open Science.

Results. In this work we observe and document some key features in this evolving landscape such as the changing and extended roles of research team members; the need to develop a new citing and referencing culture for articles, but specially for research software and data; the rising relevance of open access (to publications, software, data) policies all over the world; the existence of some barriers and difficulties like the regulations concerning academic research close to industry, or other technological applications; the need to develop standards for the "right to be forgotten"; the need to consider the impact of Open Science costs for less favored communities, countries, institutions…

Conclusions. This calls for the urgent need to observe and depict further this changing Open Science ecosystem, and to propose –as we have partially attempted in this work– new concepts to analyze this context as well as to contribute to ongoing research-on-research and to improve the implementation of Open Science practices, in order to foster better ways towards a sound, inclusive and fairer Open Science landscape.

Keywords. Open Science, research on research, scholar publications, research software, research data, research teams, citation, open access.

Recommended Reference Format:
Teresa Gomez-Diaz and Tomas Recio. 2023. Articles, software, data: An Open Science ethological study. *Maple Trans.* 3, 4, Article 17132 (December 2023), 19 pages. https://doi.org/10.5206/mt.v3i4.17132

1 Introduction

According to the Merriam-Webster Dictionary[1], *Ethology* is defined as:

(1) *a branch of knowledge dealing with human character and with its formation and evolution,*

*Both authors contributed equally to this research.
[1]https://www.merriam-webster.com/dictionary/ethology

Authors' addresses: Teresa Gomez-Diaz, University Gustave Eiffel-LIGM-CNRS, Est of Paris, France, Teresa.Gomez-Diaz@univ-eiffel.fr; Tomas Recio, University Antonio de Nebrija, Madrid, Spain, trecio@nebrija.es.

https://doi.org/10.5206/mt.v3i4.17132

(2) *the scientific and objective study of animal behavior especially under natural conditions.*

Lets us shortly explain the role of this term in our work.

Traditionally, scientists, working alone or as part of a research team, have had to consult documents, to collect and analyze data, with the help of technological (in a broad sense: from the invention of writing to printed material, from abacus to slide rules…) tools. Currently, this is still so, but we are now in a context in which the concepts of "documents, data and technological tools" have adopted new and extended forms. The observation and comparative analysis of some of the novel behaviors of the scientific community, reacting to such changes, is the goal of the present work, that is centered in the scholar context. Naturally, we acknowledge that science is also developed out of the academic world, but reflecting on the changing, Open Science influenced, conduct of the corresponding actors and productions, is out of the scope of this work.

Certainly, the behavior of scientists has always been conditioned by the ultimate goal of producing, sharing and disseminating scientific results. In the last years, the ways to achieve this goal have been the object of deep and revolutionary changes for different reasons. Let us think, for instance, on the impact of Big Data, machine learning, and AI technological revolution. But here we would like to focus on the consequences of the emergence of Open Science movements that have turned upside-down the scientific landscape. Indeed, Open Science aims to foster, in some specific manner, the visibility, accessibility and reusability of the scientific production [29, 42, 43, 91], thus clearly impacting in the approach to the three key steps of scientists work, namely, *production, sharing and dissemination of scientific results*, that now requires to adopt novel protocols.

Main stakeholders of this landscape, such the European Commission (EC) or the UNESCO, have started to adopt these new ways, and do significant work towards installing them soundly for the future. See, for example, the EC Horizon Europe Programme Guide (p.40) [29], where it is stated that Open Science has *the potential to increase the quality and efficiency of research and accelerate the advancement of knowledge and innovation…* or the UNESCO Recommendation on Open Science [91] that declares that Open Science aims *to make scientific knowledge, methods, data and evidence freely available and accessible for everyone, increase scientific collaborations and sharing of information for the benefits of science and society, and open the process of scientific knowledge creation and circulation to societal actors beyond the institutionalized scientific community.*

In order to contribute towards the achievement of such goals, we conduct this "ethological" study within the Open Science "ecosystem" [7, 88]. We will focus on the description of the different patterns that appear in this transition period concerning the roles of academic research teams, the habits of citation and referencing of the scientific productions, the pursued methodology to achieve their visibility, accessibility and reusability and the increasing relevance of the open access in their sharing and dissemination practices.

In the next sections we will first, shortly describe, providing relevant references, our perception of the constituent subjects of our ethological study: Open Science, articles, software, data (in section 2) and research team (in section 3). Then we will develop (in section 4) our observations concerning the changing roles of academic research teams in the pursuit of research outputs, and the current trends (see section 5) concerning the citation and referencing of such results, as well as the impact of Open Science on the so-called "open access" policy (section 6). We finish our work summarizing the main ingredients of the described landscape, and our subjective impressions in view of this evolving panorama.

2 Basic notions: Open Science, articles, software, data

Of course, the most basic notion of our work here is that of Open Science. In what follows we will adopt the following definition: *Open Science is the political and legal framework where research* *outputs are shared and disseminated in order to be rendered visible, accessible, reusable* quoting our previous work on this subject [42, 43].

Although we could say that it is now a well known concept, we think a few remarks are necessary here. For instance, as highlighted in [59], Open Science involves now a deeper philosophical debate that the one addressed in [30], as the whole scientific system is henceforth under the influence of the UNESCO Open Science Recommendation adopted in November 2021 [91], and of the research funders such the European Commission [29]. Other works related to Open Science key issues that we would like to mention here are [4, 5, 12, 19, 33, 55, 60, 61, 65, 71, 75, 84, 86, 93]. In particular let us emphasize the disruptive vision, perhaps too biased towards economic issues, that is presented in Lepage's work [60], that is one of the oldest works dealing with Open Science as such, as far as we are aware of.

Concerning the notion of *article*, it is largely assumed that this concept does not need to be thoroughly discussed [39]. Yet, to be precise, we would like to fix that in this work we adopt the following terminology:

> ***Research Article (RA)*** *is a written document that disseminates a scientific result produced by a research team. It can be published in a scientific journal, published as a book or a book chapter, presented in a conference or workshop, or disseminated as a preprint.*

Indeed, it is well known that articles produced by research teams are, and have been for centuries now, the main tool to disseminate and communicate scientific results among researchers. They are usually subject to a *peer review* procedure as an essential stage in their publication process, an step that is criticized – at least in its current form with predominant anonymous reviewing – by some Open Science defenders e.g. [49, 60, 87, 93]. See also the statement in [48] (p.54):

> *In short, with the digital world, the evaluation process stands ready to be reinvented in a clear, rational way by the relevant research communities themselves.*

declaring the need *to review the reviewing* process in a new technological context.

A more extended, related concept is the one of a *scientific publication*. Here, we include in this term any publicly available documents produced by a research team to disseminate a *scientific result*. They may adopt several shapes: preprints, articles, book chapters, monographs, books, conference papers, conference presentation supports, videos, podcasts. . . These outputs take often the form of a written document, but it is becoming more usual, after the pandemic times, to communicate through videos, live documents [13] or podcasts [76] although it is true that a large number of scientists are still emotionaly attached to printed documents. We will use in this work this more general term of *Scientific Publication* to extend the concept of *Research Article* to facilitate the comparison with *Research Software (RS)* and *Research Data (RD)*.

Let us remark that nowadays, the process of achieving a scientific result involves regularly the production of a dataset, maybe organized as database, and we will use here the broad term *data* to name all these research items. Similarly, in certain contexts, to get to a scientific result involves, more and more frequently, the development of *software* by the research team, with code (or a program) that has been written to analyze the data, to do simulations, or to test some theory, for example. For the purposes of the present work, the terms software and *computer program* are considered as synonyms.

More precisely, we adopt here the following definitions of Research Software and Research Data, quoting the recently proposed formulations in [41, 45, 46]:

> ***Research Software*** *is a well identified set of code that has been written by a (again, well identified) research team. It is software that has been built and used to produce a result published or disseminated in some article or scientific contribution. . .*

> ***Research Data*** *is a well identified set of data that has been produced (collected, processed, analyzed, shared and disseminated) by a (again, well identified) research team. The data has been collected, processed and analyzed to produce a result published or disseminated in some article or scientific contribution…*

Note that both definitions have been the subject of extended discussions and analyzes in the mentioned references [41, 45, 46], emphasizing, in particular, that that RS and RD can be dealt with in a similar manner. For example [46] proposes similar dissemination and CDUR (Citation, Dissemination, Use, Research) evaluation protocols for RS and RD. Here we will describe other RD and RS similarities and differences, that will be also analyzed in a larger setting including RA.

Let us also remark that a RS can also, as a written research output, convey the dissemination of a scientific result [57], which could also happen with some kinds of RD or the documentations or other documents related to these outputs.

In summary, in this article, we will use indiscriminately the terminology RA, article or scientific publication; software, computer program or RS; data, database, dataset or RD. And we will refer to these three objects in a global manner as *Research Outputs* or *ROs*.

3 The Research Team

Research team, as such, is a complex concept that can exhibit different dimensions. In this section we would like to present several observations concerning its changing role in the development and management of research outputs, emphasizing the manifold relation of a RT with each one of the kinds of ROs mentioned in the previous section.

Certainly, the concept of *Research Team (RT)* has been the subject of recent studies involving different prespectives such as collaboration [50], communities of practice [20], management [6], multidisciplinarity [89], gender-group-based studies [3], but here we will focus on its relation with the research object under consideration (articles, software or data). Since the notion of research team has such a relevant importance concerning the goals of this work we dare to start by proposing a definition for research team that mimics somehow the formulations previously proposed for RA, RS and RD.

> ***Research Team*** *is a well identified set of persons that are involved in whatever ways to produce a result published or disseminated in some article or scientific contribution in the academic context.*

Next, as a first step to describe in the following section some issues that we consider relevant concerning the changing roles of the constitutive elements of a RT regarding ROs, in the Open Science academic ecosystem, we will start by enumerating here a (surely incomplete) list of different kinds of membership relations we can actually find in RTs.

Traditionally, perhaps with some variations depending on the research topic under study, a RT was mainly constituted by researchers with a permanent position in an academic institution, including also PhD related students, lab engineers, other technical staff and some kind of leadership in the figure of a full professor with a permanent position. Nowadays, the complexified research landscape, which may include internationally funded projects, joint academic-industry cooperations, multi-academic and/or multidisciplinary actions, etc., requires enlarging the list of constitutive elements of a RT to include one or several of the following items:

- temporary positions funded by a specific project,
- researchers, engineers or trainees, hired by industrial companies associated to a given project, working in the academic context for the duration of the project,
- members of different research institutions and countries that cooperate in stable or temporary mode in a joint research context. For example, this is quite often the case in medical research,

jointly developed by academic and health institutions. It is also the case of French research labs that have multiple Head institutions, and the members of such institutions do collaborative research daily. Surely the reader is aware of, and can find many other examples of such mixed memberships,
- visiting researchers from other academic institutions temporarily included in a RT, during sabbatical leaves, etc.,
- undergraduate, master students, during the training activities associated to their diploma,
- retired or emeritus collaborators that are willing to still participate in ongoing projects,
- ...

Let us remark that most of the previous items describe RT components that were not part of the traditional concept of RT some decades ago. For instance it was quite rare in those days to include in the academic curricula the possibility for students to do external training in business or industrial enterprises, as part of their university degree studies, or, as another example of very different kind, teams constituted through international collaborative research projects were quite rare, etc. As a final example, let us refer to [37] where the reader can find a detailed description of membership kinds in the French academic software development context in a research lab.

4 Roles of RT members in ROs

Now, what are the roles of the RT constitutive elements regarding the production of research outputs? In the following we will attempt to address this question for the three kinds of ROs under consideration.

It may seem that there is no novelty concerning the behavior of a research team regarding the publication of such a traditional output as a RA. As already mentioned, articles are the main communication tool among researchers. But this model of diffusion is currently under a complete revision, in which "all is changing". In fact, we are assisting to an evolving panorama concerning its main characteristics[2]: from the printed documents to digital formats [48, 49, 83], from closed to open access [19, 27, 35, 49, 65, 83, 84], from paying for reading to paying for publishing or through the diamond model of publication [2], from blind to open peer reviewing [17, 64, 74, 87], and the raising awareness of the role of quantitative measures of research impact and their overwhelming influence in the academic life as well as the need to change the evaluation system [10, 21, 28, 32, 41, 46, 47, 49]. So it is not surprising that the demeanor of RTs concerning RAs has also evolved to adjust to these new contexts.

Thus, we acknowledge that, in present times, since the submission of a RA to a scientific journal is almost always done through digital platforms, it is a usual first step to be asked for the *affiliation* of the authors, as a way to have a certain assurance of the scientific origin of the work. Given the potential complexity of the affiliation description that could be involved in some of the above mentioned cases (section 3), it is usual to require as well some extra information to further describe and validate the declared affiliation. See for example, the F1000Research publication criteria and authorship[3]:

> **Publication criteria for research outputs presenting original data and results**
> - *Authors must be formally affiliated at an accredited institution or recognised organization.*
> - *Author affiliation is verified through institutional/organizational email address AND institutional/organizational website profile (or other means).*

[2]Note that each one of the following mentioned items have associated an extended literature. Providing a systematic list of the corresponding literature is out of the scope of this work, we will give here only few references for some items.
[3]https://f1000research.com/about/policies#publicationcriteria

- ...

Another source of verification concerning academic affiliation, widely consulted by the scientific community, could be Google Scholar[4] or ResearchGate[5], where many scholars can have an institutionally verified email address[6].

In the same line of this F1000Research example, we notice that it has became current practice in many scientific journals to ask to the *authors*, i.e. the publication *signatories*, to clarify their individual contribution to the submitted article. Indeed, it is now widely acknowledged that authors and/or signatories can have different *roles*. For example, F1000Research refers, as a guide, to The International Committee of Medical Journal Editors (ICMJE)[7] where the four criteria that should meet an *author* are defined, see also (for example) the ACM Criteria for Authorship[8] and [1, 56]. *Contributors* are qualified by ICMJE as those signatories of the publication that do not fill all the four author's requirements.

Furthermore, F1000Reseach asks that *Each author's contribution must be detailed by selecting CRediT roles on the article submission form.* The *Contributor Roles Taxonomy (CRediT)*[9] is a taxonomy that can be used to represent the roles typically played by contributors to research outputs. It is developed and maintained by non nonprofit organizations like CASRAI[10] or the National Information Standards Organization (NISO)[11] to develop and promote standards in research information management. We would like to mention Copernicus.org[12] as another example of an Open Access scientific publisher that has adopted and describes in detail the CRediT taxonomy[13].

Going a little bit further, let us remark that new technologies do raise up new challenges, for example, regarding the use of AI as a helper tool for writing an article. In particular, there are copyright challenges in determining if a work has been produced with enough human creativity[14]. In the case of F1000Research, this is addressed in this manner:

> *Authors must be aware that using AI-based tools and technologies for article content generation, e.g. large language models (LLMs), generative AI, and chatbots (e.g. ChatGPT), is not in line with our authorship criteria. All authors are wholly responsible for the originality, validity and integrity of the content of their submissions. Therefore, LLMs and other similar types of tools do not meet the criteria for authorship.*

Of course, the current notion of *author/signatory/contributor* role goes much beyond the writing, reviewing and editing a RA. For example, the above mentioned CRediT taxonomy deals as well with the diverse roles involved in addressing, among others, the following issues dealing with RD and/or RS:

- ...

[4] https://scholar.google.fr/
[5] https://www.researchgate.net
[6] See for example the user profiles in these platforms of Prof. Tomas Recio, one of the authors of this work.
[7] http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html
[8] https://www.acm.org/publications/policies/roles-and-responsibilities#author%20representations
[9] https://credit.niso.org/
[10] https://casrai.org/about/history/
[11] https://niso.org/welcome-to-niso
[12] https://copernicus.org/
[13] https://publications.copernicus.org/services/contributor_roles_taxonomy.html
[14] See for example *Using AI to Create a Work - Copyright Protection and Infringement*, June 2023, at https://www.jdsupra.com/legalnews/using-ai-to-create-a-work-copyright-6609501/.

- *Data curation: Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.*
- *Formal analysis: Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.*
- *Investigation: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.*
- *Software: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.*
- *Visualization: Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.*
- ...

In the specific case of RD, we would like to mention the EnviDat data portal[15] of the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL)[16] that proposes the Data Authorship Contributor Roles Taxonomy (DataCRediT) to distinguish six different roles in the RD production and publication[17] [24]:

- *Collection: Involvement in gathering and measuring information on targeted variables for a research dataset.*
- *Validation: Verification and cleaning of the dataset, whether as a part of the collection activity or separate.*
- *Curation: Involvement in annotating (producing metadata) and maintaining research data for use and re-use.*
- *Software: Implementation of the computer code and algorithms that assisted in the collection, validation, curation or publication of the dataset.*
- *Publication: Responsibility and involvement in activities related to the publication of the research dataset in a science data repository.*
- *Supervision: Oversight and leadership responsibility for achieving goals related to the collection, validation, curation or publication of the dataset.*

In [45, 46] we have selected this vocabulary to describe the main data activities that correspond to different roles of RT members involved in the RD: i.e. data that is collected, processed, analyzed, shared & disseminated. A larger data taxonomy can be found in the Data Management Lexicon of the Office of the Director of National Intelligence (USA) that also includes many roles involved in RD, such as Data Analyst, Data Architect, Data Custodian... [51].

In the specific case of RS, we would like to mention that one of the authors of the present work has studied in [39] (p.125) the different roles that do appear in RS development from the authorship perspective. For example many RS are disseminated as free/open source software, which may encourage the contributions of persons that are external to the RS research team. But this contribution can be maybe modified by the team in order to be integrated into the main version of the software to insure that the software remains functioning correctly after the integration of the external contribution. The RT should be aware about the management of the copyright issues regarding these contributions [39, 41].

Let us bring also the reader's attention to the important figure of *scientific responsible* that can maybe also hold the role of manager of the RS project, perhaps without contributing much to the

[15]https://www.envidat.ch/
[16]https://www.wsl.ch/en/
[17]https://www.wsl.ch/datacredit/

code writing. This figure also appears in [54] as the expert that will answer the scientific questions that are at the heart of the RS development:

Scientific software is defined by three characteristics:

1. *it is developed to answer a scientific question;*
2. *it relies on the close involvement of an expert in its scientific domain; and*
3. *it provides data to be examined by the person who will answer that question…*

In [41] (section 2.2) we have proposed a somehow simplified vision of the roles that appear in RS development, where role limits can be fuzzy depending on the RS, the RT and/or the size of the development project:

…we have selected three main roles:

(i) *RS responsible or leader,*
(ii) *main contributor,*
(iii) *minor contributor.*

Minor contributions to the RS can refer to bug correction, documentation, maintenance, testing, installation procedures… or many other tasks that are no directly related to code writing, and thus they are not directly involved in RS copyright issues, but these contributions may have a direct and important impact in the RS development (for example they can be the cause of big corrections or the generation of new code writing) and can be acknowledged by the RT by granting some percentage of the code writing to minor contributors, as proposed in [39, 41]. This is a RT decision.

The reader can find an extended description of several RS contribution roles in [80].

Similarly to the RS case, in RD there can also be this figure of the scientific expert or responsible of the project that maybe it is not directly involved in the RS code writing or in the main RD activities (collected, processed, analyzed, shared and disseminated), but without their scientific guidance, the RS or the RD will simply not exist. For the same reasons, and in the the case that there are author's rights associated to the RD, the persons that are not involved in the direct generation of these rights can be granted with a percentage of authorship, which is as well a RT decision. In here we can find a real difference with the RT behavior in the article production, as this scientific expert does usually appear among the authors of the publication, and so this person is directly granted with a percentage of authorship.

To conclude these descriptions of the current behaviors of RTs regarding the different kinds of research outputs, let us assert that we observe similar contributor roles in the case of RD as the ones already remarked for RS, that is, we can refer to:

(i) RD responsible or leader,
(ii) main contributor,
(iii) minor contributor.

And, similarly, for RA we can also refer to:

(i) RA responsible or leader,
(ii) main contributor or author (i.e. satisfying similar requirements as those of ACM or ICMJE),
(iii) other contributors.

As for RS, role limits can also be fuzzy for RA and RD.

5 Citation and referencing in ROs

As explained in [90]:

…la différence entre référence et citation : l'acte de référence relève d'un auteur donné alors que la citation est une nouvelle propriété, éventuellement calculable, du texte source. Selon

> *P. Wouters (1999), ce renversement a radicalement modifié les pratiques de référencement et littéralement créé une nouvelle "culture de la citation".*
>
> *[... the difference between reference and citation: the act of reference is the responsibility of a given author while the citation is a new property, possibly calculable, of the source text. According to P. Wouters (1999), this reversal has radically altered the practice of referral and has literally created a new "culture of citation".]*

When an article is published by a journal or a preprint is made available in platforms like arXiv[18] or Zenodo[19] there is a reference or citation form that is naturally settled and the scientific community at large knows how to deal with it.

On the other hand, it is usual practice in RAs to include, typically at the end of the document, a list of *references* to list the consulted bibliography that has been used to produce a new publication. This list of what we can call here *input information* may be completed with footnotes or endnotes to give links, comments or access to aside information, and that could be sometimes classified in a different scientific level as the one associated to articles, books and other scientific material mentioned in the bibliography.

This is the practice that we follow in the present work, but it varies among the different academic communities, publications and authors. For example in [65] each chapter has its own list of references and footnotes, or compare how these issues are dealt with in [41, 48, 49, 72] or in many other of the references listed at the end of the present work.

There are recent works to discuss and propose solutions for RS citation [23, 40, 41, 53, 79, 81, 82] and proposed tools like the JATS4R recommendation[20], Software Citation Files[21] [22], The CodeMeta Project[22] and the biblatex-software tool proposed by Software Heritage[23]. In [41] we select three easy to implement ways to propose citation forms or references to your own RS:

- *the reference to a research software paper or other kind of scientific publication that includes, and relies on, a software peer review procedure, or*
- *the reference to a standard research article that includes a description of the RS and the implemented algorithms, explaining motivations, goals and results, or*
- *a typical label, associated to the RS itself, and that identifies it as a research output, specifying its title, authors, version, date, and the place the software can be recovered from.*

And it is left to the RT to choose the preferred formulae, choice that can vary on time depending on versions and the evolution of the project.

Therefore, the RT should propose how to cite their RS, but they should also introduce and implement best citation practices to refer to other RS or other software components that have been included or used, following thus the usual behavior of scientific work referencing & citation in RAs. To have a complete list of the used external components is in particular important in order to verify their licences, and to confirm the rights of use, copy, modify and redistribute software, as proposed in the our RS dissemination procedure [37, 38]. It may also happen that the licences of included external components have consequences upon the licence that can be choosen to release your own RS, so to have settled this list may have capital importance. It is also the main way to recognise and acknowledge other's work, as it is the usual practice in RAs.

[18] https://arxiv.org/
[19] https://zenodo.org/
[20] https://jats4r.org/software-citations/
[21] https://citation-file-format.github.io/
[22] https://codemeta.github.io/
[23] https://www.softwareheritage.org/2020/05/26/citing-software-with-style/

Much work has already deal with data citation issues [11, 67], where we can find the following definition:

> *Data citation is a reference to data for the purpose of credit attribution and facilitation of access to the data.*

There are also tools like the DataCite Metadata Schema[24] proposed by the Metadata Working Group [18] which also includes propositions to cite RS.

Following with our comparison behavior between RS and RD, we can also propose here a similar set of citation forms for RD as the ones proposed to RS:

- the reference to a data paper or other kind of scientific publication that includes, and relies on, some kind of peer view procedure, or
- the reference to a standard research article that includes a description of the RD and that explains motivations, goals and results, or
- a typical label, associated to the RD itself, and that identifies it as a research output, specifying its title, persons involved in the RD production, version, date, and the place the RD can be recovered from.

Note that the proposition to cite documents (articles, documentation) associated to the RD can also be found in [9].

Similarly to RS and RA, the RD should have a list of references to other external components that have been included or used to build the RD. Similarly to RS, it is also needed to have a clear understanding of these external components' licences that allow to reproduce, distribute, display, and make adaptations, and in particular to include them in the RD under consideration, as remarked in the proposed dissemination procedures [37, 38, 46].

In order to achieve a complete and global vision of these citation and referencing issues for ROs, the reference list of any RO could (and should) include the citation of other research outputs, namely articles, software and data. In our vision, the relevance of best citation practices is reflected in the first step of the proposed CDUR evaluation protocole(s) for RS and RD, mainly its (C) Citation step [41, 46]. The use of such protocols will enhance best citation practices and contribute to spread them among the scientific communities.

Other initiatives related to evaluation tools such as peer review to promote the sharing of data and code can be found in [17, 64], that is, the Peer Reviewers' Openness Initiative[25]. On the other way around, UK "The hidden REF"[26] initiative, where REF refers to the UK Research Excellence Framework, is working towards increasing the amount on *non-traditional research outputs*, e.g. software and data, to be submitted by the Higher Education Institutes for the next research assessment in the UK in 2028, contributing in this way towards the recognition of every role that participates in the elaboration of these outputs.

6 Closed and open access in ROs

Open Science is much more than open access [83, 84], although this is one of its most relevant components. Thus, in what follows here we will reflect on some of the behavior changes of the scientific community regarding this particular aspect of Open Science.

As remarked above, open access is a crucial issue for Open Science. Indeed, we consider that Open Science stands over three foundational pillars, all involving in some manner open access:

- the Budapest Open Access Initiative (BOAI)[27] that defines open access to publications [42, 43],

[24]https://schema.datacite.org/
[25]https://www.opennessinitiative.org/
[26]https://hidden-ref.org/
[27]https://www.budapestopenaccessinitiative.org/

- the Free software definition[28] from the Free Software Foundation[29] to set the freedom to run, copy, distribute, study, change and improve software [42, 43],
- the Open Data concept as defined in [26, 65]:
  Open data as a concept is generally understood to denote data in an open format that can be freely used, re-used and shared by anyone for any purpose.

In order to recognize its relevance, we can mention, for example, the European Commission guidelines to fund open cooperative projects as key incentives that drive the behavior of the scientific community towards open access. These guidelines have evolved from pilot requirements of open access to publications and research data [27] to systematic request for open data and open access to research outputs [29], with mandatory practices to implement early and open sharing of expected outputs, including detailed instructions such as:

> *Open access to generated research data is required under the premise 'as open as possible as closed as necessary' meaning that there can be exceptions to this. Data management plans are mandatory for all projects generating or reusing data. Additionally, we recommend that you provide open access to research outputs beyond publications and data and share them as early and openly as possible.*

This specific requirement of a representative authority is particularly relevant as the context of open access is highly political [14, 15, 34, 35, 62, 63, 65, 66]. See also [72] for a complete study of current open access policies and practices in Latin America, The Caribbean and the European Union. In this later reference, as in [29], the concept of *datos de investigación* (scientific data) is already included in the more general vision of *resultados*, that corresponds in our work to the term of research outputs.

In summary, we think that we are at world-wide level fostering a new context for open access to ROs, which promotes and collects new trends of the scientific community, producing, in particular, new behaviors of RTs.

But these efforts are still far from achieving their final goals, as there do remain numerous difficulties and barriers. For example, in [77], one of the results of the the EU funded project EU NMBP-13 H2020 Gov4Nano, it is described the TRAAC framework standing over five pillars named Transparency, Reliability, Accessibility, Applicability and Completeness, in the Nanomaterials context, a framework

> *... to quantify the readiness of different tools and methods towards their wider regulatory acceptance and downstream use by different stakeholders.*

emphasizing that:

> *The framework diagnoses barriers which hinder regulatory acceptance and wider usability of a tool/method...In the future, the use of the TRAAC framework is planned to be further extended to consider data and knowledge readiness... This effort on data is currently being undertaken within the framework of EU NMBP-13 NANORIGO and Gov4Nano projects.*

In particular, we remark that the open access policy required by the EU to funded projects does facilitate a fast and easy access to Nanomaterial literature, which was, as far as we know, unthinkable few years ago.

As another consequence of open access policies and requirements, there are more and more scientific literature to understand, help, foster and accompany best dissemination practices of ROs, see for example [4, 16, 37, 38, 39, 40, 46, 49, 52, 64, 69, 73]. Since the communities of Free/Open Source Software are well established since long, best practices to develop and disseminate software

[28] https://www.gnu.org/philosophy/free-sw.en.html.
[29] https://www.fsf.org/

are also well know for long time [31], but sometimes not well enough in the context of RS [16, 37, 38, 39, 40, 52]. More references on the subject can be found on [41] and the cited work there.

Monitoring Open Science landscape has become one of the tools to observe the advances in the implementation and the adoption of best open access practices. The Registry of Open Access Repositories Mandatory Archiving Policies (ROARMAP)[30] is a searchable international registry charting the growth of open access mandates adopted by universities, research institutions and research funders that require their researchers to provide open access to their peer-reviewed research article output by depositing it in an open access repository. ROARMAP shows the growth of the number of OA policies adopted since 2005 by the institutions that are willing to register there their open access policies. ROARMAP is developed by the School of Electronics and Computer Science at the University of Southampton.

In France, the French Ministry of Higher Education and Research is developing the French Open Science Monitor[31] [8] to follow Open Science practices (in beta version for data and software) as part of the first French National Plan for Open Science [62, 63]. It is also part of the UNESCO Working Group on Open Science Monitoring [92], that is one of the five UNESCO Working Groups on Implementation of Open Science[32]. For instance, the Laboratoire d'informatique Gaspard-Monge, where one of the authors of the present article works, has a catalogue of RS productions, declaring a percentage over 75% of items disseminated as free/open source software [44].

The French Open Science Monitor shows a 61,5% open access rate by December 2021 of the 172167 French publications in 2020, and a 66,6% open access rate by December 2022 of the 158060 French publications in 2021. In UK, as a result of the Research Excellence Framework (REF) 2021 open access open mandate for journal articles and conference proceedings has had greatest impact in driving open access engagement by researchers. As a result, around 80% of the UK's research output can now be made open access [78], showing that *Cultural change is at the heart, which is slow, incremental, and involves winning hearts and minds.* In [66], the White House Office of Science and Technology Policy mentions that *the American public has experienced great benefits: more than 8 million scholarly publications have become accessible to the public. Over 3 million people read these articles for free every day.*

Nevertheless, although more and more ROs are declared open access, some are declared *open access if requested*, which does not really mean that they are easily accessible, as shown in [36], a study over 3556 analyzed articles, where 42% indicated that the datasets are available on reasonable request, but only 6.8% provided the requested data. Other studies on data practices [68] indicate that

> *... we found that scientists frequently sought data from public collections and from other researchers for comparative purposes such as "ground-truthing" and calibration. When they sought others; data for reanalysis or for combining with their own data, which was relatively rare, most preferred to collaborate with the data creators. [...] These researchers have learned through experience that interpreting data for reanalysis requires more knowledge of the context and purposes for which those data were created than is available through public documentation.*

This suggests that, in order to achieve a real RD reusability, it could maybe easier, or even necessary, to establish collaborations with the data producers.

To end up this section, we observe the raising of some new issues concerning open access in ROs. One of these issues concerns the "right to be forgotten" and academic publishing in an Open

[30] https://roarmap.eprints.org/
[31] https://frenchopensciencemonitor.esr.gouv.fr/
[32] https://www.unesco.org/en/open-science/implementation#open-science-working-groups

Science framework [85]. As mentioned there, the European Union (EU) adopted the General Data Protection Regulation (GDPR) in 2016 [25]. Its Article 17 deals with the *Right to erasure ('right to be forgotten')* (see also Articles 19 and 21), and [85] reflects on how this right is compatible with Open Science principles and questions about the pertinence of the debate regarding its local versus global applicability when considering the transnational nature of some collaborative research. The second issue we mention here is related to the economic cost of open access and the potential problems for less favored communities [59, 75, 94], yielding different problems regarding potential inequities, or inclusion/exclusion problems.

7 Conclusion

All along this article, we have described a very active, large and evolving panorama reaching practically all areas of scientific research in the academic context. It is a rising tide yielding a somehow uncontrolled flood. This calls for the urgent need to observe and depict this changing landscape, and to propose new concepts to analyze this phenomenon and to contribute to ongoing research-on-research [34] and to improve the implementation of Open Science practices.

In particular, in this work we have described and documented:

- the very changing, different, extended roles of RT members concerning contribution at articles, software and data (see section 4), including references to many documents describing these new situations and proposing new terminology),
- the need to develop a new culture for citing and referencing, mainly regarding research software and research data, but also for articles to refer to RS and RD outputs (section 5). As a side conclusion we highlight the perceived similarity between RS and RD with respect to these issues, a conclusion that has been supported through an important number of references in the same section,
- finally, in section 6 we have argued the rising relevance of the open access policies, through references describing the situation in The Caribbean, Latin America and also in some European countries and other contexts regarding open access to publications, software and data,
- but also we refer in that section 6 to some barriers and difficulties that should be taken into account in order to advance towards a complete open access and reusability of research outputs. Barriers deriving from some regulatory aspects impacting in scientific areas closed to industrial applications. Or difficulties arising from the urgent need to develop new laws and practices, to improve the conciliation of open dissemination goals and the "right to be forgotten",
- and last, but not least, our detailed observation of the Open Science ecosystem allows us to conclude warning the scientific community to pay attention to the undesired consequences of this changing Open Science landscape concerning the inclusion or exclusion, in this movement, of the less favored countries, institutions and scientific areas, in order to avoid the establishment of two Open Science worlds: North/South, Rich/Poor, Open/Closed…

It is our hope that this work humbly contributes somehow to foster better ways towards a sound and fairer Open Science landscape.

Acknowledgments

This work has been partially funded by the Vice-Présidence International of the Gustave Eiffel University (OII1-LIGM-Espagne 2023). It is also in debt to the Departamento de Matemáticas, Estadística y Computación (MATESCO) de la Universidad de Cantabria (Spain) for their kind hospitality.

This work has also been supported by the European Commission through the Erasmus+ Program under agreement 2021-1-DE01-KA220-HED-000032031 (Authoring Online Material with Multimodal, Dynamic and Interactive Applets and Automated Feedback for Learning Math (AuthOMath)[33]. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

[33]https://www.authomath.org/

References

[1] L. Allen, J. Scott, A. Brand, et al. (2014). Publishing: Credit where credit is due, Nature 508, 312–313, http://dx.doi.org/10.1038/508312a

[2] Z. Ancion, L. Borrell-Damián, P. Mounier, et al. (2022). Action Plan for Diamond Open Access, Zenodo Preprint, https://doi.org/10.5281/zenodo.6282403

[3] J. Astegiano, E. Sebastián-González, C. T. Castanho (2019). Unravelling the gender productivity gap in science: a metaanalytical review, Royal Society Open Science, 6:181566, http://dx.doi.org/10.1098/rsos.181566

[4] E. Barr, C. Bird, E. Hyatt, et al. (2010). On the Shoulders of Giants, Proceedings of the FSE/SDP workshop on Future of software engineering research (FOSER), https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/paper-68.pdf

[5] M. G. Bertram, J. Sundin, D. G. Roche, et al. (2023). Open science. Current Biology, 33(15), R792-R797, https://doi.org/10.1016/j.cub.2023.05.036

[6] C. Boardman, B. Ponomariov (2014). Management knowledge and the organization of team science in university research centers, The Journal of Technology Transfer, 39, 75-92, https://doi.org/10.1007/s10961-012-9271-x

[7] A. Boiko, O. Kramarenko, T. Mayboroda (2022). Open Research Data in the Open Science Ecosystem and business Environment, Business Ethics and Leadership, Vol 6, Issue 4, https://essuir.sumdu.edu.ua/bitstream/123456789/90516/1/Boiko_bel_4_2022.pdf

[8] L. Bracco, A. L'Hôte, E. Jeangirard, D. Torny (2022). Extending the open monitoring of open science: A new framework for the French Open Science Monitor (BSO), https://hal.science/hal-03651518/

[9] P. Buneman, G. Christie, J. A. Davies, et al. (2020). Why data citation isn't working, and what to do about it, Database, Vol 2020, https://doi.org/10.1093/databa/baaa022

[10] Coalition for Advancing Research Assessment (COARA) (2022). Agreement on reforming research assessment, https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf

[11] CODATA-ICSTI Task Group on Data Citation Standards and Practices (2013). Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data, Data Science Journal, Vol 12, https://doi.org/10.2481/dsj.OSOM13-043

[12] J. Cohoon, J. Howison (2021). Norms and Open Systems in Open Science. Information & Culture, 56(2), 115-137, https://doi.org/10.7560/IC56201

[13] R. Corless (2021). Maple Transactions-The Early Years, https://mapletransactions.org/index.php/maple/article/view/14508

[14] Council of the European Union (2022). Research assessment and implementation of Open Science - Council conclusions (adopted on 10 June 2022), https://www.consilium.europa.eu/media/56958/st10126-en22.pdf

[15] Council of the European Union (2023). High-quality, transparent, open, trustworthy and equitable scholarly publishing - Council conclusions (approved on 23 May 2023), https://data.consilium.europa.eu/doc/document/ST-9616-2023-INIT/en/pdf

[16] S. Crouch, N. C. Hong, S. Hettrick, et al. (2013). The Software Sustainability Institute: Changing Research Software Attitudes and Practices, Computing in Science & Engineering, 15(6): 74-80, http://dx.doi.org/10.1109/MCSE.2013.133

[17] M. Dahrendorf, T. Hoffmann, M. Mittenbühler, et al. (2019). Because it is the Right Thing to Do: Taking Stock of the Peer Reviewers' Openness Initiative, PsyArXiv Preprint, https://psyarxiv.com/h39jt/download/?format=pdf

[18] DataCite Metadata Working Group (2021). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs, Version 4.4., https://doi.org/10.14454/3w3z-sa82

[19] P. A. David (2008). The Historical Origins of 'Open Science': An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution, Capitalism and Society, Vol 3, Issue 2, Article 5, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2209188&gt

[20] L. Degn, T. Franssen, M. P. Sørensen, S. de Rijcke (2018). Research groups as communities of practice – a case study of four high-performing research groups. Higher Education, 76, 231-246. https://link.springer.com/article/10.1007/s10734-017-0205-2

[21] DORA (2012). San Francisco Declaration on Research Assessment, https://sfdora.org/read/, http://www.ascb.org/dora/

[22] S. Druskat (2019). The Citation File Format (CFF): Why, what, how?. https://elib.dlr.de/131409/1/2019-11-13-cff-ascl-registries-workshop.pdf

[23] C. Du, J. Cohoon, P. Lopez, J. Howison (2022). Understanding progress in software citation: a study of software citation in the CORD-19 corpus, PeerJ Computer Science 8:e1022, https://doi.org/10.7717/peerj-cs.1022

[24] I. I. Enescu, G. K. Plattner, L. E. Pernas, et al. (2020). Workflow for an Improved FAIR Environmental Data Publication in EnviDat. Revue électronique Suisse de science de l'information, pp. 60-66, http://www.ressi.ch/sites/default/files/No_special_DLCM.pdf

[25] European Parliament and of the Council (2016). Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04
[26] European Parliament and of the Council (2019). Directive (EU) 2019/1024 of 20 June 2019 on open data and the re-use of public sector information, https://eur-lex.europa.eu/eli/dir/2019/1024/oj
[27] European Commission, Directorate-General for Research and Innovation (2017). Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Version 3.2, 21 March 2017 (Version 1.0 dated 11 December 2013). https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
[28] European Commission, Directorate-General for Research and Innovation (2021). Towards a reform of the research assessment system – Scoping report, Publications Office, https://data.europa.eu/doi/10.2777/707440
[29] European Commission (2021-2023). Horizon Europe (HORIZON) Programme Guide, (Consulted in September 2023), Version 3.0, 1 April 2023 (Version 1.0 dated 17 june 2021), https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf
[30] B. Fecher, S. Friesike (2014). Open Science: One Term, Five Schools of Thought, In: Bartling, S.. Friesike, S. (Eds.) Opening Science, https://link.springer.com/chapter/10.1007/978-3-319-00026-8_2
[31] K. Fogel (2005-2022). Producing Open Source Software. How to Run a Successful Free Software Project, https://producingoss.com/
[32] Foro Latinoamericano sobre Evaluación Científica (FOLEC-CLACSO) (2022). Una nueva evaluación académica y científica para una ciencia con relevancia social en América Latina y el Caribe, V3 del 6 de junio del 2022, https://www.clacso.org/una-nueva-evaluacion-academica-y-cientifica-para-una-ciencia-con-relevancia-social-\en-america-latina-y-el-caribe/
[33] G6 statement on Open Science (2021). https://www.cnrs.fr/sites/default/files/download-file/G6%20statement%20on%20Open%20Science.pdf
[34] G7 Science and Technology Ministers (2023). G7 Science and Technology Ministers' Meeting in Sendai Communiqué https://www8.cao.go.jp/cstp/kokusaiteki/g7_2023/230513_g7_communique.pdf. And Annex1: G7 Open Science Working Group (OSWG) https://www8.cao.go.jp/cstp/kokusaiteki/g7_2023/annex1_os.pdf
[35] G8 London Summit (2013). G8 Open Data Charter and Technical Annex, https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex
[36] M. Gabelica, R. Bojčić, L. Puljak (2022). Many researchers were not compliant with their published data sharing statement: a mixed-methods study. Journal of Clinical Epidemiology, 150, 33-41. https://www.sciencedirect.com/science/article/abs/pii/S089543562200141X
[37] T. Gomez-Diaz (2010). Diffuser un logiciel de laboratoire : recommandations juridiques et administratives, Publication for the French PLUME project (2007-2013), https://zenodo.org/record/7096216
[38] T. Gomez-Diaz (2014). Free software, Open source software, licenses. A short presentation including a procedure for research software and data dissemination, Zenodo Preprint, Septembre 2014, http://zenodo.org/record/11709/. Presented at the Workshop on open licenses: Data licencing and policies, EGI Conference 2015, Lisbon, May 2015, https://indico.egi.eu/event/2452/sessions/1522/#20150520.
Spanish version: Software libre, software de código abierto, licencias. Donde se propone un procedimiento de distribución de software y datos de investigación, Zenodo Preprint, Septembre 2015, https://zenodo.org/record/31547/
[39] T. Gomez-Diaz (2015). Article vs. Logiciel : questions juridiques et de politique scientifique dans la production de logiciels. 1024 - Bulletin de la société informatique de France, N. 5 - March 2015, http://www.societe-informatique-de-france.fr/wp-content/uploads/2015/04/1024-5-gomez-diaz.pdf. First version published for the French PLUME project, October 2011, https://zenodo.org/record/7063154
[40] T. Gomez-Diaz (2019). Le Projet PLUME et le paysage actuel des logiciels de la recherche dans la science ouverte, Zenodo Preprint, https://zenodo.org/record/1405591
[41] T. Gomez-Diaz, T. Recio (2019). On the evaluation of research software: the CDUR procedure [version 2; peer review: 2 approved], F1000Research 2019, 8:1353. https://doi.org/10.12688/f1000research.19994.2
[42] T. Gomez-Diaz, T. Recio (2020-2021). Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs, Version 2 published in POLIS N. 19, december 2020, https://doi.org/10.58944/yuro5734. Version 3 available in Zenodo https://zenodo.org/record/4577066
[43] T. Gomez-Diaz, T. Recio (2021). The future of Open Science asks for a common understanding, Poster session, EGI Virtual Conference 2021, https://indico.egi.eu/event/5464/contributions/15729/
[44] T. Gomez-Diaz (2021). Logiciels libres/Open source dans l'ESR, Science Ouverte, Évaluation de la recherche, Open Source Experience Conference, Séance Open Science, Palais des congrès, Paris, 9-10 novembre 2021, https://www.youtube.com/watch?v=KPrE2RkovL8, http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/Internacional/2021novOSExp_TGD_ScienceOuverte.pdf

[45] T. Gomez-Diaz, T. Recio (2022). Research Software vs. Research Data I: Towards a Research Data definition in the Open Science context [version 2; peer review: 3 approved], F1000Research 2022, 11:118, https://doi.org/10.12688/f1000research.78195.2

[46] T. Gomez-Diaz, T. Recio (2022). Research Software vs. Research Data II: Protocols for Research Data dissemination and evaluation in the Open Science context [version 2; peer review: 2 approved], F1000Research 2022, 11:117, https://doi.org/10.12688/f1000research.78459.2

[47] N. Gras, L. Rovelli, D. Babini, P. Vommaro (2022). Forms of evaluation of research proposals oriented to development problems: practices and perspectives from national science and technology organizations and higher education institutions in Latin America and the Caribbean, CLACSO-IRDC, https://idl-bnc-idrc.dspacedirect.org/server/api/core/bitstreams/dcf8234f-fb57-43d0-8102-34d407187618/content

[48] J.-C. Guédon (2001). In Oldenburg's long shadow: librarians, research scientists, publishers, and the control of scientific publishing. Association of Research Libraries, https://www.arl.org/wp-content/uploads/2001/12/in-oldenburgs-long-shadow.pdf

[49] J.-C. Guédon, M. Jubb, B. Kramer, et al. (2019). Future of Scholarly Publishing and Scholarly Communication, Report of the Expert Group to the European Commission, January 2019, https://op.europa.eu/en/publication-detail/-/publication/464477b3-2559-11e9-8d04-01aa75ed71a1

[50] K. L. Hall, A. L. Vogel, G. C. Huang, et al. (2018). The science of team science: A review of the empirical evidence and research gaps on collaboration in science, American psychologist, 73(4), 532. Manuscript version available at https://psycnet.apa.org/manuscript/2018-23205-018.pdf

[51] The Intelligence Community (2022). Office of the Director of National Intelligence, Data Management Lexicon, https://www.dni.gov/files/documents/FOIA/DF-2022-00351-IC-Data-Management-Lexicon-January-2022.pdf

[52] R. C. Jiménez, M. Kuzak, M. Alhamdoosh, et al. (2017). Four simple recommendations to encourage best practices in research software, F1000Research 6:876, http://dx.doi.org/10.12688/f1000research.11407.1

[53] D. S. Katz, N. P. Chue Hong, T. Clark, et al. (2021). Recognizing the value of software: a software citation guide, F1000Research, 9:1257, https://doi.org/10.12688/f1000research.26932.2

[54] D. Kelly (2011). An Analysis of Process Characteristics for Developing Scientific Software, Journal of Organizational and End User Computing (JOEUC), 23(4), 64-79, https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.526.5076&rep=rep1&type=pdf

[55] M. Laakso (2022). Open Science and Open Access Publishing in Social Sciences. In T. Engels, & E. Kulczycki (Eds.), Handbook on Research Assessment in the Social Sciences Edward Elgar. https://haris.hanken.fi/ws/portalfiles/portal/28189171/Laakso_2021.pdf

[56] V. Larivière, D. Pontille, C. R. Sugimoto (2021). Investigating the division of scientific labor using the Contributor Roles Taxonomy (CRediT), Quantitative Science Studies, 2 (1): 111-128, https://doi.org/10.1162/qss_a_00097

[57] D. Lazard, R. Rioboo (1990). Integration of Rational Functions: Rational Computation of the Logarithmic Part, Journal of Symbolic Computation, 9, 113-116, https://core.ac.uk/download/pdf/81948988.pdf

[58] S. Leonelli (2023). Why the Current Insistence on Open Access to Scientific Data? Big Data, Knowledge Production, and the Political Economy of Contemporary Biology, Bulletin of Science, Technology & Society, Vol 33, Issue 1-2, https://doi.org/10.1177/0270467613496768

[59] S. Leonelli (2023). Philosophy of Open Science, Published online by Cambridge University Press, Series: Elements in the Philosophy of Science, https://www.cambridge.org/core/elements/philosophy-of-open-science/0D049ECF635F3B676C03C6868873E406

[60] C. Lepage (2010). L'open science, un défi pour les États. Cahiers Droit, Sciences & Technologies [online], https://doi.org/10.4000/cdst.277

[61] E. Méndez (2021). Open science por defecto. La nueva normalidad para la investigación, Arbor, Vol 197, n. 799, a587, https://doi.org/10.3989/arbor.2021.799002

[62] Ministry of Higher Education, Research and Innovation (2018). National Plan for Open Science, https://www.enseignementsup-recherche.gouv.fr/fr/plan-national-pour-la-science-ouverte-87868

[63] Ministry of Higher Education, Research and Innovation (2021-24). Second National Plan for Open Science, https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2021-10/second-french-plan-for-open-science-13715.pdf

[64] R. D. Morey, C. D. Chambers, P. J. Etchells, et al. (2016). The peer reviewers' Openness Initiative: Incentivizing open research practices through peer Review, Royal Society Open Science, 3: 150547, https://doi.org/10.1098/rsos.150547

[65] OECD (2015). Making Open Science a Reality, OECD Science, Technology and Industry Policy Papers, Vol No. 25, OECD Publishing, Paris, https://doi.org/10.1787/5jrs2f963zs1-en

[66] Office of Science and Technology Policy (OSTP) (2022). Memorandum for the heads of executive departments and agencies for Ensuring Free, Immediate, and Equitable Access to Federally Funded Research, August 25, 2022, https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf

[67] M. A. Parsons, R. E. Duerr, M. B. Jones (2019). The History and Future of Data Citation in Practice, Data Science Journal, Vol 18, https://doi.org/10.5334/dsj-2019-052

[68] I. V. Pasquetto, C. L. Borgman, M. F. Wofford (2019). Uses and Reuses of Scientific Data: The Data Creators' Advantage, Harvard Data Science Review, 1(2), https://doi.org/10.1162/99608f92.fc14bf2d

[69] J. M. Perkel (2023). How to make your scientific data accessible, discoverable and useful, Nature Vol 618 Issue 7967, https://media.nature.com/original/magazine-assets/d41586-023-01929-7/d41586-023-01929-7.pdf

[70] M. Potter, T. Smith (2016). Making code citable with Zenodo and GitHub, https://www.software.ac.uk/blog/2016-09-26-making-code-citable-zenodo-and-github. Also as a Zenodo Preprint. https://zenodo.org/record/45042/

[71] P. Rico-Castro (2019). ¿Amigos o enemigos? Cómo la open science pone a las políticas de open access frente al espejo. RUIDERAe: Revista de Unidades de Información (ISSN 2254-7177), (15) https://revista.uclm.es/index.php/ruiderae/article/view/2166

[72] P. Rico-Castro, L. Bonora (2023). European Commission, Directorate-General for Research and Innovation, Open Access Policies in Latin America, the Caribbean and the European Union: Progress Towards a Political Dialogue, https://data.europa.eu/doi/10.2777/162

[73] F. Rios (2016). Preserving and Sharing Software for Transparent and Reproducible Research: A Review v1.4, https://osf.io/fhz27/

[74] T. Ross-Hellauer (2017). What is open peer review? A systematic review [version 2; peer review: 4 approved]. F1000Research, 6:588, https://f1000research.com/articles/6-588/v2

[75] T. Ross-Hellauer (2022). Open science, done wrong, will compound inequities, Nature 603, 363, https://www.nature.com/articles/d41586-022-00724-0

[76] P. Schmidt (2021-2023). Podcast Code for Thought, https://codeforthought.buzzsprout.com/

[77] N. Shandilya, M. S. Barreau, B. Suarez-Merino, et al. (2023). TRAAC framework to improve regulatory acceptance and wider usability of tools and methods for safe innovation and sustainability of manufactured nanomaterials, NanoImpact, Vol 30, https://www.sciencedirect.com/science/article/pii/S2452074823000125

[78] T. Shaw, A. Barker (2023). Open access in scholarly publishing: Where are we now? Research Information Yearbook 2022/2023, https://www.researchinformation.info/feature/open-access-scholarly-publishing-where-are-we-now

[79] A. M. Smith, D. S. Katz, K. E. Niemeyer, FORCE11 Software Citation Working Group (2016). Software citation principles. PeerJ Computer Science 2:e86, https://doi.org/10.7717/peerj-cs.86

[80] J. Smithies (2019). Research Software (RS) Careers: Generic Learnings from King's Digital Lab, King's College London, Version 6.2, Zenodo Preprint, http://dx.doi.org/10.5281/zenodo.2564790

[81] L. Soito, L. J Hwang. (2016). Citations for Software: Providing Identification, Access and Recognition for Research Software, International Journal of Digital Curation, Vol 11, No 2, http://www.ijdc.net/article/view/11.2.48

[82] S. Stall, G. Bilder, M. Cannon, et al. (2023). Journal Production Guidance for Software and Data Citations, Scientific Data 10, 656, Nature, https://www.nature.com/articles/s41597-023-02491-7

[83] A. Swan (2007). Macroscope: Open Access and the Progress of Science, American Scientist, 95(3), 197-199, http://www.jstor.org/stable/27858953

[84] A. Swan (2012). Policy Guidelines for the Development and Promotion of Open Access, UNESCO, Paris, 2012, http://unesdoc.unesco.org/images/0021/002158/215863e.pdf

[85] J. A. Teixeira da Silva, S. Nazarovets (2023). Can the principle of the 'right to be forgotten' be applied to academic publishing? Probe from the perspective of personal rights, archival science, open science and post-publication peer review, Learned Publishing, Vol 36, Issue 4, pp. 651-666, https://doi.org/10.1002/leap.1579

[86] J. P. Tennant (2018). Open Science is just good science, DARIAH Annual Event 2018, https://www.youtube.com/watch?v=UEEcwRUgQu8

[87] J. P. Tennant, T. Ross-Hellauer (2020). The limitations to our understanding of peer review, Research Integrity and Peer Review, 5(1), 6, https://doi.org/10.1186/s41073-020-00092-1

[88] R. T. Thibault, O. B. Amaral, F. Argolo, et al. (2023). Open Science 2.0: Towards a truly collaborative research ecosystem, PLoS biology Essay, 21(10), e3002362, https://doi.org/10.1371/journal.pbio.3002362

[89] A. P. Thurow, C. W. Abdalla, J. Younglove-Webb, B. Gray (1999). The Dynamics of Multidisciplinary Research Teams in Academia, The Review of Higher Education 22(4), 425-440, https://muse.jhu.edu/article/30085, https://www.proquest.com/docview/1308045191?pq-origsite=gscholar&fromopenview=true&imgSeq=1

[90] D. Pontille, D. Torny (2013). La manufacture de l'évaluation scientifique: algorithmes, jeux de données, outils bibliométriques, 31: 25-61, https://minesparis-psl.hal.science/hal-00821956

[91] UNESCO (2021). UNESCO Recommendation on Open Science, https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en

[92] UNESCO (2022). Working Group on Open Science Monitoring, https://unesdoc.unesco.org/ark:/48223/pf0000383805

[93] M. Watson (2015). When will 'open science' become simply 'science'? Genome Biology 16, 101, https://doi.org/10.1186/s13059-015-0669-2

[94] M. A Winker, T. Bloom, S. Onie, J. Tumwine (2023). Equity, transparency, and accountability: open science for the 21st century, The Lancet, Vol 402, Issue 10409, pp. 1206-1209, https://doi.org/10.1016/S0140-6736(23)01575-1