The *Conundrum Challenges* for Research Software in Open Science

Teresa Gomez-Diaz [1,*,†] and Tomas Recio [2,*,†]

1 Laboratoire d'Informatique Gaspard-Monge, University Gustave Eiffel, 5 Boulevard Descartes, 77420 Champs-sur-Marne, France
2 Departamento de Matemáticas y Física, Escuela Politécnica Superior, Universidad Nebrija, C/Santa Cruz de Marcenado 27, 28015 Madrid, Spain
* Correspondence: teresa.gomez-diaz@univ-eiffel.fr (T.G.-D.); trecio@nebrija.es (T.R.)
† These authors contributed equally to this work.

Abstract: In the context of Open Science, the importance of *Borgman's conundrum challenges* that have been initially formulated concerning the difficulties to share Research Data is well known: *which* Research Data *might be shared, by whom, with whom, under what conditions, why, and to what effects*. We have recently reviewed the concepts of Research Software and Research Data, concluding with new formulations for their definitions, and proposing answers to these conundrum challenges for Research Data. In the present work we extend the consideration of the Borgman's conundrum challenges to Research Software, providing answers to these questions in this new context. Moreover, we complete the initial list of questions/answers, by asking how and where the Research Software may be shared. Our approach begins by recalling the main issues involved in the Research Software definition, and its production context in the research environment, from the Open Science perspective. Then we address the *conundrum challenges* for Research Software by exploring the potential similarities and differences regarding our answers for these questions in the case of Research Data. We conclude emphasizing the usefulness of the followed methodology, exploiting the parallelism between Research Software and Research Data in the Open Science environment.

Keywords: Research Software; Research Data; Open Science; Borgman's conundrum challenges

1. Introduction

It is now widely accepted (e.g., [1–11]) that Open Science is speedily expanding, while settled in a highly political and international context (more information at the *Science, technology and innovation (STI) policies for Open Science* OECD portal (https://stip.oecd.org/stip/open-science-portal) (accessed on 11 November 2024)). On the other hand, it is recognized that there are still defiances to be addressed; see, for example, [6,12–17]. In this context, *Borgman's conundrum challenges* are now of increasing relevance. These *conundrum challenges* have been initially formulated in the context of the study of the difficulties to share Research Data (RD) [18]:

> *[. . .] research data take many forms, are handled in many ways, using many approaches, and often are difficult to interpret once removed from their initial context. Data sharing is thus a conundrum. [. . .]*
>
> *The challenges are to understand which data might be shared, by whom, with whom, under what conditions, why, and to what effects. Answers will inform data policy and practice.*

Indeed, the *conundrum challenges* arise quite naturally, and formulate the basic problems that the scientific community may face when sharing RD. Hence, policy makers and researchers are called to furnish answers to these questions, bearing in mind the involved actors, as well as the different aspects (scientific, legal, best practices. . .) that could be implicated. Answers should be provided in such a way as to produce performing and efficient formulations, leading policy makers, and the scientific community at large, to set forth, regulate and to adopt best practices for data sharing and dissemination.

In [19,20], we have dared to propose answers to these conundrum questions as a consequence of a precise analysis of the different issues involved in defining, sharing, and dissemination conditions for RD. It is the intention of our present work to bring answers to the same questions for Research Software (RS), while highlighting the similarities and differences that arise in the production of both kinds of research outputs in the Open Science context.

Indeed, we consider that one of the difficulties implicit in these conundrum questions for RD derives from the potentially multiple, and sometimes fuzzy, meanings that may be associated to the concept of "*data*", and therefore, also to the concept of "*research data*". Yet, we consider that, even in this blurred RD definition context, answering the conundrum questions could be addressed regarding the standard practices of the scientific community for data sharing. But this approach is not possible because of the current, rapidly evolving, and novel scholar environment concerning data management in an Open Science context. A context dominated by challenging evaluation requirements and demanding necessary changes in the rewards and incentive system for researchers [1,6,7,12,13,15,20–24].

Thus, as it is complicated to bring precise answers to difficult and novel problems involving vague concepts, in our previous work we have prioritized the formulation of a new RD definition in order to tackle the Borgman's conundrum questions [19]. Let us notice that, to accomplish such task, and in order to define *what research data is*, we have left aside the difficulties arising into grasping *what data is* [25] (see page 28: *Now in its fifth century of use, the term data has yet to acquire a consensus definition. It is not a pure concept nor are data natural objects with an essence on its own*) [26] (see page 60, where answers to the question 2 *What do researchers think of the word "data" and how do they associate "data" with research materials used in their discipline?* include: *I know what [data] are until someone asks me to define them.* and *It is a complicated question [. . .] I do not know if I'm answering it or rather asking more questions*) [27] (see page 10: *While some of the interviewees appeared to be comfortable with their stance in relation to data and data-sharing, many showed signs of struggling with the notion of data, and the extent to which their research materials could be framed as data in the epistemic context of their research*) in general, as our interest relies specifically in conceptualizing the notion of data produced by a research team, rather than dealing with the general concept of data.

In our precedent work [19,20,28,29], we have developed a methodology that, among other outcomes, we have applied successfully in order to formulate such an RD definition, and we will continue to apply it here to address the conundrum challenges for RS by comparing from the already given answers for RD. Thus, we will elaborate on a parallel and simultaneous frame in which we analyze similarities and differences between the production of *data* and *software* in the research context. Hence, in Section 2, we begin by recalling the proposed formulations of the definitions of the main components of the present work, that is *research software*, *research data*, in the *Open Science* framework. Moreover, because of the importance of the producer team in the research context, we also present in this Section 2 a definition of *research team*. Then, Section 3 is devoted to revisit the conundrum challenges for RD and our previously proposed answers for this matter.

Our main contribution in this work is presented in Section 4, where we formulate answers to the conundrum challenges for the case of RS by taking advantage of our comparison methodology between RS and RD. In Section 5, we analyze the important role of research evaluation policies in the Open Science context, highlighting their relation with the conundrum challenges. Finally, in Section 6, we collect the main ideas and conclusions developed in the paper.

2. Context and Definitions

As previously discussed (and referenced) in the Introduction, Open Science is currently experiencing significant expansion, supported by international and national organisms responsible for issuing scientific research policy normative, and recommendations. International organisms are, for example, UNESCO, the G7 group that gathers Science and

Technology Ministers of several countries [4,5], or the European Commission. National organizations include, for example, the scientific institutions signing the G6 statement on Open Science [3].

This paper adopts the definition of Open Science that we have developed in [6], focusing on the related political and legal aspects:

> **Open Science** *is the political and legal framework where research outputs are shared and disseminated in order to be rendered visible, accessible and reusable.* ([6], V3, p. 2)

In these references, we have already argued the relevance of considering the characteristics of the three main constituents for Open Science research outputs: research articles, software and data. Aiming to be precise and self-contained, we adopt in the present work the following definition of Research Software, quoting our recently proposed formulation ([23]):

> **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. Each research software encloses a set of files that contains the source code and the compiled code. It can also include other elements as the documentation, specifications, use cases, a test suite, examples of input data and corresponding output data, and even preparatory material.* ([23], Section 2.1)

Likewise, we consider here the following definition of Research Data, coming from [19]:

> **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. Each research data encloses a set (of files) that contains the dataset maybe organized as a database, and it can also include other elements as the documentation, specifications, use cases, and any other useful material as provenance information, instrument information, etc. It can include the research software that has been developed to manipulate the dataset (from short scripts to research software of larger size) or give the references to the software that is necessary to manipulate the data (developed or not in an academic context).* ([19], Section 4)

Now, let us observe that, according to this definition, RS has three main characteristics:

- The goal of the RS development is to do research. As stated by D. Kelly: *it is developed to answer a scientific question* [30];
- It has been written by a research team;
- The RS is involved in the obtention of results to be disseminated through scientific articles (as the most important means for scientific exchange are still articles published in scientific journals, conference proceedings, books, etc.).

Analogously, we can notice the following four characteristics that follow from this RD definition:

- The purpose of RD collection and analysis is to do research, to answer scientific questions;
- It has been produced by a research team;
- The produced research results are intended to be published through scientific articles (or similar kind of contributions);
- Data can have associated software, which could be, or not, Research Software, for its manipulation.

Let us remark that both definitions have been the subject of extended discussions and analysis in [19,20,23,29], emphasizing, in particular, that RS and RD dissemination and evaluation issues can be dealt with in a similar manner, by using analogous protocols and methodologies [20]. See [19,23,28,29] for more discussions and references that are behind the RS and RD definitions that we have included in this Section. Notice that the Science Europe report [31] gives as first recommendation for Research Funding and Research Performing Organisations to *Provide a definition of 'research software' and other key terms.*

On the other hand, the above RS and RD definitions show the relevance of the producer Research Team (RT). Now, Research Team, as such, is a complex concept that can exhibit different dimensions whose precision is both non trivial and important, for several reasons. For example, RS involves written code that generates authorship rights, and it is crucial to understand who are the rightholders and how such rights should be managed. See, for example, ref. [28], where one of the authors of the present work studies the French legal context of the RS production. In the case of RD, the legal context is often quite different, as authorship rights, or other kind of legal rights are maybe involved in the RD manipulation, and, thus, the legal context should be analyzed on a case-by-case basis [17,19,32–34].

Moreover, the composition of an RD or an RS producer team may be simple, involving only one person that concentrates all the important activities and roles, or, as is often the case, very complex. In the case of an RS, these activities may have a more technical nature and entail dealing with a scientific question and the corresponding algorithm invention, software planning, code writing, testing, documenting, bug correction, version management, maintenance, etc., as well as other activities related to research dissemination: writing articles, presenting the research outputs in scientific conferences, etc.

In the case of RD, there also technical activities related to the data collection, process, and analysis, which can include documentation, corrections, version management, and maintenance of databases. There are also RD sharing and dissemination tasks, similar to the RS case, regarding articles to write, communications at scientific conferences, etc.

Obviously, to address such diversity of tasks usually requires an RT of complex composition, with several persons having different responsibilities and roles (team management, project management, fundraising. . .), perhaps coming from different institutions, or maybe working in a collaborative international context. See [29] for further discussions on the subject, including the following RT definition that we adopt in this work:

> **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.* ([29], Section 2)

In summary, let us end by emphasizing here that, for us, RS and RD must be regarded as the scientific contribution of an RT.

3. The Conundrum Questions for Research Data

As we have mentioned in the Introduction, we consider significant to deal with precise definitions in order to be able to handle the conundrum questions. Let us exhibit an example to illustrate this observation. One of the most mentioned RD definitions comes from the OECD [34]:

> *In the context of these Principles and Guidelines, "research data" are defined as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated.*

The OECD report also indicates *what is not* RD:

> *This term does not cover the following: laboratory notebooks, preliminary analyses, and drafts of scientific papers, plans for future research, peer reviews, or personal communications with colleagues or physical objects (e.g., laboratory samples, strains of bacteria and test animals such as mice). Access to all of these products or outcomes of research is governed by different considerations than those dealt with here.*

In our opinion, while this definition has relevant aspects, like to indicate that *A research data set constitutes a [. . .] partial representation of the subject being investigated*, we consider as well that it has two drawbacks. First, it does not discriminate the concepts of "*data*" and "*research data*" and tries to deal with *what data is or not* at the same time as it is stated *what research data is*. And, secondly, it seems to us that, in the OECD vision, there is not any

reference to the producer team of the involved data, so it is difficult to answer conundrum questions such as the second one: *who shares the RD?*

Hence, as argued in the Introduction, in order to avoid this kind of difficulty, we have developed a detailed RD definition (that has been already included in Section 2) in [19,20] in order to provide answers to Borgman's conundrum challenges [18] related to sharing RD:

> *The challenges are to understand which data might be shared, by whom, with whom, under what conditions, why, and to what effects. Answers will inform data policy and practice.*

as well as to two extra ones that we consider equally relevant, namely **how** and **where** to share RD [19]. What follows is the summary of the answers we have developed in [19].

Which data might be shared? It is a decision of the *RD's producer research team*: similarly to the stage in which the RT decides to present some research work in the form of a document for its dissemination as a preprint, or a journal article, a conference paper, a book... So, it is the team who decides which data might be shared, in which form and when (following maybe funder or institutional Open Science requirements).

By whom? By the research team who has collected, processed, analyzed the RD, and who decides to share and disseminate it, that is, by the RD producer team. Yet, data ownership issues can be sometimes a tricky question; see [32] and further references given in [19].

How? By following some kind of dissemination procedure like the one proposed in [20] in order to identify correctly the RD set of files, to set a title and the list of persons in the producer team (that can be completed with their different roles), to determine the important versions and associated dates, to write the associated documentation, to verify the legal [19,32] (and *ethical*) context of the RD, including issues like data security and privacy, and to give the license settling the sharing conditions, etc., which can include the publication of a data paper [35,36].

In order to increase the return on public investments in scientific research, RD dissemination should respect principles and follow guidelines as described in [2,34,37]. Further analysis on RD dissemination issues can be found in [20].

Where? There are different places to disseminate an RD, including the web pages of the producer team, of the funded project, in a repository like Zenodo (https://zenodo.org/ (accessed on 11 November 2024)) or in a more specific scientific area data repository or an institutional repository. Let us mention here the Registry of Research Data Repository (https://www.re3data.org/ (accessed on 11 November 2024)), funded by the https://digitalresearchservices.ed.ac.uk/resources/re3data-org (accessed on 11 November 2024) German Research Foundation (DFG) (http://www.dfg.de/ (accessed on 11 November 2024)), which is *a global registry of RD repositories that covers repositories from different academic disciplines*, and can help the RT to find the repository or repositories where they would like to disseminate their RD outputs. Note that the Science Europe report [38] provides criteria for the selection of trustworthy repositories to deposit RD.

With whom? Each act of scholar communication has its own target public and, initially, the RD sharing and dissemination strategy can target the same public as the one that could be interested on an associated research article. But it can happen that the RD is of interdisciplinary value and can attract interest in a larger context than the strictly related to the initial research goal, as observed by [18]:

> *An investigator may be part of multiple, overlapping communities of interest, each of which may have different notions of what are data and different data practices. The boundaries of communities of interest are neither clear nor stable.*

So, it can be difficult to assess the target community of interest for a particular RD, but this also happens for articles or other publications or research outputs, and it seems to us that this has never been an obstacle for sharing, for example, a publication. Thus, [18]:

> *...the intended users may vary from researchers within a narrow specialty to the general public.*

Under what conditions? The sharing conditions are to be found in the license that goes with the RD. It can be, for example, a Creative Commons license (https://creativecommons.org/ (accessed on 11 November 2024)), or other kinds of licenses, that settle the attribution, re-use, mining...conditions [33]. For example, in France, the law of 2016 for a Digital Republic Act sets, in a *Décret*, the list of licenses that can be used for RS or RD release [39].

Why and to what effects? There may be different reasons to release some RD, from the contribution to build more solid, and easy to validate, scientific results, to simply react to the recommendations or requirements of the funder of a project, of the institutions supporting the research team, or those of a scientific journal, or as a response to Open Science issues [6]. The work [18] gives a thorough analysis on this subject. As documented there:

> *"The value of data lies in their use. Full and open access to scientific data should be adopted as the international norm for the exchange of scientific data derived from publicly funded research."*

Let us remark that, in the case of RD, the Data Management Plans (DMPs) can be a useful tool to deal with the answers of these conundrum challenges. As Harvard University tell us, a Data Management Plan (https://datamanagement.hms.harvard.edu/plan-design/data-management-plans (accessed on 11 November 2024)), is a formal, living document that outlines how data will be handled during and after a research project. In our vision, a DMP can be the right tool to help researchers to ask, possibly answer, and manage the associated information to the RD conundrum questions. DMPs are increasingly required by research funding organizations (see, for example, [1,2,8,38]) and there are more and more institutions providing tools and help to deal with these documents; see, for example the UK Digital Curation Centre (DCC) (https://www.dcc.ac.uk/ (accessed on 11 November 2024)), the French DMP OPIDoR (https://dmp.opidor.fr/ (accessed on 11 November 2024)), or the California Digital Library DMP Tool (https://dmptool.org/ (accessed on 11 November 2024)).

As a final observation, we consider that the above answers to the *conundrum challenges* have, among other consequences, the benefit of facilitating the improvement of RD sharing and dissemination best practices, enhancing thus trustworthiness and transparency in the research endeavor.

4. The Conundrum Questions for Research Software

In this Section we propose original (as far as we know) answers to the conundrum challenges for RS, by highlighting the similarities and differences that appear in the production, sharing and dissemination of RS and RD. Note that previous work regarding RS sharing [28,40,41] do raise other sets of questions that bring out a parallel and complementary study to these conundrum challenges. Questions include *what means research software?*, *which is the list of research software of the lab?*, *how to reference and cite the RS?*, *how to provide free/open acces?*...

Which RS might be shared? As we have seen in Section 2, to produce RS or RD may involve similar activities (documentation, corrections, version management, project management...), as well as others that are of different technical nature, like software development or data collection.

In the case of software, the Agile Principles for software development (http://agilemanifesto.org/principles.html, https://en.wikipedia.org/wiki/Agile_software_development (accessed on 11 November 2024)) promote, among others, the practice of *Deliver working software frequently*, principle that also appears in the Free/Open Source Software (FOSS) (https://en.wikipedia.org/wiki/Free_and_open-source_software (accessed on 11 November 2024)) [42] movements as *release early, release often* [43].

These development practices may prompt the RT to disseminate early versions of a software project, with the intention, for example, of communicate a new RS project, or to attract collaborators external to the initial RT.

But this early dissemination may have little interest in the case of RD. As mentioned in [18]:

> *If the rewards of the data deluge are to be reaped, then researchers who produce those data must share them, and do so in such a way that the data are interpretable and reusable by others.*

Thus, it seems that potential RD users do usually expect data objects just when they are mostly ready for reuse, while RS users or external collaborators can be interested in software that is yet unready for reuse, or far away from a *final* form, having in this way the possibility to participate in the development and to, maybe, influence the evolution of future versions, to fit their own interests.

On the other hand, an RS may have different development branches, with versions that can be easily shared with potential users, as well as other more experimental ones, where the RT explore different options for the software evolution. Moreover, an RS can also have different components, such as, for example, a computing kernel of interest for the Computer Algebra community, and a user graphical interface, more related to the Computer Graphics scholars. As remarked by one of the reviewers of this work, *creating branches for RD is an interesting idea, technically possible as in the case of code.* As far as we know, this is not an usual practice, but, yes, it is technically possible, showing, thus, the benefits of our comparison methodology, as in this case, we can learn from usual RS practices that can also be taken into consideration for RD.

Therefore, while in RD, the RT decision is about whether the research output is already in its final form, ready to be shared and reused, the decision for RS can be a much complex one: which components, which versions to share, to share early versions or not, experimental branches or not...

So, *Which RS might be shared?* corresponds to an involved RT decision: the team decides which RS might be shared, in which form and when, maybe following funder or institutional Open Science requirements.

By whom? By the RS development RT that takes the decision to share and disseminate it.

How? An important similarity between RS and RD is that, currently, they have not got a publication procedure as widely accepted as the one existing for articles published in scientific journals (see [20]), despite the fact that *data papers* and *software papers* [35,36] are becoming increasingly popular and there are more and more suggestions about where to publish these kinds of papers (see, for example, the Software Sustainability Institute list of Journals in which it is possible to publish software at https://www.software.ac.uk/top-tip/which-journals-should-i-publish-my-software (accessed on 11 November 2024) or the CIRAD (French Agricultural Research Centre for International Development) list of Scientific Journals and book editors at https://ou-publier.cirad.fr/revues? (accessed on 11 November 2024) that compiles, at the time of writing of this article, 171 entries and where it is possible to select the data papers (113 at the consultation date) or the Software papers (80)). Even in this case, the journals do not usually deal with the publication of RS or RD as such stand-alone objects. As a consequence, the RTs may be a bit disoriented concerning methods for these outputs' dissemination, and procedures, like the one proposed in [20], may be an effective help in order to face this problem.

Moreover, in [20], we have shown that it is possible to use the same dissemination protocol for RS and RD, but carefully recalling the steps in which it is important to take into consideration that data and software are objects of different nature, for example, in the legal (or ethical) aspects that may be involved, and that should be closely considered by the RTs *before the output dissemination*.

So, *How RS might be shared?* Our answer is as follows: by following a dissemination procedure like the one proposed in [20], where the RT should identify correctly the

RS set of files, as well as setting a title, identifying the list of persons in the producer team (that can be completed with their different roles), determining the important versions and associated dates, giving documentation, verifying the legal context of the RS, and giving a license to settle the sharing conditions, etc. Of course, this protocol can include steps towards the publication of a software paper.

Where? There are different places to disseminate an RS, including the web pages of the producer team, of the funded project, or in a repository like Zenodo (https://zenodo.org/ (accessed on 11 November 2024)). Note that the Zenodo repository can provide with a Digital Object Identifier (DOI) (https://en.wikipedia.org/wiki/Digital_object_identifier (accessed on 11 November 2024)) for the RS or RD in the case there is not already one, as well as with a *citation form*.

FOSS development communities [42] use collaborative platforms or *Forges* (https://en.wikipedia.org/wiki/Forge_(software) (accessed on 11 November 2024)), that is, web-based collaborative platforms that provide tools to manage different tasks of the software development and facilitate the collaboration (internal to the RT or with external collaborators). These forges are also a popular tool to share and disseminate the produced software, and some are well know among the RS developer teams.

Criteria helping to select code hosting facilities can be found, for example, on the popular website Wikipedia (https://en.wikipedia.org/wiki/Comparison_of_source-code-hosting_facilities (accessed on 11 November 2024)), or in private counsel companies like *Rewind* (https://rewind.com/blog/github-vs-bitbucket-vs-gitlab-comparison/ (accessed on 11 November 2024)). As shown in the Wikipedia statistics, GitHub (https://github.com/ (accessed on 11 November 2024)) is currently the most popular in terms of number of users and projects, but, as mentioned in the Wikipedia GitHub page (https://en.wikipedia.org/wiki/GitHub (accessed on 11 November 2024)) (see also the included references), it is currently owned by *Microsoft*, one of the richest and most powerful companies in the world.

It seems unclear to us what the limits that are *Microsoft* may have for exploring all these software projects that are at its disposal in the GitHub forge, considering its capacity to reuse all the collected information for AI produced software, a matter that is still subject to some very preliminary legal dispositions [44,45].

We think that the research community should be a bit more aware of the strategic value of the RS produced, usually within publicly funded projects, or at least have some kind of reflection on this delicate subject, demanding the existence of public RS and RD repositories [3], and avoiding the indiscriminate use of privately owned platforms. Institutions can use FOSS like GitLab (https://gitlab.com/gitlab-org/gitlab-foss (accessed on 11 November 2024)) to install their own facilities to help RS developers to manage their software projects on platforms internal to the institution (see, for example, the GitLab platform of the University Gustave Eiffel at https://gitlab.univ-eiffel.fr/ (accessed on 11 November 2024)), while reflecting on *digital sovereignty* [46–48].

So, *Where RS might be shared?* RS can be shared in repositories like Zenodo, in forges like GitHub, or in institutional repositories. A relevant example of a specific scientific area that provides tools for data and software sharing in Life Sciences is ELIXIR (https://elixir-europe.org/ (accessed on 11 November 2024)), an intergovernmental organization that brings together Life Sciences resources from across Europe.

With whom? As previously discussed for RD (and applicable to all kinds of research outputs: publications. . .), each act of scholar communication has its own target public, and initially, the RS sharing and dissemination strategy can be conceived to target just the public hat could be interested by a research article. But it might happen that the RS is of interdisciplinary value, and could rise interest in a larger context that the strictly related to the initial research goal. Therefore, the intended users may vary from researchers within a narrow specialty to researchers in other scientific disciplines or even to the general public.

Under what conditions? Software licenses are the needed tools to ensure the legal conditions for use, copy, modify, and redistribute software. Over these *legally protected actions* stand the usual *scientific actions*, as researchers, as part of their daily activities, do use, contribute to, write, share and disseminate, modify, include and re-distribute RS components.

In this context, we would like to refer to our previous work, devoted to advising RTs regarding RS sharing and dissemination practices with free/open source licenses [20,23,28], in order to ensure and clarify the context in which the legal and the scientific actions find no obstacle, remarking that the sharing and dissemination conditions are to be found in the license that goes with the RS.

Free/Open Source licenses can be found at the Fre Software Foundation (https://www.gnu.org/licenses/license-list.html (accessed on 11 November 2024)), the Open Source Initiative (https://opensource.org/license (accessed on 11 November 2024)) or at the Software Package Data Exchange (SPDX) (https://spdx.org/licenses/ (accessed on 11 November 2024)). See [42] for more information on licenses. Note that the popular Creative Commons licenses that can be used for RD sharing [33] are to be avoided in the case of RS, as recommended by Creative Commons (https://creativecommons.org/faq/#can-i-apply%20a-creative-commons-license-to-software (accessed on 11 November 2024)).

Why and to what effects? There maybe different reasons to release some RS, from the contribution to build more solid, and easy to validate science, to simply answer to the recommendations or requirements of the funder of a project, of the RT supporting institutions, or those of a scientific journal, or regarding Open Science issues [6,23]. Since Jon Claerbout raised the first concerns (as far as we know) about reproducibility issues [49,50]:

> *an article about computational science in a scientific publication is not the scholarship itself, it's merely scholarship advertisement. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.*

more and more voices are raised in order to improve the reproducibility conditions of published research results; see, for example, [17,51,52]. As a consequence, the research community is creating national organizations that are federated in the Global Reproducibility Network (https://www.ukrn.org/global-networks/ (accessed on 11 November 2024)) in order to *bring together different communities across the higher education research ecosystem of their nation, with the aim of improving rigour, transparency and reproducibility.*

Finally, let us recall the case of the Image Processing On Line (IPOL) journal, founded in 2009 [53] as a contribution to implement reproducible research in the Image Processing field, and then expanded to more general signal-processing algorithms, such as video or physiological signal processing, among others. They propose re-defining the concept of publication, which is no longer just the article, but the combination of the article, its associated source code, and any associated data needed to reproduce the results, that is, the research article and associated RD and RS as a whole. Our wish would be that more and more journals work in this way, providing simultaneous publication of these three important outputs, while remaining well aware of the technical difficulties (as well as of the necessary human and finance resources) arising in keeping software produced since 2009 in a working status.

Following the analysis of the parallelism between RS and RD conundrum answers, let us remark here too that, unlike the case of Data Management Plans (DMPs) mentioned at the end of Section 3, Software Management Plans for RS are still not widely required by research funders, although new recommendations are emerging in the landscape [31]. Yet, some models are available such as PRESOFT [54], or the Software Sustainability Institute (https://www.software.ac.uk/ (accessed on 11 November 2024)) Checklist for a

Software Management Plan [55]. More information and references on Research Software Management Plans (RSMPs) can be found at [56] or at the web page (https://igm.univ-mlv.fr/~teresa/presoft/ (accessed on 11 November 2024)) maintained by one of the authors of the present work. Notice that, while DMPs are usually centered over the research funded project in which maybe some data will be produced, RSMPs, like the PRESOFT model, are centered in the RS as a research output, maybe funded under several, different projects. As a consequence, RSMPs may be more useful than DMPs in order to handle the information required by the *conundrum questions* for a given RS, which has, as a consequence, the possibility of improving RS sharing and dissemination best practices for RTs, thus collaborating, as in the RD context, on augmenting trustworthiness and transparency in the research realm.

5. The Conundrum Challenges and Research Evaluation

As previously discussed, we consider that providing answers to these conundrum challenges has, among other consequences, the possibility of improving RD and RS sharing and dissemination best practices for RTs, thus enhancing integrity, correctness, rigor, reproducibility, reusability and transparency in the research endeavor.

Nevertheless, to reach this end, the need of new and adapted evaluation protocols and practices in this changing Open Science context is increasingly widely recognized. For example, the DORA San Francisco Declaration on Research Assessment states the following:

> *There is a pressing need to improve the ways in which the output of scientific research is evaluated by funding agencies, academic institutions, and other parties. [...]*
> *Outputs from scientific research are many and varied, including: research articles reporting new knowledge, data, reagents, and software; intellectual property; and highly trained young scientists. Funding agencies, institutions that employ scientists, and scientists themselves, all have a desire, and need, to assess the quality and impact of scientific outputs. It is thus imperative that scientific output is measured accurately and evaluated wisely.* ([21])

and high-level advisory groups for the European Commission have produced reports to draw up recommendations such as the following:

> *Funders, research institutions and other evaluators of researchers should actively develop/adjust evaluation practices and routines to give extra credit to individuals, groups and projects who integrate Open Science within their research practice. [...]*
> *Public research performing and funding organisations (RPOs/RFOs) should provide public and easily accessible information about the approaches and measures being used to evaluate researchers, research and research proposals.* ([22])

or recognizing that the evaluation of research is the keystone to boost the evolution of the Open Science policies and practices in one of its most important pillars, the scholar publication system, as follows:

> *The report views research evaluation as a keystone for scholarly communication, affecting all actors. Researchers, communities and all organisations, in particular funders, have the possibility of improving the current scholarly communication and publishing system: they should start by bringing changes to the research evaluation system.* ([7])

Therefore the evolution of the research evaluation system is an important enabler in order to improve the adoption of Open Science best practices, and it is currently a big issue and subject of debate at many levels and by different stakeholders, as shown in the vast number of references that, just considering the references in the present work, somehow address this question: [1,2,4–7,9–13,15,16,20,22,23,28], and this is only to mention a few of the large number of those addressing a topic that is becoming increasingly popular in the scientific literature. . .

Concerning the subject of RS and RD, we remark that, as analyzed in our previous work [20,23], it is in the interest of the research communities and institutions to adopt clear and transparent protocols for the evaluation of RD and RS as such stand-alone objects, such

as sharing and disseminating high-quality RS and RD outputs requests for time, work and hands willing to verify the quality of these artifacts, writing documentation, etc. Incentives are needed to motivate the RTs.

Moreover, as we have seen in the **How?** question of Section 4, one of the similarities between RS and RD is that they have not a widely accepted publication procedure, which hinders, among other effects, the fact of being correctly identified as research outputs with a *citation form*. Note that, in the case of research articles, this citation form is usually provided by the scientific journals in the publication stage of their dissemination, and note as well that the citation form is a step that has been included in our RD and RS dissemination procedures. To deposit RS and/or RD in a repository like Zenodo provides with a citation form for the output, but this is still a practice that should be more widely implemented and adopted. This is why we think that the formulation of such evaluation protocols should include, as well, as a first step, the establishment of best RD and RS citation practices.

In [20,23], we have proposed the CDUR protocol for RS and RD evaluation, a protocol which is flexible enough to be applied in different evaluation contexts as it is to be adapted by the evaluation committees to each particular evaluation situation. We include here a resumed version, but its detailed presentation and extended discussion can be found in [23]. There are four steps in the CDUR protocol, namely **(C)** Citation, **(D)** Dissemination, **(U)** Use, **(R)** Research, that are to be applied in the following chronological order:

(C) **Citation.** This step measures if the RS or RD are well identified as a research output, i.e., if there is a good citation form (including title, authors and/or producers, dates...), which could be extended up to require a good metadata set. We can look here for best citation practices applied to RS or RD coming from other teams. This is a legal related point where we ask for authors (if any) to be well identified, whats their affiliations are, and, for example, the associated % of their participation in software writing.

(D) **Dissemination.** In this step, we look to evaluate best dissemination practices in agreement with the scientific policy of the evaluation context. The dissemination of RS and RD needs a license to set the sharing conditions. For RD, there are maybe further legal issues to look at (personal data, *sui generis* database rights...). This is a policy point in which we look at Open Science requirements.

(U) **Use.** This point examines "software" or "data" aspects, in particular, the correct results that have been obtained, and we can also look if their reuse has been facilitated, the output quality, best software/data practices such as documentation, testing, installation or reuse protocols, up to read the code, launch the RS, use examples... This is the reproducibility point that looks at the validation of the scientific results obtained with the RS and/or the RD.

(R) **Research.** This point examines the research aspects associated with the RS and/or RD production: the quality of the scientific work, the proposed and coded algorithms and data structures, which are the related publications, the collaborations, the funded projects... This point measures the impact of the RD and/or RS related research.

Consequently, the RS and RD that are to be positively evaluated by the CDUR protocols are to be correctly identified as research outputs, as well as correctly disseminated; they are disseminated in such a way that their reuse is facilitated, as well as the reproducibility of the obtained research results, which are, in turn, to be evaluated positively.

Thus, the advocacy to adopt new evaluation methods is clear, but any change in the evaluation system requires researchers to adapt to these new rules and practices, that is, to adopt new ways to share and disseminate their RD and RS outputs. Figure 1 shows the interconnections that link both stages, as the RS and RD sharing and dissemination practices would need to adjust to the new evaluation rules that are in place and only suitable practices for RS and RD sharing and dissemination (maybe in a restricted context if necessary) would successfully and positively be evaluated.

More precisely, when evaluation rules evolve, the evaluated researchers do change their sharing and dissemination practices in order to adapt to the new rules. But the evaluation of research is not a system to be changed in a short period of time; this a

progressively and nonhomogeneous process that will still take time in order to adapt to new Open Science policies. For example, evaluation rules can increasingly evolve in order to enhance the reusability of research outputs like RS and RD, which is the object of the **(U) Use** step of the CDUR protocols [20,23], thus the loop.
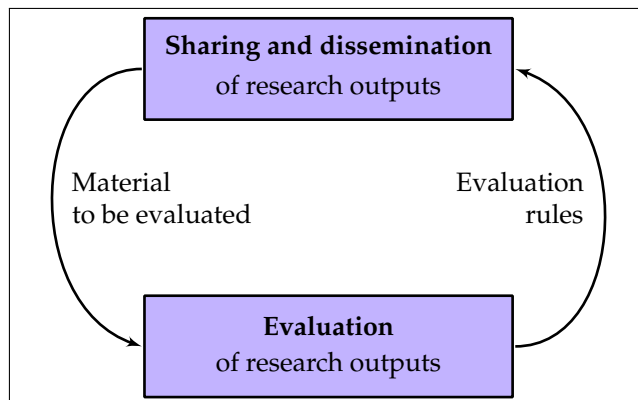

Figure 1. The research outputs dissemination/evaluation loop.

Yet, to change the research evaluation system requires a clear vision of what *"good research is"* and that is to be evaluated positively, which will be reflected in the (for example) Open Science policies. And the *new* evaluation system will provide the control tool that will be applied to verify that the policies are adequately uptaken, as remarked in [6]:

> *Funder and institution evaluations and research community evaluations are therefore a powerful tool to enhance effective Open Science evolutions [. . .] But,* as the cat biting its own tail, *the evaluation wave can only play fully its role if policies and laws are well into place.*

Indeed, another loop. . .

We can conclude as follows: to provide correct and clear answers to the conundrum questions will have, as a consequence, the improvement of RS and RD sharing and dissemination practices, which, in turn, will enhance trustworthiness, correctness, rigour, reproducibility, reusability and transparency in the research endeavor. But control mechanisms, maybe new in this dynamic Open Science context, are necessary in order to boost and improve these best practices. That is, in order to stimulate the adoption of new, best practices, it is necessary to put into place suitable evaluation mechanisms, which require for clear and precise answers to the corresponding conundrum questions.

6. Conclusions

To share and disseminate RS and/or RD, and maybe other research outputs, is a conundrum. It may require to understand a complex context, as has been studied, for example, in the Computer Science Gaspard-Monge Lab [28,40,41]. RS and RD sharing involves tasks of different nature and several levels of decision. It asks for time and resources (funding, persons). It is also necessary to decide if these resources would be maybe better allocated in new research projects or if they should be reserved for maintenance and bug corrections, or user support related to already disseminated RS and/or RD.

The aim of this work is to provide answers to the conundrum questions that appear in RS (and RD) sharing and dissemination practices, with the intention of helping researchers and RTs to better understand the issues that arise at this stage and how to deal with them.

In order to enunciate these RS conundrum answers, we have taken advantage of the comparison methodology that we have already applied successfully in previous works, that is, to compare the different issues that arise in the RS and RD production, sharing and dissemination, which helps us to identify and to decide whether we can propose similar or different answers to the conundrum challenges in both cases.

One of the identified differences is associated with the possibly expected *final status* of RD, rather different to the potential sharing of RS in the early stages of the software development, or related to the issues involved in deciding which development branches, which versions are to be shared.

In both cases, that is, in RS and RD, the role of the producer RT is major, but the roles and activities of the RT members may be very different, and that can be associated with the different activities (code writing, data collection and analysis...), but also related to the different legal contexts that may appear in RS and in RD. For example, to write software yields copyright issues, and software licenses should be used in order to ensure the use, copy, modification and redistribution of the RS, while RD may have a rather different legal (and *ethical*) context and its dissemination could ask for other kind of licenses or to undertake other kind of legal considerations. Other detected differences between RS and RD are connected to management plans, for example.

On the other hand, we have exposed similarities corresponding to the missing publication procedure, and the similar protocols that we can use for their dissemination and evaluation, which takes into account that they are objects of different nature, as highlighted, for example, in the **(U) Use** step of the CDUR protocols.

The evaluation step is an important enabler in order to improve the adoption of Open Science best practices and to increasingly render RS and RD visible, accessible and reusable. Providing correct and clear answers to the questions that the conundrum poses, as a consequence, enhances RS and RD sharing and dissemination best practices, which, in turn, requires adapted evaluation practices; finally, it fosters integrity, correctness, rigour, reproducibility, reusability and transparency in the research endeavor.

All these issues should be considered by researchers as well as by their Research Performing Institutions and Research Funders in order to adopt policies and practices that correspond to the researchers and science's best interests. As an example of currently ongoing work, we would like to mention the UNESCO Open Science Monitoring Initiative [57] and the recent Science Europe report [31], where we can observe consequences of our propositions.

We hope the present work will be of further help to this end.

References

1. Burgelman, J.-C.; Pascu, C.; Szkuta, K.; Von Schomberg, R.; Karalopoulos, A.; Repanas, K.; Schouppe, M. Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Front. Big Data* **2019**, *2*, 43. [CrossRef] [PubMed]
2. European Commission, Directorate-General for Research and Innovation. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Version 4.1. 2024. Available online: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf (accessed on 11 November 2024).
3. G6 Statement on Open Science. 2021. Available online: https://www.cnrs.fr/sites/default/files/download-file/G6%20statement%20on%20Open%20Science.pdf (accessed on 11 November 2024).

4.   G7 Science and Technology Ministers. G7 Science and Technology Ministers' Meeting in Sendai Communiqué. 2023. Available online: https://www8.cao.go.jp/cstp/kokusaiteki/g7_2023/230513_g7_communique.pdf (accessed on 11 November 2024).

5.   Annex1: G7 Open Science Working Group (OSWG). Available online: https://www8.cao.go.jp/cstp/kokusaiteki/g7_2023/annex1_os.pdf (accessed on 11 November 2024).

6.   Gomez-Diaz, T.; Recio, T. Towards an Open Science Definition as a Political and Legal Framework: On the Sharing and Dissemination of Research Outputs. Version 3. 2021. Available online: https://zenodo.org/doi/10.5281/zenodo.4577065 (accessed on 11 November 2024).

7.   European Commission: Directorate-General for Research and Innovation. Future of Scholarly Publishing and Scholarly Communication. Report of the Expert Group to the European Commission. 2019. Available online: https://op.europa.eu/publication-detail/-/publication/464477b3-2559-11e9-8d04-01aa75ed71a1 (accessed on 11 November 2024).

8.   Office of Science and Technology Policy (OSTP). Memorandum for the Heads of Executive Departments and Agencies for Ensuring Free, Immediate, and Equitable Access to Federally Funded Research. 2022. Available online: https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf (accessed on 11 November 2024).

9.   Rico-Castro, P.; Bonora, L. 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. 2023. Available online: https://data.europa.eu/doi/10.2777/162 (accessed on 11 November 2024).

10.   Swan, A. *Policy Guidelines for the Development and Promotion of Open Access*; UNESCO: Paris, France, 2012. Available online: http://unesdoc.unesco.org/images/0021/002158/215863e.pdf (accessed on 11 November 2024).

11.   UNESCO Recommendation on Open Science. 2021. Available online: https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en (accessed on 11 November 2024).

12.   Coalition for Advancing Research Assessment (COARA). Agreement on Reforming Research Assessment. 2022. Available online: https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf (accessed on 11 November 2024).

13.   European Commission—European Research Executive Agency (REA). Report on Research Assessment. 2024. Available online: https://ora.ox.ac.uk/objects/uuid:36928da0-bfba-4c61-864e-de30e216f76d/files/szs25xb180 (accessed on 11 November 2024).

14.   Fillon, A.; Maniadis, Z.; Méndez, E.; Sánchez-Núñez, P. Should we be wary of the role of scientific publishers in Open Science? [version 1; peer review: 1 approved with reservations, 1 not approved]. *Open Res. Eur.* **2024**, *4*, 127. [CrossRef]

15.   Foro Latinoamericano sobre Evaluación Científica (FOLEC-CLACSO). Una Nueva Evaluación Académica y científica para una Ciencia con Relevancia Social en América Latina y el Caribe. 2022. Available online: https://www.clacso.org/una-nueva-evaluacion-academica-y-cientifica-para-una-ciencia-con-relevancia-social-en-america-latina-y-el-caribe/ (accessed on 11 November 2024).

16.   Méndez, E.; Sánchez-Núñez, P. Navigating the Future and Overcoming Challenges to Unlock Open Science. In *Ethics and Responsible Research and Innovation in Practice*; González-Esteban, E., Feenstra, R.A., Camarinha-Matos, L.M., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2023; p. 13875. [CrossRef]

17.   Moussaoui, J.R.; McCarley, J.S.; Soicher, R.N.; McCarley, J.S. Ethical Considerations of Open Science. In *Beyond the Code: Integrating Ethics into the Undergraduate Psychology Curriculum*; Pantesco, B., Beising, R., Thakkar, V., Naufel, K., Woolf, L., Eds.; Society for the Teaching of Psychology: Washington, DC, USA, 2024. [CrossRef]

18.   Borgman, C.L. The conundrum of sharing research data. *J. Am. Soc. Inf. Sci. Technol.* **2012**, *63*, 1059–1078. [CrossRef]

19.   Gomez-Diaz, T.; Recio, T. 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. [CrossRef]

20.   Gomez-Diaz, T.; Recio, T. 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. [CrossRef]

21.   DORA—San Francisco Declaration on Research Assessment. 2012. Available online: https://sfdora.org/read/ (accessed on 11 November 2024).

22.   European Commission, Directorate-General for Research and Innovation. Open Science Policy Platform Recommendations. 2018. Available online: https://op.europa.eu/publication-detail/-/publication/5b05b687-907e-11e8-8bc1-01aa75ed71a1 (accessed on 11 November 2024).

23.   Gomez-Diaz, T.; Recio, T. On the evaluation of research software: The CDUR procedure [version 2; peer review: 2 approved]. *F1000Research* **2019**, *8*, 1353. [CrossRef] [PubMed]

24.   Oancea, A. Research governance and the future(s) of research assessment. *Palgrave Commun.* **2019**, *5*, 27. [CrossRef]

25.   Borgman, C.L. *Big Data, Little Data, No Data: Scholarship in the Networked World*; MIT Press: Cambridge, MA, USA, 2017.

26.   Gualandi, B.; Pareschi, L.; Peroni, S. What do we mean by "data"? A proposed classification of data types in the arts and humanities. *J. Doc.* **2022**, *79*, 51–71. [CrossRef]

27.   Huvila, I.; Sinnamon, L.S. When data sharing is an answer and when (often) it is not: Acknowledging data-driven, non-data, and data-decentered cultures. *J. Assoc. Inf. Sci. Technol.* **2024**, 1–16. [CrossRef]

28.   Gomez-Diaz, T. Article vs. Logiciel: Questions juridiques et de politique scientifique dans la production de logiciels. *1024—Bulletin de la Société Informatique de France* **2015**, *5*, 119–140. [CrossRef]

29.   Gomez-Diaz, T.; Recio, T. Articles, software, data: An Open Science ethological study. *Maple Trans.* **2023**, *3*, 17132. [CrossRef]

30.   Kelly, D. An Analysis of Process Characteristics for Developing Scientific Software. *J. Organ. End User Comput.* **2011**, *23*, 64–79. [CrossRef]

31. Science Europe. Developing and Aligning Policies on Research Software: Recommendations for RFOs and RPOs. 2024. Available online: https://scienceeurope.org/our-resources/recommendations-research-software/ (accessed on 11 November 2024).

32. de Cock Buning, M.; van Dinther, B.; Jeppersen de Boer, C.G.; Ringnalda, A. The Legal Status of Research Data in the Knowledge Exchange Partner Countries. *Knowledge Exchange Report*. 2011. Available online: https://repository.jisc.ac.uk/6280/ (accessed on 11 November 2024).

33. Labastida, I.; Margoni, T. Licensing FAIR Data for Reuse. *Data Intell.* **2020**, *2*, 199–207. [CrossRef]

34. Organisation for Economic Co-operation and Development (OECD). *OECD Principles and Guidelines for Access to Research Data from Public Funding*; OECD Publishing: Paris, France, 2007. [CrossRef]

35. Dedieu, L. Revues Publiant des Data Papers et des Software Papers. CIRAD. 2024. Available online: https://collaboratif.cirad.fr/alfresco/s/d/workspace/SpacesStore/91230085-7c0c-4283-b2c4-25bd792121a9 (accessed on 11 November 2024).

36. Schöpfel, J.; Farace, D.; Prost, H.; Zane, A. Data papers as a new form of knowledge organization in the field of research data. *Knowl. Org.* **2019**, *46*, 622–638. [CrossRef]

37. Wilkinson, M.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; Bonino da Silva Santos, L.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef] [PubMed]

38. Science Europe. Practical Guide to the International Alignment of Research Data Management (Extended Edition). 2021. Available online: https://www.scienceeurope.org/media/4brkxxe5/se_rdm_practical_guide_extended_final.pdf (accessed on 11 November 2024).

39. Journal Officiel de la République française, Lois et décrets: Décret n. 2017-638 du 27 avril 2017 Relatif aux Licences de Réutilisation à Titre Gratuit des Informations Publiques et aux Modalités de leur Homologation. Available online: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000034502557 (accessed on 11 November 2024).

40. Gomez-Diaz, T. Free/Open Source Research Software Production at the Gaspard-Monge Computer Science Laboratory: Lessons Learnt. In the Open Research Tools and Technologies Devroom, FOSDEM. 2021. Available online: https://archive.fosdem.org/2021/schedule/event/open_research_gaspard_monge/ (accessed on 11 November 2024).

41. Gomez-Diaz, T. *Sur la Production de Logiciels Libres au Laboratoire d'Informatique Gaspard-Monge (LIGM): Ce Que Nous Avons Appris*; Atelier Blue Hats, Etalab; Direction Interministérielle du Numérique: Paris, France, 2021. [CrossRef]

42. Fogel, K. Producing Open Source Software. How to Run a Successful Free Software Project. 2005–2022. Available online: https://producingoss.com/ (accessed on 11 November 2024).

43. Raymond, E. The cathedral and the bazaar—Musings on Linux and Open Source by an accidental revolutionary. *Knowl. Technol. Policy* **1999**, *12*, 23–49. Available online: https://link.springer.com/content/pdf/10.1007/s12130-999-1026-0.pdf (accessed on 11 November 2024). [CrossRef]

44. European Commission. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying down Harmonised Rules on Artificial Intelligence. 2024. Available online: http://data.europa.eu/eli/reg/2024/1689/oj (accessed on 11 November 2024).

45. Kamocki, P.; Bond, T.; Lindén, K.; Margoni, T.; Kelli, A.; Puksas, A. Mind the Ownership Gap? Copyright in AI-Generated Language Data. In CLARIN Annual Conference. 2024. Available online: https://www.ecp.ep.liu.se/index.php/clarin/article/download/1024/931/1053 (accessed on 11 November 2024).

46. European Commission. Global Gateway: EU, Latin America and Caribbean Partners Launch in Colombia the EU-LAC Digital Alliance. 2023. Available online: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_23_1598/IP_23_1598_EN.pdf (accessed on 11 November 2024).

47. Madiega, T.A. Digital Sovereignty for Europe. 2020. Available online: https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf (accessed on 11 November 2024).

48. Velasco Pufleau, M. La Alianza Digital Como Oportunidad para Profundizar las Relaciones Entre la Unión Europea y América Latina y el Caribe. In *Inteligencia Artificial y Diplomacia: Las Relaciones Internacionales en la era de las Tecnologías Disruptivas*; Sistema Económico Latinoamericano y del Caribe (SELA): Caracas, Venezuela, 2024. Available online: https://orbilu.uni.lu/bitstream/10993/62003/1/IA%20y%20Diplomacia_SELA_0924.pdf (accessed on 11 November 2024).

49. Buckheit, J.B.; Donoho, D.L. *Wavelab and Reproducible Research*; Springer: New York, NY, USA, 1995. Available online: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=003e123838b20cd1cfc716a1313bfd0a164a2d37 (accessed on 11 November 2024).

50. Donoho, D.L.; Maleki, A.; Rahman, I.U.; Shahram, M.; Stodden, V. Reproducible Research in Computational Harmonic Analysis. *IEEE Comput. Sci. Eng.* **2009**, *11*, 8–18. [CrossRef]

51. Baker, M. 1500 scientists lift the lid on reproducibility. *Nature* **2016**, *533*, 452–454. [CrossRef]

52. Ioannidis, J.P.A. Why Most Published Research Findings Are False. *PLoS Med.* **2005**, *2*, e124. [CrossRef]

53. Nicolaï, A.; Bammey, Q.; Gardella, M.; Nikoukhah, T.; Boulant, O.; Bargiotas, I.; Monzón, N.; Truong, C.; Kerautret, B.; Monasse, P.; et al. The approach to reproducible research of the Image Processing on Line (IPOL) journal. *Informatio* **2022**, *27*, 76–112. [CrossRef]

54. Gomez-Diaz, T.; Romier, G. Research Software Management Plan Template V3.2. *Projet PRESOFT*, Bilingual Document (FR/EN). *Zenodo Preprint*. 2018. Available online: https://zenodo.org/doi/10.5281/zenodo.1405613 (accessed on 11 November 2024).

55. The Software Sustainability Institute (SSI). Checklist for a Software Management Plan. *Zenodo Preprint*. 2018. Available online: https://zenodo.org/doi/10.5281/zenodo.1422656 (accessed on 11 November 2024).

56. Grossmann, Y.V.; Lanza, G.; Biernacka, K.; Hasler, T.; Helbig, K. Software Management Plans—Current Concepts, Tools, and Application. *Data Sci. J.* **2024**, *23*, 43. [CrossRef]

57. UNESCO. Open Science Monitoring Initiative—Principles of Open Science Monitoring. 2024. Available online: https://doi.org/10.52949/49 (accessed on 11 November 2024).