

Research Software evaluation: CDUR - new evaluation protocols

Teresa Gomez-Diaz [*], Tomas Recio [**]

[*] CNRS - Laboratoire d'informatique Gaspard-Monge (Est of Paris)

[**] Universidad Antonio de Nebrija (Madrid)

This work is licensed under
Creative Commons Attribution 4.0 International License

<http://creativecommons.org/licenses/by/4.0/>

Open Science Days 2021, 9-10 october 2021
Max Planck Digital Library (MPDL), Berlin



Context

Local context in my lab LIGM:

- LIGM's HCERES evaluation (2013-18), february 2019
- Important research software production: **66, 50** free/open source
- How to evaluate this research production ?

Global context: Free/OSS, Open Science, evolutions in evaluation

- Free Software (R. Stallman, FSF, 1985) **and** OSS (OSI, 1988)
- Budapest Open Access Initiative (BOAI, 2002), *open access definition*
- 2 Reports EC Expert Groups - evolutions for Open Science (adoption & practice)
 - ▶ [9] (2017) Cabello Valdes C, Rentier B, et al.: Evaluation of research careers fully acknowledging Open Science practices. ...
 - ▶ [11] (2019) Guédon JC, Jubb M, et al.: Future of Scholarly Publishing and Scholarly Communication. Report of the Expert Group to the EC.

[9, 11] **Software (use & prod.) is now included in Open Science best practices.**

[11] **The conclusion is actually simple: the evaluation of research is the keystone.**

Goal of this talk (1/2)

Software production is not correctly assessed in research evaluation.

It is necessary to change evaluation practices,
taking into consideration Open Science evolutions.

The **goal** of this talk is to present this work and motivate the lecture of:

Gomez-Diaz T. and Recio T.,

On the evaluation of research software: the CDUR procedure

[version 2; peer review: 2 approved] 26 Nov 2019 (V1, 5 Aug 2019).

F1000Research 2019, 8:1353, <https://doi.org/10.12688/f1000research.19994.2>

+1 year of writing, double expertise, 80 references, 85 footnotes

Choice of F1000Research:

Section *Science Policy Research* (renamed to *Research on Research, Policy & Culture*)
and open peer review procedure.

Article included in the *Mathematical, Physical, and Computational Sciences* collection.

Note1: **RS** means *Research Software*.

Note2: *[nb]* refers to the reference number *nb* in version 2.

Goal of this talk (2/2)

The article proposes the **CDUR** protocol to take into account RS production in research evaluation.

The protocol **CDUR** is flexible in order to be adapted to all different research evaluation contexts, all along the research life.

There are four steps:

- (C) Citation** clear RS identification as research production, title, authors, version, dates...
- (D) Dissemination** dissemination best practices
any dissemination has its own goals and target public
- (U) Use** evaluation of software aspects of **RS**
correct results, facilitated reuse...
- (R) Research** evaluation of research aspects of **RS**
scientific work: algorithms, related publications, impact...

Plan

1 Context

2 Goals

3 Concepts

- Research Software (RS)
- RS author
- RS publications
- Reference and citation

4 On research evaluation

- Research evaluation contexts
- Two evaluation methods

5 Protocol(s) CDUR for RS evaluation

6 Conclusions

Concept: Research Software (RS) (1/3)

- [2] (1994) Partha D, David PA: Toward a new economics of science

there may be important positive spillovers across projects in the form of "learning effects" [...] including the development of computer software for performing data processing, storage, retrieval and network transmission.

- [16] (2011) Kelly D: An Analysis of Process Characteristics for Dev. Scientific Soft.

Scientific software is defined by (1) it is developed to answer a scientific question ; (2) it relies on the close involvement of an scientific expert ; and (3) it provides data to be examined by the person who will answer that question ...

- [12] (2011) TGD: Article vs. Logiciel: questions juridiques et de politique scientifique...
(2009) TGD: Guide laboratoire pour recenser ses développements logiciels (PLUME)

logiciel du laboratoire tout programme utile pour faire avancer la recherche, qui a été produit avec la participation d'un membre du laboratoire. Il arrive souvent que des publications de recherche soient associées.

- [18] (2012) Sletholt MT, Hannay JE, et al.: What Do We Know about Scientific Software Development's Agile Practices ?

software developed by scientists for scientists

- [19] (2016) Hettrick S: Research Software Sustainability

Research software is developed within academia and used for the purposes of research: to generate, process and analyse results. This includes a broad range of software, programs written by researchers for their own use.

- [10] (2018) NASA Committee: Open Source Software Policy Options for NASA Earth and Space Sciences

Research software – that is, the software that researchers develop to aid their science...

Concept: Research Software (RS) (2/3)

- [16] (2011) Kelly D: An Analysis of Process Characteristics for Dev. Scientific Soft.
 - ▶ **exclusion** of what can be included in other definitions:
[...] control software whose main functioning involves the interaction with other software and hardware; user interface software [...]; and any generalized tool that scientists may use in support of developing and executing their software, but does not of itself answer a scientific question.
 - ▶ the importance of the **correctness**:
If the software gives the wrong answer, all other qualities become irrelevant
- [12] (2011) TGD: Article vs. Logiciel: questions juridiques et de politique scientifique...
definitions don't care about **software status**:
"project", "ended", disseminated, quality, scope, size, documented, maintained, team's internal use for an article, currently used in several labs...

The proposed definition of Research Software (section 2.1):

Research software (RS) is a well identified set of code that has been written by a 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 RS encloses a set of files containing the source code and the compiled code. It can also include other elements as the documentation, specifications, use cases...

Concept: Research Software (RS) (3/3)

Conclusion of RS definition:

- what is done: code, software as a well identified set of files,
- who does it: author(s), but also contributors or scientific expert(s),
- to make what: research, science, that is, associated articles,
- **important:** quality and correctness of the produced scientific results.

Moreover:

- **software/computer program** is a legal concept: Directive 2009/24/EC of the European Parliament & Council 23/04/2009 <https://eur-lex.europa.eu/eli/dir/2009/24/oj>
- [12] (2011) TGD: Article vs. Logiciel: questions juridiques et de politique scientifique... studies software as a legal object and as research production
- TGD and Recio T., Open comments on the Task Force SIRS report: Scholarly, Infrastructures for RS, RIO 2021, <https://doi.org/10.3897/rio.7.e63872>
- Free/Open source Research Software production at the Gaspard-Monge Computer Science laboratory, FOSDEM'21, https://fosdem.org/2021/schedule/event/open_research_gaspard_monge/
- Other works on research software, FAIR, and related concepts, see for example:
 - ▶ RDA FAIR 4 Research Software (FAIR4RS) working group, <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>
 - ▶ Research Software definition - Subgroup 3, https://docs.google.com/document/d/1PvYiYJxd7-vrmTusTvS8fYp47Wu6v-c_XMu-LjIBKio/
 - ▶ A Fresh Look at FAIR for Research Software, <https://arxiv.org/abs/2101.10883>

Concept: RS author

What means RS author ?

[12] (2011) TGD: Article vs. Logiciel: questions juridiques et de politique scientifique...

- legal concept: the author writes the code
- scientific concept: expert contributions, maybe no writing code
without the scientific expert, the RS will not exist
- maybe other contributions:
documentation, bug fixing, test, maintenance, translations...

Definition of a RS author:

- in the article we select three rôles (limits can be fuzzy):
 - ▶ (i) RS leader,
 - ▶ (ii) main or important contributor (code writing),
 - ▶ (iii) minor contributor (code writing or other contribution).

Persons with no code writing can be assigned with some participation percentage of code writing by the team.

Concept: RS publications

The article studies current situation regarding RS papers with *software peer review*.

- Journal of Open Research Software (JORS)
- The Journal of Open Source Software (JOSS)
- Research Ideas and Outcomes (RIO)
- Software Impacts
- SoftwareX
- also: (2010) Image Processing On Line Journal (IPOL)
- also: Peer Community in... could be well adapted for RS review

See N. Chue Hong list at Software Sustainability Institute (SSI)

<https://www.software.ac.uk/resources/guides/which-journals-should-i-publish-my-software>

In France, Projet PLUME (2006-2013), <https://projet-plume.org/>:

- publication of RELIER “RS description cards”, with links to articles, stats: 358 RS in French, where 116 also presented in English
- publication of “validated cards in the sense of PLUME”, stats: 96 RS out of 406
- theme classification, keywords, search interface

Concepts: reference and citation

[39] (2013) Pontille D, Torny D: La manufacture de l'évaluation scientifique ...

[...] 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".

A reference **sets** title, author(s), date, and identifies RS as a scientific object.

The article considers three different types of reference:

- the one related to the RS paper (with *software peer review*),
- the one related to a classic research article describing the RS,
- a "reference": author(s), RS title, short description, version, date, url.

Remarks:

- There can be several references associated to a RS.
- There are more complete identifications: metadata, CITATION files...
- Software Citation Group, Software Citation Implementation Working Group...
Recognizing the value of software: a software citation guide (2021),

<https://f1000research.com/articles/9-1257>

Research evaluation contexts

Research evaluation contexts, all along the research life:

- PhD, recruitment, career evolution,
- articles, publications, *peer review*
- participation to conferences, workshops, (selection)
- project funding: call answer, stages, end of the project
- set collaboration networks, usually in an international context

In general, the first evaluation coming into play: **self evaluation**.

Any dissemination has its own goals and target public:

- this result will be in a preprint or a journal article ?
- this project will be funded ?
- these researchers will collaborate in this project or publication ?
- these decisions evolve in time, following research evolution, but also when facing a new evaluation, for example I need more articles...

Two evaluation methods

Roughly speaking, there are two evaluation methods:

[8] (2016) Mårtensson P, Fors U, et al.: Evaluating research: A multidisciplinary approach to assessing research practice and quality (62 references)

- the quality method: what criteria ?
- indicators, metrics: which ones ?
 - ▶ should be used with careful attention [11, 39, 51, 52, 53, 54]
- social factor

[55] (1999) Martin U: Computers, Reasoning and Mathematical Practice
*[...] the community's "social knowledge":
the methods of checking the proof are social rather than formal.*

See the "EC Expert reports" [9, 11]:

- plenty of recommendations, take into account RS
- [9] Open Science Career Assessment Matrix (OS-CAM)
- [11] how to establish evaluation committees

Protocol(s) CDUR for RS evaluation

Designed to help evaluated researchers, evaluation committees, decision makers.

- (C) Citation** measure if RS is well identified as a research output:
good citation form, but also metadata, best citation practices...
legal point: authors, affiliations, participation %
- (D) Dissemination** best dissemination practices, in agreement with
the scientific policy of the evaluation context
[14] (2014) TGD: Free software, Open source software, licenses...
policy point: Open Science, **legal point:** licenses
- (U) Use** “software aspects” **of RS:** correct results, facilitate reuse, good softw.
practices: doc, test, install, up to read the code, launch RS...
point reproducibility: validation of scientific results
- (R) Research** “research aspects”: quality of the scientific work, proposed and
coded algorithms & data structures, related publications,
collaborations...
point research: impact

Flexibility of application: each decision maker or evaluation committee
sets its own CDUR protocol adapted to the evaluation context and goals.

Conclusions (1/2)

A call for action to foster a debate on RS evaluation protocols, CDUR will be adopted ?
For success in Open Science and RS best practices' & adoption of evolutions...

it is necessary to change research evaluation methods.

Expected consequences:

In agreement with [11] (2019) Guédon JC, Jubb M, et al.: *world brain vision*

- maximize [RS] accessibility and usability,
- support and expand range of contributions (equity, diversity, inclusivity criteria),
- support community building, and
- promote high-quality research with heightened integrity.

But also increased transparency in RS evaluation methods.

For success in Open Science & adoption of its evolutions...

it is necessary to understand better what Open Science is:

(Gomez-Diaz & Recio, 2020-21) **Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs**, POLIS N. 19, 2020, <http://uet.edu.al/polis/images/1.pdf>
V3 dated February 28th 2021 at: <https://zenodo.org/record/4577066>

Clearly, a policy is only as good as its enforcement.

[4] (2016) Howison J, Bullard J: Software in the scientific literature (p.15)

With our sincere thanks to...

- Martin Boosen and Open Science days organization

and also to

- LIGM directors
- F1000Research
- Support LIGM
- Referees: 1 confidential, 2 public (32 proposed in Summer)
- RS producers at LIGM
- You, listening here...

- *This work is partially funded by the CNRS-IEA PREOSI project (2021-22).*