

L'évaluation des logiciels de la recherche : CDUR - nouveaux protocoles d'évaluation

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

[*] CNRS - Laboratoire d'informatique Gaspard-Monge

[**] Universidad Antonio de Nebrija (Madrid)

Cette œuvre est mise à disposition selon les termes de la licence
Creative Commons Attribution 4.0 International

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

Journées Mathrice 2020 et 2022

Institut de Mathématiques de Toulouse, 22-24 mars 2022



Plan

1 Contexte

2 Objectif de cette présentation

- Protocole CDUR

3 Les concepts

- Logiciel de la recherche (RS)
- Auteur d'un logiciel de la recherche
- Publication d'un logiciel de la recherche (RS papers)
- Référence et citation

4 Sur l'évaluation

- Contextes d'évaluation
- Deux méthodes d'évaluation

5 Protocole(s) CDUR

6 Sur la Science Ouverte

7 Conclusions

Préambule :

Évolution du cadre juridique en décembre 2021

Cette présentation s'inscrit désormais dans le cadre du récent Décret N. 2021-1572 du 3 décembre 2021 **relatif au respect des exigences de l'intégrité scientifique** par **les établissements publics contribuant** au service public de la recherche et les fondations reconnues d'utilité publique ayant pour activité principale **la recherche publique**.

Décrète dans son Article 2, item 3 :

*Les établissements publics et fondations reconnues d'utilité publique mentionnés au troisième alinéa de l'article L. 211-2 du code de la recherche : **Promeuvent la diffusion des publications en accès ouvert et la mise à disposition des méthodes et protocoles, des données et des codes sources associés aux résultats de la recherche afin d'en garantir la traçabilité et la reproductibilité.***

<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044411360>

Contexte (1/2) : Logiciels LIGM

(Mars 2020)

- Évaluation HCERES (2013-18) au LIGM (février 2019)
- Mission Logiciels LIGM (juin 2006) :
*étudier la mise en place de services (suivi de versions, publication, etc.)
associés au développement logiciel au sein du laboratoire,
avec pour objectif de favoriser leur visibilité*
- Production importante de logiciels : **66** listés pour 2013-18
- Nombre important de logiciels diffusés sous licence libre : **50/66**

Question : comment sont pris en compte les logiciels de la recherche dans l'évaluation : d'un laboratoire, d'un chercheur, d'un candidat... ?

Contexte (2/2) : Évaluation de la science et la Science Ouverte

(Mars 2020) Pourquoi les évolutions sont si lentes pour la Science Ouverte ?

- Free Software (R. Stallman, FSF, 1985) **et** OSS (OSI, 1988)
- Déclaration de Budapest (BOAI, 2002), définition d'*accès libre*
- 2 Conférences en France avec présentations J.C. Guédon :
 - ▶ Journées Open Access et évaluation de la recherche (Toulouse, 10/2016) <https://openeval2016.sciencesconf.org/>
 - ▶ Carrefour de l'IST (Nancy, 03/2017) <https://carist.sciencesconf.org/>
- 2 Rapports de groupes d'experts pour l'EC :
 - ▶ [9] (2017) Cabello Valdes C, Rentier B, et al. : Evaluation of research careers fully acknowledging Open Science practices. Rewards, incentives and/or recognition for researchers practicing Open Science.
 - ▶ [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 (utilisation & production) fait partie de la Science Ouverte.**

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

Objectifs de la présentation (1/2)

Après l'étude de la situation, on constate que la production de logiciels n'est pas correctement prise en compte dans l'évaluation de la recherche, pourtant cette production peut prendre une place importante.

Il est nécessaire de **faire évoluer les pratiques d'évaluation**, et cela dans le cadre des évolutions de la Science Ouverte.

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>

Revue F1000Research, avec *open peer review* + section *Science Policy Research* (renommée *Research on Research, Policy & Culture*), inclus dans *Mathematical, Physical, and Computational Sciences collection*

+1 an d'écriture, double expertise, 80 références, 85 notes pied de page

Note1 : RS veut dire *research software* o logiciel de la recherche

Note2 : [nb] indique la référence dans l'article (en V2) avec ce *nb*.

Objectifs de la présentation (2/2)

L'article propose le protocole **CDUR** pour prendre en compte la production des logiciels de la recherche dans l'évaluation.
Il est flexible pour s'adapter aux différents contextes d'évaluation, tout au long de la vie scientifique.

Il y en a quatre étapes :

- (C) Citation** mesure la bonne identification du RS en tant que production de la recherche, implique nom, auteurs, version, dates... informations bien établies
- (D) Dissemination** toute diffusion a ses objectifs et un public cible, ce point regarde les bonnes pratiques suivies
- (U) Use** point dédié aux aspects purement "logiciel" **du RS** : produit des résultats corrects, facilite la réutilisation
- (R) Research** point dédié aux aspects purement "recherche" : qualité du travail scientifique, publications associées au RS, impact

Plan

1 Contexte

2 Objectif de cette présentation

3 Les concepts

- Logiciel de la recherche (RS)
- Auteur d'un logiciel de la recherche
- Publication d'un logiciel de la recherche (RS papers)
- Référence et citation

4 Sur l'évaluation

5 Protocole(s) CDUR

6 Sur la Science Ouverte

7 Conclusions

Concept : logiciel de la recherche (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 : logiciel de la recherche (RS) (2/3)

- [16] (2011) Kelly D : An Analysis of Process Characteristics for Dev. Scientific Soft.
 - ▶ **exclut** ce qui peut être inclus dans d'autres définitions :
[...] 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.
 - ▶ l'importance de l'**exactitude** :
If the software gives the wrong answer, all other qualities become irrelevant
- [12] (2011) TGD : Article vs. Logiciel : questions juridiques et de politique scientifique ...
les définitions ne tiennent pas compte de l'**état du RS** :
*"en projet", "fini", diffusé, qualité, portée, taille, documenté, maintenu,
utilisé uniquement par une équipe pour réaliser une publication, ou utilisé dans plusieurs laboratoires*

La définition de 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 : logiciel de la recherche (RS) (3/3)

Conclusion de la définition de RS :

- ce qui est fait : code, c'est-à-dire un ensemble de fichiers bien identifié,
- qui le fait : auteur(s), mais aussi contributeurs et/ou experts scientifiques,
- pourquoi faire : recherche, science, ie. article(s) associés,
- **important** : qualité et exactitude des résultats scientifiques produits.

Quelques notes :

- **Software/computer program** est un concept juridique : 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... étudie les logiciels comme objets juridiques et comme production scientifique
- 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>
- Sur la production de logiciels libres au LIGM, Atelier BlueHats, avril 2021 <https://communs.numerique.gouv.fr/ateliers/motivation-contributions-open-source/>
- D'autres références, par exemple :
 - RDA FAIR 4 Research Software (FAIR4RS) working group, <https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg>
 - Gruenpeter, M. et al. (2021) Defining Research Software : a controversial discussion <https://zenodo.org/record/5504016>
 - Katz, D. et al. (2021), A Fresh Look at FAIR for Research Software, Patterns, <https://doi.org/10.1016/j.patter.2021.100222>

Concept : auteur d'un logiciel de la recherche

Que veut dire auteur d'un RS ?

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

- concept légal : écriture du code

l'auteur écrit le code

- concept scientifique : apport d'expertise

sans l'expert scientifique le code n'existerait pas

- il peut y avoir d'autres contributions

documentation, correction de bugs, test, maintenance...

Définition d'auteur d'un RS :

- selection de 3 rôles dans l'article (limites peuvent être floues) :
 - ▶ (i) responsable du RS,
 - ▶ (ii) contributeur principal ou important (écriture du code),
 - ▶ (iii) contributeur mineur (écriture du code ou autre participation).

Les personnes qui n'écrivent pas du code peuvent avoir un % de participation décidé par l'équipe.

Concept : publication d'un RS (RS papers)

L'article étudie certaines publications dédiées aux RS, présentent les logiciels, avec des procédures de *software peer review* :

- Journal of Open Research Software (JORS)
- The Journal of Open Source Software (JOSS)
- Research Ideas and Outcomes (RIO)
- Software Impacts
- SoftwareX
- et aussi : (2010) Image Processing On Line Journal (IPOL)
- et aussi : (2011) swMATH publications de zbMATH qui mentionnent scientific softw.

Voir la liste de N. Chue Hong, Software Sustainability Institute (SSI)

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

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

- publication de fiches RELIER (description courte de RS), avec lien vers les publications associées, 358 RS (fr), 116 RS (en)
- publication de fiches de *logiciel validé par la communauté de l'ESR au sens de PLUME* et mis en production sur au moins 3 sites, stats : 96 RS sur 406
- pas de *software peer review*
- classification thématique, mots clés, interface de recherche

Concepts : référence et citation

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

[...] 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"

Une référence **fixe** nom, auteur, date, et identifie le RS tant qu'object scientifique.

L'article retient 3 types de référence :

- RS paper, avec *software peer review*,
- un article de recherche "classique" qui décrit le logiciel,
- une "référence" : auteurs(s), nom du RS, description courte, version, date, url.

À noter que :

- Il peut avoir plusieurs références associées à un RS.
- Identifications plus complètes : metadonnées, CITATION files [48] ...
- Software Citation Group, Software Citation Implementation Working Group ...

Plan

1 Contexte

2 Objectif de cette présentation

3 Les concepts

4 **Sur l'évaluation**

- Contextes d'évaluation
- Deux méthodes d'évaluation

5 Protocole(s) CDUR

6 Sur la Science Ouverte

7 Conclusions

Contextes d'évaluation

Les différents contextes d'évaluation :

- thèse, recrutement, évolution de carrière
- articles, publications, *peer review*
- participation à conférences, workshops, (sélection)
- financement de projets : demande, étapes, fin de projet
- établir des réseaux de collaboration, contexte international
- laboratoire

En général, la première évaluation qui entre en jeu est l'**autoévaluation**.

Toute diffusion à des objectifs, un public cible :

- ce résultat sera dans un preprint ou dans une publication ?
- ce projet sera-t-il financé ?
- ces chercheurs voudront participer à ce projet, à cet article ?
- les décisions évoluent dans le temps... et peuvent évoluer aussi selon la prochaine évaluation en vue (par ex. j'ai besoin de + d'articles...)

Deux méthodes d'évaluation

En gros, il y en a deux méthodes d'évaluation :

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

- étudier la qualité : quels critères ?
- indicateurs et métriques : lesquels ? [11, 39, 51, 52, 53, 54]
 - ▶ l'utilisation du facteur d'impact des revues
 - ▶ l'utilisation de métriques en général

- facteur social

[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.

Voir les "EC Expert reports" [9, 11] :

- ils donnent plein de recommandations
- il faut tenir compte de la production de RS
- [9] Open Science Career Assessment Matrix (OS-CAM)
- [11] comment établir les comités d'évaluation (par ex.)

Plan

5 Protocole(s) CDUR

6 Sur la Science Ouverte

7 Conclusions

Protocole(s) CDUR

Conçu pour aider les chercheurs évalués, les comités d'évaluation, les décideurs.

(C) Citation mesure si RS bien identifié en tant que produit de la recherche, référence, ou bien metadata, citer les autres RS...

point légal : auteurs, affiliations, % de participation

(D) Dissemination les bonnes pratiques de diffusion suivies en fonction de la politique scientifique du contexte d'évaluation ;

[14](2014) TGD : Free software, Open source software, licenses. A short presentation including a...

point politique scientifique : Science Ouverte, **point légal** : licences

(U) Use point dédié aux aspects purement "logiciel" **du RS** : produit des résultats corrects, facilite la réutilisation ; peut regarder aussi bonnes pratiques logiciel : exemples, docs., test, facilite install., voir le code, lancer le RS, normes de qualité, dépôt APP, valorisation, startup...

point reproductibilité, validation des résultats obtenus

(R) Research point dédié aux aspects purement "recherche" : qualité du travail scientifique, algorithmes et structures de données proposées et programmées, publications, utilisateurs, collaborations et projets...

point impact de la recherche

Flexibilité d'application : chaque ensemble décideurs/comité d'évaluation établi **sa propre procédure**, adaptée aux objectifs fixés et au contexte d'évaluation.

Sur la Science Ouverte

The future of Open Science asks for a common understanding

Open Science is the political and legal framework where research outputs are shared and disseminated in order to be rendered visible, accessible and reusable.

I Three selected pillars

- BOAI (2002)
- Free Software Foundation (1985)
- CODATA (1966)

II Towards a political and legal framework

III Enablers:

- Institutional policies
- Infrastructures
- Research evaluation

EGI Virtual Conference 2021, Lisbon, 19-21 october 2021

<https://padlet.com/gwenfranck/EGI2021Posters>

(2020-21) TGD, T. Recio : Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs, POLIS N. 19 2020, <https://uet.edu.al/polis/wp-content/uploads/2022/01/polis-19.pdf>
V3 du 28/02/2021 : <https://zenodo.org/record/4577066>

The future of Open Science asks for a common understanding

Teresa Gomez-Diaz, CNRS/IGM, Est de Paris
Tomas Recio, University Nebrija, Madrid

A definition is missing
Definition of Open Science remains vague, probably because definitions are vague. (Hager, 2019) 2019
... there is no single, accepted, unified definition or notion of 'open science'. (Therrien, 2018)
... there is a lack of consensus about what Open Science is, mostly due to the fact that there is no formal definition of Open Science. (Gomez-Diaz & Recio, 2020-21)

Recent, inclusive & complete visions

Our contribution

Goal
To contribute to the adoption of a common, unified vision.

Definition proposal
Open Science is the political and legal framework where research outputs are shared and disseminated in order to be rendered visible, accessible and reusable (Gomez-Diaz & Recio, 2020-21).

Three steps supporting this proposal
I - Three selected pillars for a common understanding
II - Towards a political and legal framework
III - Enablers: three cornerstones to get to a working framework

I - Three selected pillars for a common understanding
Free Software Foundation, 1985
The free software definition provides the criteria for whether a particular software program qualifies as free. It is not a program as such, the freedom to study how the program works, and change it, is a requirement of freedom.

II - Towards a political and legal framework
Towards a political and legal framework
Towards a political and legal framework

III - Enablers: three cornerstones to get to a working framework
Institutional policies: the required evolution of policies of universities and RPOs.
Infrastructures: the development of Open Science-related infrastructures and services.
Research evaluation: the transformation of evaluation policies and practices.

Conclusions

Ouvrir le débat, est-ce que ce type de procédure sera adopté ?

Si on veut réussir à faire évoluer les pratiques en Science Ouverte, RS ...

il est nécessaire de faire évoluer les méthodes d'évaluation pour en tenir compte.

Conséquences attendues :

en accord avec [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

Mais aussi :

Pour un labo : liste de RS fait partie de la production scientifique après évaluation.

Augmenter la transparence dans les méthodes d'évaluation de RS, d'un laboratoire.

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

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