

# Les logiciels de la recherche dans le cadre de la Science ouverte : définition, diffusion, évaluation.

Teresa Gomez-Diaz

CNRS - Laboratoire d'informatique Gaspard-Monge

Contient des travaux en collaboration avec Prof. T. Recio (Univ. Nebrija, Madrid)

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License  
(CC-BY-NC-ND 4.0) – <http://creativecommons.org/licenses/by-nc-nd/4.0/>

**Formation : Les Plans de Gestion des Logiciels de la Recherche**

Université Gustave Eiffel, 23 mai 2024



# Objectifs

(\*) (2010) TGD, Diffuser un logiciel de laboratoire : recommandations juridiques et administratives

(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

Notre objectif ici est de présenter ces travaux sur les logiciels de la recherche, dans le cadre de la Science ouverte, où nous proposons :

- Définitions
- Procédure de diffusion
- Protocoles CDUR d'évaluation

Ces travaux sont réalisés en collaboration avec le Prof. Tomas Recio, Universidad Antonio de Nebrija, Madrid, (<http://www.recio.tk/>).

Ils ont été adaptés aux données de la recherche, voir :

(2022) TGD, T. Recio. Research Software vs Research Data I (Definition)

(2022) TGD, T. Recio. Research Software vs Research Data II (Dissemination, CDUR)

(2024) TGD, On the sharing and dissemination of Research Software and Research Data in the OS context

(\*) Les références complètes sont à la fin de la présentation.

# Plan

## 1 Sur la Science ouverte

- Définition proposée de Science ouverte
- Recommandation de l'UNESCO sur une science ouverte

## 2 Les logiciels de la recherche - Research Software

- Le concept, la définition
- Le logiciel est un objet juridique
- Auteurs d'un logiciel de la recherche
- Référence et citation

## 3 Procédure de diffusion

## 4 Protocole(s) **CDUR** : évaluation de la recherche et RS

## 5 Conclusions

# Définition proposée de Science ouverte - article

(2020-21) TGD, T. Recio. Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs

La Science ouverte est :

[FR] *le cadre politique et juridique dans lequel les productions scientifiques sont partagées et diffusées afin d'être rendues visibles, accessibles et réutilisables.*

[EN] *the political and legal framework where research outputs are shared and disseminated in order to be rendered visible, accessible and reusable*

Version	Titre	Date	Publication
V3	Towards an Open Science definition as a political and legal framework: on the sharing and dissemination of research outputs	02/2021	<a href="https://zenodo.org/record/4577066">https://zenodo.org/record/4577066</a> Avec la réf. : Alma Swan, UNESCO, 2012
V2	Towards an Open Science definition as a political and legal framework...	12/2020	POLIS N. 19, pp. 36-56 - <a href="#">PDF</a>
V1	A policy and legal Open Science ...	09/2020	<a href="https://zenodo.org/record/4075106">https://zenodo.org/record/4075106</a>

**Objectifs** : comprendre ce qu'est la Science ouverte, la motivation qui nous mène à proposer cette définition et voir comment structurer l'information.

# Définition proposée de Science ouverte - poster

## 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

**The future of Open Science asks for a common understanding**  
Teresa Gomez-Diaz, CNRS/LIGM, Est of Paris  
Tomas Recio, University Nebrija, Madrid

**A definition is missing**  
Definition: [an Open Science] consists, roughly, probably because objectives are vague... (Bouvier, 2020) 2021  
...there is no single accepted, unified definition or vision of Open Science... (Gomez-Diaz & Recio, 2020-21)  
...Research outputs are shared and disseminated in order to be rendered visible, accessible and reusable (Gomez-Diaz & Recio, 2020-21)  
[Open Science] - (Gomez-Diaz & Recio, 2020-21)

**Research, inclusive & complete visions**

**For the purpose of this Recommendation, Open Science is defined as an inclusive ecosystem... allowing broader scientific knowledge (scientific products, accessible and reusable for everyone)... [2] It includes all scientific, disciplinary and cross-disciplinary activities... [3] It builds on the following key pillars: open access to scientific knowledge; open access to scientific publications; open access to scientific data; open access to scientific instruments... (OECD, 2021)**

**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**

- BOAI (2002)**  
The Open Access Initiative (BOAI) was created in 2002 to promote open access to research outputs.
- Free Software Foundation, 1985**  
The Free Software Foundation (FSF) was created in 1985 to promote free software.
- CODATA (1966)**  
CODATA was created in 1966 to promote open access to scientific data.

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

- European Open Access Initiative, 2002**  
The European Open Access Initiative (EOAI) was created in 2002 to promote open access to research outputs.
- Committee on Data for Science and Technology, 1966**  
The Committee on Data for Science and Technology (CODATA) was created in 1966 to promote open access to scientific data.

**III - Enablers: three cornerstones to get to a working framework**

- Institutional policies**  
The required evaluation of policies of universities and RPOs.
- Infrastructures**  
The development of Open Science-oriented infrastructures and services.
- Research evaluation**  
The transformation of evaluation policies and practices.

**References**

**EGI Virtual Conference 2021**  
19-21 October 2021

# Recommandation de l'UNESCO sur une science ouverte

- À la demande des 193 pays participant à la Conférence générale de l'UNESCO, 40th session, 11/2019
- Prise de conscience, conséquence de la pandémie de COVID-19
- Une année de préparation, consultation, commentaires au rapport préliminaire  
<https://unesdoc.unesco.org/ark:/48223/pf0000374409>
- Texte adopté par les 193 pays, 41st session, 11/2021  
<https://unesdoc.unesco.org/ark:/48223/pf0000379949>

Aux fins de la présente Recommandation, la **science ouverte** s'entend comme un **concept inclusif qui englobe différents mouvements et pratiques** visant à **rendre les connaissances scientifiques** multilingues, **librement accessibles à tous et réutilisables par tous**, à renforcer la collaboration scientifique et le partage des informations au profit de la science et de la société, ainsi qu'à ouvrir les processus de création, d'évaluation et de diffusion des connaissances scientifiques aux acteurs de la société au-delà de la communauté scientifique traditionnelle.

Elle **inclut toutes les disciplines scientifiques et tous les aspects des pratiques savantes, y compris les sciences fondamentales et appliquées, les sciences naturelles et les sciences sociales et humaines**, et repose sur les piliers essentiels suivants : les connaissances scientifiques ouvertes ; les infrastructures de la science ouverte ; la communication scientifique ; la participation ouverte des acteurs de la société ; et le dialogue ouvert avec les autres systèmes de connaissances.

Suivi des groupes de travail,

<https://www.unesco.org/en/open-science/implementation#open-science-working-groups>

# Plan

- 1 Sur la Science ouverte
- 2 Les logiciels de la recherche - Research Software
  - Le concept, la définition
  - Le logiciel est un objet juridique
  - Auteurs d'un logiciel de la recherche
  - Référence et citation
- 3 Procédure de diffusion
- 4 Protocole(s) **CDUR** : évaluation de la recherche et RS
- 5 Conclusions

# Définition de logiciel de la recherche

(2007) TGD. Autour de la valorisation de logiciels développés dans un laboratoire de recherche

(2009) TGD. Guide laboratoire pour recenser ses développements logiciels

(2011, 2015) TGD. Article vs. Logiciel : questions juridiques et de politique scientifique dans la production de logiciels

**Un logiciel du laboratoire** est un 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.*

- finalité : recherche
- un membre du laboratoire participe à l'écriture du logiciel (idem aux publications)
- la production importante sont les articles, les logiciels sont des objets associés

(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

(2021) TGD, T. Recio. Open comments on the Task Force SIRS report: Scholarly Infrastructures for Research Software

**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...*

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*

# Le logiciel est un objet juridique

(2011, 2015) TGD. Article vs. Logiciel : questions juridiques et de politique scientifique dans la production de logiciels  
(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

- Code de la propriété intellectuelle (CPI), Article L. 112-2 :  
un *logiciel* est une œuvre de l'esprit protégée par le droit d'auteur.
- Arrêté du Ministère de l'Industrie du 22 décembre 1981 (vocabulaire de l'informatique) :  
un *logiciel* est un ensemble des programmes, procédés et règles, et éventuellement de la documentation, relatifs au fonctionnement d'un ensemble de traitement de données.
- Directive 2009/24/CE du Parlement européen et du Conseil du 23 avril 2009 concernant la protection juridique des programmes d'ordinateur :  
les termes *programme d'ordinateur* visent les programmes sous quelque forme que ce soit, y compris ceux qui sont incorporés au matériel. Ces termes comprennent également les travaux préparatoires de conception aboutissant au développement d'un programme...

---

(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

**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...*

# Auteurs d'un logiciel de la recherche

(2011, 2015) TGD. Article vs. Logiciel : questions juridiques et de politique scientifique dans la production de logiciels  
(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

## Que veut dire auteur d'un RS ?

- 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, install, 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.

# Référence et citation

(2013) D. Pontille, D. Torny. La manufacture de l'évaluation scientifique : algorithmes, jeux de données...

<https://minesparis-psl.hal.science/hal-00821956>

*[...] 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"*

(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

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...
- Software Citation Group, Software Citation Implementation Working Group...

# Plan

- 1 Sur la Science ouverte
- 2 Les logiciels de la recherche - Research Software
- 3 Procédure de diffusion**
- 4 Protocole(s) **CDUR** : évaluation de la recherche et RS
- 5 Conclusions

# Procédure de diffusion des logiciels de la recherche

S'adapte à chaque situation, valable pour les données.

- Choisir un nom, éviter les noms déjà utilisés, les marques.  
(2018) DoRANum or Harvard, Bien nommer ses fichiers/File Naming Conventions.
- (\*) Établir la liste des auteurs (avec % de participation), leurs affiliations.  
Plan de gestion de logiciels de la recherche (PGLR) (2018, TGD, G. Romier, Modèle de PGLR V3.2, PRESOFT)
- (\*) Établir la liste des fonctionnalités principales.
- (\*) Établir la liste des briques logicielles ou les données utilisées, avec licences.
- **Choisir et mettre en place une licence**, accord de tous les détenteurs des droits.  
Si possible : un accord signé. Attention à la compatibilité et héritage des licences.
- Choisir un site web, forge, dépôt pour la distribution. Indiquer les licences et les conditions d'utilisation, copie..., comment citer l'œuvre.
- Créer et indiquer une adresse courriel de contact.
- (\*) La traçabilité est importante, archiver en .tar.gz régulièrement. Réviser PGLR.
- Informer la direction des laboratoires et les tutelles (si pas fait au point licence).
- **Diffuser le logiciel** et/ou les données.
- Informer la communauté cible, considérer les data ou les software papers.

(\*) À revoir à chaque nouvelle version du logiciel.

(2010) TGD. Diffuser un logiciel de laboratoire : recommandations juridiques et administratives

(2014) TGD. Free software, Open source software, licenses. A short presentation including a procedure for RS and data diss...

# Plan

- 1 Sur la Science ouverte
- 2 Les logiciels de la recherche - Research Software
- 3 Procédure de diffusion
- 4 Protocole(s) **CDUR** : évaluation de la recherche et RS
- 5 Conclusions

# Protocole(s) **CDUR** : évaluation de la recherche et RS

(2019) TGD, T. Recio. On the evaluation of research software: the CDUR procedure

Le protocole **CDUR** est proposé pour prendre en compte les logiciels de la recherche dans l'évaluation. 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 metadataset avec nom, auteurs, version, dates...  
aussi bonnes pratiques pour citer les autres RS...  
**point** légal : auteurs, affiliations, participation
- (D) Dissemination** les bonnes pratiques de diffusion suivies en fonction de la politique scientifique du contexte d'évaluation. Plan de Gestion de RS (PGLR).  
**point** Science ouverte, 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 proposés et programmés, publications, utilisateurs, collaborations et projets...  
**point** impact de la recherche

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

# Conclusions

Nous avons construit un cadre d'étude pour les logiciels de la recherche, pour les comprendre, les expliquer et pour valoriser leur contribution à la Science ouverte.

Il se construit en trois étapes :

- Définition
- Procédure de diffusion
- Protocoles CDUR d'évaluation

Nous avons aussi adapté ces travaux aux données de la recherche.  
Et d'autres équipes ont utilisé, par exemple, le protocole CDUR dans le cadre des nanomatériaux :

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

Ceci montre une des forces de nos travaux sur les logiciels : leur adaptabilité à d'autres contextes et à d'autres productions scientifiques.

Notre objectif est de porter des éléments de discussion pour contribuer à cet édifice, complexe et en construction, de la Science ouverte.

# Références (1/2)

- 2007 TGD. Autour de la valorisation de logiciels développés dans un laboratoire de recherche, LIGM.
- 2009 TGD. Licence & copyright pour les développements de logiciels libres de laboratoires de recherche, PLUME, <https://zenodo.org/record/7063146>
- 2010 TGD. Diffuser un logiciel de laboratoire : recommandations juridiques et administratives, PLUME, <https://zenodo.org/record/7096216>
- 2011 TGD. Article vs. Logiciel : questions juridiques et de politique scientifique dans la production de logiciels, PLUME, <https://zenodo.org/record/7063154>
- 2013 TGD. Articles, logiciels, données : étude de la diffusion de la production scientifique, Poster, Journées FRéDoc 2013, [http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/JourneesDR/2013octFREDOC\\_A0\\_Portrait.pdf](http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/JourneesDR/2013octFREDOC_A0_Portrait.pdf)
- 2014 TGD. Articles, software, data: a study of the (French) scientific production, Poster, EUDAT 2014, [http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/Internacional/2014septeuadat\\_70x100.pdf](http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/Internacional/2014septeuadat_70x100.pdf)
- 2014 TGD. Free software, Open source software, licenses. A short presentation including a procedure for research software and data dissemination, Zenodo, <https://zenodo.org/record/11709>
- 2015 TGD. Article vs. Logiciel : questions juridiques et de politique scientifique dans la production de logiciels, Société Informatique de France, <https://zenodo.org/record/18993>
- 2018 TGD, G. Romier. Modèle de Plan de Gestion de Logiciel de la Recherche (PGLR), V3.2, PRESOFT, <https://zenodo.org/record/1405614>
- 2019 TGD. Le Projet PLUME et le paysage actuel des logiciels de la recherche dans la science ouverte, Zenodo, <https://zenodo.org/record/2591474>
- 2019 TGD. Les logiciels de la recherche et leurs licences : trois visions sur un objet, Cours CERFACS, <https://hal.science/hal-02434287v1>
- 2019 TGD, T. Recio. On the evaluation of research software: the CDUR procedure, F1000Research, Research on Research, <https://doi.org/10.12688/f1000research.19994.2>
- 2020-1 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://doi.org/10.58944/yuro5734>, **V3 du 28/02/2021**, <https://zenodo.org/record/4577066>

## Références (2/2)

- 2021 TGD. Sur la production de logiciels libres au Laboratoire d'informatique Gaspard-Monge (LIGM) : ce que nous avons appris, Atelier BlueHats (Etalab)  
[http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/Seminaires/2021avrilBlueHats\\_TGD.pdf](http://igm.univ-mlv.fr/~teresa/logicielsLIGM/documents/Seminaires/2021avrilBlueHats_TGD.pdf)
- 2021 TGD, T. Recio. Open comments on the Task Force SIRS report: Scholarly Infrastructures for Research Software (EOSC Executive Board, EOSCArchitecture), RIO 7: e63872, <https://doi.org/10.3897/rio.7.e63872>
- 2021 TGD, T. Recio. The future of Open Science asks for a common understanding, Poster, EGI Virtual Conference 2021, <https://indico.egi.eu/event/5464/contributions/15729/>
- 2022 TGD, T. Recio. Research Software vs Research Data I: Towards a Research Data definition in the Open Science context, F1000Research, Research on Research, <https://f1000research.com/articles/11-118/v2>
- 2022 TGD, T. Recio. Research Software vs Research Data II: Protocols for Research Data dissemination and evaluation in the Open Science context, F1000Research, Research on Research, <https://f1000research.com/articles/11-117/v2>
- 2022 TGD, T. Recio. Research Software and Research Data: dissemination, evaluation and reusability in the Open Science context, Poster, IDCC22, <https://zenodo.org/record/6778872>
- 2023 TGD. Les logiciels de la recherche et leurs licences : trois visions sur un objet, Cours Université Gustave Eiffel, <https://hal.science/hal-02434287v2>
- 2023 TGD, T. Recio. How to achieve FAIRER research data by studying evaluation & assessment protocols, Poster, Open Science FAIR 2023, <https://zenodo.org/records/8398431>
- 2023 TGD, T. Recio. Articles, software, data: An Open Science ethological study, Maple Transactions, <https://doi.org/10.5206/mt.v3i4.17132>
- 2024 TGD, T. Recio. Open comments on the "Open Letter: Establishing a national research software award", Zenodo, <https://zenodo.org/records/10493797>
- 2024 TGD, T. Recio. Open comments for the NIH Request for Information (RFI): Best Practices for Sharing NIH Supported Research Software (NIH, 2023), Zenodo, <https://zenodo.org/records/10617585>
- 2024 TGD. On the sharing and dissemination of Research Software and Research Data in the Open Science context, Open Science Days 2024, Max Planck Digital Library, <https://zenodo.org/records/10617691>