



**HAL**  
open science

## Sur la diffusion des données et des logiciels de la recherche dans le cadre de la Science ouverte

Teresa Gomez-Diaz, Tomas Recio

### ► To cite this version:

Teresa Gomez-Diaz, Tomas Recio. Sur la diffusion des données et des logiciels de la recherche dans le cadre de la Science ouverte. ANF Science ouverte 2023. Pratiquer la Science ouverte : des services et des compétences en action, Jul 2023, Paris, France. , 2023. hal-04401471

**HAL Id: hal-04401471**

**<https://hal.science/hal-04401471>**

Submitted on 17 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



# Sur la diffusion des **données** et des **logiciels de la recherche** dans le cadre de la **Science ouverte**

**Teresa Gomez-Diaz**, [teresa.gomez-diaz@univ-eiffel.fr](mailto:teresa.gomez-diaz@univ-eiffel.fr)  
**Université Gustave Eiffel-LIGM-CNRS, Est of Paris**  
<https://siteigm.univ-mlv.fr/>

**Tomas Recio**, [trecio@nebrija.es](mailto:trecio@nebrija.es)  
**Universidad Antonio de Nebrija, Madrid**  
<https://www.nebrija.com/>

**Objectif :** rendre les productions scientifiques visibles et accessibles dans le contexte évolutif actuel de la **Science ouverte** demande une diffusion réalisée dans un cadre maîtrisé, avec une licence qui permet, entre autres, la réutilisation et la modification de cette production, et en traitant correctement les questions juridiques, s'il y en a.

**Mots clés :** logiciels de la recherche, données de la recherche, définition, diffusion, Science ouverte.

## Définitions

**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...

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

**Research data (RD)** is a well identified set of data that has been produced (collected, processed, analyzed, shared & disseminated) by a 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... 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 (RS or other).

## Procédure de diffusion pour les logiciels de la recherche

- Choisir un nom pour le **RS**, éviter les noms déjà utilisés, les marques.  
(2018) DoRANum. Bien nommer ses fichiers.
- (\*) (*research team step*) Établir la liste des auteurs et des contributeurs (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 briques logicielles et/ou les données utilisées, avec leurs licences. Suivre des bonnes pratiques de citation pour les données et les logiciels de la recherche ?
- Choisir une licence, avec l'accord des auteurs et propriétaires des droits. Si possible, obtenir un accord signé. Attention à la compatibilité et à l'héritage des licences.
- Choisir un site web, forge, dépôt pour la distribution. Indiquer les licences et les conditions d'utilisation, copie..., indiquer comment citer l'œuvre, utiliser des identifiants pérennes (PIDs...)
- (\*) (*research work step*) Faciliter la réutilisation : établir la liste des fonctionnalités principales. La traçabilité est importante, archiver en .tar.gz régulièrement, garder une trace des articles... Réviser la documentation, le PGLR, le site web...
- Informer la direction des laboratoires et les tutelles (si cela n'a pas été fait au point licence).
- Créer et indiquer une adresse courriel de contact.
- Diffuser les logiciels de la recherche (**RS**).
- Informer la communauté cible, considérer les software papers.

(\*) A revoir à chaque nouvelle version.

(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 research software and data dissemination

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

## Procédure de diffusion pour les données de la recherche

- Choisir un nom pour le **RD**, éviter les noms déjà utilisés, les marques.  
(2018) DoRANum. Bien nommer ses fichiers.
- (\*) (*research team step*) Établir la liste des personnes qui ont participé à la production, avec leurs affiliations.  
Traiter les droits d'auteur, s'il y en a. Plan de gestion de données (PGD) ?  
★ Traiter les autres contextes juridiques (ou éthiques), s'il y en a, (avec les experts juridiques ?).
- (\*) Établir la liste des briques logicielles et/ou les données utilisées, avec leurs licences. Suivre des bonnes pratiques de citation pour les données et les logiciels de la recherche ?
- Choisir une licence, avec l'accord des auteurs et propriétaires des droits. Si possible, obtenir un accord signé. Attention à la compatibilité et à l'héritage des licences.
- Choisir un site web, forge, dépôt pour la distribution. Indiquer les licences et les conditions d'utilisation, copie..., indiquer comment citer l'œuvre, utiliser des identifiants pérennes (PIDs...)
- (\*) (*research work step*) Faciliter la réutilisation : comme les données ont été utilisées pour vos recherches. La traçabilité est importante, archiver en .tar.gz régulièrement, garder une trace des articles... Réviser la documentation, le PGD, le site web...
- Informer la direction des laboratoires et les tutelles (si cela n'a pas été fait au point licence).
- Créer et indiquer une adresse courriel de contact.
- Diffuser les données de la recherche (**RD**).
- Informer la communauté cible, considérer les data papers.

(\*) A revoir à chaque nouvelle version.

 **Université Gustave Eiffel**



**ANF Science ouverte 2023**

Pratiquer la Science ouverte :  
des services et des compétences en action  
4 - 6 juillet 2023

Sorbonne Université – Campus Pierre et Marie Curie  
4 place Jussieu, Paris

