

Free/Open source Research Software production at the Gaspard-Monge Computer Science laboratory (LIGM) Lessons learnt

Teresa Gomez-Diaz

CNRS - Laboratoire d'informatique Gaspard-Monge

This work is licensed under
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

FOSDEM21, February 6th 2021





– A case study: the LIGM Lab

- LIGM stands for “*Laboratoire d'informatique Gaspard-Monge*”, that is, the Gaspard-Monge Computer Science laboratory
- Co-funded: CNRS, ENPC, Univ. Gustave Eiffel, located est of Paris
- <http://ligm.u-pem.fr/>
- around 150 members, where 80 are permanent researchers
- research themes: *image analysis, image processing, geometry in computer science, learning, networks, real-time algorithms, combinatorics, algorithms, bioinformatics, computational logic and automata, automatic language processing, signal processing...*
- Important research software production, mostly FLOSS:
 - ▶ **goal:** mainly research
 - ▶ 2008-13 : 61 software items, where 53 FLOSS
 - ▶ 2013-18 : 66 software items, where 50 FLOSS

LIGM research software mission

- 2002: CNRS research engineer
- 2006: LIGM software mission

to study the implementation of services (versioning, publication...) associated with the laboratory software development to promote their visibility.

- 2008-13: PLUME project, CNRS national project
 - (*)(2019) TGD: Le Projet PLUME et le paysage actuel des logiciels de la recherche...
 - ▶ publication of research software descriptions
 - ▶ reports on FLOSS, legal issues: licenses & author's rights, dissemination procedure, dissemination & FLOSS policies...
- 2017-18: PRESOFT project, Software Management Plan template
 - (2018) TGD, G. Romier: Research Software management Plan Template V3.2...
- since 2018: ongoing work on research software evaluation, Open Science, Infrastructures for Open Science..., see for example
 - (2019) TGD, T. Recio: On the evaluation of research software: the CDUR procedure
 - (2020) TGD, T. Recio: A policy and legal Open Science framework: a proposal
 - (2021) TGD, T. Recio: Open comments on the Task Force SIRS report: Scholarly...

(*) See references at the end of the talk.

Problems detected in 2006-07: 2021's vision

Open up !	Which is the software production of the Lab ? → produce the list of the software production
Definition	What I do is a lab's software ? → need to define the object "Lab's software" or "research software"
Reference	No name, no authors, no dates, no version... → need to establish a reference or citation form
Dissemination	Some on personal web pages, some on forges, some with web page... → need to establish a dissemination procedure
FLOSS	Real interest on free software, but often no licence → need to understand FLOSS concepts and licensing issues
Legal issues	Who decides of licence ? → need to understand legal issues and author's rights
Policies	Definition, licensing, dissemination, decisions... → need to clarify Lab's software policies
Value	Which is the value of this scientific production ? → need to propose evolutions for research evaluation procedures

**Problems of different nature: legal, policies, publication, scientific, best practices...
requiring different approaches, but alike in every scientific field, in many labs.**

Research software (RS) definition

(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...

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.

- goal: research
- a member of the lab participates on the code writing (as in a publication)
- the important production are publications: RS as an associated object

(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 Infra...

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...

References, Publications, List of Lab's software

- LIGM 2008-2013 (PLUME): 61 RS identified, where 53 FLOSS

The screenshot shows the PLUME software repository interface. At the top, there is a search bar and a navigation menu with links like 'Home', 'About', 'Help', 'Contact', and 'Logout'. Below the search bar, there is a section titled 'Search' with a dropdown menu for 'Software' and a search input field. A 'Search' button is next to it. To the right, there is a 'Software' section with a list of items, each with a thumbnail, title, and a brief description. At the bottom, there is a 'Software' section with a list of items, each with a thumbnail, title, and a brief description.

Metadata, keywords, classification,
search interfaces... and links to publications.

The screenshot shows the LIGM research software repository interface. At the top, there is a search bar and a navigation menu with links like 'Home', 'About', 'Help', 'Contact', and 'Logout'. Below the search bar, there is a section titled 'Search' with a dropdown menu for 'Software' and a search input field. A 'Search' button is next to it. To the right, there is a 'Software' section with a list of items, each with a thumbnail, title, and a brief description. At the bottom, there is a 'Software' section with a list of items, each with a thumbnail, title, and a brief description.

- LIGM 2013-2018 (internal document): 66 RS identified, where 50 FLOSS

- Evolutions

- within the Open Science ongoing work at University Gustave Eiffel, see (2020) TGD, T. Recio: A policy and legal Open Science framework: a proposal
- (since 2011) Image Processing On Line (IPOL) Journal: article+RS
P. Monasse, Extraction of the Level Lines of a Bilinear Image, IPOL 2019, <https://www.ipol.im/pub/art/2019/269/>
- (since 2015) SoftwareX Journal with RS in GitHub
B. Perret, et al., Higra: Hierarchical Graph Analysis, SoftwareX, Vol. 10, 2019, <https://www.sciencedirect.com/science/article/pii/S235271101930247X>

Dissemination procedure, Software Management Plans

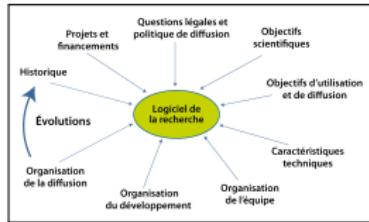
(2010) TGD: Diffuser un logiciel de laboratoire : recommandations juridiques et administratives

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

Flexible, to be adapted to different situations, also valid for research data.

- Choose a name, avoid trademarks and proprietary names, you can associate date, version, target platform...
 - (*) Establish list of authors (with % of participation) and affiliations. Include minor contributors.
 - (*) Establish the list of main functionalities.
 - (*) Establish the list of included software & data components, their licences or documents with rights to access, copy...
 - Choose a license, have an agreement (signed) with rightholders and authors, consider FLOSS licences and CC licences, for example. Beware of licence compatibility and inheritance issues.
 - Use a website, forge, deposit for dissemination, indicate licences and how to cite the work. Use PIDs if possible.
 - (*) Archive a tar.gz or similar regularly to keep track of added functionalities.
 - Inform your laboratories and head institutions (if not done in the license step).
 - Set and indicate clearly a contact address.
 - Distribute the software or data component. Inform the target community. Consider Software or Data papers...
- (*) Review for each new version.

(2018) TGD, G. Romier: Research Software management Plan Template V3.2...



SMP title with RS name,
authors, dates...

1. Metadata
2. Software context →
3. Software features →
4. Team organisation
5. Development organisation
6. Distribution organisation
7. SMP management

References

- 2.1 History
- 2.2 Project(s) related to the RS
- 2.3 Legal issues and distribution policy
- 3.1 Scientific goals
- 3.2 Usage and distribution objectives
- 3.3 Technical features

Studies on legal issues & scientific policies

- (2009) TGD: Licence & copyright pour les développements de logiciels libres de laboratoires...
- (2009) TGD: Guide laboratoire pour recenser ses développements logiciels
- (2011, 2015) TGD: Article vs. Logiciel: questions juridiques et de politique scientifique...

To understand and to explain research software legal and scientific policy issues by comparison with the well understood production of research papers.

Aspects légaux		
	Article	Logiciel
Droit auteur	droits moraux, droits patrimoniaux	droits moraux réduits droits pat. dévolus à l'employeur
Oeuvre	article	code source, code objet, doc., ...
Auteurs	signataires, même %	notion complexe, pb. légal, établir % de participation
Propriétaires	auteurs, même % cession des droits	tutelles en général, mais dépend du régime salariés, des contrats, ...
Dates	soumission, publication	matériel de conception, versions
Évolution	œuvre indépendante	œuvre indépendante ? il faut revoir auteurs, dates, lic., ...
Travaux préc.	références, citations	briques : compatibilité, héritage lic.
Diffusion	éditeur, web	web, forges, besoin de licence
Droits	lire, citer, ne pas copier	lire, ne pas utiliser, ..., besoin lic.
Licences	droits et obligations, CC (web)	droits et obligations, libres, propriétaires

Aspects relatifs à la politique scientifique		
	Article	Logiciel
Définition (L, T)	ok	à définir
Signature (C, T)	ok, déf. par tutelles	à définir (copyright) associer les laboratoires
Références (L, T)	HAL	PLUME
Liste des œuvres (L, T)	document à jour	document inconnu, PLUME peut être utile
Libre accès (C, L, T, CSI)	politique (+/-) ok, dépôt ok (HAL)	politique (lic.) à définir, dépôt à établir
Validation (C, L, T, CSI)	procédure referee, reproductibilité	à définir, validé au sens PLUME
Qualité/évaluation (C, L, T, CSI)	nb. citations	articles associés, attirer utilisateurs, contrats
Motivation (C, L, T, CSI)	recherche, article	recherche, pas le logiciel
Objet (C, L, T, CSI)	scientifique	3D : scientifique, potentiel de transf. de tech., obj. industriel

<https://www.societe-informatique-de-france.fr/wp-content/uploads/2015/04/1024-5-gomez-diaz.pdf>

All references given at the end of this talk include studies of different aspects of legal and scientific policy issues.

How to give **value** to the RS production ?

By increasing RS producers' scientific reputation, see

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

This work proposes the **CDUR** protocol(s) 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
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

Open up !	working on infrastructures and services for RS → ongoing work at University Gustave Eiffel working group
Definition	What I do is a lab's software ? → definitions proposed for "Lab's software" or "research software"
Reference	No name, no authors, no dates, no version... → we have now dissemination procedures, SMPs...
Dissemination	Some on personal web pages, some on forges, some with web page... → we have now dissemination procedures, SMPs...
FLOSS	Real interest on free software, but often no licence → need to continue "evangelism": new members, students...
Legal issues	Who decides of licence ? → need to continue "evangelism": new members, students...
Policies	Definition, licensing, dissemination, decisions... → ongoing work at University Gustave Eiffel working group
Value	Which is the value of this scientific production ? → research software evaluation procedures CDUR are proposed

Problems of different nature but **alike in every scientific field, in many labs**:
→ general but adaptable solutions are proposed.

References

- 2007 TGD: Autour de la valorisation de logiciels développés dans un laboratoire de recherche.
- 2009 TGD: Licence & copyright pour les développements de logiciels libres de laboratoires de recherche, PLUME, <https://projet-plume.org/fr/ressource/faq-licence-copyright>
- 2009 TGD: Guide laboratoire pour recenser ses développements logiciels, PLUME, <https://projet-plume.org/ressource/guide-laboratoire-recensement-developpements-logiciels>
- 2010 TGD: Diffuser un logiciel de laboratoire : recommandations juridiques et administratives, PLUME, <http://www.projet-plume.org/fr/ressource/diffuser-logiciel-recomm-juridiques-admin>
- 2014 TGD: Free software, Open source software, licenses. A short presentation including a procedure for research software [...] 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> (2011) V1, PLUME, <http://www.projet-plume.org/fr/ressource/article-vs-logiciel>
- 2018 TGD, G. Romier: Research Software management Plan Template V3.2, Projet PRESOFT, Zenodo, <https://zenodo.org/record/1405614>. Also in DMP OPIDoR.
- 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, T. Recio: On the evaluation of research software: the CDUR procedure, F1000Research, <https://doi.org/10.12688/f1000research.19994.2>
- 2020 TGD, T. Recio: A policy and legal Open Science framework: a proposal, Zenodo, <https://zenodo.org/record/4075106>
- 2021 TGD, T. Recio: Open comments on the Task Force SIRS report: Scholarly Infrastructures for Research Software, RIO 7: e63872, <https://doi.org/10.3897/rio.7.e63872>