Research software at Gaspard-Monge — FOSDEM 2021

Teresa Gomez-Diaz · FOSDEM 2021, Open Research

Slides from FOSDEM, used under CC BY 2.0 BE.

Source: https://video.fosdem.org/2021/D.research/open_research_gaspard_monge.webm

Converted with Video2Any · video2any.com

Not affiliated with or endorsed by FOSDEM.

Free/Open source Research Software production at the Gaspard-Monge Computer Science laboratory

Lessons learnt

Teresa Gomez-Diaz

D.research
2021-02-06





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



– 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

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

Research software (RS) definition

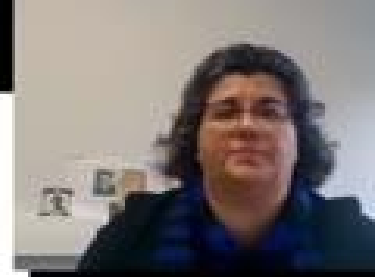
(2007) TGD: Autour de la valorisation de logiciels développés dans un laboratoire de recherche
(2009) TGD: Guide laboratoire pour recenser ses développements logiciels
(2011, 2015) TGD: Article vs. Logiciel: questions juridiques et de politique scientifique...

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
(2020) 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...



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
(2020) 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...



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

Dissemination procedure, Software Management Plans

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

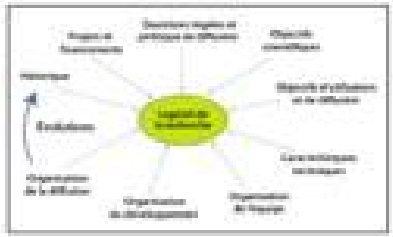
(2014) TGD: Free software, Open source software, licenses. A short presentation including a...

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

(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
Œuvre	article	code source, code objet, doc.
Auteurs	signataires, même %	notion complexe, pb. légal, multiple % en participation
Propriétaires	auteurs, même % cession des droits	tutelles en général, mais dépend du régime salarial, des contrats, ...
Dates	soumission, publication	matériel de conception, versions
Évolution	œuvre indépendante	œuvre indépendante ? ⇒ 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)	oui	à définir
Signature (C, T)	oui, déf. par curateu	à définir (copyright) associer les laboratoires
Références (L, T)	HAL	PLUME
Liste des œuvres (L, T)	document à jour	document inutile, PLUME peut être utile
Libre accès (C, L, T, CSI)	politique (+/-) oui, dépôt oui (HAL)	politique (NC) 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)	oui, citations	articles associés, retour 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

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