

Research Software landscape in France

Teresa Gomez-Diaz, Geneviève Romier

CNRS/LIGM, CNRS/CC-IN2P3

International RSE leaders meeting

The Alan Turing Institute, London, 30th-31st January 2018

This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International License.

<http://creativecommons.org/licenses/by-nc-nd/4.0/>



Plan

- 1 Research organisation and structures
- 2 Research Software Engineers
 - Some stats
 - Different organisations in labs and units
- 3 Professional networks
- 4 Funding sources
- 5 National scope initiatives for software management
 - SourceSup
 - CeCILL licences
 - PLUME project (2006-2013)
 - Two ongoing projects : PRESOFT and FG-SOL

Research organisation and structures

- All scientific fields
- 20 state-owned research institutions
 - ▶ conduct 54% of public research
 - ▶ CNRS, CEA, INRA, INSERM, CNES, INRIA...
- Higher education
 - ▶ conduct 40% of public research
 - ▶ +200 universities and other institutions
- Research units depend on institutions/universities (funding, human resources)
- Research software producers work in the units (researchers and engineers)
- Example : CNRS has around 30 000 members, with +1 000 research units shared with other institutions

Higher education and research in France, facts and figures

https://publication.enseignementsup-recherche.gouv.fr/eesr/10EN/EESR10EN_RESUME-higher_education_and_research_in_france_facts_and_figures_summary.php

Research Software Engineers : some stats

- Part of the ancillary research staff
 - ▶ permanent positions are regularly open once per year
 - ▶ catalog of professions, carrier path...
- CNRS stats
 - ▶ around 1 000 permanent RSE / 2 000 computer science, statistics & scientific computing
 - ▶ total 10 000 permanent engineers, all disciplines
 - ▶ about 360 non-permanent RSE (mainly project funding)
 - ▶ 20% female
- Universities' stats
 - ▶ 2 400 permanent computer science, statistics & scientific computing
 - ▶ total 9 000 permanent engineers, all disciplines

Bilan social et parité 2016 - http://bilansocial.dsi.cnrs.fr/pdf/BSP_2016.pdf

Bilan social du Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche - http://cache.media.enseignementsup-recherche.gouv.fr/file/Bilan_social/39/7/Bilan_social_MESR_2014-2015_659397.pdf

RSE in labs and units : different organizations

- Laboratories, units and services can have different organizations according to size, disciplines and head institutions
- Laboratories can have several localizations
- Several possibilities :
 - ▶ RSEs in a centralized service internal to a lab that provides resources to research projects or teams (ex. CNRS/IN2P3 labs)
 - ▶ RSEs in a “development service unit” (ex. INRIA)
 - ▶ RSEs are in the lab and can be part of a research team, or to be involved temporally in a project
 - ▶ Participation into teams/projects can evolve on time
- Many labs have few, isolated RSEs, so it is important to participate in the professional networks

Professional networks

- Officially supported by the CNRS, open to all
- DevLOG, <http://devlog.cnrs.fr/>, **software developers**
 - ▶ officially supported by CNRS since 2011
 - ▶ mailing list : 1 350 subscribers
 - ▶ main actions : JDEV (biannual conf), topic workshops, training (**software development**, dissemination, IPR & lics...)
 - ▶ community building
- RBDD, <http://rbdd.cnrs.fr/>, **databases practioners**
 - ▶ officially supported by CNRS since 2012
 - ▶ mailing list : 350 subscribers
 - ▶ main actions : recommendations, workshops and training sessions (databases and **related tools dev.**, distribution, IPR & lics...)
- Calcul, <http://calcul.math.cnrs.fr/>, computing
- Cross-disciplines experience sharing
- Other professional networks : region, disciplines, institutions...

Funding sources for software projects and non-permanent RSE

- Institutions, universities
- Different organisations, industry...
- Local, regional, national
- Europe (H2020 and others)
- Other international calls...

- 1 Research organisation and structures
- 2 Research Software Engineers
- 3 Professional networks
- 4 Funding sources
- 5 National scope initiatives for software management
 - SourceSup
 - CeCILL licences
 - PLUME project (2006-2013)
 - Two ongoing projects : PRESOFT and FG-SOL

SourceSup : national academic forge

- <https://sourcesup.renater.fr/>
- Provided by CRU then Renater (NREN), since 2004
- Hosts software development projects for academic research and universities, accepts external collaborators
- Currently 3 068 hosted projects and 8 147 registered users
- Free/open source technologies :
subversion, git, Jenkins, Sonar, Nexus...
- Authentication with the French identity federation
- A complete set of functionalities :
 - ▶ link with Renater mailing lists server Sympa(*)
 - ▶ project websites
 - ▶ building, deploying and automating tools (Jenkins)
 - ▶ code quality (Sonar)
 - ▶ documentation management (Nuxeo)
 - ▶ ...

(*) Sympa is an open source mailing list server created by CRU and provided as a service by CRU then Renater since 1997

CeCILL licences

- <http://www.cecill.info/index.en.html>
- Developed by CEA, CNRS and INRIA since 2005
- CeCILL, CeCILL-B and CeCILL-C
- CeCILL is fully compatible with GNU GPL

Today Free Software is important in the scientific community as well as in administrations and in the entreprise. Nevertheless, the use of licenses created in the US, such as the GNU General Public License raises some legal issues. These issues may lead to uncertainties that may prevent some companies and organisations to contribute Free Software.

To provide a better legal safety while keeping the spirit of these licenses, three French public research organisations, CEA, CNRS and Inria launched a projet to write Free Software licenses conforming to French law.



- PLUME project (2006-2013)

Promote economical, Useful and Maintained softwarE For the Higher Education And THE Research communities

- <http://www.projet-plume.org/>
- Initiated and supported by CNRS
- Share software expertise in academics
- Improve the visibility of software produced in research laboratories
- Community building
- Main results :
 - ▶ The portal provides a software catalogue (mainly FLOSS)
 - ▶ 1 000 software description cards, with several keyword levels
 - ▶ about 300 other documents and pages
 - ▶ publication procedure, classified in 18 sections
 - ▶ 2 200 members, 950 contributors
 - ▶ Actions : topic workshops and training sessions oriented to software development (tools, dissemination, IPR & licenses...) or targeting community building around software interests, cross-disciplines experience sharing...



- PLUME and research software

- Several research units asked for research software descriptions
- 3 pilots (ICJ, LAAS, LIGM), extended progressively
- (fr) RELIER <http://www.projet-plume.org/relier>
- (fr) https://projet-plume.org/fiches_dev_ESR
- (en) PLUME-FEATHER <http://www.projet-plume.org/en>
- keywords include labs and head institutions
- **Results :**
 - ▶ more than **350** research software descriptions, 1/3 also in English
 - ▶ actions on software dissemination :
 - ▶ (fr) <https://projet-plume.org/patrimoine-logiciel-laboratoire>
 - ▶ documentation, guidelines, 10 workshops...
 - ▶ work on legal issues (French IPR law and licenses)

- **PRESOFT :**
Preservation for REsearch SOFTWARE
- Projet CNRS/IN2P3, 2017-18
- IdGC, CC-IN2P3, LIGM
- Software management Plan (SMP)
- Results :
 - Template
 - Firsts SMPs
 - Analysis in progress

PRESOFT : Preservation for REsearch SOFTWARE Pérennisation de logiciels de la recherche

Vincent Breton Teresa Gomez-Diaz Geneviève Romier
CNRS/IdGC CNRS/LIGM CNRS/CC-IN2P3

Objectif : étudier l'implémentation de plans de gestion de logiciels
dans les unités de recherche afin d'améliorer leur pérennisation.

Le projet PRESOFT

Projet SMP
Dates : janvier 2017 - décembre 2018
Partenaires : deux unités IN2P3 (IdGC, CC-IN2P3) et un laboratoire (LIGM)

Objectifs scientifiques

- Développer des pratiques et modèles adaptés pour les Plans de Gestion de Logiciels ou Software Management Plans (SMP)
- Étudier la réalisation de plans de gestion
- Étudier les obstacles, la faisabilité, les conditions d'adoption
- Étudier l'impact
- Publier les nouveaux résultats

Résumé

- Janvier 2017 : mise en place initiale des modèles et procédures
- Juin 2017 - décembre 2017 : réalisation des plans de gestion de logiciels
- Recueil d'expériences, des utilisations du modèle proposé
- 2018 : évaluer l'impact dans les laboratoires et les projets

Motivation

Data Management Plan (DMP)

- Regroupe les programmes de Recherche - IDRS, H2020...
- Document formel écrit sur les données
- Contenu variable selon les exigences du financement
- Modèle de routine, en document constant
- L'objectif est souvent un seul document ou données
- Exemple : le programme cadre de recherche européenne Horizon 2020
- DMP se rapporte au 2015, obligatoires depuis 2016

Software Management Plan (SMP)

- Document centré sur le logiciel
- Document déjà demandé dans certains appels d'offres (IdGC)
- Pas de document opérationnel/déterminé a priori
- Le projet...

Plan de gestion du logiciel ou SMP

- Document écrit pour l'implémentation du logiciel
- Permet de mieux gérer le cycle de vie du logiciel
- Peut être adapté au projet
- Peut être adapté aux besoins de l'équipe et des chercheurs
- Peut être évolutif, de planification
- Permet d'identifier les étapes, les actions, les responsabilités
- Peut être utilisé pour la production scientifique
- Il ne se confond pas avec un plan de développement

Bénéfices du SMP

- Un seul document de référence
- Un seul fichier dans une équipe
- Gestion des versions (éviter le conflit de gestion)
- Identification de ce qui marche bien
- La connaissance des producteurs scientifiques
- Une meilleure gestion des productions, des données
- Le lien entre la science et une politique d'équipe, de collaboration, d'innovation

Cycle de vie du logiciel de la recherche



Procédure

- Les étapes suivent le cycle de vie du logiciel
- Objectifs du SMP : Planifier, Produire, Maintenir, Archiver, Diffuser, Utiliser
- Organiser les nouvelles informations
- Développer les nouvelles pratiques
- Mettre à jour le document à l'issue de la phase de planification

Références

- Article de Legrand : opérations juridiques et de politique scientifique dans la production de logiciels
- T. Gomez-Diaz, 2016 : Rôle des logiciels dans la recherche scientifique et la production de logiciels
- Free software, Open source software, Services, et autres présentations including a procédure pour le logiciel scientifique
- Article de Gomez-Diaz, T. Gomez-Diaz, septembre 2016 : présentation de la Conférence ESE, mai 2016
- Research Software Sustainability Report on Knowledge Software and Science, février 2016
- The Software Sustainability Institute, Checklist for a Software Management Plan, 2016



Mots clés : logiciel de la recherche, pérennisation, préservation, accès libre, reproductibilité, science ouverte.

Journées JDEV, Marseille, 4-7 juillet 2017

http://devlog.cnrs.fr/_media/jdev2017/poster_jdev2017_presoft_teresa_gomezdiaz.pdf

FG-SOL : Research Software services

Prototype level, ongoing work

Services :

- Access
- Consulting
- Preservation
- Testing&Computing

... to make research software discoverable,
accessible, (re)usable, which will lead to
better research.

Research software services

FG-SOL : access & preservation for research software

Teresa Gomez-Diaz
CNRS/LIGM

Vincent Legoll
CNRS/IPHC

Jérôme Pansanel
CNRS/CREATIS

Sorina Pop
CNRS/CC-IN2P3

Geneviève Romier
CNRS/CC-IN2P3

***Vision:** to create the necessary environment where research software can be discovered, studied and re-used by other researchers in order to validate and reproduce published results, which will produce new knowledge.*

Résumé

collaborative design: CNRS, ESI, CNRS, University of Edinburgh

- user driven design approach
- user driven design approach
- user driven design approach
- user driven design approach
- user driven design approach

Access

• based on open existing components: ESI, Apache, R, etc.

• providing software descriptions, with links to publications and data

• providing metadata, search interface

Preservation

• based on existing components: ESI, Apache, R, etc.

• based on existing components: ESI, Apache, R, etc.

• based on existing components: ESI, Apache, R, etc.

Consulting

• software sustainability: ESI, CNRS, University of Edinburgh

• software sustainability: ESI, CNRS, University of Edinburgh

• software sustainability: ESI, CNRS, University of Edinburgh

SOL2020

March 2016

• based on existing components: ESI, Apache, R, etc.

• based on existing components: ESI, Apache, R, etc.

• based on existing components: ESI, Apache, R, etc.

Testing & Computing

• user interface for ESI, Apache, R, etc.

• user interface for ESI, Apache, R, etc.

• user interface for ESI, Apache, R, etc.

FG-SOL

with an open source

Aims

- to implement SOL2020 research software services
- to implement SOL2020 research software services
- to implement SOL2020 research software services

Technical Architecture

• based on ESI, Apache, R, etc.

• based on ESI, Apache, R, etc.

• based on ESI, Apache, R, etc.

France Grilles

• French National Grid Initiative

• French National Grid Initiative

• French National Grid Initiative

VIP architecture adopted by FG-SOL

VIP: Virtual Imaging Platform

Developed by CNRS/ITIS

Keywords: infrastructures, open science, research software, reproducibility, accessibility, free/open access.

EGI Conference 2017, Catania, 9-12 May 2017
<https://indico.egi.eu/indico/event/3249/contribution/1>

Thanks for your attention !

Any questions ?