



## Overview of Higher Education and Research Forges

Daniel Le Berre (University of Artois/CNRS)

Open Science Committee (CoSO)

Source Code and Software College

WG2 Tools and technical and social best practices

National contact point for HER forges

EOSC France 2024: Événement National Tripartite - Sept. 12, 2024

Context

## Context of the presentation: software forges

### A software forge: what are we talking about?

- ▶ Tools for collaborative development (version control, issues and contributions management, forum, project management, quality assurance, etc)
- ▶ Designed for software, can also be used for managing publications and data

## Context of the presentation: software forges

### A software forge: what are we talking about?

- ▶ Tools for collaborative development (version control, issues and contributions management, forum, project management, quality assurance, etc)
- ▶ Designed for software, can also be used for managing publications and data

### Many open source self hosted solutions available

- ▶ Easy installation and maintenance on commodity hardware for basic functionalities
- ▶ More complicated installation et maintenance on more demanding hardware for advanced functionalities (continuous intégration, continuous deployment, quality assurance, etc.)

### Why is it important?

- ▶ Domination of a US commercial actor (GitHub), a developer social network
- ▶ Risk for researchers to rely on such private, non European solution for their code
- ▶ Many self hosted solutions installed by institutions in the non coordinated way

# Context of the presentation : software @ Open Science Committee (CoSO)

## Source Code and Software College @ CoSO

The college was nominated in April 2022 by HER ministry after an open call for interest.  
Head: Roberto Di Cosmo et François Pellegrini

## Five working groups

- ▶ WG 1: Identification and highlight of software production in HER
- ▶ **WG 2: Tools and technical and social best practices**
- ▶ WG 3: Valorisation and sustainability
- ▶ WG 4: Skills Network / Charter / National and International
- ▶ WG 5: Recognition and careers

The report: [made available of May 16, 2023](#)

*Higher Education and Research Forges in France - Definition, uses, limitations encountered and needs analysis*

# Higher Education and Research Forges in France

## Broad consultation

Result of an interaction between the members of the WG with many colleagues (Research Software Engineers networks, Software Engineering Researchers networks but also many people involved in the forges of their institution or their community)

## Historique

05/16/2023 V1 39 forges  
06/04/2023 V2 minor spelling fixes  
09/06/2023 V3 40 forges, English version 1  
11/15/2023 V4 67 forges, English version 2  
01/25/2024 V5 73 forges  
05/15/2024 V6 E4S, English version 3



Report (HAL)



Git repository (IN2P3)

## National contact point for HER forges

Appointment of a national contact point for HER forges within the CoSO source code and software college on May 20, 2024.

His role:

- ▶ Referent and source of information
- ▶ Analysis of the specific needs of HER
- ▶ Federation of staff and collaborative development
- ▶ Functional mission without dedicated budget

Contact: `contact-national-forges@groupes.renater.fr`

### First actions

- ▶ Event with GitLab Inc. to learn how to contribute to GitLab and present to GitLab Inc. some use cases specific to HER
- ▶ Public national forum about software forges in HER  
`forum-national-forges@groupes.renater.fr` (82 members Sept. 2024)

## Forges used in HER



## Various options of forges available

**Commercials** github.com, gitlab.com, bitbucket

**FLOSS communities** Apache, Eclipse, FSF, OW2, etc

**National (HER)** SourceSup (RENATER)

**Self-Hosted** 73 public forges available

## Various options of forges available

Commercials github.com, gitlab.com, bitbucket

FLOSS communities Apache, Eclipse, FSF, OW2, etc

National (HER) SourceSup (RENATER)

Self-Hosted 73 public forges available

A wide choice of free self-hosted solutions available

Forgejo, Gerrit, Gitea, GitLab, Gogs, Redmine, SourceHut, Trac, Tuleap, ...

## Various options of forges available

Commercials github.com, gitlab.com, bitbucket

FLOSS communities Apache, Eclipse, FSF, OW2, etc

National (HER) SourceSup (RENATER)

Self-Hosted 73 public forges available

A wide choice of free self-hosted solutions available

Forgejo, Gerrit, Gitea, GitLab, Gogs, Redmine, SourceHut, Trac, Tuleap, ...

One solution seems to be preferred

64 instances of GitLab

## What market shares for each solution?

According to the source of the registered public repositories on Software Heritage:

- ▶ **GitHub: 226M**
- ▶ GitLab.com: 5.3M
- ▶ Self hosted GitLab instances: 0,3M
- ▶ BitBucket: 2,6M
- ▶ SourceForge: 0,4M

Source: <https://archive.softwareheritage.org>

# The french national forge SourceSup (RENATER)

## Features

- ▶ Maintained by the national infrastructure RENATER
- ▶ Opened in 2004 using GForge software
- ▶ Migrated to FusionForge in 2009
- ▶ New features added in 2015 using side dedicated software
- ▶ 762 public projects, 5200 private projects and 13 000+ users in December 2022
- ▶ Access available by identity federation (Shibboleth) at the national level.

## Observation

The technical and functional evolution of SourceSup is decided by the RENATER management committee. No major evolution of SourceSup is planned to date.

## Point of attention

The national forge, which provided the HER community with widely used and requested tools, is moving away from the practices of a majority of developers.

# Self hosted forges in HER (73 public forges listed)

## Some good reasons to self host

- ▶ Possibility to install locally a “professional” forge with no license fee;
- ▶ With a low effort of installation and maintenance;
- ▶ For teaching or research (or both);
- ▶ By teams, labs, universities, institutes;
- ▶ Public or private/internal

## More importantly . . .

To have a tool adapted to your needs, while protecting your data.

## However . . .

- ▶ Integration into an existing information system can be complex (identity federation, shared file system, etc.).
- ▶ Advanced services (continuous integration, containers) require more expertise.

# Pros and cons of an ecosystem of self-hosted forges

## Sovereignty, organization, robustness, expertise

- ▶ Controlling where the data is stored
- ▶ Each structure manages its forge with its own policy
- ▶ The forge becomes a software catalog of the structure
- ▶ The forge can be a showcase of the structure's software production
- ▶ Unbreakable: decentralized like the internet
- ▶ Strong national expertise on these tools

# Pros and cons of an ecosystem of self-hosted forges

## Sovereignty, organization, robustness, expertise

- ▶ Controlling where the data is stored
- ▶ Each structure manages its forge with its own policy
- ▶ The forge becomes a software catalog of the structure
- ▶ The forge can be a showcase of the structure's software production
- ▶ Unbreakable: decentralized like the internet
- ▶ Strong national expertise on these tools

## Scope, maintenance, functionalities, expertise

- ▶ **Potentially limited scope of users**
- ▶ Duplication of installation and maintenance effort
- ▶ Services offered vary depending on the forges
- ▶ Dispersed software production, may be less visible
- ▶ Specific expertise difficult to find, not available everywhere



# The importance of being identified in a forge

## Creating an issue, participating in a discussion

The notion of external, guest account allows this without allowing to create projects.

**It is a kind of communication account.**

## Proposing a code contribution

To do this, you must have the right to create projects (to make a copy of the project, and modify the code to then propose your modifications).

**It is a real developer account.**

## The challenge

Facilitate contributions of all kinds, when the contributor wishes.

## The main problem of HER forges: their access

<b>Expectations</b>	<b>Means</b>	<b>Difficulties</b>
Open to all	Free creation of accounts and projects	Management, moderation, IP, spam: time-consuming activities
Open to the greatest number	Delegation of identification to third-party providers	Moderation, IP, spam: time-consuming activities
Open to international HER	Identity federation (eduGAIN, Orcid)	Shibboleth support in GitLab, IP
Open to national HER	RENATER identity federation	Shibboleth support in GitLab, IP
Restricted access	Invitation to existing projects by project members	Limits contributions, interaction with users
Closed access	Authentication by internal directory	No possibility of interaction

# About the predominance (88%) of GitLab instances in HER

## Advantages

- ▶ Free/Libre Open Source Software
- ▶ State-of-the-art features
- ▶ Easy training on a tool widely used in the community
- ▶ Contributors outside HER are in a known environment

## Points of attention

- ▶ Heavy reliance on software developed by a commercial company
- ▶ The conditions of use and available features may change over time

# Can we improve the current situation?

## Choose your forge according to your needs

- ▶ Local for the start of the project
- ▶ Academic (national or international) to broadcast in the academic world
- ▶ FLOSS community to go beyond the academic world
- ▶ **Commercial if no other solution is possible**

## Open existing forges using identity federation

- ▶ RENATER for national academics
- ▶ eduGAIN for international academics

## Train HER members on forge basic usage

- ▶ for all profiles
- ▶ in all disciplines

## Call to (European) action

We believe that our findings on forges in HER in France are also applicable to most countries in Europe.

### An academic forge for every researcher in Europe

- ▶ Provide a state-of-the-art academic forge to HER in Europe
- ▶ Accessible to people outside academia
- ▶ Which would foster collaboration inside Europe
- ▶ Which can also help building a European research software developer network

### An alternative to identity federation: forge federation

- ▶ ForgeFed specification based on activity pub (e.g. mastodon) allows forge federation
- ▶ ForgeFed Reference Implementation in ForgeJo
- ▶ GitLab implementation is based on a community effort: **academia can help making it happen**