



## Forges de l'ESR : état des lieux

Collège Codes sources et logiciels du Comité pour la science ouverte  
GT2 Outils et bonnes pratiques techniques et sociales  
Daniel Le Berre, co-pilote du GT2

Gitlab 2023 @CNRS, 29 juin 2023

## Contexte du rapport

### Collège codes sources et logiciels du CoSO

Le collège a été nommé en avril 2022 par le MESR suite à un AMI.

Co-pilotes : Roberto Di Cosmo et François Pellegrini

### Le collège est constitué de 5 groupes de travail

- ▶ GT 1 : Identification et mise en avant de la production logicielle de l'ESR
- ▶ **GT 2 : Outils et bonnes pratiques techniques et sociales**
- ▶ GT 3 : Valorisation et durabilité
- ▶ GT 4 : Réseau compétences / Charte / National et international
- ▶ GT 5 : Reconnaissance et carrières

### Le rapport : rendu public le 16 mai 2023

Fruit d'une interaction entre les membres du GT avec de nombreux collègues (Réseaux Calcul, DevLog, GDR GPL mais aussi de nombreuses personnes impliquées dans les forges de leur établissement ou de leur communauté)

# Pour cette présentation

## Focus sur le thème du jour

- ▶ Gitlab

## Présentation des problématiques communes

- ▶ L'accès à la forge est sans doute le critère le plus important en valorisation
- ▶ Les fonctionnalités offertes sont un critère de choix au départ du projet logiciel
- ▶ La robustesse et la pérennité sont des critères supplémentaires pour être une alternative crédible aux solution commerciales

## Pas de présentation

- ▶ de la forge nationale SourceSup
- ▶ des forges communautaires libres
- ▶ des forges commerciales

mais ces sujets peuvent être traités lors des questions.

# Un écosystème morcelé de 39 forges publiques auto-hébergées

## Les raisons

- ▶ Possibilité d'installer en local une forge libre "professionnelle" ;
- ▶ Avec un faible effort d'installation et de maintenance ;
- ▶ Pour l'enseignement ou la recherche (ou les deux) ;
- ▶ Par des équipes, des labos, des universités, des instituts ;
- ▶ Publiques ou privées/internes

## Mais surtout . . .

Pour avoir un outil adapté à ses besoins, en protégeant ses données.

# Omniprésence de gitlab

37 des 39 forges détectées sont des instances de gitlab.

## Avantages

- ▶ Environnement connu, formations possibles indépendamment de l'instance
- ▶ On peut facilement transférer un projet d'une instance à l'autre
- ▶ Fonctionnalités offertes très riches, de la version communautaire à la version payante

## Points d'attention

- ▶ Dépendance envers une solution technique libre gérée par une entreprise privée
- ▶ Pas de moyen facile de transférer des projets (avec tickets, MR, etc) sur un autre socle technique

# Fonctionnalités offertes par les 37 instance gitlab publiques de l'ESR

- ▶ Dépôts git publics avec MR (toutes)
- ▶ Usage de runners privés (24)
- ▶ Usage de runners partagés (20)
- ▶ Gitlab pages (16)
- ▶ Gestion des images de conteneurs (9)
- ▶ Gestion des artefacts (8)

# Les conséquences de la fragmentations

1. Accès difficile à ces différentes forges (dans l'ESR et hors ESR)
2. Fonctionnalités offertes variables
3. Maintenance/support de ces forges variables

# Les conséquences de la fragmentations

1. Accès difficile à ces différentes forges (dans l'ESR et hors ESR)
2. Fonctionnalités offertes variables
3. Maintenance/support de ces forges variables

## Visibilité, fonctionnalités

- ▶ Tout le monde ne connaît pas leur existence.
- ▶ Laquelle choisir quand on a le choix ?

## Créer un compte sur une forge est nécessaire pour ...

### Créer un ticket, participer à une discussion

La notion de compte externe, invité permet cela sans avoir le droit de créer des projets.

**C'est une sorte de compte de communication.**

### Proposer une contribution de code

Il faut pour cela avoir le droit de créer des projets (pour faire une copie du projet, et modifier le code pour ensuite proposer ses modifications).

**C'est un véritable compte de développeur.**

### L'enjeu

Faciliter les contributions de toutes natures, quand le contributeur le souhaite.

## Le principal problème des forges de l'ESR : leur accès

Attentes	Moyens	Difficultés
Accès ouvert à tous	Création libre des comptes et des projets	Gestion, modération, PI, spam : activités chronophages
Accès ouvert au plus grand nombre	Délégation de l'identification à des fournisseurs tiers	Modération, PI, spam : activités chronophages
Accès ouvert à l'ESR international	Fédération d'identité (eduGAIN, Orcid)	Support Shibboleth dans GitLab, PI
Accès ouvert à l'ESR national	Fédération d'identité RENATER	Support Shibboleth dans GitLab, PI
Accès restreint	Invitation sur les projets par des membres du projet	Limite les contributions, l'interaction avec les usagers
Accès fermé	Authentification par annuaire interne	Pas de possibilité d'interaction

## Exemples de politiques de forges de l'ESR

Forge	Politique d'ouverture
CIRAD <a href="https://gitlab.cirad.fr">gitlab.cirad.fr</a>	ESR National (RENATER)
CNRS <a href="https://src.koda.cnrs.fr">src.koda.cnrs.fr</a>	Fermé (Janus)
IN2P3 <a href="https://gitlab.in2p3.fr">gitlab.in2p3.fr</a>	ESR International (eduGAIN)
Huma-Num <a href="https://gitlab.huma-num.fr">gitlab.huma-num.fr</a>	Au plus grand nombre
IMT <a href="https://gitlab.ev.imtbs-tsp.eu">gitlab.ev.imtbs-tsp.eu</a>	Fermé
INRA <a href="https://forgemia.inra.fr">forgemia.inra.fr</a>	ESR National (RENATER)
INRIA <a href="https://gitlab.inria.fr">gitlab.inria.fr</a>	Restreint
IRD <a href="https://forge.ird.fr">forge.ird.fr</a>	ESR National + autres (RENATER, CRU)
IRSTEA <a href="https://gitlab.irstea.fr">gitlab.irstea.fr</a>	Restreint
SourceSup <a href="https://sourcesup.renater.fr">sourcesup.renater.fr</a>	ESR National + autres (RENATER, CRU)
U. Bordeaux <a href="https://gitub.u-bordeaux.fr">gitub.u-bordeaux.fr</a>	Fermé
U. Caen <a href="https://git.unicaen.fr">git.unicaen.fr</a>	ESR National (RENATER)
U. Gustave Eiffel <a href="https://gitlab.univ-eiffel.fr">gitlab.univ-eiffel.fr</a>	Fermé

# Évolution fonctionnelle des forges

## Constat

- ▶ Les forges comme SourceForge, GForge, FusionForge intégraient un forum et un espace de téléchargement
- ▶ Les forges type GitHub, GitLab sont centrées autour du code source
- ▶ Les forges permettent de facilement contribuer aux logiciels qu'elles hébergent

## Interprétation

La forge devient un réseau social de développeurs

## Conséquence

L'intérêt d'une forge croît avec son ouverture, son nombre d'utilisateurs.

## Point d'attention

Les forges auto-hébergées ne sont pas des réseaux sociaux de développeurs.

## Conséquence : utilisation importante des forges commerciales

Le besoin de collaborations favorise l'utilisation des forges commerciales

Le manque d'ouverture sur la société des forges actuelles de l'ESR force les logiciels à vocation internationale, avec interaction avec la société, d'utiliser des forges commerciales.

Exemples : lauréats du premier prix du logiciel libre de recherche

- ▶ Coq : <https://github.com/coq/coq>
- ▶ Scikit-learn : <https://github.com/scikit-learn/scikit-learn>
- ▶ Faust : <https://github.com/grame-cncm/faust>
- ▶ Gammapy : <https://github.com/gammapy/gammapy>

Point d'attention

L'utilisation des forges commerciales nécessite que les membres de l'ESR acceptent individuellement les conditions d'utilisation de ces plateformes.

# Maintenance/support

## Gestion de projets git

Installer une instance gitlab de base pour ses personnels est relativement simple.

## Intégration continue

Activer l'intégration continue est plus complexe :

- ▶ consommation de ressources plus importante ;
- ▶ utilisation de solutions d'orchestration pour s'adapter à la demande ;
- ▶ des enjeux de sécurité spécifiques.

## Mais aussi

- ▶ Problème de gestion de ressources (taille des dépôts git, contenu) ?
- ▶ Comment les sauvegarder ?
- ▶ Comment nettoyer les instance des contenus obsolètes ?

# Les problèmes à résoudre pour les forges existantes

## Ouverture vers un plus large public

Sur les multiples forges existantes dans l'Enseignement supérieur et de la Recherche, tant pour l'interaction entre personnels qu'avec la société, un problème essentiel à résoudre est la définition d'une politique d'accès cohérente qui maximise les interactions sans mettre en danger les infrastructures.

## Confiance de la communauté

Il existe un besoin de confiance dans une plateforme qui soit robuste et pérenne, sur laquelle les usagers puissent compter dans le temps pour diffuser leurs travaux de recherche, quelle qu'en soit la forme.

## Maintenance mutualisée

Être capable de maintenir une architecture de forge avec intégration continue sur la durée nécessite des ressources et des compétences particulières à mutualiser.

# Les problèmes à résoudre au niveau des projets

## Cycle de vie des projets (et visibilité) : vitrine vs outil

L'absence de contrôle (ou le faible contrôle) des créations de projets peut rendre leur gestion compliquée. Il est difficile, voire impossible, pour les administrateurs de déterminer quels projets doivent être conservés ou pas.

## Organisation des projets sur la forge

L'une des principales difficultés dans la gestion d'une forge est de savoir comment organiser les projets.

## Conclusion

De nombreuses instances de gitlab disponibles dans l'ESR

- ▶ Facilite la formation des personnels
- ▶ Souveraineté des projets hébergés
- ▶ De nombreuses compétences sur le territoire

Un problème récurrent : l'accès à ces instances est limité

La forge ne peut pas jouer son rôle de créateur de communauté.

Besoin d'ouverture

Pour faire de la valorisation basée sur du contenu libre, il faut une forge la plus ouverte possible (cf. retour d'expérience du rapport disponible sur une forge).

Besoin de mutualiser nos efforts

Donner accès à des forges robustes, pérennes, et dotées de fonctionnalités équivalentes aux forges commerciales nécessite sans doute une mutualisation des efforts.