



Les forges logicielles : une question de souveraineté et de visibilité

Daniel Le Berre

Comité pour la science ouverte

Collège codes sources et logiciels

GT2 Outils et bonnes pratiques techniques et sociales

JSO CNRS 2023, 22 novembre 2023

Contexte

Contexte de cette présentation : les forges logicielles

Une forge logicielle, de quoi parle-t-on ?

- ▶ Outil de développement collaboratif (gestion de versions, de tickets, de contributions, forum, gestion de projets, assurance qualité, etc)
- ▶ Pour le logiciel au départ, mais aussi utilisé pour les données et les publications

Contexte de cette présentation : les forges logicielles

Une forge logicielle, de quoi parle-t-on ?

- ▶ Outil de développement collaboratif (gestion de versions, de tickets, de contributions, forum, gestion de projets, assurance qualité, etc)
- ▶ Pour le logiciel au départ, mais aussi utilisé pour les données et les publications

De nombreuses solutions libres auto-hébergeables disponibles

- ▶ Installation et maintenance facile avec des ressources matérielles modestes pour les besoins de base
- ▶ Installation et maintenance plus complexe avec des ressources matérielles plus gourmandes pour les fonctionnalités avancées (intégration continue, déploiement continu, assurance qualité)

Contexte de cette présentation : le logiciel au CoSO

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

Forges de l'Enseignement supérieur et de la Recherche - Définition, usages, limitations rencontrées et analyse des besoins

État des lieux des forges de l'ESR

Consultation large

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)

Historique

16/05/2023 V1 39 forges
04/06/2023 V2 correction de typos
06/09/2023 V3 40 forges, version anglaise
15/11/2023 V4 67 forges



Rapport (HAL)



Dépôt git (IN2P3)

Les forges utilisées dans l'ESR

Les différentes offres de forges disponibles pour l'ESR

Commerciales github.com, gitlab.com, bitbucket

Communautaires libres Apache, Eclipse, FSF, OW2, etc

Nationale ESR SourceSup (RENATER)

Auto-hébergées 67 forges publiques disponibles

Les différentes offres de forges disponibles pour l'ESR

Commerciales github.com, gitlab.com, [bitbucket](https://bitbucket.com)

Communautaires libres Apache, Eclipse, FSF, OW2, etc

Nationale ESR SourceSup (RENATER)

Auto-hébergées 67 forges publiques disponibles

Un large choix de solutions auto-hébergeables libres disponible

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

Les différentes offres de forges disponibles pour l'ESR

Commerciales [github.com](#), [gitlab.com](#), [bitbucket](#)

Communautaires libres Apache, Eclipse, FSF, OW2, etc

Nationale ESR SourceSup (RENATER)

Auto-hébergées 67 forges publiques disponibles

Un large choix de solutions auto-hébergeables libres disponible

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

Une solution semble privilégiée

58 instances de GitLab

Quelles parts de marché pour chaque solution ?

Selon les données de Software Heritage, en nombre de dépôts/projets enregistrés :

- ▶ **GitHub : 193M**
- ▶ GitLab.com : 3,9M
- ▶ Autres instances de GitLab : 0,2M
- ▶ BitBucket : 2,5M
- ▶ SourceForge : 0,4M

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

La forge nationale SourceSup (RENATER)

Caractéristiques

- ▶ Opérée par RENATER
- ▶ Ouverte en 2004 en se basant sur GForge
- ▶ Migration à FusionForge en 2009
- ▶ Ajout de nouvelles fonctionnalités à l'aide d'outils dédiés en 2015
- ▶ 762 projets publics, 5200 projets privés et 13 000+ utilisateurs en décembre 2022
- ▶ Accès par fédération d'identité RENATER, comptes réseau universel (CRU)

Observation

L'évolution technique et fonctionnelle de SourceSup est décidée par le comité de direction de RENATER. Aucune évolution majeure de SourceSup n'est prévue à ce jour.

Point d'attention

La forge nationale, qui fournissait à la communauté ESR des outils largement utilisés et demandés, s'éloigne des pratiques d'une majorité de développeurs.

Les forges auto-hébergées de l'ESR (67 forges publiques répertoriées)

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.

Points d'attention

- ▶ L'intégration dans un système d'information existant peut être complexe (fédération d'identité, système de fichier partagé, etc.).
- ▶ Les services avancés (intégration continue, conteneurs) demandent plus d'expertise.

Forges publiques du CNRS dans le rapport

- ▶ src.koda.cnrs.fr
- ▶ forge.in2p3.fr
- ▶ gitlab.in2p3.fr
- ▶ plmlab.math.cnrs.fr
- ▶ gitlab.mbb.cnrs.fr
- ▶ gitlab.huma-num.fr

CC IN2P3

CC IN2P3

CC IN2P3

CC IN2P3

Pros et cons d'un écosystème de forges auto-hébergées

Souveraineté, organisation, robustesse, expertise

- ▶ On maîtrise où se trouvent les données
- ▶ Chaque structure gère sa forge avec sa propre politique
- ▶ La forge devient un catalogue de logiciels de la structure
- ▶ La forge peut être une vitrine de la production logicielles de la structure
- ▶ Incassable : décentralisé comme internet
- ▶ Expertise nationale forte sur ces outils

Pros et cons d'un écosystème de forges auto-hébergées

Souveraineté, organisation, robustesse, expertise

- ▶ On maîtrise où se trouvent les données
- ▶ Chaque structure gère sa forge avec sa propre politique
- ▶ La forge devient un catalogue de logiciels de la structure
- ▶ La forge peut être une vitrine de la production logicielles de la structure
- ▶ Incassable : décentralisé comme internet
- ▶ Expertise nationale forte sur ces outils

Périmètre, maintenance, fonctionnalités, expertise

- ▶ Périmètre d'utilisateurs potentiellement limité
- ▶ Duplication de l'effort d'installation et maintenance
- ▶ Services offerts variables selon les forges
- ▶ Production logicielle dispersée, peut être moins visible
- ▶ Expertise spécifique difficile à trouver, pas disponible partout

A propos de la prédominance (87%) des instances GitLab dans l'ESR

Les avantages

- ▶ Logiciel libre
- ▶ Fonctionnalités à l'état de l'art
- ▶ Facilité de formation à un outil largement utilisé dans la communauté
- ▶ Les contributeurs hors ESR sont dans un environnement connu

Les points d'attention

- ▶ Forte dépendance sur un logiciel développé par une société commerciale
- ▶ Les conditions d'utilisation, les fonctionnalités disponibles peuvent évoluer
- ▶ L'ESR n'est pas actuellement sa priorité (cf problèmes Shibboleth)
- ▶ L'auto-hébergement n'est pas actuellement sa priorité (cf tarifs *SaaS* et *Self Managed* identiques)

Choisir sa forge

Comment choisir sa forge : souveraineté

Les besoins de souveraineté

- ▶ Respecter la législation qui s'applique (hébergement pour l'Europe en Europe)
- ▶ Disposer en permanence, en autonomie et en totale indépendance des projets
- ▶ Maîtriser quels projets sont diffusés
- ▶ Contrôler le moissonnage des projets
(utilisé pour entraîner des IA génératives de code par exemple)

Points d'attentions

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

L'enjeu de l'interaction avec une communauté

De nombreuses formes de contributions sont possibles

- ▶ Reporter un problème, un bug
- ▶ Proposer des améliorations
- ▶ Proposer de nouvelles fonctionnalités
- ▶ Proposer de la documentation, des tutoriels
- ▶ Aider les autres utilisateurs
- ▶ Contribuer au code
- ▶ ...

Toutes les contributions nécessitent un compte sur la forge qui héberge le projet

L'enjeu de l'interaction avec une communauté

De nombreuses formes de contributions sont possibles

- ▶ Reporter un problème, un bug
- ▶ Proposer des améliorations
- ▶ Proposer de nouvelles fonctionnalités
- ▶ Proposer de la documentation, des tutoriels
- ▶ Aider les autres utilisateurs
- ▶ Contribuer au code
- ▶ ...

Toutes les contributions nécessitent un compte sur la forge qui héberge le projet

Les contributeurs sont rares

Il faut éviter les freins à la contribution :

- ▶ Empêcher la création de projets, donc les contributions de code
- ▶ Demander un compte à l'aide d'un garant
- ▶ Devoir créer un compte sur une forge

Comment choisir sa forge : visibilité

Une forge est désormais un réseau social de développeurs

Pour augmenter sa visibilité, les chances d'avoir des contributeurs, aller sur la forge avec le plus d'utilisateurs.

La fédération d'identité : un accès à des réseaux de contributeurs

- ▶ RENATER : réseau académique national
- ▶ eduGAIN : réseau académique international

Sortir du monde académique

- ▶ Rejoindre une communauté de logiciels libres (Apache, Eclipse, FSF, OW2, etc) : pour les projets mûrs, de certaines thématiques
- ▶ Rejoindre une forge commerciale : ouverte à tous

Comment choisir sa forge : les fonctionnalités

Utiliser des outils connus des développeurs

- ▶ Désintérêt de la forge nationale SourceSup
- ▶ Augmentation du nombre d'instances de GitLab auto-hébergées
- ▶ Utilisation massive de GitHub

Bénéficier de fonctionnalités avancées, à la carte

- ▶ Intégration continue
- ▶ Pages
- ▶ Assurance qualité

Comment choisir sa forge : le support

Longévité de la forge

- ▶ *Too big to fail?* choix de GitHub souvent par défaut
- ▶ Les forges commerciales ne sont pas éternelles (cf. Google code)
- ▶ Les forges communautaires non plus (cf. Gitorious)

Disponibilité, support

- ▶ Forges communautaires libres
- ▶ Forges commerciales
- ▶ Grandes instances auto-hébergées
- ▶ Forge nationale

En résumé

On peut/doit choisir sa forge

Choisir sa forge selon ses besoins

- ▶ Une forge nationale avec un socle technique spécifique
- ▶ De nombreuses forges auto-hébergées disponibles dans l'ESR
- ▶ Des forges communautaires libres pour certaines thématiques
- ▶ Le recours par défaut aux forges commerciales pose de nombreuses questions de souveraineté

Un choix n'est pas définitif

La plupart des forges sont basées sur git : interopérabilité pour le partage du code/données/article.

On peut changer de forge selon l'évolution des besoins !

Forge commerciale → *Auto-Hébergée*, *Forge Nationale* : gain de souveraineté

Auto-hébergée → *Forge commerciale* : gain de visibilité (hors académique)

Auto-Hébergée → *Auto-Hébergée* : gain de visibilité, de fonctionnalités, de support

Le plus important : la bonne forge, au bon moment

Pendant la recherche : souveraineté

Il peut s'agir d'une forge interne, privée.

Choisir la forge la plus "proche" de son laboratoire.

Recherche publiée : visibilité

- ▶ Cible académique : choisir une forge permettant un accès RENATER ou eduGAIN
- ▶ Cible économique :
 - ▶ si une communauté libre existe sur la thématique, meilleure option
 - ▶ sinon forge commerciale

Dans un monde idéal

Une forge nationale (ou un réseau, une fédération de forges), souveraine, ouverte sur la société, offrant des fonctionnalités et un support comparable aux forges commerciales