

Publications & données de la recherche: quels services pour les chercheurs et quel rôle pour HAL?

Rencontres Huma-Num 2021 Nathalie Fargier, CCSD

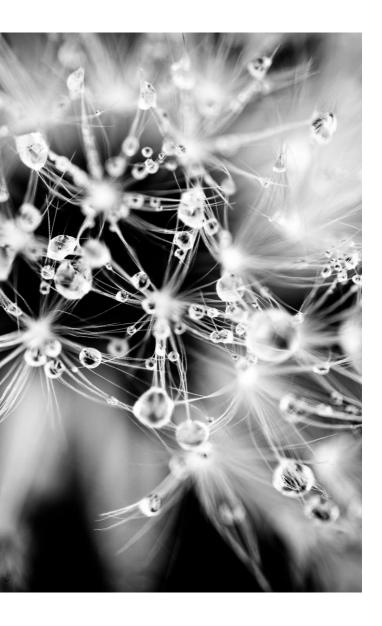












## Contexte

### Infrastructure de recherche dédiée à la Science Ouverte

- AO multidisciplinaire
- Portails (Documents, Disciplines, Institutions)
- Ressources techniques

HAL



- Modèle diamant
- HAL, arXiv, CWI, Zenodo
- (Open) peer review
- Maths, Informatique, SHS

Revues



- Gestion scientifique et administrative
- Site web personnalisé
- Versement dans HAL

Conférences

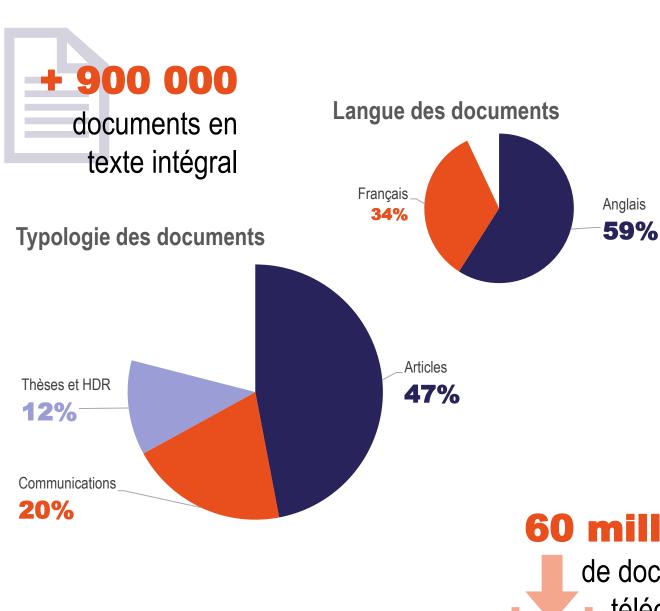


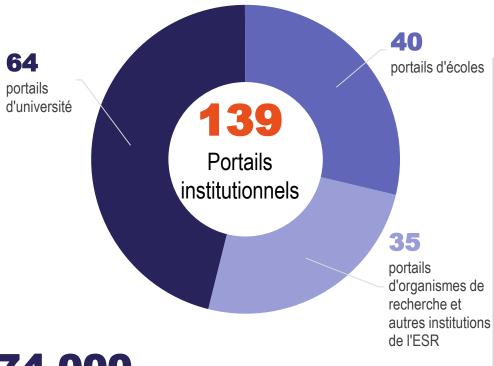


## Originalité de HAL

- ▶ Une dimension nationale
- Une couverture pluridisciplinaire
- Une plateforme qui héberge l'archive HAL et une centaine d'archives institutionnelles d'universités, organismes de recherche et grandes écoles
- Un système d'information avec une alimentation mutualisée
- Un engagement de qualité (gestion des **affiliations** et de données de référence)
- Différentes possibilités de **valorisation** (conférence, laboratoire, évènement, etc.)
- Un archivage à long terme des documents déposés
- Un écosystème (Episciences, SciencesConf)

### Chiffres clefs





74 000 comptes utilisateurs actifs

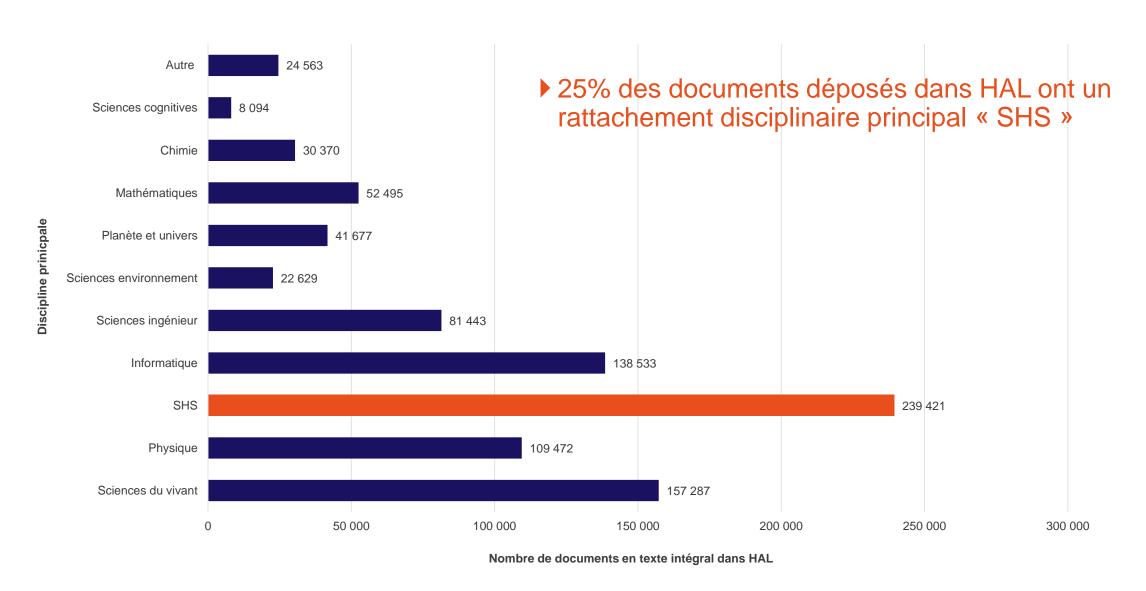
**60 millions** 

de documents téléchargés

1 080 personnes formées dont **86%** de chercheurs



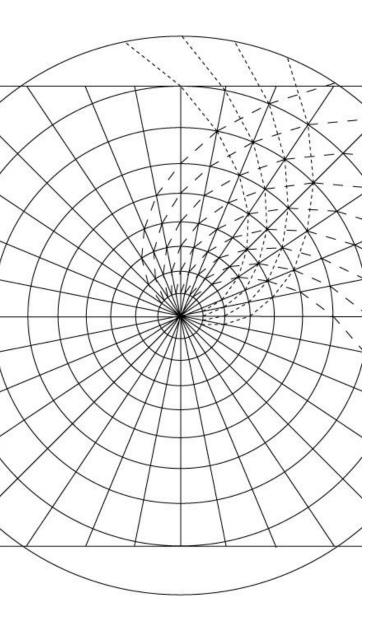
### Cartographie des disciplines





### Orientations stratégiques

- Technique : assurer la performance de l'infrastructure matérielle et logicielle | renforcer l'interopérabilité et la sécurité
- Services : mettre en œuvre un HAL NG | développer les overlay services | articuler publications, données de la recherche et logiciels
- Visibilité: être acteur/moteur de la dynamique en faveur de la SO | renforcer l'ancrage du CCSD dans son écosystème national et conforter son internationalisation
- ▶ Organisation : positionner le CCSD comme opérateur et animateur de réseaux | développer une démarche centrée Utilisateurs | aligner son modèle de pilotage, de financement de RH



## POC EOSC Pillar



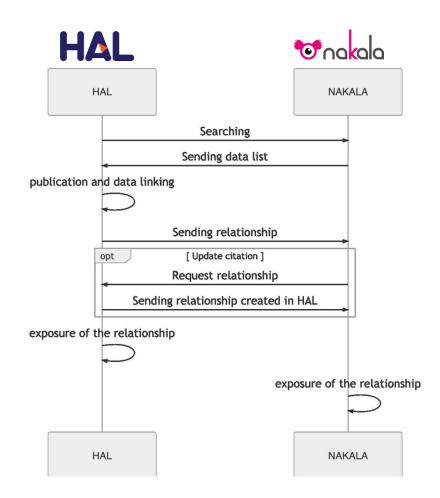
## EOSC Pillar Use Case

- WP6 « EOSC in action: Use cases and community-driven pilots » > Task 6.5 « FAIR principles in data life-cycles for Humanities »
- ▶ Partenaires : CCSD HumaNum
- Objectif : lier les publications SHS déposées dans HAL aux données présentes dans Nakala
- Calendrier : POC livré en préproduction en juin 2021



### Cas d'usages Workflow

- Je veux lier une publication dans HAL à un jeu de données dans Nakala
- Je veux visualiser/accéder aux données liées à une publication





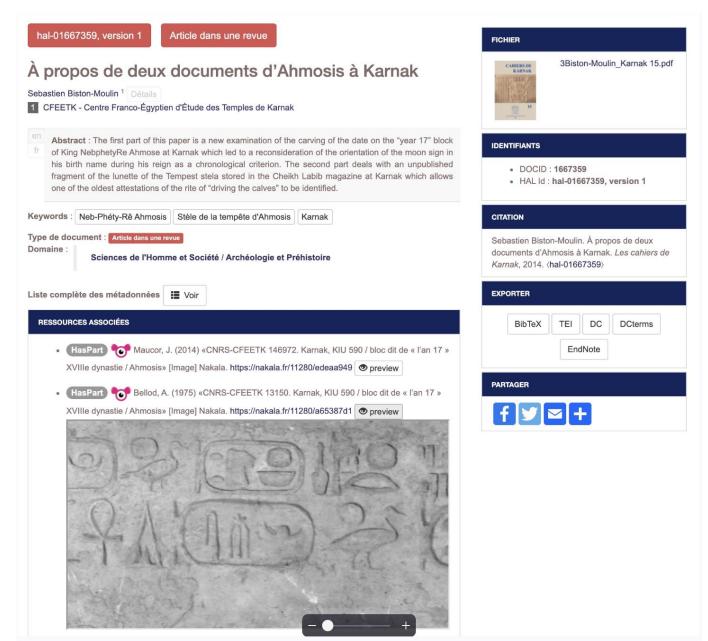
### Lier publications et données

- Disposer d'un compte HAL et des droits nécessaires sur les publications concernées
- Indiquer l'identifiant Nakala du jeu de données
- Chercher le jeu de données dans Nakala
- Utiliser un vocabulaire reconnu pour qualifier la relation entre la publication et les données



### Exposition des relations publications-données dans HAL

- Citation des ressources associées à une publication
- Possibilité de visualiser les ressources associées (player Nakala)





# Exposition dans les outils d'export de HAL

- ▶ Triplestore Data HAL (RDF)
- ► Entrepôt OAI-PMH (DC)





### **HALiance**

HAL Interoperability, Accessibility and Network Centric Environment



### Enjeux

- ▶ Répondre aux **attentes des utilisateurs** actuels et à venir
  - Simplification et facilitation du dépôt de texte intégral
  - Prise en compte des publications présentes dans d'autres SI
  - Continuum publications données de la recherche
  - Amélioration des outils de gestion et d'affiliation

## Principaux défis

- Des flux de métadonnées et de données en croissance constante qui nécessitent des mécanismes de synchronisation et d'automatisation
- Une collecte ciblée de publications à organiser
- Une articulation publications-données à industrialiser
- ▶ Des composants logiciels à migrer
- Une architecture modulaire à déployer

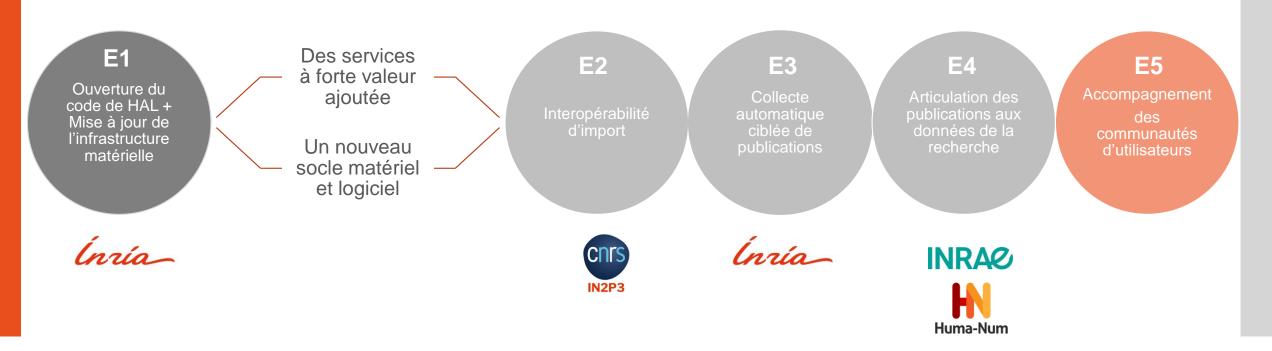
## Fiche d'identité du projet

- ▶ Etablissement coordinateur : CNRS
- Partenaires : INRAE, Inria, in2p3
- Opérateur : CCSD
- Durée : 60 mois
- Communauté d'utilisateurs de l'équipement : ensemble des communautés de l'ESR
- Axes de mutualisation avec l'Equipex+ Commons



### 1 équipement, 5 éléments

"Re-designing HAL's code and its core services and in aligning the latter with the recommendations and criteria of excellence defined in the framework of national and international initiatives that foster open science in order to offer researchers and their institutions a high-quality work environment."





### Elément 4: méthode mise en œuvre

#### Existant

- Possibilité de déposer des annexes à une publication et des données
- Possibilité d'ajouter le DOI d'un jeu de données disponible sur un autre entrepôt
- Workflow HAL Software Heritage
- ▶ 2 cas d'usage
  - SHS: Nakala
  - Sciences du Vivant : Data INRAE

### Attendus

- Mise en place de lien publications-données,
  affichage/exposition de ces liens et réciprocité entre entrepôts
- Simplification du dépôt pour les chercheurs avec HAL comme point d'entrée



### Verrous scientifiques et techniques

- Définition d'un noyau commun de métadonnées (DC) et de métadonnées spécifiques aux communautés de recherche
- Reconnaissance des chercheurs-déposants par HAL et par les entrepôts de données partenaires
- Mise en place d'un système d'authentification unique
- Mise en œuvre d'un mécanisme dépôt pour les jeux de données





- nathalie.fargier@ccsd.cnrs.fr
- contact@ccsd.cnrs.fr
- http://blog.ccsd.cnrs.fr
- @ ccsd\_fr et @hal\_fr
- Newsletter ActuHAL