

HAL : Plateforme d'archives ouvertes

Laurent Capelli
Directeur Technique CCSD

Plan de la présentation

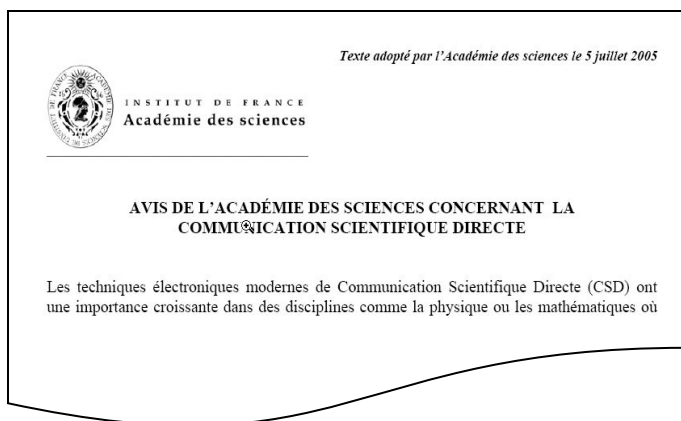
1. Le modèle français et la BSN
2. L'archive ouverte HAL
 1. La plateforme
 2. Portails et collections
 3. Identifiants
 4. Interopérabilité
 5. Référentiels
3. Questions, réponses

Le modèle français

- Protocole inter établissements
- Bibliothèque Scientifique Numérique

HAL, un service pour la recherche

- Un accord inter-établissements (2006-2008)
 - CEA, CEMAGREF, CGE, CIRAD, CNRS, CPU, IFREMER, INERIS, INRA, INRIA, INRETS, INSERM, Institut PASTEUR, IRD
- Un engagement du CNRS à offrir et développer en partenariat un service pour toute la recherche française



HAL, une plateforme commune et partagée

- Une convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL (2 avril 2013)



Une convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL (Hyper Articles en Ligne) a été signée, mardi 2 avril 2013, à l'Académie des sciences. Elle s'inscrit dans le cadre de la politique nationale en faveur de l'Open Access, du partage des résultats de la recherche, de leur diffusion et de leur conservation.

HAL, un outil avec les financeurs

- L'ANR incite les chercheurs à intégrer leurs publications dans le système HAL

Convention ANR/CCSD

➤ 24 juin 2014

- OpenAire
- HAL première archive compatible

La diffusion des publications scientifiques liées aux projets financés par l'ANR dans les archives ouvertes, en particulier HAL contribue à renforcer la visibilité et l'attractivité de la recherche française.

Elle peut aussi aider à simplifier le suivi et l'évaluation en évitant la saisie multiple des informations et en rendant les documents aisément accessibles à tous les chercheurs impliqués dans les différentes étapes de ces processus.

Dans le cadre de la préparation de son système d'information, l'ANR demande donc que, dans le respect des règles relatives à la propriété intellectuelle (propriété littéraire et artistique et propriété industrielle), et des règles de confidentialité inhérentes à des recherches, toutes les publications consécutives aux projets financés par elle soient d'ores et déjà intégrées par les chercheurs au système d'archives ouvertes HAL avec lequel elle collaborera.



14 Novembre 2007

La Bibliothèque Scientifique Numérique

- BSN 1 : **Acquisitions** des revues (courant) et des archives
- BSN 2 : Dispositif d'**accès** et d'hébergement
- BSN 3 : Dispositifs de **signalement**
- BSN 4 : **Archives ouvertes**
- BSN 5 : **Numérisation**
- BSN 6 : **Archivage pérenne**
- BSN 7 : **Édition scientifique**
- BSN 8 : **PEB**, Fourniture de documents, prêt
- BSN 9 : **Formation**, compétences et usages
- BSN 10 : **Données de la recherche**

BSN 4

- Comment assurer une bonne représentation de l'ensemble des acteurs dans le dispositif ?
- Comment stabiliser l'outil qui doit être portable et ne pas reposer sur quelques personnes ressources ?
- Quels types d'évolution technique pour asseoir dans la durée et assurer la croissance de Hal en lien avec le développement des archives institutionnelles ?
- Quelles relations avec les éditeurs ?
- Comment faire évoluer les relations avec les scientifiques ?



BIBLIOTHÈQUE SCIENTIFIQUE
NUMÉRIQUE
DIGITAL SCIENTIFIC
LIBRARY

HAL v3



[Accueil](#)
[Dépôt](#)
[Consultation](#)
[Recherche](#)
[Documentation](#)
[Mon espace](#)
[Privilèges](#)

Accueil

Archive ouverte HAL

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion d'articles scientifiques de niveau recherche, publiés ou non, et de thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

À l'attention du déposant

- Le dépôt doit être effectué en accord avec les co-auteurs et dans le respect de la politique des éditeurs
- La mise en ligne est assujettie à une modération, la direction de HAL se réservant le droit de refuser les articles ne correspondant pas aux critères de l'archive
- Tout dépôt est définitif, aucun retrait ne sera effectué après la mise en ligne de l'article
- Les fichiers textes au format pdf ou les fichiers images composant votre dépôt sont maintenant envoyés au CINES dans un contexte d'archivage à long terme.

À l'attention des lecteurs

- Dans un contexte de diffusion électronique, tout auteur conserve ses droits intellectuels, notamment le fait de devoir être correctement cité et reconnu comme l'auteur d'un document.

LES DERNIERS DOCUMENTS DE HAL

- Romain Reullion, Mathieu Leclaire, Jonathan Passerat-Palmbach. Model Exploration Using OpenMOLE - a workflow engine for large scale distributed design of experiments and parameter tuning. IEEE High Performance Computing and Simulation conference 2015, Jun 2015, Amsterdam, Netherlands. IEEE High Performance Computing and Simulation conference 2015, 2015, <http://hpcs2015.cisedu.info/>. <hal-01163457>
- Véronique Maume-Deschamps, Didier Rullière, Khalil Said. A risk management approach to capital allocation. 2015. <hal-01163180>
- Patrick Rambourg. "Des spécialistes du prêt-à-manger à Paris : cuisine publique et modes de restauration à la fin du Moyen Âge". Les établissements de restauration dans le monde, Oct 2009, Paris, France. L'Harmattan, Paris, 2012, p. 113-125, Les établissements de restauration dans le monde. <hal-01162992>
- Fabrice Gamboa, Thierry Klein, Agnès Lagnoux. Sensitivity analysis based on Cramér von Mises distance. 2015. <hal-01163393>

RECHERCHE

Termes de recherche

NOMBRE DE RESSOURCES

1 009 559

NOMBRE DE RESSOURCES AVEC FICHIERS

369 347

BLOG CCSD

SCIENCESCONF.ORG INSIDE (22/05/15)
 Discrètement mais sûrement, la plate-forme de gestion de congrès Sciencesconf.org élargit son audience. Notre précédent billet de janvier 2013 relevait 328 conférences ouvertes, aujourd'hui le compteur en affiche 1974. Depuis début 2015, ce sont en moyenne 90 nouvelles conférences qui sont créées par mois. La majorité utilise le module de dépôt : cela représente plus de ... Continuer la lecture de Sciencesconf.org inside →

[LIRE LE BILLET](#)

Le modèle français, une archive centralisée

- En 2001, **HAL** est développé par le Centre pour la Communication Scientifique Directe (CCSD)
- Archive nationale de dépôt de la recherche
- Multidisciplinaire
- Inter-établissements

HAL, centralisation apparente, mutualisation certaine ! une plate forme commune

- Garantir l'accessibilité au texte intégral
- Permettre l'indexation global du texte intégral contenu dans la base
- Préserver un niveau scientifique homogène et de qualité
- Offrir une meilleure visibilité à l'international
- Interconnecter avec les bases mondiales de référence
- Fournir des URL pérennes et ainsi permettre la citation numérique
- « Horodater » les dépôts et ainsi valider une antériorité scientifique.
- Enrichir automatiquement les référentiels institutionnels (Labintel, CRAC)
- Limiter la saisie des références à un seul système
- Effectuer l'archivage à long terme
- *Et ... limiter l'effort à une seule unité spécialisée : le model économique français pour les archives ouvertes et reconnu pour être le plus économique et le plus efficace*

Déposer dans HAL

- Modalités de dépôt
 - Identification simplifiée
 - Le compte créé est actif immédiatement
 - Contributeurs : auteurs, documentalistes, etc.
 - Niveau scientifique requis
 - **« Tout article qui, soumis à une revue à comité de lecture, serait envoyé à un rapporteur »**
 - Vérification avant mise en ligne
 - Vérification technique
 - Modération scientifique
 - Format des documents libre
 - Génération automatique de documents PDF pour une accessibilité universelle

Dépôt sans retrait ?

- Le retrait n'est pas possible par le contributeur : ceci contribue à le responsabiliser et le motive pour ne déposer qu'un travail de qualité
- Cela évite des tentatives d'utilisations scientifiques frauduleuses :
 - dépôt de résultats contradictoires pour prendre date avec retrait postérieur du dépôt scientifiquement en erreur, etc.
- On ne peut pas citer dans une publication scientifique un texte qui risque à tout moment d'être retiré par son auteur ! La stabilité des dépôts et des adresses est donc cruciale
- L'alternative au dépôt sans retrait est la possibilité de déposer de nouvelles versions
 - La dernière version est proposée par défaut au lecteur, mais celui-ci peut toujours accéder aux versions antérieurs

Questions juridiques

Vous avez besoin de l'accord de l'éditeur si vous voulez déposer la version numérisée du document édité.

En revanche vous pouvez déposer la version manuscrite de votre document si vous respectez ces règles :

- **AVANT publication** : l'auteur reste détenteur des droits sur son texte
- **APRES publication SOUS contrat** :
 - Sauf interdiction explicitement précisée dans le contrat, l'auteur a le droit de diffuser son document dans une archive institutionnelle électronique.
 - En cas de clause explicite autorisant l'exploitation électronique du document par l'éditeur : l'auteur doit demander l'autorisation de dépôt à l'éditeur.
- **APRES publication SANS contrat** : l'auteur a le droit de diffuser son document

HAL, citabilité électronique

- Un « ressource » dans l'archive ouverte c'est :
 - une ou plusieurs versions d'un même article comportant chacune
 - la notice descriptive
 - Incluant éventuellement des références bibliographiques
 - un ou plusieurs formats du même fichier (PDF, TeX, XML, DOC, etc.)
 - éventuellement des fichiers annexes
 - images, sons, vidéos, posters, présentation « ppt », etc.

- L'identifiant pérenne, « citable », référence un « container » qui abrite tous les composants de l'article scientifique
 - une URL simplifiée qui garantit au lecteur d'arriver toujours sur la version la plus récente de l'article et lui permet l'accès à l'historique
 - une URL stable ne dépendant pas du stockage physique du document

HAL, identifiants

hal.archives-ouvertes.fr/hal-00000001

hal.archives-ouvertes.fr/hal-00000001v2

hal.archives-ouvertes.fr/hal-00000001/document

halId_id

- Alignement avec des identifiants internationaux
 - DOI (**doId_id**), arXiv (**arxivId_id**), PubMed (**pubmedId_id**), ADS (**bibcodeId_s**)
- Relation avec les bases « partenaires »
 - Horizon (IRD, **irdId_id**), STAR (ABES, **nntId_id**), SAM (ensam, **ensamId_id**), etc.

La plate-forme HAL

- Typologies étendues

Publications

- Article dans des revues
- Communication dans un congrès
- Poster
- Ouvrage (y compris édition critique et traduction)
- Chapitre d'ouvrage
- Direction d'ouvrage, Proceedings
- Brevet
- Autre publication

Documents non publiés

- Pré-publication, Document de travail
- Rapport

Travaux universitaires

- Thèse
- HDR
- Cours

Données de recherche

- Image
- Vidéo
- Son
- Carte

La typologie de HAL est calquée sur celle de l'AÉRES (HCERES) permettant ainsi aux laboratoires et institutions d'extraire automatiquement des données exploitables directement dans les quadriennaux

La plate-forme HAL

- Formats et embargos liés à chaque fichier

Fichier	Format	Actions
<p>PC.pdf (625.69 Ko)</p> <p>Visibilité: immédiatement (2015-06-14)</p>	<p>Fichiers auteur</p> <ul style="list-style-type: none"> ✓ Les fichiers que je dépose sont des fichiers produits par l'(les) auteur(s) <p>Fichiers éditeur</p> <p>J'ai vérifié que cet éditeur fait partie de ceux qui autorisent le dépôt des fichiers éditeurs sur une archive ouverte J'ai obtenu l'accord explicite de cet éditeur pour ce dépôt particulier L'institution a financé les frais de publication pour que cet article soit en libre accès</p>	
<p>PC.pdf (625.69 Ko)</p> <p>Visibilité: immédiatement (2015-06-14)</p> <p>Commentaire:</p> <p>Type d'annexe:</p> <p>Annexe principale: Non</p>	<p>Données supplémentaires</p>	<p>✎ 🗑️</p>
<p>PC.pdf (625.69 Ko)</p> <p>Visibilité:</p> <p>Commentaire:</p>	<p>Les fichiers que je dépose sont des fichiers produits par</p> <p>Document (pdf, jpg, ...)</p>	<p>✎ 🗑️</p>
	<p>immédiatement (2015-06-14)</p> <p>dans 15 jours (2015-06-29)</p> <p>dans 1 mois (2015-07-14)</p> <p>dans 3 mois (2015-09-14)</p> <p>dans 6 mois (2015-12-14)</p> <p>dans 1 an (2016-06-14)</p> <p>à partir d'une date spécifique</p> <p>dans 2 ans (2017-06-14)</p>	

Supprimer tous les fichiers 🗑️

La plate-forme HAL

- Licences pour le dépôt

Licence

- Paternité  
- Paternité - Pas d'utilisation commerciale   
- Paternité - Pas de modifications   
- Paternité - Partage selon les Conditions Initiales   
- Paternité - Pas d'utilisation commerciale - Pas de modification    
- Paternité - Pas d'utilisation commerciale - Partage selon les Conditions Initiales    
- Marque du Domaine Public 
- CC0 - Transfert dans le Domaine Public  
- Licence Ouverte - etalab 
- Copyright (Tous droits réservés)
- Domaine public

La plate-forme HAL

- Champs multi-valués et multi-langues

Titre * *Vous pouvez renseigner le titre en plusieurs langues : choisir la langue et cliquer sur + pour ajouter une nouvelle langue*

My title anglais ▾ 🗑️

Le titre français ▾ 🗑️

allemand ▾ +

us-Titre *Vous pouvez renseigner le sous-titre en plusieurs langues : choisir la langue et cliquer sur + pour ajouter une nouvelle langue*

allemand
italien
espagnol
abkhaze
afar +

La plate-forme HAL

- Métadonnées automatiquement renseignées
 - depuis le PDF : GROBID
 - dans les images : IPTC, XMP
 - depuis les identifiants
 - internationaux : DOI, PubMed, ADS, arXiv
 - nationaux : IRD



Identifiants

Ajoutez l'identifiant DOI ou les identifiants de ce dépôt dans les bases arXiv, PubMed ou ADS

Recharger les métadonnées

10.1103/PhysRevB.67.184421	DOI ▾	↻	🗑️
	arXiv Id ▾	+	

Les métadonnées

- Fondée sur des référentiels communs

Titre

Forme auteur¹, IdHAL

1 - Structure de recherche, partenaire privé

- Identifiant : DOI, PubMed, ADS, bases partenaires, etc.
- Type de document
- Publication : nom de la revue
- Indexation : disciplines scientifiques
- Financeur : projet ANR, projet Européen
- Déposant : compte utilisateur CCSD

AURéHAL :

Accès Unifié aux Référentiels HAL

<https://aurehal.archives-ouvertes.fr>

- Auteurs (formes auteurs)
- Structures de recherche
- Revues
- Projets ANR
- Projets Européens
- Disciplines

AURéHAL, Caractéristiques

- Les référentiels peuvent être consultés, même sans authentification
- L'historique de chaque entrée est visible par toute personne authentifiée : date de création, de mise à jour
- Chaque entrée d'un référentiel permet de lancer une requête dans HAL pour consulter les dépôts liés
- Une mise à jour dans un référentiel génère automatiquement une mise à jour dans les dépôts liés

AURéHAL, Principes

- Alimentation
 - le dépôt dans HAL
 - les producteurs
 - AURéHAL
- Un administrateur de portail a les droits de création et modification sur tous les référentiels, à l'exception de celui des disciplines
- Toute personne qui a les droits « référent structure » peut modifier les informations de la structure de recherche dont il est référent

IdHAL, identifiant chercheur

- Identifiant **unique** d'un auteur sur HAL
URI : <https://cv.archives-ouvertes.fr/IdHAL>
- **Créé** par un auteur (lié à l'utilisateur connecté)
- ≠ de fusion des formes auteur
- Rassemble les formes auteur co-existant dans HAL (avec choix publication par publication)
- Choix d'une forme auteur « par défaut »

IdHAL, identifiant chercheur

- Éléments de l'identité numérique du chercheur
- Lien avec des identifiants auteur d'autres plateformes :
 - IdRef, arXivid, Orcid, ResearcherId, VIAF, ISNI
- et réseaux sociaux :
 - Twitter, LinkedIn, blog de recherche, ...

La plate-forme HAL

- Documents importés dans HAL liés à la source d'origine

IDENTIFIANTS

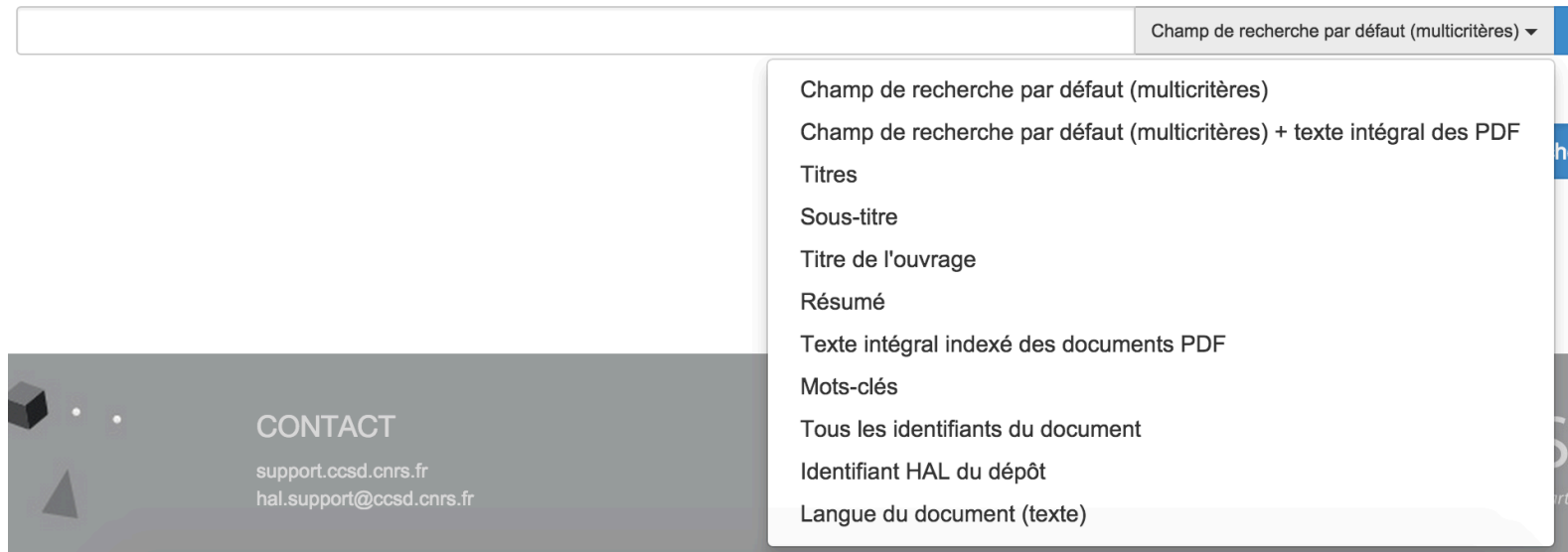
- DOCID : 1162059
- HAL Id : hal-00992192, version 2
- DOI : 10.1016/j.conengprac.2010.01.004
- ENSAM : <http://hdl.handle.net/10985/8155>

- Relation entre dépôt HAL

Relation	Identifiant HAL	Commentaire
<div style="border: 1px solid #ccc; padding: 5px; width: 150px;"> <ul style="list-style-type: none"> ✓ illustre est illustré par est requis par requiert est une partie de a comme partie est référencé par référence est un autre format de a pour autre format est conforme à </div>	<input type="text" value="Identifiant HAL"/>	<input type="text" value="Commentaire"/>
<input type="button" value="Annuler"/> <input type="button" value="↓ Déposer"/>		<input type="button" value="+"/>

La plate-forme HAL

- Consultation et recherche des dépôts
 - Solr
 - plus de 100 champs de recherche
 - recherche plein texte



The screenshot shows a search bar with a dropdown menu open. The dropdown menu lists the following search criteria:

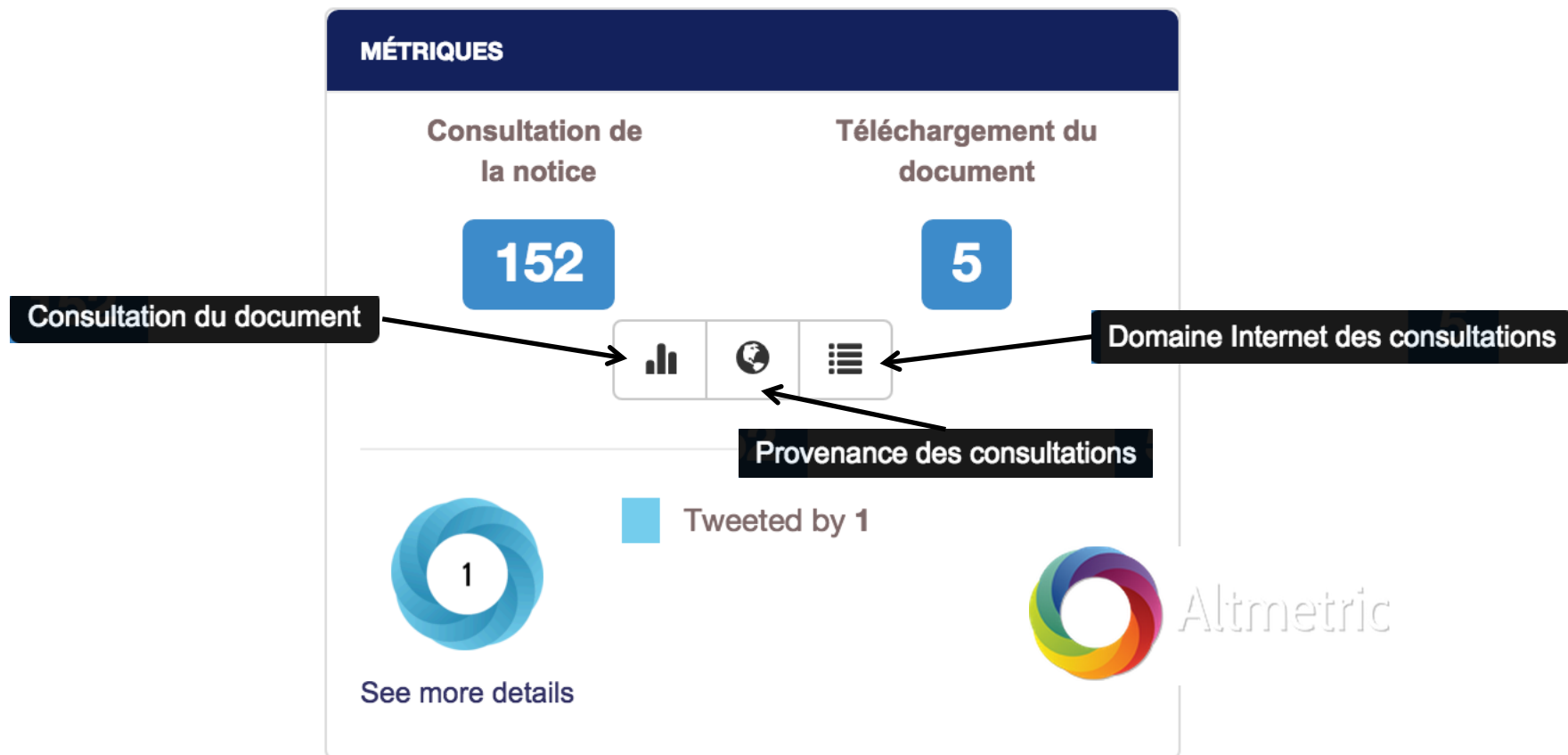
- Champ de recherche par défaut (multicritères)
- Champ de recherche par défaut (multicritères) + texte intégral des PDF
- Titres
- Sous-titre
- Titre de l'ouvrage
- Résumé
- Texte intégral indexé des documents PDF
- Mots-clés
- Tous les identifiants du document
- Identifiant HAL du dépôt
- Langue du document (texte)

Below the search bar, there is a contact information section:

CONTACT
 support.ccsd.cnrs.fr
 hal.support@ccsd.cnrs.fr

La plate-forme HAL

- Statistiques de consultation et d'usage



HAL, un format pivot

- Customisation des directives de la TEI
 - `teiHeader`, `bibliFull`, `bibliStruct`

- Transformations XSL
 - BibTeX
 - Dublin Core
 - Endnote

- Format import SWORD

Une appropriation facile

Dans une base de données commune, HAL

- des portails institutionnels personnalisés
 - intégration automatique des dépôts quelque soit l'origine du portail, en fonction de l'affiliation des auteurs
 - préservation de l'identité de l'établissement
 - adresses URL, graphisme, métadonnées spécifiques, etc.
- Plus de 3000 portails et collections
 - Toutes les grandes EPST, le CEA, des COMUE, de nombreuses universités et grandes écoles, etc.

The screenshot shows the HAL-Inria website interface. At the top, there are logos for UPMC (Université Paris Descartes) and Inria. The main navigation bar includes 'Accueil', 'Déposer', 'Consulter', 'Recherche', 'Services', 'Documentation', and 'tutoriels'. Below the navigation, there is a search bar and a welcome message: 'Bienvenue sur HAL-Inria'. A sidebar on the right contains a 'RECHERCHE' section with a search input field and a 'LIENS UTILES' section with links to documentation and support. At the bottom, there is a 'NOMBRE DE DOCUMENTS FULLTEXT' section.

A grid of logos for various ParisTech institutions. The logos are arranged in a grid-like fashion. From top to bottom and left to right, the logos include: AgroParisTech, Arts et Métiers ParisTech, Chimie Paris ParisTech, École des Ponts ParisTech, École Polytechnique ParisTech, ENSAE ParisTech, ENSTA ParisTech, ESCI ParisTech, HEC Paris, Institut d'Optique Graduate School, and Mines ParisTech. The logos are diverse in color and design, representing different scientific and technical fields.

HAL, des collections

Des "collections" pour présenter des sélections de documents

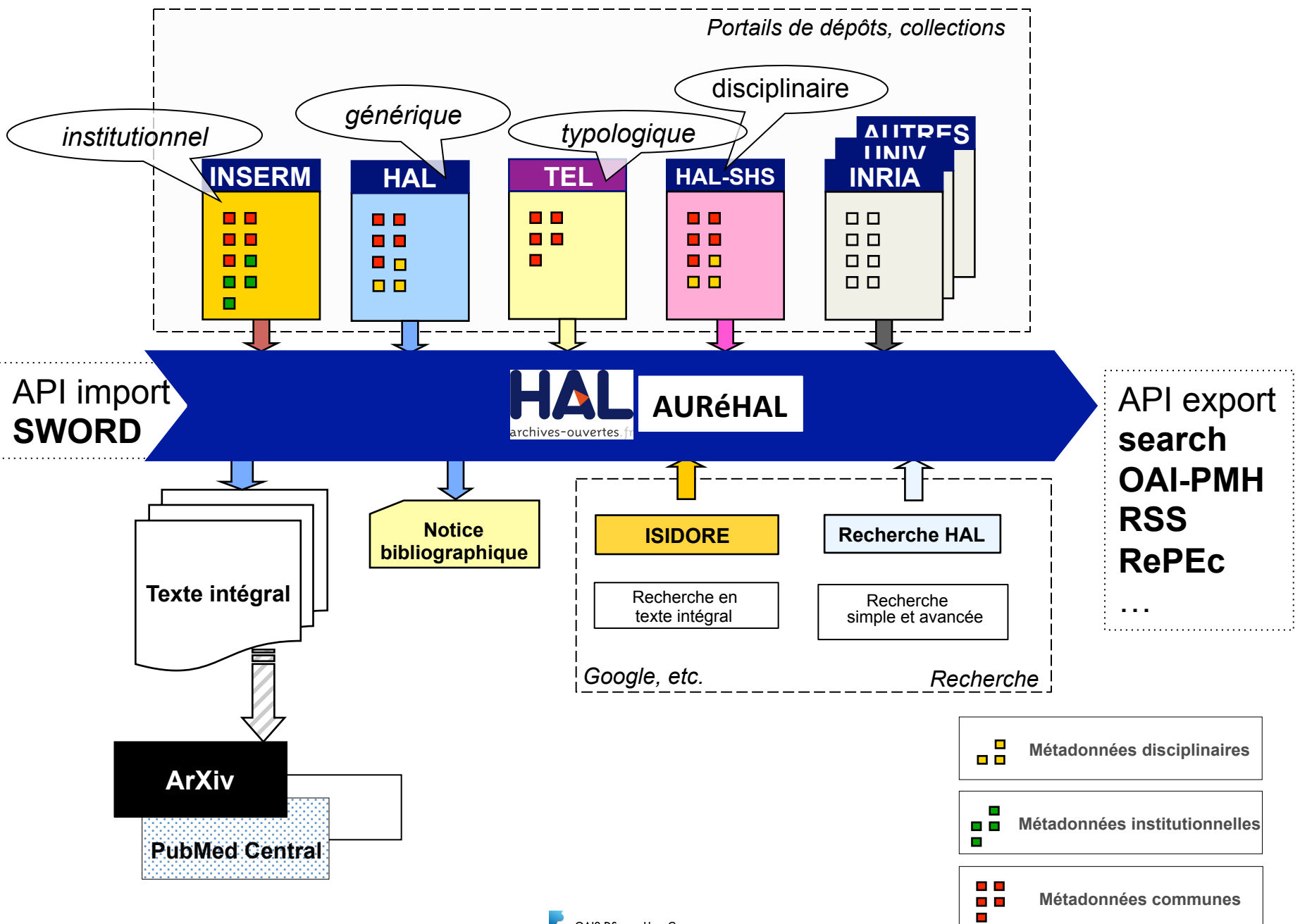
- Mise en oeuvre à l'aide d'un système de "tampons", apposé manuellement ou automatiquement
- Un site web personnalisable pour le gestionnaire
- Exemples de collections :
 - Articles d'une institution, d'un laboratoires
 - Communications à un colloque
 - Articles d'une revue,
 - etc.

Une URL pérenne

HAL passerelle vers l'international

- Depuis l'origine HAL est interconnecté avec arXiv,
 - arXiv « la » première archive ouverte créée par Paul Ginsparg en 1991
 - L'archive ouverte de référence en physique, mathématiques, informatique

- En 2007 HAL s'est interconnecté avec PubmedCentral
 - PMC, créé par les NIH, sous l'impulsion d'Harold Varmus
 - collaboration entre le CCSD et l'INSERM



HAL, des APIs

Recherche Hal

Rechercher les ressources de HAL via une API.

[Documentation »](#)

Recherche référentiels

Rechercher dans les référentiels de HAL via une API.

[Documentation »](#)

OAI-PMH

Serveur OAI-PMH de HAL.

[Documentation »](#)

Import SWORD

API SWORD.

[Documentation »](#)

Import SOAP

Web Service SOAP de dépôt dans HAL.

Obsolète à partir de juin 2015, merci d'utiliser l'import SWORD

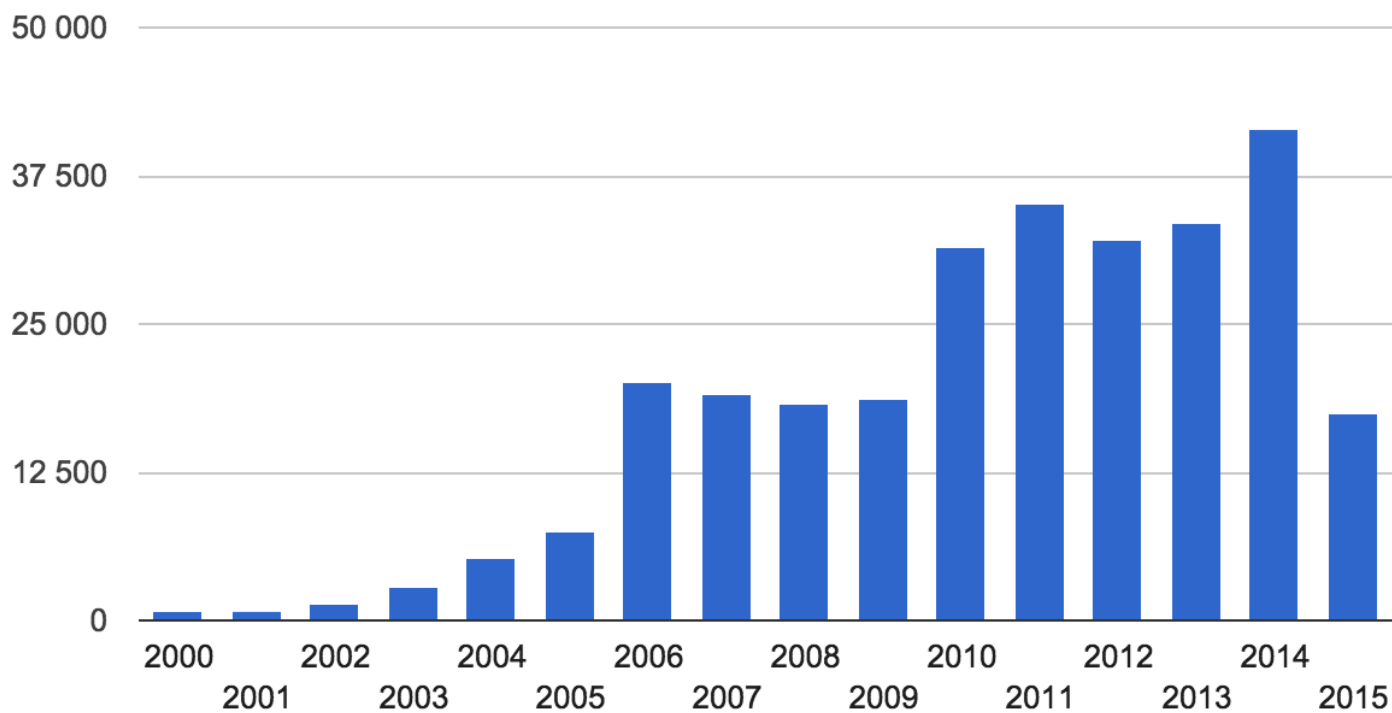
[Documentation »](#)

HAL, des APIs

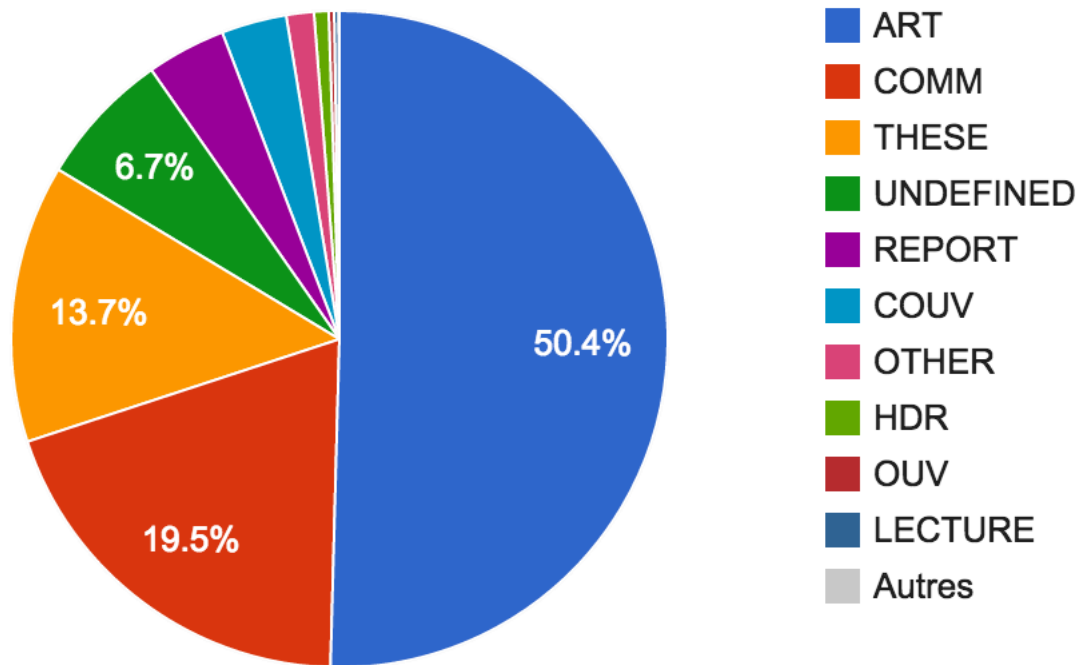
- **SWORD** : <https://api.archives-ouvertes.fr/sword/>
 - API REST HTTP
 - verbe GET, POST, PUT et DELETE
 - entête HTTP
 - On-Behalf-Of
 - **Packaging** ← **TEI, JATS**
 - Content-Type
 - Content-Disposition
 - Content-MD5
 - Export-To-Arxiv, Export-To-PMC, Hide-For-RePEc, Hide-In-OAI

Statistiques

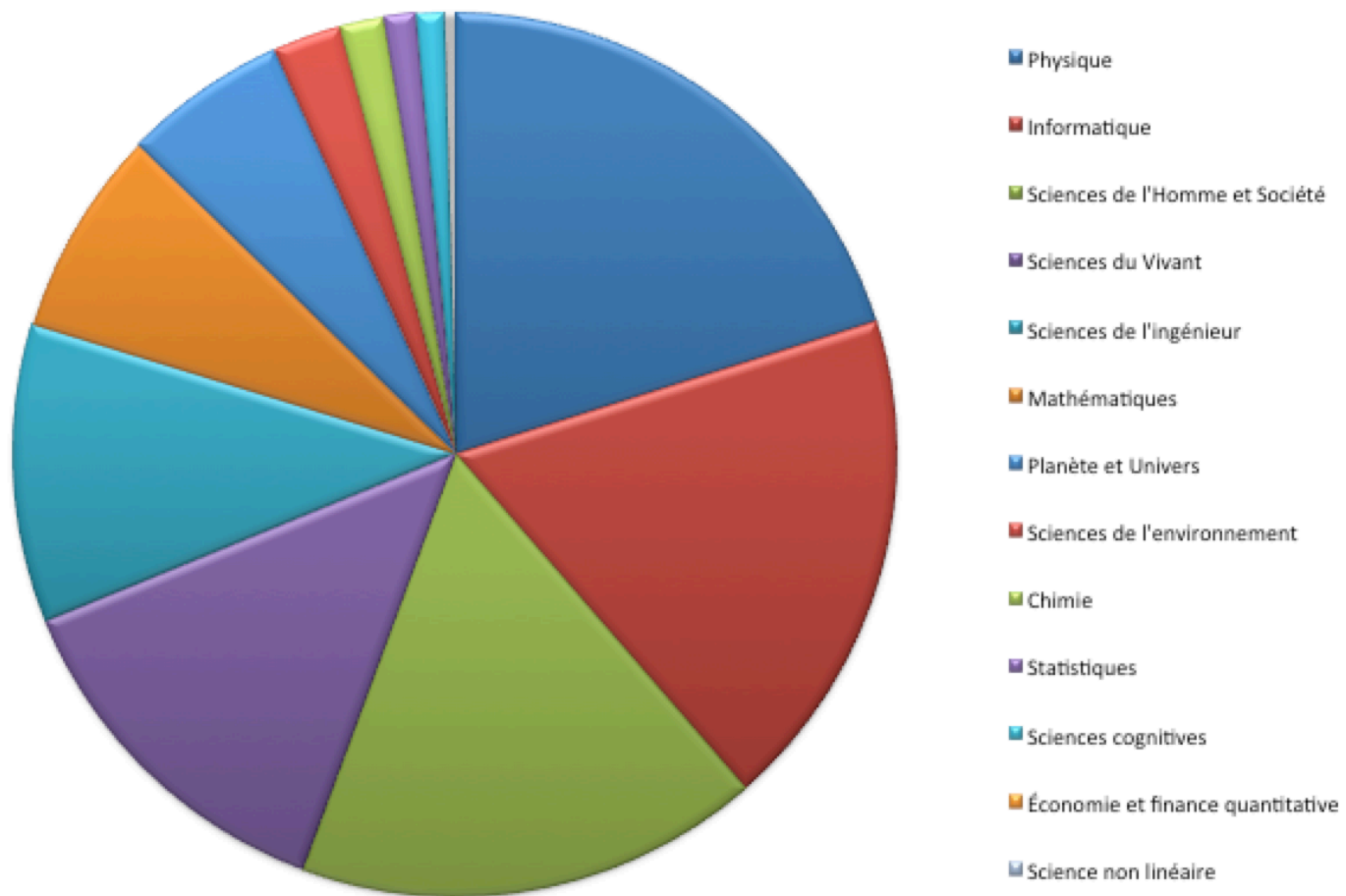
Date de dépôt : année



Statistiques



Statistiques



Webometrics

Le classement mondial des archives ouvertes en 2015 place HAL

- au 4° rang des archives ouvertes institutionnelles (HALSHS)
- au 4° rang des archives ouvertes

Top Portals

World Rank	Portal	Country	Size	Visibility	Files Rich	scholar
1	ResearchGate		2	2	1	2
2	Academia.edu		6	1	25	11
3	Scientific Electronic Library Online Brazil SciELO Brazil					
4	HAL Hyper Article en Ligne					
5	China National Knowledge Infrastructure					
6	DIALNET					
7	DIVA Digitala Vetenskapliga Arkivet Academic Archive On-line					
8	Berkeley Electronic Press Bepress					
9	Revue.org					
10	Scientific Electronic Library Online Chile SciELO Chile					

repositories.webometrics.info

Top Institutionals

World Rank	Portal	Country	Size	Visibility	Files Rich	scholar
1	(1) Smithsonian/NASA Astrophysics Data System		5	1	1353	1
2	CERN Document Server		2	8	1	2
3	NASA Technical Reports Server		278	2	4	40
4	HAL Sciences de l'Homme et de la Société		57	3	53	14
5	HAL Institut National de Recherche en Informatique et en Automatique Archive Ouverte		61	4	2	12
6	Repositório Digital Universidade Federal do Rio Grande do Sul LUME		4	15	5	6
7	Digital CSIC		68	6	27	20
8	Universidade de São Paulo Biblioteca Digital de Teses e Dissertações		123	14	13	3
9	University of California eScholarship Repository		386	9	12	4
10	MIT Institutional Repository		129	5	19	43

Pour conclure...

Le chercheur est l'acteur principal du système :

- Il détient la version électronique de sa publication
- Il est le seul à pouvoir donner immédiatement un certain nombre d'informations sur sa publication
- Il est le mieux à même d'en assurer le suivi
- Il reste responsable et acteur du devenir de sa production scientifique

Des usages encore à imaginer...

Adresses et URL

- <https://hal.archives-ouvertes.fr>
- <https://aurehal.archives-ouvertes.fr>
- <https://api.archives-ouvertes.fr>
- <https://cv.archives-ouvertes.fr>

- *Support utilisateurs*
 - hal.support@ccsd.cnrs.fr
 - <http://support.ccsd.cnrs.fr>