



# JOURNÉE DE "FAMILIARISATION" AVEC LES NOUVELLES API DE HAL V3

Organisée par le réseau métier ARAMIS

Villeurbanne - 2015-02-25

<http://aramis.resinfo.org/wiki/doku.php?id=ateliers:halv3>

<http://ccsd.cnrs.fr/>



**HAL** - <https://hal.archives-ouvertes.fr>

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.

**TEL** - <http://tel.archives-ouvertes.fr>

Le serveur TEL (thèses-en-ligne) promotion de l'auto-archivage en ligne des thèses de doctorat et habilitations à diriger des recherches (HDR).

**Sciencesconf.org** - <http://www.sciencesconf.org>

Plateforme Web ouverte aux organisateurs de colloques, workshops ou réunions scientifiques.

**MédiHAL** - <https://medihal.archives-ouvertes.fr>

MédiHAL, archive ouverte d'images scientifiques et des documents iconographiques de science.

**ISIDORE** - <http://www.rechercheisidore.fr>

Moteur de recherche unifié des SHS, moissonnant une grande quantité de liens émanant des bibliothèques numériques. MOA : TGE Adonis.

**Episciences.org** - <http://episciences.org>

Plateforme d'"épi-revues". (en cours)

**HELOISE** - <http://heloise.ccsd.cnrs.fr>

Le système Héloïse permet d'informer sur la politique des éditeurs en matière de diffusion en open access des publications scientifiques.

Créé en 2000

Missions :

Développement des archives  
ouvertes pour la communauté ESR  
ainsi que des services connexes

Basé dans les locaux du [Centre de  
Calcul de l'IN2P3](#)

Partenaire de projets européens :  
MedOANet, DARIAH-EU, PEER  
Equipex DILOH, ANR Campus AAR





- HAL
- HALSHS
- TEL
- MédiHAL
- Liste des portails
- AURéHAL
- API
- Documentation

## Documentation API-HAL

Les APIs (Application Programming Interface) sont des interfaces qui permettent une communication machine à machine. Cette page vous donne accès à la documentation des APIs de HAL v3.0.

### 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 »](#)



#### CONTACT

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



# LES API DE RECHERCHE

- API Recherche **HAL** = documents en ligne \*
- API Recherche **Référentiels** :
  - Auteurs\*
  - Structures de recherche\*
  - Domaines\*
  - Liste de métadonnées\*
  - Revues\*
  - Projets ANR\*
  - Projets européens\*
  - Affiliations des auteurs
  - Types de documents
  - Métadonnées

\* Servis par Apache Solr

Via les API de type Solr vous accédez aux même données/fonctionnalités que le CCSD



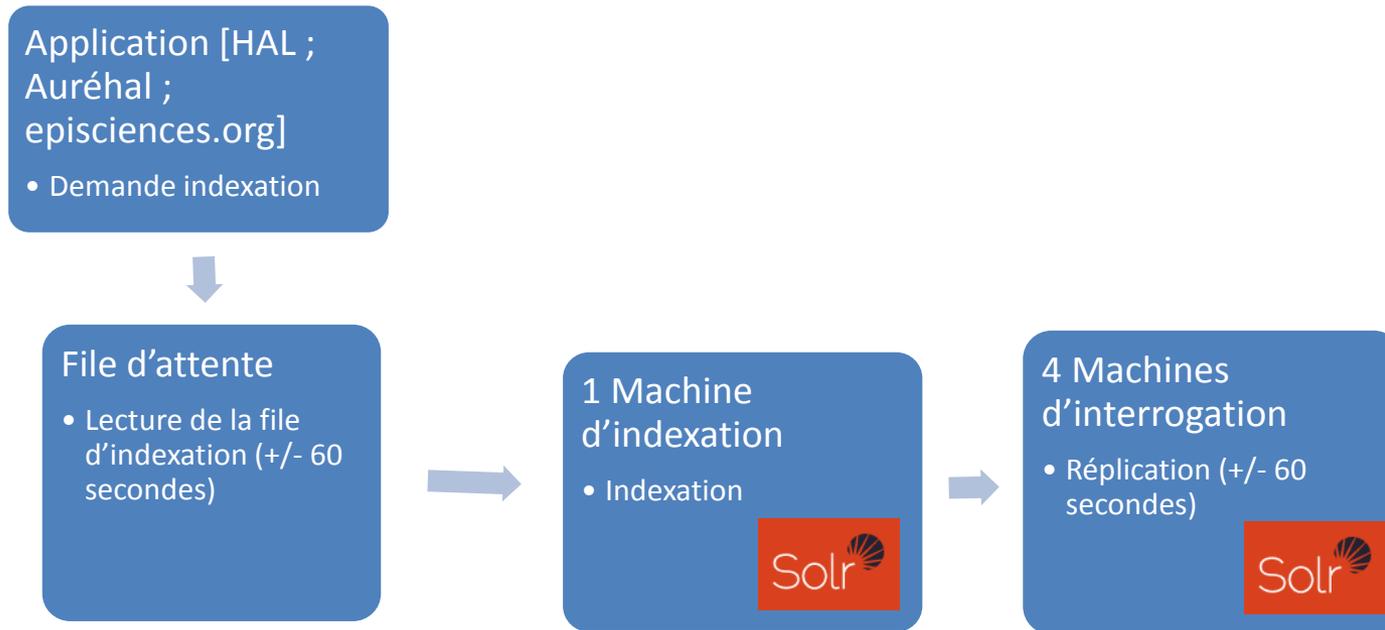
## ▶ Apache Solr

- ▶ Serveur recherche/indexation avec API REST
- ▶ Recherche texte intégral + facettes
- ▶ Réponses aux formats JSON ; XML ; CSV ; etc.
- ▶ Logiciel libre
- ▶ <https://lucene.apache.org/solr/>



- ▶ Solr au CCSD :
  - ▶ HAL
    - ▶ Recherche / pages consultation
    - ▶ Autocomplétion
    - ▶ Exports
    - ▶ Flux syndication
    - ▶ Statistiques
    - ▶ API
  - ▶ Auréhal
    - ▶ Recherche / pages consultation
  - ▶ Episciences
    - ▶ Recherche / pages consultation

# FONCTIONNEMENT INDEXATION





## API DE RECHERCHE HAL

[HTTPS://API.ARCHIVES-OUVERTES.FR/DOCS/SEARCH](https://api.archives-ouvertes.fr/docs/search)

- ▶ Tous les dépôts de HAL en ligne
  - ▶ +/- toutes les métadonnées d'un dépôt sous différentes formes
  - ▶ + version des métas aux formats :
    - ▶ Bibtex
    - ▶ XML-TEI
    - ▶ COinS
    - ▶ Endnote
  - ▶ Texte intégral indexé mais pas stocké

## Exemple de métadonnées renvoyées par l'API de recherche HAL au format XML

```
<str name="halId_s">in2p3-00007503</str>
<str name="uri_s">http://hal.in2p3.fr/in2p3-00007503</str>
<int name="version_i">1</int>
<int name="status_i">11</int>
<str name="instance_s">democrite</str>
<int name="sid_i">6</int>
<str name="submitType_s">file</str>
<str name="docType_s">ART</str>
<int name="thumbId_i">4930912</int>
▼ <str name="citationRef_s">
  <i>Nuclear Physics A</i>, Elsevier, 1998, 643, pp.331-346
</str>
▼ <str name="citationFull_s">
  B. Roussiere, P. Kilcher, A. Wojtasiewicz, J. Genevey, A. Gizon, et al.. Isomeric M3 transitions in
   $^{183}\text{Pt}$  and  $^{181}\text{Os}$ . <i>Nuclear Physics A</i>, Elsevier, 1998, 643, pp.331-346. <a
  target="_blank" href="http://hal.in2p3.fr/in2p3-00007503">&lt;in2p3-00007503&gt;</a>
</str>
▼ <str name="label_s">
  B. Roussiere, P. Kilcher, A. Wojtasiewicz, J. Genevey, A. Gizon, et al.. Isomeric M3 transitions in
   $^{183}\text{Pt}$  and  $^{181}\text{Os}$ . Nuclear Physics A, Elsevier, 1998, 643, pp.331-346. &lt;in2p3-00007503&gt;
</str>
▼ <arr name="collaboration_s">
  <str>ISOLDE</str>
</arr>
<int name="length_i">0</int>
▼ <arr name="localReference_s">
  <str>IPNO DRE 9705,ISN 9712</str>
</arr>
<str name="page_s">331-346</str>
<str name="volume_s">643</str>
<str name="modifiedDate_s">1998-12-21 11:01:39</str>
```

- ▶ Types de métadonnées accessibles
  - ▶ Pour la recherche (indexées)
  - ▶ Pour l'affichage (stockées)
  - ▶ Pour les facettes (listes de valeurs)
  - ▶ Pour le tri (pas de champs multivalués pour le tri)

# API RECHERCHE – TYPES DE CHAMPS

<http://api.archives-ouvertes.fr/docs/search/schema/field-types>

Nom	Description	Suffixe de champ	Pour affichage	Pour facette	Pour recherche	Pour Tri
alphaOnlySort	Tri lexicographique	_s	non	non	non	oui
boolean	Booléen (true ou false)	_bool	oui	oui	oui	oui
double	Nombre en virgule flottante (double-precision 64-bit IEEE 754 floating point)	_i	oui	oui	oui	oui
facetString	Texte pour générer des facettes	_fs	oui	oui	non	non
identifier	Texte pour une recherche sur des identifiants	_id	non	non	oui	non
int	Nombre entier ( $-2^{31}$ à $2^{31}-1$ )	_i	oui	oui	oui	oui
location	Coordonnées géographiques)	0_coordinate (latitude) et 1_coordinate(longitude)	oui	oui	oui	non
long	Nombre entier ( $-2^{63}$ à $2^{63}-1$ )	_i	oui	oui	oui	oui
string	Texte pour affichage ou facettes	_s	oui	oui	non	oui
stringCaseInsensitive	Texte pour recherche exacte insensible casse/accents	_sci	oui	oui	non	oui
tdate	Date ISO 8601 (eg 2013-01-31T23:12:01Z)	_tdate	oui	non	oui	oui
text	Texte pour recherche	_t	non	non	non	non
textStopWords	Texte pour une recherche sans mots-vides	_st	non	non	oui	non
text_autocomplete	Texte pour une recherche de type auto-complétion	_autocomplete	non	non	oui	non
tint	Nombre entier (-128 à +127)	_i	oui	oui	oui	oui

# API RECHERCHE – EXEMPLES DE CHAMPS

<http://api.archives-ouvertes.fr/docs/search/schema/fields/#fields>

<code>producedDateD_i</code>	Date de publication : jour	<ul style="list-style-type: none"><li>• 0</li><li>• 1</li><li>• 15</li><li>• 12</li><li>• 2</li></ul> <a href="#">+ plus d'exemples...</a>	tint Nombre entier (-128 à +127)	Oui	Oui	
<code>producedDateM_i</code>	Date de publication : mois	<ul style="list-style-type: none"><li>• 0</li><li>• 6</li><li>• 9</li><li>• 10</li><li>• 5</li></ul> <a href="#">+ plus d'exemples...</a>	tint Nombre entier (-128 à +127)	Oui	Oui	
<code>producedDateY_i</code>	Date de publication : année	<ul style="list-style-type: none"><li>• 2010</li><li>• 2011</li><li>• 2009</li><li>• 2012</li><li>• 2008</li></ul> <a href="#">+ plus d'exemples...</a>	tint Nombre entier (-128 à +127)	Oui	Oui	
<code>producedDate_s</code>	Date de publication	<ul style="list-style-type: none"><li>• 2006</li><li>• 2007</li><li>• 2008</li><li>• 2010</li><li>• 2009</li></ul> <a href="#">+ plus d'exemples...</a>	string Texte pour affichage ou facettes	Oui	Oui	
<code>producedDate_tdate</code> <span>Non vide</span>	Date de publication (ISO 8601)	<ul style="list-style-type: none"><li>• 2006-01-01T00:00:00Z</li><li>• 2007-01-01T00:00:00Z</li><li>• 2008-01-01T00:00:00Z</li><li>• 2010-01-01T00:00:00Z</li><li>• 2009-01-01T00:00:00Z</li></ul> <a href="#">+ plus d'exemples...</a>	tdate Date ISO 8601 (eg 2013-01-31T23:12:01Z)	Oui	Oui	

- ▶ Pré-requis client http
- ▶ Point d'entrée pour HAL :  
<http://api.archives-ouvertes.fr/search/>

- ▶ Point d'entrée pour un portail :  
`http://api.archives-ouvertes.fr/search/<instance>`
- ▶ Exemple :
- ▶ <http://api.archives-ouvertes.fr/search/in2p3/>
- ▶ NB : Dans le cas d'un portail/instance, si un dépôt a plusieurs versions, on ne trouve que la dernière.

- ▶ Point d'entrée pour une collection :  
`http://api.archives-ouvertes.fr/search/<TAMPON>`
- ▶ Exemple :
- ▶ <http://api.archives-ouvertes.fr/search/FRANCE-GRILLES/>
- ▶ NB : Dans le cas d'une collection, si un dépôt a plusieurs versions, on ne voit que la version tamponnée

# PRINCIPAUX PARAMÈTRES

## Valeurs + paramètres sensibles à la casse

Nom paramètre/usage	Description	exemple	Valeur par défaut
q	Requête de recherche	q=exemple	* :*
Chercher dans un champ	NomDuChamp:valeur	q=title_t:test	text
wt	Format de réponse		Json
Equation de recherche	Termes + opérateurs booléens avec des parenthèses	Journal AND (Histoire OR History)	
Intervalles	Recherche d'intervalles	submittedDateY_i:[2000 TO *]	
Opérateur booléens	AND OR NOT + - &&		AND
fq	Filtres	fq=submitType_s:file	
rows	Nombre de résultats	rows=5	30
start	Décalage des résultats	Start=10	0
fl	Champs à retourner	fl=auth*,docid,label_s	docid,label_s
sort	Tri : nom du champ + asc ou desc	sort=submittedDateY_i desc	score
facet	Construire des facettes	facet=true	false
facet.field	Un champ pour faire des facettes	facet.field=docType_s	
facet.mincount	Nombre minimum de valeurs pour retourner une facette	facet.mincount=1	0
facet.sort	Tri des facettes [count ou index]	facet.sort=index	count





Exemples en ligne de requêtes sur l'API de recherche HAL