



Utilisation des API HAL pour affichage des données sur un site Web

Mission OSI - Marion Étasse - Samuel Libot

Décembre 2018

Préalable : points d'entrée dans HAL

Requêtes par API = plusieurs points d'entrée

- portail et collection : on raisonne en ensembles/sous-ensembles
- affiliations aux structures : les auteurs et donc les publications sont liés à un labo ou une équipe

Interroger par ensemble directement dans l'URL

- Portail UGA (saga) : https://api.archives-ouvertes.fr/search/saga/?q=*&wt=json →
- Collection du labo 3S-R : https://api.archives-ouvertes.fr/search/3S-R/?q=*&wt=json →

Un code de portail se saisit en minuscules, celui d'une collection en majuscules

Préalable : points d'entrée dans HAL

Interroger par structure d'affiliation (sur les ID)

Les différents types de structures et champs correspondants :

- laboratory / labStructId_i
- researchteam / rteamStructId_i
- department / deptStructId_i
- regrouplaboratory / rgrpLabStructId_i

Il existe d'autres types de structures à des échelles supérieures, qui n'ont pas d'intérêt pour nous à ce stade.

Avantages/inconvénients

Avantages	Inconvénients
Portail/Collection	
Syntaxe de la requête simplifiée	Nécessite que les publications soient déposées et tamponnées au bon endroit.
Structures d'affiliation	
Précision des requêtes	Syntaxe plus complexe et nécessite une connaissance plus précise du modèle de données de HAL

Changer les champs et format de sortie

Champs retournés par défaut : **docid**, **label_s** et **uri_s**

Format JSON, pour du XML on ajoute
&wt=xml&indent=true

https://api.archives-ouvertes.fr/search/saga/?q=*&wt=xml&indent=true

```
<result name="response" numFound="137519" start="0">
  <doc>
    <int name="docid">677269</int>
    <str name="label_s">Onur Ozturk, Andras Sebo, Marie-Laure Espinouse, Maria Di
    <str name="uri_s">https://hal.archives-ouvertes.fr/hal-00677269</str></doc>
```

Spécifier les champs à retourner :
&fl=Champ1,Champ2,Champ3...

Comment afficher les 5 dernières publications de mon laboratoire ?

Par code de collection :

https://api.archives-ouvertes.fr/search/3S-R/?q=*&rows=5&wt=xml&indent=true&sort=producedDate_tdate desc →

Par Id de laboratoire

https://api.archives-ouvertes.fr/search/?q=labStructId_i:706&rows=5&wt=xml&indent=true&sort=producedDate_tdate desc →

Comment afficher les publications d'une personne donnée ? 1/2

Champs interrogeables pour une personne (par ordre de pertinence) :

- **authIdHal_i** : identifiant num rique IdHal
- **authIdHal_s** : cha ne de caract re IdHal (forme variable : prenom-nom, pnom, nomp, nom-prenom, etc.)
- **authId_i** : identifiant num rique "simple" dans le r f rentiel Auteur
- **authId_i** : identifiant num rique "simple" dans le r f rentiel Auteur
authLastNameFirstName_s : Nom de famille Pr nom

Possibilit  de grouper les publications par ann es, pour un affichage de type CV. On peut  galement grouper ou facetter sur le type de publications.

Comment afficher les publications d'une personne donnée ? 2/2

Requête pour afficher les publications de **Lucie Bailly**, classées par **ordre décroissant de publication**, **groupées par type de publication**. Avec les facettes : volumétrie des **publications par année**, **type de publication** et **nature de notice dans HAL**.

```
https://api.archives-ouvertes.fr/search/?q=authIdHal_s:lucie-bailly&wt=xml&indent=true&sort=producedDate_tdate desc&groupe=true&group.field=docType_s&fl=docid,label_s,uri_s,producedDateY_i&facet=true&facet.pivot=submittedDateY_i,docType_s,submitType_s →
```


Comment afficher les publications d'une équipe de mon laboratoire ?

Interrogation sur Id structure ou sigle de collection dans l'URL si la collection existe.

Équipe **ADELE** du LIG :

https://api.archives-ouvertes.fr/search/?q=rteamStructId_i:49632&wt=xml&indent=true&sort=producedDate_tdate desc ➔


Comment afficher les publications pour une année donnée ?

Gestion des dates dans HAL : dates de d p t, de publication, de parution du doc  lectronique...

Pour chaque type de date, un champ au format ISO, un champ pour le jour, pour le mois et pour l'ann e.

Pour une ann e donn e aux yeux d'un chercheur, on utilisera le champ **producedDateY_i**

Publications sur le portail UGA pour 2017 :

https://api.archives-ouvertes.fr/search/saga/?q=producedDateY_i:2017&wt=xml&indent=true&sort=producedDate_tdate desc 

Comment afficher les publications pour un type de publication donné ?

Utilisation du champ `docType_s`. Valeurs possibles [ICI](#).

Exemple sur conférences, articles, ouvrages et chapitres d'ouvrage du labo 3SR :

`https://api.archives-ouvertes.fr/search/?q=labStructId_i:706 AND docType_s:(ART OR COUV OR COMM OR OUV)&wt=xml&indent=true&sort=producedDate_tdate desc` [→](#)

Spécifier le nombre de lignes retournées dans une requête.

Par d faut, 30 lignes max. Se modifie avec le param tre **&rows=X**

Possibilit  de naviguer sur des pages de r sultats avec **&rows + &start** + utilisation d'un marqueur alphanum rique g n r  par Solr.

Ou script avec requ te en deux temps pour transformer la valeur de numFound en variable sp cifiant la valeur de rows lors de la seconde requ te :

1) https://api.archives-ouvertes.fr/search/saga/?q=producedDateY_i:2017&wt=xml&indent=true

```
<result name="response" numFound="8023" start="0">
```

2) https://api.archives-ouvertes.fr/search/saga/?q=producedDateY_i:2017&wt=xml&indent=true&rows=8023