

API initiation

Brigitte Bidégaray-Fesquet

Laboratoire Jean Kuntzmann, Grenoble



Journées CasuHAL 2024
13 juin 2024



API initiation

- 1 Introduction
- 2 Structure d'une interrogation
 - Structure générale
 - Opérateurs
 - Filtres
 - Facettes
 - Affichage
- 3 Les types de champ
- 4 Exemples
 - Pour la gestion d'un portail
 - Pour la gestion d'une collection
 - Pour la «gestion» d'un auteur
- 5 Pour finir...

Qu'est-ce qu'une API ?

- **API = Application Programming Interface**
ou «interface de programmation d'application»
c'est une interface logicielle qui permet l'échange de données entre logiciels.
- **Comme nous ne sommes pas des logiciels. . .**
Nous allons ici apprendre à **utiliser le point d'entrée de l'API** de recherche de HAL, c'est-à-dire à **écrire des requêtes** dans la barre de recherche d'un navigateur et à **formater les sorties** pour que nous, humains, puissions les analyser.

Les API de HAL

- Il y a plusieurs API dans HAL.
- Nous allons nous concentrer sur l'API de recherche HAL.
- Les autres API
 - API SWORD de dépôt sur HAL
 - Les API de recherche dans les référentiels

anrproject	doctype	metadata
author	domain	metadatalist
authorstructure	instance	structure
europeanproject	journal	
 - seront traitées dans l'atelier «API avancé»



API initiation

1 Introduction

2 Structure d'une interrogation

- Structure générale
- Opérateurs
- Filtres
- Facettes
- Affichage

3 Les types de champ

4 Exemples

- Pour la gestion d'un portail
- Pour la gestion d'une collection
- Pour la «gestion» d'un auteur

5 Pour finir...

Structure générale d'une interrogation

- Le point d'entrée de l'API

<https://api.archives-ouvertes.fr/search/>

- On précise ensuite

- le champ sur lequel on veut chercher : **q**
- des filtres sur les réponses : **fq**
- des facettes (détaillées dans l'atelier «API avancé»)
- les champs retournés dans la réponse : **fl**
- l'ordre dans lequel on trie les réponses : **sort**
- le rang des résultats retournés : **start** et **rows**
- le format de sortie : **wt** (JSON par défaut)

Format d'une requête

- Le point d'entrée de l'API

<https://api.archives-ouvertes.fr/search/>

- Une requête est de la forme (q comme *query* (requête))

?q=<champ> :<valeur>

- Pour chercher tous les documents dont le titre contient la chaîne de caractères «Bloch» :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"](https://api.archives-ouvertes.fr/search/?q=title_t:)

```
{
  "response": {
    "numFound": 498,
    "start": 0,
    "maxScore": 7.5637684,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "3140297",
        "label_s": "Julie Mazaleigue-Labaste. Bloch (Iwan). Christian Lacombe. Dictionnaire Sade, L'Harmattan, 2021, 978-2-343-22243-1. &#x27E8;halshs-03140297&#x27E9;",
        "uri_s": "https://shs.hal.science/halshs-03140297"
      }
    ]
  }
}
```

Raw Parsed

Opérateurs binaires AND (et) et OR (ou)

- L'opérateur par défaut est AND.
- Documents dont le titre contient les chaînes de caractères «Bloch» et «Maxwell» :

[https://api.archives-ouvertes.fr/search/?q=title_t:\(\"Bloch\" \"Maxwell\"\)](https://api.archives-ouvertes.fr/search/?q=title_t:(\)

```

{
  "response": {
    "numFound": 18,
    "start": 0,
    "maxScore": 12.95817,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "3255337",
        "label_s": "Brigitte Bidégaray, Antoine Bourgeade, Didier Reignier, Richard Ziolkowski. Multi-level Maxwell-Bloch simulations. Fifth International Conference on Mathematical and Numerical Aspects of Wave Propagation (Waves 2000), 2000, Santiago de Compostela, Spain. pp.221-225. &#x27E8;hal-00319999&#x27E9;",
        "uri_s": "https://hal.science/hal-00319999"
      }
    ]
  }
}

```

Raw Parsed

- Documents dont le titre contient les chaînes de caractères «Bloch» ou «Maxwell» :

[https://api.archives-ouvertes.fr/search/?q=title_t:\(\"Bloch\" OR \"Maxwell\"\)](https://api.archives-ouvertes.fr/search/?q=title_t:(\)

Opérateur unaire NOT (non)

- Documents dont le titre contient la chaîne de caractères «Bloch» mais pas «Maxwell» :

[https://api.archives-ouvertes.fr/search/?q=title_t:\(\"Bloch\" NOT \"Maxwell\"\)](https://api.archives-ouvertes.fr/search/?q=title_t:(\)

```
{
  "response": {
    "numFound": 480,
    "start": 0,
    "maxScore": 7.559031,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "3140297",
        "label_s": "Julie Mazaleigue-Labaste. Bloch (Iwan). Christian Lacombe. Dictionnaire Sade, L'Harmattan, 2021, 978-2-343-22243-1. &#x27E8;halshs-03140297&#x27E9;",
        "uri_s": "https://shs.hal.science/halshs-03140297"
      }
    ]
  }
}
```

Raw

Parsed

Quelques caractères spéciaux

- + - && || ! () { } [] ^ " ~ * ? :
- Il faut les échapper (avec \) si ils sont dans une chaîne sur laquelle on veut requêter
- + et - : le terme qui suit doit être présent ou absent
- &&, || et ! : équivalent de AND, OR et NOT
- () : pour regrouper
- [] : pour les intervalles (de dates par exemple)
- " " : encadre une chaîne de caractères
- ~ : recherche approchante
- ? et * : remplacent un ou plusieurs caractères
- : : séparateur champ-valeur

Filtres (1/3)

- On filtre avec **fq** (comme *filter query*)
`?fq=<champ> :<valeur>`
- Documents dont le titre contient les chaînes de caractères «Bloch» et «Maxwell» :

`https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"`

`&fq=title_t:"Maxwell"`

```
{
  "response": {
    "numFound": 18,
    "start": 0,
    "maxScore": 6.783948,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "3255242",
        "label_s": "Brigitte Bidégaray. Time Discretizations for Maxwell-Bloch Equations. Numerical Methods for Partial Differential Equations, 2003, 19 (3), pp.284-300. &#x27E8;10.1002/num.10046&#x27E9;. &#x27E8;hal-00319990&#x27E9; &#x27E9;",
        "uri_s": "https://hal.science/hal-00319990"
      },
    ]
  }
}
```

Raw Parsed

Filtres (2/3)

- On peut bien sûr filtrer par rapport à un autre champ que la requête initiale.
- Documents dont le titre contient les chaînes de caractères «Bloch», déposés entre 2016 et 2020 :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"](https://api.archives-ouvertes.fr/search/?q=title_t:)
&fq=submittedDateY_i:[2016 TO 2020]

```
{
  "response": {
    "numFound": 174,
    "start": 0,
    "maxScore": 7.179823,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "1802988",
        "label_s": "Tatiana Kovalevich. Tunable Bloch surface waves devices. Optics / Photonic. Université Bourgogne-Franche-Comté, 2017. English. &#x27E8;NNT : 2017UBFC0226&#x27E9;. &#x27E8;tel-01802988&#x27E9;",
        "uri_s": "https://theses.hal.science/tel-01802988"
      }
    ]
  }
}
```

Raw Parsed



Filtres (3/3)

- On peut cumuler plusieurs filtres et donc avoir plusieurs fois `&fq=...`
- Thèses et HDR dont le titre contient les chaînes de caractères «Bloch», déposées entre 2016 et 2020 :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"](https://api.archives-ouvertes.fr/search/?q=title_t:)
[&fq=submittedDateY_i:\[2016 TO 2020\]](#)
[&fq=docType_s:\(THESE OR HDR\)](#)

```

{
  "response": {
    "numFound": 7,
    "start": 0,
    "maxScore": 7.180143,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "1802988",
        "label_s": "Tatiana Kovalevich. Tunable Bloch surface waves devices. Optics / Photonic. Université Bourgogne Franche-Comté, 2017. English. &#x27E8;MNT : 2017UBFCD022&#x27E9;. &#x27E8;tel-01802988&#x27E9;",
        "uri_s": "https://theses.hal.science/tel-01802988"
      }
    ]
  }
}

```

Raw Parsed



Facettes (1/3)

Un autre moyen de filtrer en complément à une requête est d'utiliser des facettes.

- Pour générer des facettes, il faut ajouter le paramètre **facet=true** à une requête
- Ensuite, on précise
 - le champ qui sert pour la facette
facet.field=<champ> :<valeur>
 - le type de tri
facet.sort=index (tri lexicographique)
facet.sort=count (tri par nombre d'occurrence)
- Il faut aussi ajouter **&rows=0** (pour n'avoir que les facettes)



Facettes (2/3)

- Répartition par domaine dans la collection LJK :

[https://api.archives-ouvertes.fr/search/LJK/?q=*&rows=0
&facet=true&facet.field=level0_domain_s&facet.sort=count](https://api.archives-ouvertes.fr/search/LJK/?q=*&rows=0&facet=true&facet.field=level0_domain_s&facet.sort=count)

```
{
  "response": {
    "numFound": 6787,
    "start": 0,
    "maxScore": 1,
    "numFoundExact": true,
    "docs": []
  },
  "facet_counts": {
    "facet_queries": {},
    "facet_fields": {
      "level0_domain_s": [
        "math",
        3138,
        "info",
        3115,
        "stat",
        1056,
        "spi",
        584,
        "phys",
        453,
        "sdv",
        291,
        "sdu",
        238,
        "sde",
        224,
        "sbe"
      ]
    }
  }
}
```

Raw Parsed

Facettes (3/3)

- On peut se servir de champs non multi-valués pour servir de pivot

`facet.pivot=<pivot>`

- Répartition par type de documents des types de dépôts dans la collection LJK :

`https://api.archives-ouvertes.fr/search/LJK/?q=*&rows=0
&indent=true&facet=true&facet.pivot=docType_s,submitType_s`

```
  "docType_s,submitType_s": [
    {
      "field": "docType_s",
      "value": "COMM",
      "count": 2765,
      "pivot": [
        {
          "field": "submitType_s",
          "value": "file",
          "count": 1523
        },
        {
          "field": "submitType_s",
          "value": "notice",
          "count": 1188
        },
        {
          "field": "submitType_s",
          "value": "annex",

```


Affichage (1/6)

- On choisit le format de sortie avec `&wt=<format>`
 - où on a une réponse de l'API pour `<format>` égal à `xml-tei`, `bibtex`, `endnote`, `rss` ou `atom`
 - ou une réponse de Apache Solr pour `<format>` égal à `json`, `xml` ou `csv`.
 - Les champs retournés par défaut sont `docid`, `label_s` et `uri_s`
`https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"`

```
{
  "response": {
    "numFound": 498,
    "start": 0,
    "maxScore": 7.5637684,
    "numFoundExact": true,
    "docs": [
      {
        "docid": "3140297",
        "label_s": "Julie Mazaleigue-Labaste. Bloch (Iwan). Christian Lacombe. Dictionnaire Sade, L'Harmattan, 2021, 978-2-343-22243-1. &#x27E8;halshs-03140297&#x27E9;",
        "uri_s": "https://shs.hal.science/halshs-03140297"
      }
    ]
  }
}
```

Raw

Affichage (2/6)

- Sortie XML :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"&wt=xml](https://api.archives-ouvertes.fr/search/?q=title_t:)

```

<response>
  <result name="response" numFound="498" start="0" maxScore="7.5523005" numFoundExact="true">
    <doc>
      <str name="docid">3140297</str>
      <str name="label_s">Julie Mazaleigue-Labaste. Bloch (Iwan). Christian Lacombe. Dictionnaire Sade, L'Harmattan, 2021, 978-2-343-22243-1. &#x27E8;halshs-03140297&#x27E9;</str>
      <str name="uri_s">https://shs.hal.science/halshs-03140297</str>
    </doc>
    <doc>
      <str name="docid">4140211</str>
      <str name="label_s">Yves Gomas, Christian Nitschelm. Bloch, Marie. Biographical Encyclopedia of Astronomers, 2022, pp.1-3. &#x27E8;10.1007/978-1-0716-0738-1_100785-1&#x27E9;. &#x27E8;hal-04140211&#x27E9;</str>
      <str name="uri_s">https://hal.science/hal-04140211</str>
    </doc>
  </result>
</response>

```

- Sortie XML-TEI :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"&wt=xml-tei](https://api.archives-ouvertes.fr/search/?q=title_t:)

Affichage (3/6)

- **Sortie BibTex :**

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"&wt=bibtex](https://api.archives-ouvertes.fr/search/?q=title_t:)

```
@incollection{mazaleiguelabaste:halshs-03140297,  
  TITLE = {{Bloch (Iwan)}},  
  AUTHOR = {Mazaleiguelabaste, Julie},  
  URL = {https://shs.hal.science/halshs-03140297},  
  BOOKTITLE = {{Dictionnaire Sade}},  
  EDITOR = {Christian Lacombe},  
  PUBLISHER = {{L'Harmattan}},  
  YEAR = {2021},  
  MONTH = Jan,  
  PDF = {https://shs.hal.science/halshs-03140297/file/Mazaleiguelabaste%20-%20Bloch%20%28Iwan%29.pdf},  
  HAL_ID = {halshs-03140297},  
  HAL_VERSION = {v1},  
}
```

- **Sortie RSS :**

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"&wt=rss](https://api.archives-ouvertes.fr/search/?q=title_t:)

- **Sortie Atom :**

[https://api.archives-ouvertes.fr/search/?q=title_t:"Bloch"&wt=atom](https://api.archives-ouvertes.fr/search/?q=title_t:)

- Les sorties **csv** et **endnote** retournent un fichier.



Affichage (4/6)

- On peut choisir d'autres champs pour la sortie avec `&fl=<champ>`

- Mise en évidence des auteurs publiant sur Maxwell-Bloch :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Maxwell-Bloch"](https://api.archives-ouvertes.fr/search/?q=title_t:)

`&fl=authLastName_s`

Raw

Parsed

```
{
  "response": {
    "numFound": 16,
    "start": 0,
    "maxScore": 39.034706,
    "numFoundExact": true,
    "docs": [
      {
        "authLastName_s": [
          "Bidégaray",
          "Bourgeade",
          "Reignier",
          "Ziolkowski"
        ]
      }
    ]
  }
}
```

Affichage (5/6)

- Les résultats peuvent être triés (cf. infra pour le choix du champ)
- Le tri se fait avec le paramètre
`&sort=<champ> <sens>`
- où `<sens>` vaut `asc` ou `desc` (pour un tri par ordre croissant ou décroissant).

- Tri par année de publication décroissante :

[https://api.archives-ouvertes.fr/search/?q=title_t:"Maxwell-Bloch"&sort=publicationDateY_i desc&wt=xml](https://api.archives-ouvertes.fr/search/?q=title_t:)

```

<response>
  <result name="response" numFound="16" start="0" numFoundExact="true">
    <doc>
      <str name="docid">4192379</str>
      <str name="label_s">Brigitte Bidégaray-Fesquet. Régimes particuliers pour l'équation de Maxwell-Bloch. Mathematical and Numerical Methods for Electromagnetism Mini-Workshop, Jul 2023, Montpellier, France. &#x27E8;hal-04192379&#x27E9;</str>
      <str name="uri_s">https://hal.univ-grenoble-alpes.fr/hal-04192379</str>
    </doc>
    <doc>
      <str name="docid">3224248</str>
      <str name="label_s">Olivier Peyrusse, P Jonnard, J.-M André. Maxwell-Bloch modeling of an x-ray pulse amplification in a one-dimensional photonic crystal. Physical Review A : Atomic, molecular, and optical physics [1990-2015], 2021, 103, pp.043508. &#x27E8;10.1103/PhysRevA.103.043508&#x27E9;. &#x27E8;hal-03224248&#x27E9;</str>
      <str name="uri_s">https://hal.science/hal-03224248</str>
    </doc>
  </result>
</response>

```



Affichage (6/6)

- Le nombre de résultats est systématiquement retourné.
- Par défaut seuls les 30 premiers résultats sont affichés.
- Dans le cas où il y a beaucoup de résultats, on peut découper les réponses par lots

&rows=<nombre de résultats>

&start=<rang du premier résultat>

- Publications avec le mot «Ukraine» :

https://api.archives-ouvertes.fr/search/?q=title_t:Ukraine

&sort=publicationDateY_i desc&start=100&rows=50

API initiation

- 1 Introduction
- 2 Structure d'une interrogation
 - Structure générale
 - Opérateurs
 - Filtres
 - Facettes
 - Affichage
- 3 Les types de champ**
- 4 Exemples
 - Pour la gestion d'un portail
 - Pour la gestion d'une collection
 - Pour la «gestion» d'un auteur
- 5 Pour finir...

Des champs pour tous les usages

Les champs sont dupliqués pour différentes utilisations : affichage, facettes, recherche ou tri.

On les reconnaît à leur suffixe, parmi lesquels

suffixe	type	affichage	facette	recherche	tri
_bool	booléen	x	x	x	x
_fs	facette	x	x		
_i	nombre	x	x	x	x
_id	identifiant			x	
_sort	alpha (lexicogr.)				x
_s	chaîne de caractères	x	x		x
_sci	chaîne (sans casse)	x	x	x	x
_t	texte			x	
_tdate	date (ISO 8601)	x		x	x

Attention à la variante utilisée !

Certains champs, la plupart des champs, existent donc en plusieurs versions.

Par exemple :

				Indexé	Stocké	Multivalué
authLastName_s	Auteur : Nom de famille	 Exemples...	string Texte pour affichage, facettes ou recherche exacte	 Oui	 Oui	 Oui
authLastName_sci	Auteur : Nom de famille Contient une copie de ce champ : <ul style="list-style-type: none"> authLastName_s 		stringCaseInsensitive Texte pour recherche exacte mais insensible à la casse/accents	 Oui		 Oui
authLastName_t	Auteur : Nom de famille Contient une copie de ce champ : <ul style="list-style-type: none"> authLastName_s 		text Texte pour recherche, insensible à la casse et aux accents	 Oui		 Oui

Requête	Nombre
?q=authLastName_t:"Bidegaray"	98
?q=authLastName_t:"Bidégaray"	98
?q=authLastName_s:"Bidegaray"	6
?q=authLastName_s:"Bidégaray"	16

Liste des champs

- **On en trouve une ici :**

<https://api.archives-ouvertes.fr/docs/search/?schema=field-types>

- **Elle est dans l'ordre alphabétique.**

Pratique si on sait ce que l'on cherche, moins pratique sinon.

- **Elle n'est pas exhaustive.**

La preuve : certains champs agglomèrent des champs qui ne sont pas dans la liste.

- **Il y a des champs multi-critères**

- très généraux comme **text** (par défaut) ou **text_fulltext** ;
- plus spécialisés, par exemple **europeanProject_t**.

- **Il y a des champs dynamiques**

qui en génèrent d'autres,
par exemple pour le multi-linguisme.



Thèmes des champs

- Type (de document, de dépôt)
- Auteur (par nom/prénom, mail, identifiants, organisme,...)
- Titre (aussi titre d'ouvrage)
- Journal / conférence (titre, éditeur, Sherpa, ...)
- Dates (diverses)
- Pays, ville,...
- Structures (de l'équipe au regroupement d'institution, par nom, identifiant,...)
- Classifications diverses (ACM, MeSH,
- Projets ANR (par titre, référence, acronyme, appel,...), européens, financements
- Autres identifiants (doi, arXiv, PubMed, d'institutions,)
- Collections (nom, identifiants, mère, fille,...)
- ...et beaucoup d'autres

API initiation

- 1 Introduction
- 2 Structure d'une interrogation
 - Structure générale
 - Opérateurs
 - Filtres
 - Facettes
 - Affichage
- 3 Les types de champ
- 4 Exemples**
 - Pour la gestion d'un portail
 - Pour la gestion d'une collection
 - Pour la «gestion» d'un auteur
- 5 Pour finir...

Évaluer l'évolution du dépôt en texte intégral des articles dans un portail

```
https://api.archives-ouvertes.fr/search/saga/?q=docType_s:ART&rows=0  
&facet=true&facet.pivot=submittedDateY_i,submitType_s  
&wt=xml
```

Cette interrogation comprend :

- la restriction à un portail,
- la requête uniquement sur les articles,
- l'utilisation de facettes,
- et les deux pivots utiles : la date et le type de soumission.

Si pour l'utilisation par un logiciel c'est la même chose, un humain ne voit pas les résultats de la même manière si on inverse l'ordre des pivots :

[https://api.archives-ouvertes.fr/search/saga/?q=docType_s:ART
&rows=0&facet=true&facet.pivot=submitType_s,submittedDateY_i
&wt=xml](https://api.archives-ouvertes.fr/search/saga/?q=docType_s:ART&rows=0&facet=true&facet.pivot=submitType_s,submittedDateY_i)

```
<?xml version="1.0" encoding="UTF-8" ?>
<response>
  <result name="response" numFound="103364" start="0" maxScore="0.3195819">
    <lst name="facet_counts">
      <lst name="facet_queries"/>
      <lst name="facet_fields"/>
      <lst name="facet_ranges"/>
      <lst name="facet_intervals"/>
      <lst name="facet_heatmaps"/>
      <lst name="facet_pivot">
        <arr name="submittedDateY_i,submitType_s">
          <lst>
            <str name="field">submittedDateY_i</str>
            <int name="value">2019</int>
            <int name="count">12341</int>
          </lst>
          <arr name="pivot">
            <lst>
              <str name="field">submitType_s</str>
              <str name="value">notice</str>
              <int name="count">9804</int>
            </lst>
            <lst>
              <str name="field">submitType_s</str>
              <str name="value">file</str>
              <int name="count">2529</int>
            </lst>
            <lst>
              <str name="field">submitType_s</str>
              <str name="value">anex</str>
              <int name="count">8</int>
            </lst>
          </arr>
        </lst>
      </lst>
    </result>
  </response>
```

```
<?xml version="1.0" encoding="UTF-8" ?>
<response>
  <result name="response" numFound="103365" start="0" maxScore="0.3173635">
    <lst name="facet_counts">
      <lst name="facet_queries"/>
      <lst name="facet_fields"/>
      <lst name="facet_ranges"/>
      <lst name="facet_intervals"/>
      <lst name="facet_heatmaps"/>
      <lst name="facet_pivot">
        <arr name="submitType_s,submittedDateY_i">
          <lst>
            <str name="field">submitType_s</str>
            <str name="value">notice</str>
            <int name="count">6725</int>
          </lst>
          <arr name="pivot">
            <lst>
              <str name="field">submittedDateY_i</str>
              <int name="value">2019</int>
              <int name="count">9804</int>
            </lst>
            <lst>
              <str name="field">submittedDateY_i</str>
              <int name="value">2018</int>
              <int name="count">6212</int>
            </lst>
            <lst>
              <str name="field">submittedDateY_i</str>
              <int name="value">2009</int>
              <int name="count">4715</int>
            </lst>
          </arr>
        </lst>
      </lst>
    </result>
  </response>
```

Trouver les articles co-publiés avec les Etats-Unis en 2021

[https://api.archives-ouvertes.fr/search/LJK/?q=instStructCountry_s:us
&fq=publicationDateY_i:2021&wt=xml](https://api.archives-ouvertes.fr/search/LJK/?q=instStructCountry_s:us&fq=publicationDateY_i:2021&wt=xml)

Cette interrogation comprend :

- la restriction à une collection (idem portail mais les collections sont en majuscules),
- la requête uniquement sur les publications avec le pays 'us' (code ISO 3166),
- et un filtre sur la date de publication.

En utilisant des facettes, on peut lister les pays avec lesquelles il y a des co-publications pour une collection :

[https://api.archives-ouvertes.fr/search/LJK/?q=*&rows=0
&facet=true&facet.pivot=instStructCountry_s&wt=xml](https://api.archives-ouvertes.fr/search/LJK/?q=*&rows=0&facet=true&facet.pivot=instStructCountry_s&wt=xml)

```

▼<lst name="facet_pivot">
  ▼<arr name="instStructCountry_s">
    ▼<lst>
      <str name="field">instStructCountry_s</str>
      <str name="value">fr</str>
      <int name="count">6784</int>
    </lst>
    ▼<lst>
      <str name="field">instStructCountry_s</str>
      <str name="value">us</str>
      <int name="count">619</int>
    </lst>
    ▼<lst>
      <str name="field">instStructCountry_s</str>
      <str name="value">gb</str>
      <int name="count">241</int>
    </lst>
    ▼<lst>
      <str name="field">instStructCountry_s</str>
      <str name="value">de</str>
      <int name="count">230</int>
    </lst>
    ▼<lst>
      <str name="field">instStructCountry_s</str>
      <str name="value">cn</str>
      <int name="count">160</int>
    </lst>
    ▼<lst>
      <str name="field">instStructCountry_s</str>
      <str name="value">it</str>
      <int name="count">152</int>
    </lst>
  </arr>
</lst>

```


Identifier ce qui dépend de la collection

Ce qui précède suppose que la collection est déjà bien alimentée.

Pour récupérer ce devrait être dans une collection (en cas de tamponnage manuel par exemple), on peut aussi interroger sur les identifiants des structures d'affiliation :

- équipe de recherche : `rteamStructId_i`
- département : `deptStructId_i`
- laboratoire : `labStructId_i`
- regroupement de laboratoire : `rgrpLabStructId_i`

Comment identifier un auteur ?

(plus ou moins par ordre de pertinence décroissante)

- `authIdHal_i` : identifiant numérique IdHal
- `authIdHal_s` : chaîne de caractère IdHal (normalement prenom-nom, mais pas toujours)
- `authId_i` : identifiant numérique "simple" dans le référentiel Auteur
- `authLastNameFirstName_s` : chaîne de caractère (Nom de famille Prénom)
- `authLastName_s` : chaîne de caractère (Nom de famille)

Pour sensibiliser au dépôt en texte intégral

```
https://api.archives-ouvertes.fr/search/  
?q=authIdHal_s:brigitte-bidegaray-fesquet&fq=docType_s:ART  
&fl=journalSherpaCondition_s,journalSherpaPrePrint_s,  
journalSherpaPostPrint_s,submitType_s,uri_s
```

Cette interrogation comprend :

- la requête uniquement sur un auteur via son idHAL textuel,
- un filtre sur les articles dans des revues seulement,
- l'affichage d'un certain nombre données Sherpa/Romeo de la revue
- détail des conditions
- autorisation des pre- et post-prints
- du type de soumission (texte intégral, notice)
- et du lien vers la notice.

```
--  
  {  
    "journalSherpaPrePrint_s": "can",  
    "journalSherpaPostPrint_s": "can",  
    "journalSherpaCondition_s": "Author's post-print on pre-print servers, including arXiv]  
[Publisher's version/PDF and author's post-print on author's personal website, institutional  
website or open access repository][Non-commercial use][Publisher copyright must be acknowledged]  
[Publisher's version/PDF may be used][Publisher last reviewed on 04/01/2018",  
    "uri_s": "https://hal.science/hal-00319991",  
    "submitType_s": "file"  
  },  
  {  
    "journalSherpaPrePrint_s": "can",  
    "journalSherpaPostPrint_s": "restricted",  
    "journalSherpaCondition_s": "Author's pre-print on any website or open access repository]  
[Author's post-print on author's personal website, institutional repository, subject repository  
or funding agency designated repository][Publisher's version/PDF cannot be used][Set statement to  
accompany pre-print and authors post-print - see policy][Must link to publisher version with  
DOI",  
    "uri_s": "https://hal.science/hal-02294954",  
    "submitType_s": "file"  
  },  
}
```

API initiation

- 1 Introduction
- 2 Structure d'une interrogation
 - Structure générale
 - Opérateurs
 - Filtres
 - Facettes
 - Affichage
- 3 Les types de champ
- 4 Exemples
 - Pour la gestion d'un portail
 - Pour la gestion d'une collection
 - Pour la «gestion» d'un auteur
- 5 Pour finir...

La documentation

- La documentation de l'API de recherche HAL
<https://api.archives-ouvertes.fr/docs/search>
- Les types de champs et la liste des champs
<https://api.archives-ouvertes.fr/docs/search/?schema=field-types>
- Les autres référentiels
<https://api.archives-ouvertes.fr/docs/ref>

Connaissez-vous Bruno ?

- Vous n'aimez pas taper une longue requête dans un navigateur ?
- Vous ne savez pas comment conserver vos requêtes préférées ?
- Connaissez-vous Bruno ? (un petit outil discuté au GT API)

The screenshot shows the Bruno application interface. On the left, a sidebar displays a collection named 'HAL-LJK' with a sub-collection 'GET Pays'. The main area shows a REST client configuration for a GET request to the URL: `https://api.archives-ouvertes.fr/search/LJK/?q=instStructCountry_sus&fq=publicationDateY_i:2023&wt=xml`. Below the configuration, a table lists the query parameters:

Name	Value	Actions
q	instStructCountry_sus	✓ 🗑️
fq	publicationDateY_i:2023	✓ 🗑️
wt	xml	✓ 🗑️

Below the table is a '+ Add Param' button. The right pane shows the 'Response' tab with an XML document. The XML content is as follows:

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <response>
3   <result name="response" numFound="35" start="0" maxScore
4     =1.64169" numFoundExact="true">
5     <doc>
6       <str name="docid">4026006</str>
7       <str name="label_s">Jean-Guillaume Dumas, Aude M
8         aignan, Clément Fernet, Daniel S. Roche. VESPO: Verified Eva
9         luation of Secret Polynomials: with application to low-stora
10        ge dynamic proofs of retrievability. Privacy Enhancing Techn
11        ologies Symposium, Jul 2023, Lausanne (CH), Switzerland. pp.
12        354--374. &amp;#x27E8;10.56553/gopets-2023-0085&amp;#x27E9;.
13        &amp;#x27E8;hal-03365854v5&amp;#x27E9;</str>
14       <str name="uri_s">https://hal.science/hal-033658
15        54v5</str>
16     </doc>
17     <doc>
18       <str name="docid">4256824</str>
19       <str name="label_s">Buy Nguyen, Pedram Akbarian,
20        Trungtin Nguyen, Nhat Ho. A General Theory for Softmax Gatin
21        g Multinomial Logistic Mixture of Experts. 2023. &amp;#x27E8;
22        ;hal-04256824&amp;#x27E9;</str>
23       <str name="uri_s">https://hal.science/hal-042568
24        24</str>
25     </doc>
26     <doc>
27       <str name="docid">3952063</str>
28       <str name="label_s">Anatoli B. Juditsky, Arkadi
29        Nemirovski, Michael Zibulevsky. Radiation design in computed
30        tomography via convex optimization. 2023. &amp;#x27E8;hal-03
  
```