

Les API de HAL

Découverte et exemples d'utilisation

Journées CASUHAL 2019 - Université de Rouen Normandie

Mardi 18 juin 2019

Hélène Jouguet - Université d'Orléans



Une Interface utilisée par des Programmes pour interagir avec une Application

1

API

→ code

→ application

2

API

→ extraction / requête

→ traitement / analyse manuel

<https://www.youtube.com/watch?v=0FQ6w4CO5Nw>

2

<https://www.reseau-tao.fr/>

ITINÉRAIRE

Itinéraire Horaires A proximité

Nouvelle recherche - Modifier

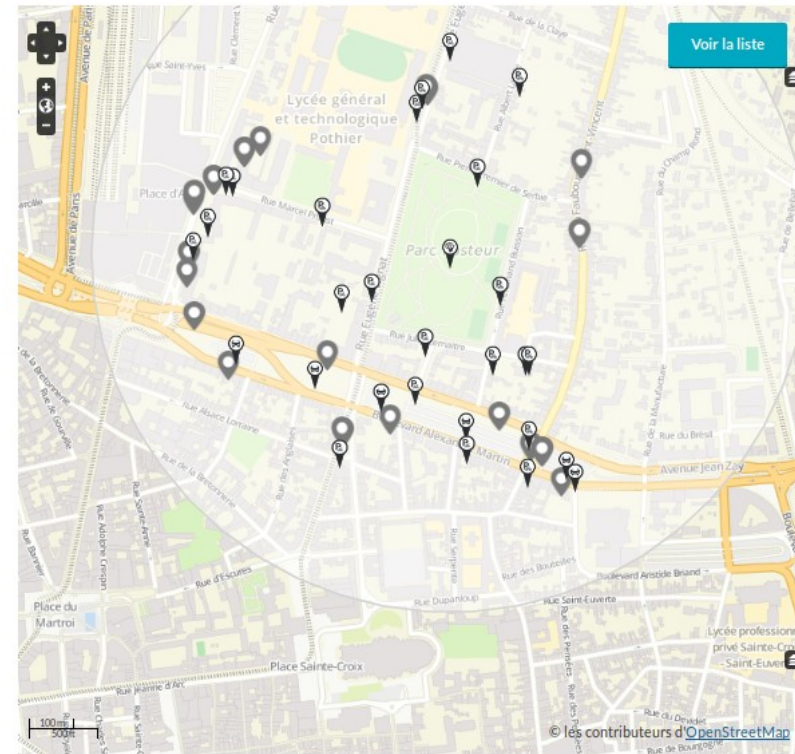
Les services autour de

Parc Pasteur (Orléans) Dans un rayon de 400 m

[Itinéraire depuis ce lieu](#) | [Itinéraire vers ce lieu](#)

Mobilité & Services de transport

<input checked="" type="checkbox"/> Aires de stationnement	<input checked="" type="checkbox"/> Equipement Vélo
<input checked="" type="checkbox"/> Station VLS	<input checked="" type="checkbox"/> Aires de services
<input checked="" type="checkbox"/> Points de vente	



Résultats proposés par Navitia  navitia

API Navitia the open API for building cool stuff with transport data <https://www.navitia.io/>

<https://www.openstreetmap.org/>

Les API archives-ouvertes.fr

Recherche HAL

Rechercher les ressources de HAL via une API.



Documents

Moteur de recherche



Référentiels HAL

Rechercher dans les référentiels de HAL via une API.



anproject Projets ANR

Référentiel des projets ANR

documentation

URL : /ref/anproject

authorstructure Affiliation des auteurs

Affiliation des auteurs aux structures dans les ressources de HAL

documentation

URL : /search/authorstructure

doctype Types de documents

Référentiel des types de documents

documentation

URL : /ref/doctype

instance Instances de portail

Référentiel des instances de portail de HAL

documentation

URL : /ref/instance

metadata Métadonnées

Référentiel des métadonnées

documentation

URL : /ref/metadata

structure Structures de recherche

Référentiel des structures de recherche

documentation

URL : /ref/structure

author Auteurs

Référentiel des auteurs

documentation

URL : /ref/author

europeanproject Projets européens

Référentiel des projets européens

documentation

URL : /ref/europeanproject

domain Domaines

Référentiel des domaines

documentation

URL : /ref/domain

journal Revues

Référentiel des revues

documentation

URL : /ref/journal

metadatalist Liste de métadonnées

Référentiel des listes de métadonnées

documentation

URL : /ref/metadatalist

Extraits d'un dépôt HAL

Format du dépôt **Fichier**

Type de document **Article dans des revues**

Titre **en** Dual Neural Routing of Visual Facilitation in Speech Processing

Sous-Titre **en** Visual facilitation of speech processing

Auteur(s) **en** Luc Amal^{1,*}, Benjamin Morillon¹, Christian A. Kell^{1,2}, Anne-Lise Giraud^{1,3}
* Auteur correspondant

1 Laboratoire de Neurosciences cognitives **81517** - France
Ecole normale supérieure - Paris **59784** ; INSERM **12560** ; **30002**

2 Brain Imaging Center, Department of Neurology **103723** - D-60590 Frankfurt am Main - Allemagne
Goethe-University **324905**

3 CENIR - Centre de Neuro-Imagerie de Recherche **166232** - CENIR - Centre de Neuroimagerie de Recherche Bâtiment de Médecine Physique et Réadaptation niveau -1 47-83 boulevard de l'Hôpital 75651 PARIS Cedex 13 - France
Assistance publique - Hôpitaux de Paris (AP-HP) **300068** ; Hôpital de la Pitié Salpêtrière [Paris] **322144**

Langue du document anglais

Nom de la revue **en** **J Neurosci - Journal of Neuroscience (ISSN : 0270-6474, EISSN : 1529-2401)**
— Published by Society for Neuroscience
— SHERPA/ReMeO **yellow** Can archive pre-print (ie pre-refereeing) **(plus d'infos sur ce journal)**

Audience Internationale

Date de publication 2009-11-28

Page/Identifiant 13445-13453

Domaine **Sciences du Vivant [q-bio]Neurosciences [q-bio.NC]**

Projet(s) ANR **ANR-06-NEURO-0047 - TEMPOFRONT** Physiopathologie des relations temporo-frontales dans le cerveau humain (2006)
— Physiopathologie des relations temporo-frontales dans le cerveau humain
— Neurosciences, neurologie et psychiatrie

Collaboration/Projet Neurophysiology of audiovisual speech integration

Référentiel auteurs

ind	idHAL	nom	prenom
444746	0	Amal	Luc
509623	0	Amal	Luc Henri

Référentiel structures

ind	name
81517	Laboratoire de Neurosciences cognitives

Référentiel revues

ind	titre	issn
6330	Journal of Neuroscience	0270-6474

Référentiel domaines

📁 Sciences du Vivant [q-bio]
▶ 📁 Neurosciences [q-bio.NC]

Référentiel projets ANR

ind	titre	acronyme	reference
2014	Physiopathologie des relations temporo-frontales dans le cerveau humain	TEMPOFRONT	ANR-06-NEURO-0047



API documents

Interroger l'API

Point d'accès	https://api.archives-ouvertes.fr
Chemin	https://api.archives-ouvertes.fr/search/
Critère de recherche et options	https://api.archives-ouvertes.fr/search/?q={query}
Limiter à un portail	https://api.archives-ouvertes.fr/search/univ-orleans/
	">https://api.archives-ouvertes.fr/search/univ-orleans/?q=>*

Réponse par défaut

Le format	Json
3 champs retournés	Docid : Identifiant unique interne du document label_s : Référence bibliographique du document Uri : URI
<pre>▼ response: numFound: 1 start: 0 ▼ docs: ▼ 0: docid: 1778296 ▼ label_s: "Emmanuel Chateau Dutier, Maud Ingarao, Jean-Philippe Magué, Philippe Pons, Séverine Gedzelman, et al.. Towards an XML Corpora Exposition as LOD with the Lightweight Xquery-Based Framework SynopsX . Digital Humanities 2016, Jul 2016, Cracovie, Poland. 2016, &#x27E8;http://dh2016.adho.org/&#x27E9;. &#x27E8;halshs-01778296&#x27E9;" ▼ uri_s: "https://halshs.archives-ouvertes.fr/halshs-01778296"</pre>	
Le nombre de résultat	30 lignes

<http://hal.archives-ouvertes.fr/halshs-01778296>

<https://api.archives-ouvertes.fr/search/?q=halshs-01778296>

Base de documents : recherches simples

Chercher un critère simple :

<https://api.archives-ouvertes.fr/search/?q=casuhal>

Index de recherche par défaut

« text » : Permet de chercher en même temps dans plusieurs champs

Chercher un critère dans un champ :

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

Chercher un critère et filtrer le résultat :

https://api.archives-ouvertes.fr/search/?q=open science&fq=docType_s:THESE

Base de documents : recherches simples

Se repérer par rapport à une interrogation du portail HAL

The screenshot displays the HAL search interface. The browser's address bar shows the URL: `https://hal.archives-ouvertes.fr/search/index/?qa[text][]=open+science&qa[docType_s][]=THESE&submit_advanced=Rechercher&docType_`. A blue box highlights the search criteria in the address bar. Below the browser, the HAL logo and navigation menu are visible. The search results page shows a search for "open science" and "THESE". The filters section on the left indicates 126 results for "Thèse". The search results table shows the following entries:

Champ de recherche par défaut (multicritères)	Type de document
open science	Type de document
THESE	Type de document
Champ de recherche par défaut (multicritères)	+

L'option format

Format

https://api.archives-ouvertes.fr/search/univ-orleans/?q=*&wt=?

<https://api.archives-ouvertes.fr/docs/search/?#wt>

L'option nombre de résultats

rows=12345	NumFound ...
https://api.archives-ouvertes.fr/search/ANR-07-CORP-009/?q=*&rows=134	
Utiliser rows=0	Ne retourne que le nombre de résultats et enlève les documents du corps de la réponse. Vous pouvez par exemple utiliser ce paramètre à des fins de statistiques ou si vous voulez retourner uniquement des facettes.

Les champs retournés

Champs retournés	Docid : Identifiant unique interne du document label_s : Référence bibliographique du document uri_s : URI	fl=champ (champs stockés) fl=champ,champ,champ
https://api.archives-ouvertes.fr/search/?q=title_t:casuhal&fl=*		
https://api.archives-ouvertes.fr/search/?q=title_t:open science&fl=*tdate,title_s,auth*		
https://api.archives-ouvertes.fr/search/?q=title_t:open science&fl=*tdate,title_s,authFullName_s		

Les champs, les types de champs

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

Anr
auth
Citation
Classification
Collection
Conference
Contributor
Dates
Deptstruct
Doctype
Docid
Europeanproject
...

Type	Usage prévu	Exemple
Indexées	recherche	title_t
Stockées	affichage	label_s
Facettes	listes de valeurs	submitType_s
Tri	tri de valeurs multivalués	author_sort

<https://api.archives-ouvertes.fr/search/?q=text:casuhal>

https://api.archives-ouvertes.fr/search/?q=text_fulltext:casuhal

Les options de l'API 2

Ordonner les résultats	<code>https://api.archives-ouvertes.fr/search/?q*:*&sort=docid asc</code>
Paginer le résultat	Pour parcourir des résultats très nombreux (limite max des résultats étant à 10000)
Grouper les résultats	<p>Permet de grouper des résultats en fonction d'un champ. Permet de ne rapporter qu'un certain nombre de résultat par groupe</p> <p>Les 5 premiers documents de la collection FRANCE-GRILLES de chaque groupe de type de documents : triés par date de publication descendante, groupés par type de documents :</p> <pre>http://api.archives-ouvertes.fr/search/ ?q*:* &fq=collCode_s:FRANCE-GRILLES &group=true &group.field=docType_s &group.limit=5 &wt=xml &sort=producedDate_tdate desc &fl=docid,label_s,uri_s,producedDate_tdate</pre>

Modalités d'interrogation

Point d'accès	https://api.archives-ouvertes.fr	
Chemin	https://api.archives-ouvertes.fr/search/	
	Booléens, guillemets, troncature	
	AND par défaut	
	?q=title_t:digital humanities	201
	?q=title_t:(digital humanities)	101
	?q=title_t:(digital OR humanities)	5267
	?q=title_t:«digital humanities»	201
	?q=title_t:humanities NOT digital	108

Critères de recherche

Avoir des listes des contenus des champs pour créer ses requêtes.

Interrogation pour une structure :

Limiter sa recherche dans le chemin par le portail ou la collection OU chercher dans le champ de structure d'affiliation

- authorityInstitution_s : organisme de délivrance pour les thèses
 - https://api.archives-ouvertes.fr/search/?q=*&rows=0&wt=xml&facet=true&facet.field=authorityInstitution_s

- DocType_s : Type de document
 - https://api.archives-ouvertes.fr/search/?q=*&rows=0&wt=xml&facet=true&facet.field=docType_s

- CollCategory_s : types de catégories de collections
 - https://api.archives-ouvertes.fr/search/?q=*&rows=0&wt=xml&facet=true&facet.field=collCategory_s

- JournalTitle_S : Revue titre
 - https://api.archives-ouvertes.fr/search/?q=*&rows=0&wt=xml&facet=true&facet.field=journalTitle_s

Filtrer les réponses

	Pour limiter les résultats	fq=champ:valeur
Filtres sur un critère	<pre>https://api.archives-ouvertes.fr/search/univ-orleans/ ?q=*&wt=xml&fq=collCategory_s:COLLOQUE</pre> <pre>https://api.archives-ouvertes.fr/search/?q=*</pre> <pre>%3A*&rows=0&wt=xml&indent=true&facet=true&facet.field=collCategory_s</pre> <pre>?q=*&wt=xml&fq=collCategory_s:COLLOQUE NOT collCategory_s:(INSTITUTION OR UNIV OR LABO OR AUTRE OR THEME OR ECOLE OR PRES OR REVUE OR SET)</pre>	
Filtres sur les dates (intervalle de dates)	<pre>https://api.archives-ouvertes.fr/search/univ-orleans/ ?q=*&fq=submittedDateY_i:[2019 TO *]</pre>	
Recherche de documents dont l'un des champs est vide ou non	<pre>Exemple de syntaxe pour avoir les documents sans DOI fq=-doiId_s:[" TO *]</pre> <pre>http://api.archives-ouvertes.fr/search/?q=*&wt=xml&fq=-doiId_s:[" TO *]</pre> <hr/> <pre>Exemple de syntaxe pour avoir les documents avec un DOI fq=doiId_s:[" TO *]</pre> <pre>http://api.archives-ouvertes.fr/search/?q=*&wt=xml&fq=doiId_s:[" TO *]&fl=label_s,doiId_s</pre>	













Utiliser les facettes

Pour grouper les résultats

Indique le nombre de document correspondant de chaque groupe

Accueil Dépôt Consultation ▼ Recherche Documentation

Consultation / Par année de publication

Années	Documents	Notices	Annexes	Tous
2015 - 2020	 174286	306528	 4201	485015
2010 - 2015	 227403	451846	 2410	681659
2005 - 2010	 155954	361350	 476	517780
2000 - 2005	 58192	147518	 382	206092
1995 - 2000	 26642	48608	 441	75691
1990 - 1995	 20826	23103	 351	44280

Utiliser les facettes

Activer les facettes	facet=true
Champ sur lequel appliquer la facette	facet.field={champ}
Connaître le nombre de document dans son portail et répartir la réponse par types de documents	
<pre>https://api.archives-ouvertes.fr/search/univ-orleans ?q=:*: &rows=0 &facet=true&facet.field=docType_s</pre>	

Facettes à plusieurs dimensions

Activer les facettes	facet=true	
Champ sur lequel appliquer la “facette à pivot”	facet.pivot={champ1,champ2}	
Connaître le nombre de document par type dans mon portail, et pour chaque type de document, distinguer les types de dépôt (notice, fichier, annexe)		
<pre>https://api.archives-ouvertes.fr/search/univ-orleans/ ?q=:*:* &rows=0 &facet=true&facet.pivot=docType_s,submitType_s</pre>		

Facettes à intervalles

Utilisation des options “range”

Requête qui affiche par intervalle de 2 ans le nombre de dépôt cumulés :

```
https://api.archives-ouvertes.fr/search/univ-orleans/  
?q=*:*  
&rows=0  
&facet=true  
&facet.range=submittedDateY_i  
&facet.range.start=1990&facet.range.end=2017  
&facet.range.gap=2
```



API référentiels

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

Référentiels listes

Instances de portail	Liste toutes les instances de portail de HAL	4 champs 2 options de format
	Liste des instances https://api.archives-ouvertes.fr/ref/instance	
Types de documents	Liste des types de document par portail	3 champs Option instance_s option langue (lang) 2 options de format
	Types de documents accepté dans mon portail https://api.archives-ouvertes.fr/ref/doctype/?instance_s=univ-orleans&wt=xml Types de documents acceptés sur TEL https://api.archives-ouvertes.fr/ref/doctype/?instance_s=tel&wt=xml	
Métadonnées		

Référentiels ouverts/complexes

	Liste contenu	Champs
Auteurs		
Structures		
Disciplines/Domain		
Revue/Journal		
Projets ANR		
Projets européens		



Traitement post-extraction

Outils pour exploiter les fichiers





Requêtes

“

Objectif :

Récupérer le nombre de documents dans mon portail, rangé par année, type de dépôt et type de document

”

Objectif :

Récupérer le nombre de documents dans mon portail, rangé par année, type de dépôt et type de document

```
http://api.archives-ouvertes.fr/search/  
univ-orleans/  
?q=*:*  
&rows=0  
&facet=true&facet.pivot=submittedDateY_i,submitT  
ype_s,docType_s  
&wt=xml
```

“

Objectif :

Récupérer la liste des thèses de doctorat de mon université. Récupérer les champs : identifiant HAL, date de soutenance, auteur, école doctorale, acronyme du laboratoire, titre.

Au format CSV

”

Objectif :

Récupérer la liste des thèses de doctorat de mon université. Récupérer les champs : identifiant HAL, date de soutenance, auteur, école doctorale, acronyme du laboratoire, titre.

Au format CSV

```
http://api.archives-ouvertes.fr/search/  
?q=docType_s:THESE AND authorityInstitution_s:  
Université d'Orléans  
&wt=csv  
&sort=producedDate_s%20desc  
&rows=2000  
&fl=halId_s,defenseDate_s,authFullName_s,thesis  
School_s,labStructAcronym_s,instStructAcronym_s  
,title_s
```

<https://hal-upec-upem.archives-ouvertes.fr/page/api>

Liste
authorityInstitution

“

Objectif :

Récupérer la liste complète des dépôts liés à l'UMS CLEO avec les champs : titre, auteurs, résumé, type de document, date de publication, date de dépôt. Afficher au format xml puis télécharger en csv.

”

Objectif :

Récupérer la liste complète des dépôts liés à l'UMS CLEO avec les champs : titre, auteurs, résumé, type de document, date de publication, date de dépôt. Afficher au format xml puis télécharger en csv.

```
https://api.archives-ouvertes.fr/  
search/  
?q=structId_i:84824  
&fl=uri_s,title_s,authFullName_s,abstract_s  
&rows=70  
&wt=csv
```

“
Objectif :

Récupérer la liste des auteurs qui ont mon institution comme établissement d'appartenance. Retourner les champs qui se terminent par ...Name_s et trier le résultat alphabétiquement par nom de famille.

”

Objectif :

Récupérer la liste des auteurs qui ont mon institution comme établissement d'appartenance. Retourner les champs qui se terminent par ...Name_s et trier le résultat alphabétiquement par nom de famille.

```
http://api.archives-ouvertes.fr/ref/author  
?q=structureId_i:300009  
&sort=lastName_s asc  
&fl=*Name_s
```

“

Objectif :

Récupérer la liste des collections pour lesquelles au moins 100 dépôts liés à mon institution se trouvent

”

Objectif :

Récupérer la liste des collections pour lesquelles au moins 100 dépôts liés à mon institution se trouvent

```
https://api.archives-ouvertes.fr/search/univ-orleans/  
?q=*:*  
&rows=0  
&wt=xml  
&facet=true  
&facet.field=collsParentOfCategoryColl_fs  
&facet.mincount=100
```

“

Objectif :

Récupérer la liste des auteurs affiliés à mon institution avec tous les champs possibles.

”

Objectif :

Récupérer la liste des auteurs affiliés à mon institution avec tous les champs possibles.

```
https://api.archives-ouvertes.fr/  
Ref/author/  
?q=structureId_i:300297  
&fl=*
```

```
&rows=99
```


“
Objectif :
Récupérer la liste des auteurs qui ont mon institution comme établissement d'appartenance. Rapporter les champs se terminant par Name_s triés par nom de famille.”

Objectif :

Récupérer la liste des auteurs qui ont mon institution comme établissement d'appartenance. Rapporter les champs se terminant par Name_s triés par nom de famille.

```
http://api.archives-ouvertes.fr/  
Ref/author/  
?q=structureId_i:300009  
&wt=json  
&sort=lastName_s asc  
&fl=*Name_s
```

```
&rows=99
```

“

Objectif :

Récupérer la liste des instances existantes dans HAL qui comprend l'indication du nombre de dépôts effectués dans chacune des instances

”

Objectif :

Récupérer la liste des instances existantes dans HAL qui comprend l'indication du nombre de dépôts effectués dans chacune des instances

- Chercher le champ instance dans la documentation HAL
- instance_s → lien plus d'exemple
- [https://api.archives-ouvertes.fr/search/](https://api.archives-ouvertes.fr/search/?q=*:*&rows=0&wt=xml&facet=true&facet.field=instance_s)
- [?q=* :*](https://api.archives-ouvertes.fr/search/?q=*:*&rows=0&wt=xml&facet=true&facet.field=instance_s)
- [&rows=0](https://api.archives-ouvertes.fr/search/?q=*:*&rows=0&wt=xml&facet=true&facet.field=instance_s)
- [&wt=xml](https://api.archives-ouvertes.fr/search/?q=*:*&rows=0&wt=xml&facet=true&facet.field=instance_s)
- [&facet=true](https://api.archives-ouvertes.fr/search/?q=*:*&rows=0&wt=xml&facet=true&facet.field=instance_s)
- [&facet.field=instance_s](https://api.archives-ouvertes.fr/search/?q=*:*&rows=0&wt=xml&facet=true&facet.field=instance_s)

“
Objectif :
Récupérer les affiliations de l’auteur : Marin Dacos
”

Objectif :

Récupérer les affiliations de l'auteur : Marin Dacos

```
https://api.archives-ouvertes.fr/  
ref/  
affiliation/  
?lastName_t=Dacos&firstName_t=Marin
```

“
Objectif :
Afficher la liste des revues dans lesquels les dépôts de ma structure sont publiés. Pour chaque revue avoir le nombre.
”

Objectif :

Afficher la liste des revues dans lesquels les dépôts de ma structure sont publiés. Pour chaque revue avoir le nombre.

- [https://api.archives-ouvertes.fr/](https://api.archives-ouvertes.fr/search/univ-orleans/?q=*:*&rows=0&facet=true&facet.field=journalTitle_s&wt=xml)
- [search/univ-orleans/](https://api.archives-ouvertes.fr/search/univ-orleans/)
- [?q=* :*](https://api.archives-ouvertes.fr/search/univ-orleans/?q=*:*)
- [&rows=0](https://api.archives-ouvertes.fr/search/univ-orleans/?q=*:*&rows=0)
- [&facet=true&facet.field=journalTitle_s](https://api.archives-ouvertes.fr/search/univ-orleans/?q=*:*&facet=true&facet.field=journalTitle_s)
- [&wt=xml](https://api.archives-ouvertes.fr/search/univ-orleans/?q=*:*&facet=true&facet.field=journalTitle_s&wt=xml)

“
Objectif :
Récupérer la liste des dépôts du mois dernier qui ont été effectués dans mon portail
”

https://api.archives-ouvertes.fr/
search/univ-orleans/
?q=* :*
&fq=submittedDate_tdate:[NOW-1MONTHS/DAY TO NOW/HOUR]

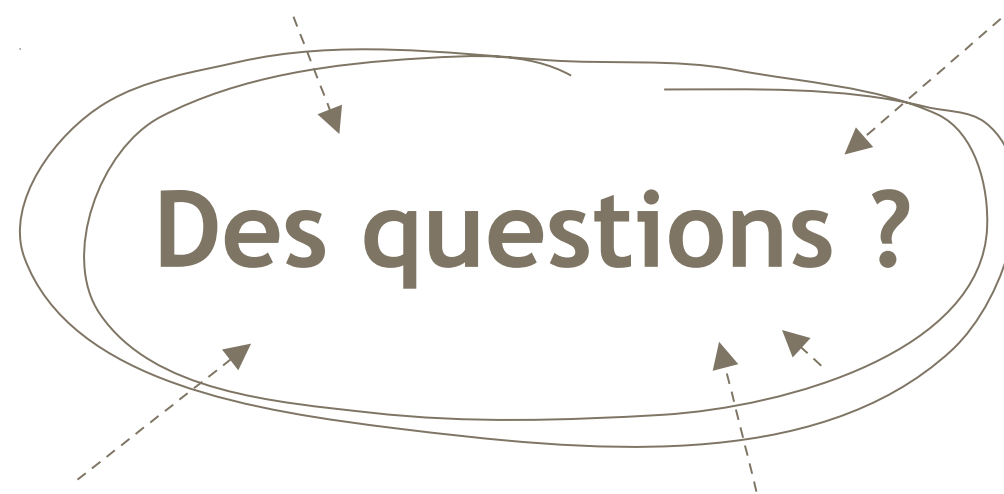
“
Objectif :
Récupérer la liste des dépôts dans mon portail liés à des projets ANR
commencés en 2016
”

Objectif :

Récupérer la liste des dépôts dans mon portail liés à des projets ANR commencés en 2016

- [https://api.archives-ouvertes.fr/search/univ-orleans/](https://api.archives-ouvertes.fr/search/univ-orleans/?q=anrProjectYearDate_s:%222016%22&wt=xml)
- [&q=anrProjectYearDate_s:%222016%22](https://api.archives-ouvertes.fr/search/univ-orleans/?q=anrProjectYearDate_s:%222016%22&wt=xml)
- [?wt=xml](https://api.archives-ouvertes.fr/search/univ-orleans/?q=anrProjectYearDate_s:%222016%22&wt=xml)

Merci !



Wiki de CasuHAL (rubrique API) : <https://wiki.ccsd.cnrs.fr/wikis/hal/index.php/Accueil>