



## Works-magnet - Jeux de données - Tutoriel

Laetitia Bracco, Eric Jeangirard, Anne L'Hôte

### ► To cite this version:

Laetitia Bracco, Eric Jeangirard, Anne L'Hôte. Works-magnet - Jeux de données - Tutoriel. 2026. <hal-05650426>

**HAL Id: hal-05650426**

**<https://hal.univ-lorraine.fr/hal-05650426v1>**

Preprint submitted on 9 Jun 2026

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# Works-magnet - Jeux de données

## Tutoriel

L'objectif de cette partie de l'outil Works-magnet est de faciliter la recherche et la sélection de **jeux de données** dans la base de données du Baromètre français de la Science Ouverte (qui moissonne notamment DataCite, l'agence internationale attribuant des identifiants aux jeux de données de la recherche). Ce tutoriel permet de générer des **indicateurs sur les jeux de données produits par un établissement** et disposant d'**identifiants DataCite**.

<https://works-magnet.esr.gouv.fr/>

## 1. La recherche

### 1. 1. Recherche de jeux de données

- Choisir l'entrée "*Les jeux de données de mon institution*"

#### Améliorer les détections automatiques faites par IA

Des techniques d'IA sont utilisées à grande échelle pour analyser des corpus bibliographiques. En particulier, OpenAlex les utilise pour construire des liens d'affiliations dans les publications (utilisés pour le classement de Leiden) ou encore Grobid-Softcite pour détecter les mentions de données et logiciel dans le texte intégral. Contribuer à corriger ces données améliore la fiabilité et la représentativité de ces analyses. En savoir plus

##### Améliorer les affiliations ROR dans OpenAlex

Analyser les chaînes d'affiliation brutes les plus fréquentes récupérées dans OpenAlex pour votre recherche.  
Vérifier les ROR calculés automatiquement par OpenAlex. Parfois, ils peuvent être inexacts ou manquants.  
Corriger les erreurs (ROR inexacts ou manquants) et envoyez un retour d'information à OpenAlex.

##### Les mentions de logiciels ou de données de la recherche

Explorer les mentions de logiciels et jeux de données trouvés dans le texte intégral  
Corriger les erreurs (type ou caractérisations)  
Soumettre les corrections

#### Constituer un corpus de publications ou de jeux de données

##### Les publications de mon institution

Explorez les chaînes d'affiliation brutes les plus fréquentes récupérées dans les données françaises de l'Open Science Monitor et dans OpenAlex pour votre requête.  
Valider ☒ ou exclure ☒ chacune d'entre elles, qu'elle corresponde effectivement à votre institution ou non.  
Enregistrer (exporter vers un fichier) ces décisions et le corpus de publications que vous venez de construire.

##### Les jeux de données de mon institution

Explorer les signatures d'affiliation brutes les plus fréquentes récupérées dans les données françaises de l'Open Science Monitor et dans OpenAlex pour votre requête (jeux de données uniquement)  
Valider ☒ ou exclure ☒ chacun d'entre eux, qu'il corresponde effectivement à votre institution ou non.  
Sauvegarder (exporter vers un fichier) ces décisions et le corpus de datasets que vous venez de construire.

<https://works-magnet.esr.gouv.fr/datasets/search>

## 1. 2. Recherche des jeux de données par structure

- Saisir le ROR de votre structure puis presser la touche Entrée pour valider
- Si pertinent, ajouter les enfants liés à cette structure dans ROR en cliquant sur "Get children from ROR"
- Sélectionner les années souhaitées
- Cliquer sur "Search datasets"

[Home](#) > Find the datasets affiliated to your institution

Affiliation name, ROR of your institution \*

Press ENTER to search for several terms / expressions. If several, an OR operator is used.

Start year: 2020, End year: 2024

04vfs2w97 X UL X University of Lorraine X Université de Lorraine X Get children from ROR

Client Id (to add a whole repository, in case of repository managed by your institution)

To find your client.id, please open the [Datacite API link](#) and replace "YOUR\_DOI" by a DOI of your repository, then look for the value of the field `relationships.client.data.id`. Multiple values should be comma separated.

Clear search Search datasets

<https://works-magnet.esr.gouv.fr/datasets/search?affiliations=https%3A%2F%2Fror.org%2F04vfs2w97&endYear=2024&getRorChildren=0&startYear=2020>

- Dans le champs "Affiliation name, ROR of your institution", il est recommandé de saisir un ROR, mais vous pouvez également saisir une chaîne de caractères. Ce champ peut-être multi-valué, dans ce cas la recherche retournera tous les jeux de données rattachés à l'une ou à l'autre des valeurs saisies.

- À cette étape, le champ "Client Id" permet de saisir un "client.id" selon la nomenclature DataCite. Ceci permet d'ajouter et de valider automatiquement tous les jeux de données issus d'un entrepôt attribuant des DOI DataCite. Ainsi, dans le cas d'un entrepôt dépendant intégralement de votre structure, tous les jeux de données de cet entrepôt seront automatiquement ajoutés dans votre corpus final.

- Pour trouver votre client.id, veuillez ouvrir le [lien de l'API Datacite](#) en remplaçant "YOUR\_DOI" par le DOI d'un de vos jeux de données, puis rechercher la valeur du champ "relationships.client.data.id". Les valeurs multiples doivent être séparées par des virgules.

### 1. 3. Résultats

L'outil va alors lister les affiliations d'une part (voir l'onglet "Select the raw affiliations for your institution"), et les jeux de données d'autre part (voir l'onglet "List of datasets"), retrouvés pour le(s) ROR et / ou le(s) libellé(s) de la recherche dans la base DataCite.

Home > Build my corpus of datasets from repositories > Select the affiliations and build the corpus

Find the datasets affiliated to your institution

Select the raw affiliations for your institution

List of datasets

Insights

Select the raw affiliations corresponding to your institution

The array below summarizes the most frequent raw affiliation strings retrieved in the French Open Science Monitor data and in OpenAlex for your query.

You can validate  or exclude  each of them, whether it actually corresponds to your institution or not. If an affiliation is validated, it will also validate all the works with that affiliation string.

The second column indicates the ROR automatically computed by OpenAlex. Sometimes, they can be inaccurate or missing. If any errors, please use the first tab to send feedback.

You can save (export to a file) those decisions, and restore them later on.

0 selected affiliations

Validate

Exclude

Reset status

To be decided (185 ie. 100 %)

Select all  Select page

Search in affiliations name

100

<<

<

1 to 100 of 185

>

>>

| Status                   | Affiliation                                                                                                            | ROR computed by OpenAlex | Examples of works                                                                                                                                                                                                 | Number of works |
|--------------------------|------------------------------------------------------------------------------------------------------------------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <input type="checkbox"/> | Any                                                                                                                    |                          |                                                                                                                                                                                                                   |                 |
| <input type="checkbox"/> | <div>TO BE DECIDED</div> university of lorraine ror 04vfs2w97 [ source: FOSM ]                                         |                          | <div>• <a href="#">10.6084/m9.figshare.c.5047189</a></div> <div>• <a href="#">10.6084/m9.figshare.c.5010101</a></div>                                                                                             | 118             |
| <input type="checkbox"/> | <div>TO BE DECIDED</div> universite de lorraine [ source: FOSM ]                                                       |                          | <div>• <a href="#">10.5281/zenodo.3609502</a></div> <div>• <a href="#">10.12763/smdgr1</a></div>                                                                                                                  | 46              |
| <input type="checkbox"/> | <div>TO BE DECIDED</div> universite de lorraine ul 34 cours leopold cs 25233 54052 nancy cedex france [ source: FOSM ] |                          | <div>• <a href="#">hal-03569077</a></div> <div>• <a href="#">hal-02963802</a></div> <div>• <a href="#">hal-02985174</a></div> <div>• <a href="#">hal-02985175</a></div> <div>• <a href="#">hal-03146301</a></div> | 40              |

<https://works-magnet.esr.gouv.fr/datasets/results?affiliations=https%3A%2F%2Fror.org%2F04vfs2w97&endYear=2024&getRorChildren=0&startYear=2020>

### 2. La sélection

La sélection correspond à l'étape durant laquelle vous allez constituer le corpus à partir duquel seront ensuite calculés les indicateurs sur les jeux de données.

## 2. 1. La sélection par affiliation

- Parmi les formes d'affiliations retrouvées, sélectionner celles relatives à votre établissement
- Cliquer sur "Validate"

2 selected affiliations

☒ Validate
 ☐ Exclude

To be decided (185 ie. 100 %)

100

<< < 1 to 100 of 185 > >>

| Status                                            | Affiliation                                                                                   | ROR computed by OpenAlex | Examples of works                                                                                                                                | Number of works |
|---------------------------------------------------|-----------------------------------------------------------------------------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <input type="checkbox"/> Any                      |                                                                                               |                          |                                                                                                                                                  |                 |
| <input checked="" type="checkbox"/> TO BE DECIDED | university of lorraine ror 04vts2w97 [ source: FOSM ]                                         |                          | <a href="#">10.6084/m9.figshare.c.5047189</a> <a href="#">10.6084/m9.figshare.c.5010101</a>                                                      | 118             |
| <input checked="" type="checkbox"/> TO BE DECIDED | universite de lorraine [ source: FOSM ]                                                       |                          | <a href="#">10.5281/zenodo.3609502</a> <a href="#">10.12763/smdgr1</a>                                                                           | 46              |
| <input type="checkbox"/> TO BE DECIDED            | universite de lorraine ul 34 cours leopold cs 25233 54052 nancy cedex france [ source: FOSM ] |                          | <a href="#">hal-03569077</a> <a href="#">hal-02963802</a> <a href="#">hal-02985172</a> <a href="#">hal-02985174</a> <a href="#">hal-02985173</a> | 40              |
| <input type="checkbox"/> TO BE DECIDED            | universite de lorraine ror 04vts2w97 [ source: FOSM ]                                         |                          | <a href="#">10.5061/dryad.hhmgknkhv</a> <a href="#">10.5061/dryad.wh70rxwnf</a>                                                                  | 39              |

- Idéalement, reproduire cette prise de décision sur toutes les formes d'affiliation retrouvées.

- Ici les jeux de données sont regroupés par affiliation afin de permettre une validation plus rapide et plus efficace. La source dans laquelle cette affiliation a été retrouvée est précisée entre crochets. Le nombre de jeux de données concernés est indiqué dans la dernière colonne. En cas de doute, des exemples de jeux de données sont listés dans la colonne "Examples of works".

## 2. 2. La sélection par jeu de données

Après avoir validé les affiliations, il faut examiner les jeux de données ne disposant pas nécessairement d'affiliations complètes et qui peuvent être ajoutés au corpus.

Les jeux de données correspondant aux affiliations validées à l'étape précédente sont déjà comptabilisés (ils ont le statut "Validated").

- Choisir l'onglet *"List of datasets"*
- Sélectionner les jeux de données qui semblent pertinents pour votre structure
- Cliquer sur *"Validate"*

Find the datasets affiliated to your institution

Select the raw affiliations for your institution

List of datasets Insights

Export datasets (2283)

1 selected dataset

Validate

Exclude

Reset status

Validate 1158 datasets without affiliations but linked to an article from my institution

Validate 237 datasets without affiliations but at least 3 authors detected from my institution

Validated (281 ie. 1... To be decided (2002 ie. 88 %)

Select all Select page

Search in any field

100 << < 1 to 100 of 2002 > >>

| Status        | Ids                          | Type    | Year | Publisher  | client.id | Affiliations                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Linked Article | My institution author ORCID |
|---------------|------------------------------|---------|------|------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|-----------------------------|
| To be d...    |                              |         |      |            |           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                |                             |
| TO BE DECIDED | 10.3410/E734885577.793569689 | dataset | 2020 | H1 Connect |           | <ul style="list-style-type: none"> <li>close Big data in IBD: a look into the future. Olivera P 1 Danese S 2 3 Jay N 4 Natoli G 3 email Peyrin-Broulet L 5 1 Gastroenterology Section, Department of Internal Medicine, Centro de Educación Médica e Investigaciones Clínicas (CEMIC), Buenos Aires, Argentina 2 IBD Center, Department of Gastroenterology, Humanitas Clinical and Research Centre, Rozzano, Milan, Italy 3 Humanitas Clinical Research Hospital, Rozzano, Milan, Italy 4 Opailleur and Department of Medical Information, LORIA and Nancy University Hospital, Vandœuvre-lès-Nancy, Nancy, France 5 INSERM U954 and Department of Hepatogastroenterology, Nancy University Hospital, Université de Lorraine, Vandœuvre-lès-Nancy, Nancy, France.</li> </ul> |                |                             |

- Un lien vers le jeu de données est disponible. Les affiliations retrouvées sont ici listées. Si une publication est liée à ce jeu de données, elle est précisée dans "Linked Article".

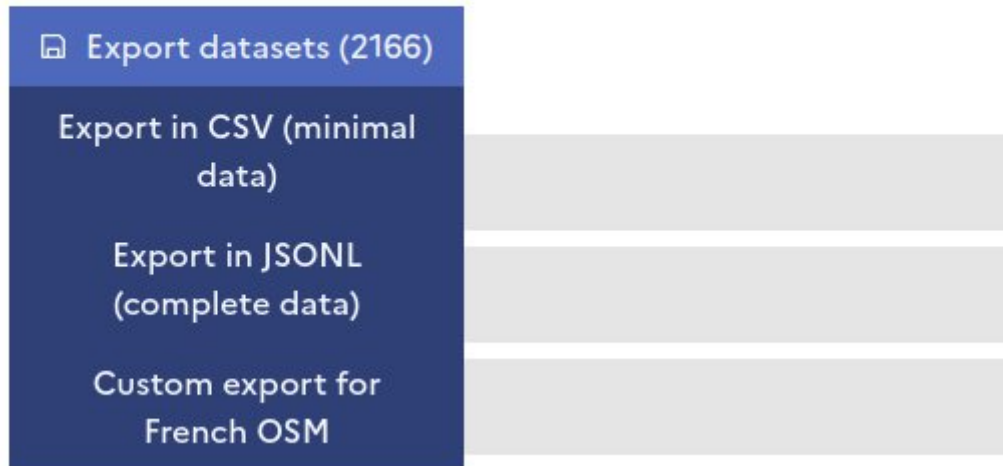
- Les actions "Validate the datasets without affiliations but linked to an article from my institution" et "Validate the datasets without affiliations but at least 3 authors detected from my institution" permettent respectivement de valider les jeux de données liés à une publication référencée dans votre baromètre local (si existant), ou de valider les jeux de données dont au moins 3 auteurs sont affiliés à votre établissement. Ces 2 actions sont conditionnées à une recherche par ROR.

## 2. 3. Export

- Placer votre souris sur *"Export datasets"*

- Dans la liste déroulante, choisir *"Custom export for French OSM"*

## Find the datasets affiliated to your institution



### 3. Baromètre de la Science Ouverte

Ce fichier est directement téléchargeable dans le "Studio" du Baromètre de la Science Ouverte afin de faire une demande de déclinaison locale de Baromètre des jeux de données, veiller à bien choisir l'action "Baromètre local des jeux de données"

<https://barometredelascienceouverte.esr.gouv.fr/declinaisons/comment-realiser-bso-local?expanded=1>

## 2. Envoyez-nous votre fichier

Vous pouvez nous envoyer votre fichier tout au long de l'année.  
Les dates de constitution et d'envoi de ce fichier n'ont pas  
d'incidence sur les graphiques générés car ceux-ci reposent sur  
des "snapshots" des outils utilisés (Unpaywall, PubMed...).

Action demandée

Baromètre local des jeux de données



Email de contact \*

Identifiant de l'établissement

Tel que communiqué lors de votre précédente demande

Nom de la structure \*

Acronyme de la structure

Fichier de publications

Fichier Excel, ODS ou CSV (séparateur point virgule ;). Merci d'inclure une ligne d'en-  
têtes avec les noms de colonnes, comme dans le fichier exemple.

**Browse...**

works\_magnet\_datasets\_bso\_1780676452979\_2020\_2024\_32849130

Ce fichier contient :

408 DOI

*Guide réalisé en juin 2026 par l'Université de Lorraine et le Ministère de  
l'Enseignement Supérieur, de la Recherche et de l'Espace, [bso@recherche.gouv.fr](mailto:bso@recherche.gouv.fr)*

