



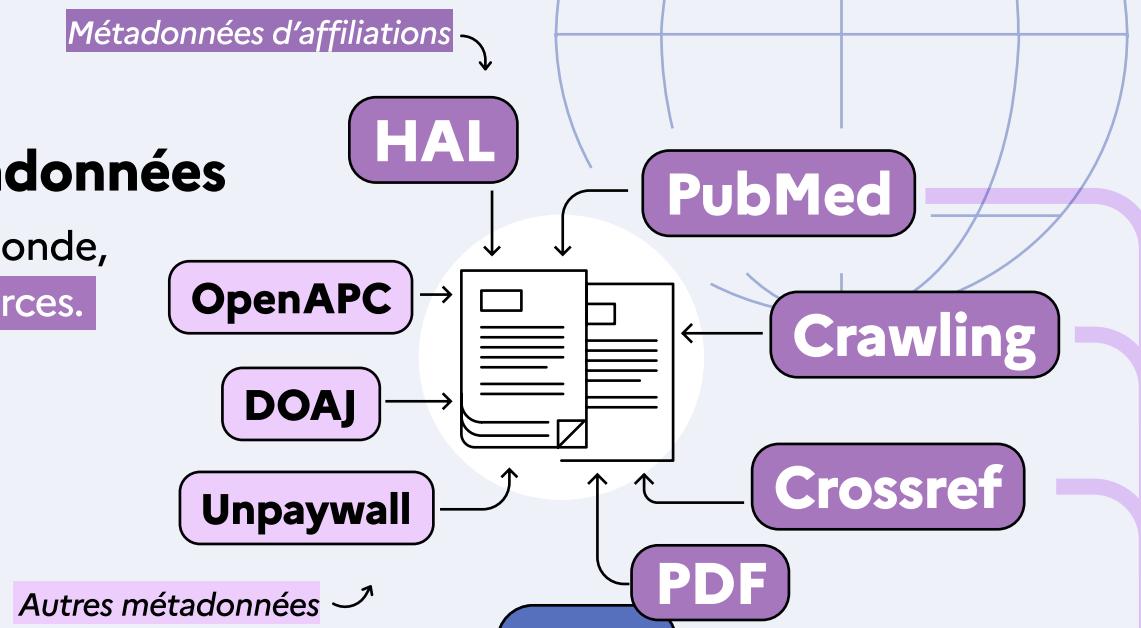
# Baromètre français de la Science Ouverte Publication : une méthodologie inédite

Nous avons donc décidé de collecter un maximum de métadonnées d'affiliations pour chaque publication du monde, à partir de **plusieurs sources libres**.

Notre originalité : aucune utilisation de bases de données propriétaires.

## #1 Collecter... un maximum de métadonnées

Pour chaque publication du monde, consolidation de plusieurs sources.



## #2 Déetecter le pays d'affiliation

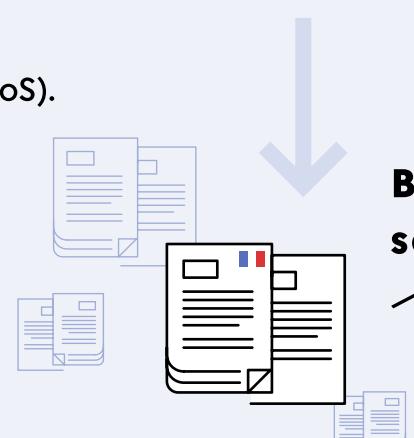
On filtre pour ne garder que les publications pour lesquelles au moins une affiliation est française.

Taux de détection des publications scientifiques françaises



La méthodologie du Baromètre a permis de constituer la **base de données des publications françaises la plus exhaustive à ce jour dans le monde\***.

60% pour un des outils mondiaux de référence, le Web of Science (WoS).



Outil libre d'extraction d'informations structurées dans des publications scientifiques

"Sorbonne Université, Paris" → France ✓

"Hotel Dieu de France, Beirut, Lebanon" → Liban ✗

Base des publications scientifiques françaises

170 000/an



## #3 Enrichir...

### ... le statut d'ouverture



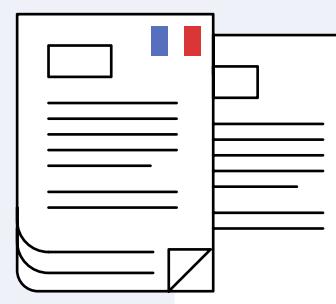
Pour les DOI crossref :

les informations viennent d'[Unpaywall](#)

Pour les publications dans HAL (sans DOI) :

les informations viennent de [HAL](#)

#### Éditeur et archive ouverte



Archive ouverte

Éditeur

Accès fermé



### ... la classification disciplinaire

Grâce à un **algorithme de machine learning de classification automatique** (fastText) basé sur le titre, le résumé et le nom de la revue.



Sciences sociales

Informatique

Biologie fondamentale

Sciences de la terre, Ecologie

Mathématiques

Sciences humaines

Sciences physiques, Astronomie

Sciences de l'ingénieur

Recherche médicale

Chimie

FastText  
entraîné sur  
les bases



Si des métadonnées sont disponibles dans HAL, on effectue **un transcodage de la nomenclature HAL vers celle du Baromètre.**

## #4 Partager...

### avec la communauté toutes ces données agrégées et calculées

Données datavisualisées sur le site du Baromètre ...

[barometredelascienceouverte.esr.gouv.fr](http://barometredelascienceouverte.esr.gouv.fr)

... et disponibles sur le portail Open Data du MESR

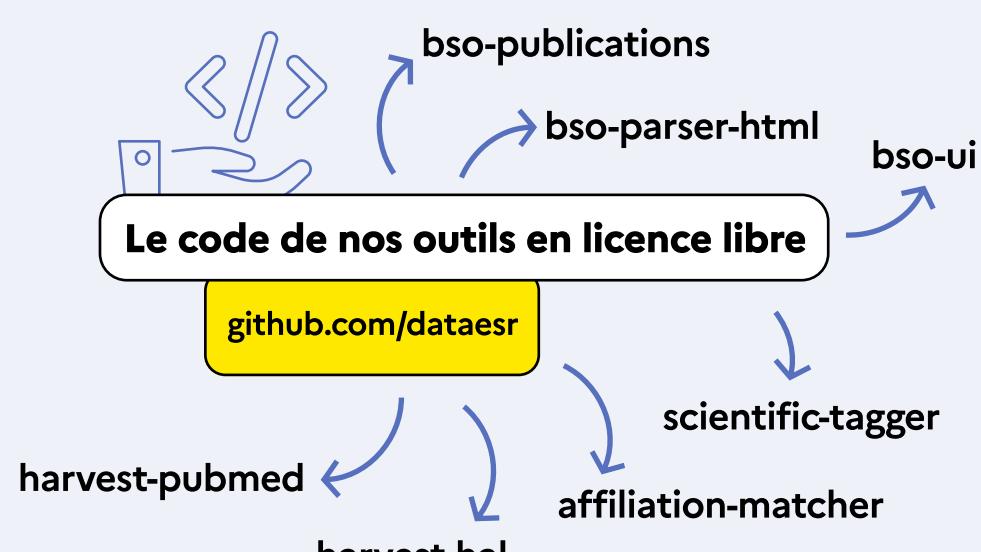
[data.enseignementsup-recherche.gouv.fr](http://data.enseignementsup-recherche.gouv.fr)

### Mais aussi...



Les déclinaisons locales avec les Baromètres locaux

[barometredelascienceouverte.esr.gouv.fr/déclinaisons/comment-realiser-bso-local](http://barometredelascienceouverte.esr.gouv.fr/déclinaisons/comment-realiser-bso-local)



\* Lauranne Chaignon, Daniel Egret; Identifying scientific publications countrywide and measuring their open access: The case of the French Open Science Barometer (BSO). Quantitative Science Studies 2022; 3 (1): 18–36. doi: [doi.org/10.1162/qss\\_a\\_00179](https://doi.org/10.1162/qss_a_00179)