



# Améliorer la qualité des données d'OpenAlex

*Regards croisés  
Sorbonne Université /  
Université de Lorraine*

**Nicolas Alarcon - Laetitia Bracco**

Journée francophone OpenAlex - 2 décembre 2025

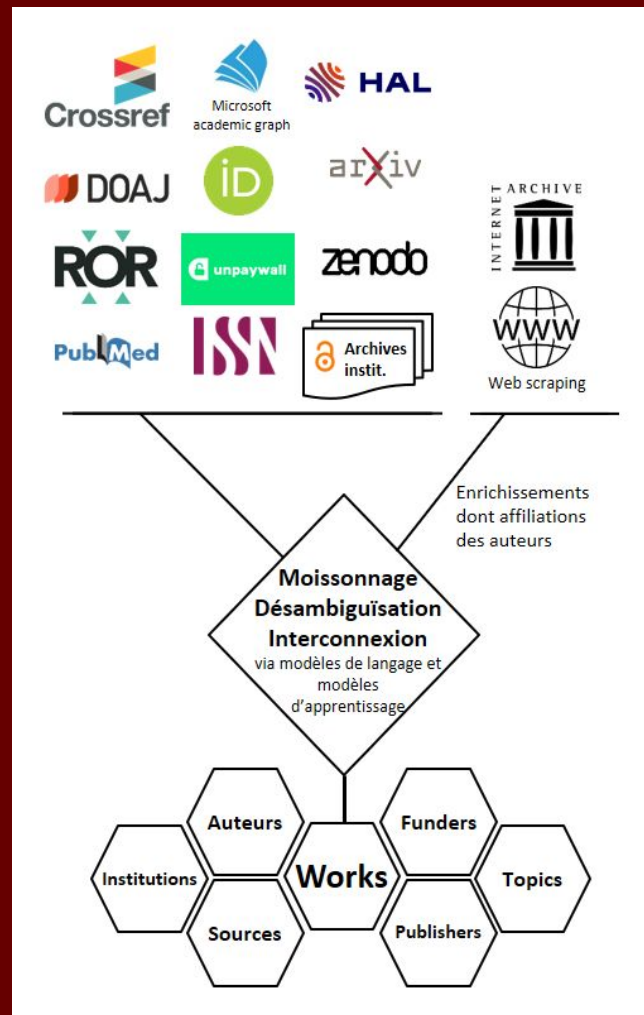


# OpenAlex, c'est quoi ?

Base de données bibliographiques et bibliométriques ouverte et gratuite

Vise l'exhaustivité de la production scientifique mondiale (type, origine géographique, langues, etc.)

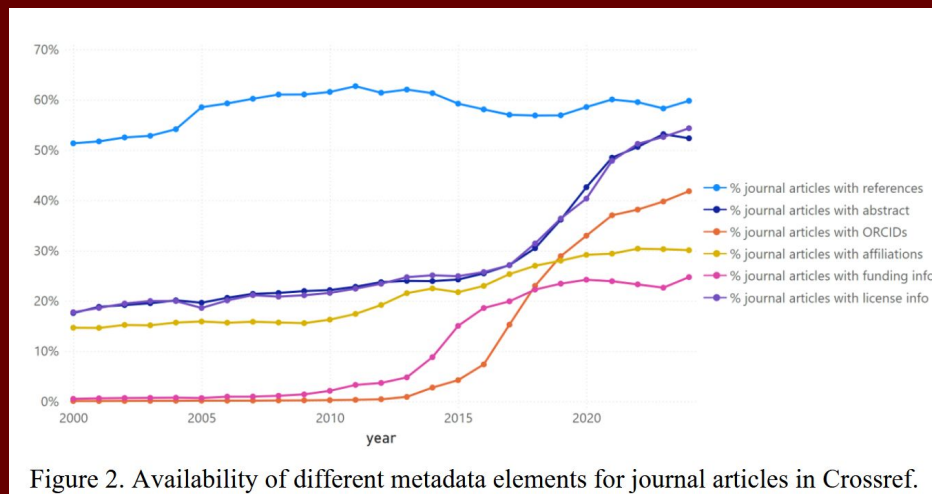
Alimentée par plusieurs sources :  
MAG, Crossref, Pubmed, ROR,  
Datacite, archives ouvertes,  
webscrapping



# Pourquoi faire de la curation dans OpenAlex ?

Applique des algorithmes pour inférer des métadonnées non récupérées via les sources

Nécessité de consolider ces métadonnées d'affiliation, en particulier pour la France



van Eck, Nees, et Ludo Waltman. « Crossref as a Source of Open Bibliographic Metadata ». OSF, 12 mai 2025. *OSF Preprints*, [https://doi.org/10.31222/osf.io/smx5\\_v2](https://doi.org/10.31222/osf.io/smx5_v2), p.4

# Mise en qualité dans ROR : une première étape essentielle

## Université de Lorraine

---

Un chantier mis en place en 2023 au sein des bibliothèques, un dossier qui touche aujourd'hui plusieurs directions

27 demandes de création, de très nombreuses corrections sur les noms, alias, acronymes

Objectifs :

- Nettoyage et alignement des données dans les référentiels AuréHAL, Idref, ROR et Wikidata pour les structures de recherche UL
- Etablir l'historique des fusions et scissions de laboratoires

# Mise en qualité dans ROR : une première étape essentielle

chantier nov 2023 → avril 2024 puis avril 2025

142

demandes  
d'alignement  
d'identifiants

132

structures  
correctement  
référéncées

60

demandes de  
création de  
structures child

6 rattachements child  
7 rattachements related  
6 détachements

## Sorbonne Université

---

Incontournable pour lancer le travail de nettoyage dans OpenAlex puisque *“chaque institution dans OpenAlex dispose d'une entrée correspondante dans ROR”* ([source](#))

# Positions politiques en faveur d'OpenAlex

## Université de Lorraine

---

En juillet 2024, parution de l'article "Pour une évaluation de la recherche transparente et de qualité" par Hélène Boulanger, présidente de l'Université de Lorraine et Nicolas Fressengeas, Vice-président en charge du numérique, des données et de la science ouverte

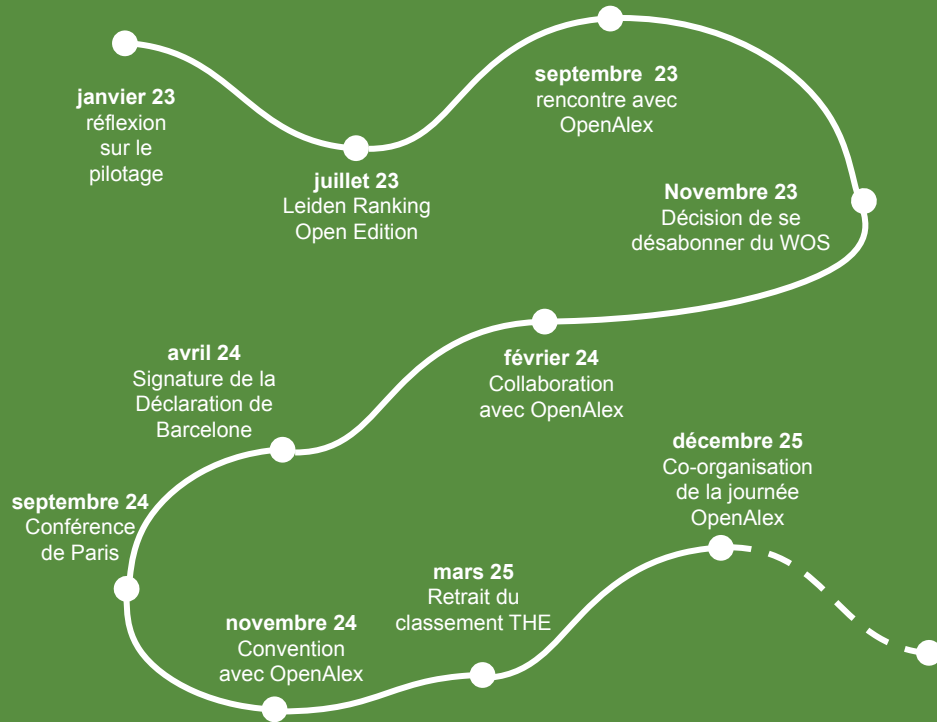
Signature de la Déclaration de Barcelone

Arrêt des corrections dans le Web of Science

Soutien financier de 10 000 € par an à OpenAlex

Retrait des classements THE et QS, plus de commentaire sur le classement de Shanghai

# Positions politiques en faveur d'OpenAlex



**Sorbonne Université**

---

# L'accompagnement et la formation des utilisateurs

## Université de Lorraine

---

Depuis décembre 2024, mise en place d'un atelier en ligne mensuel sur OpenAlex ([consulter le support](#))

Formation des doctorants dans le cadre du séminaire annuel "Approche critique des indicateurs bibliométriques"

Accompagnement individuel à l'utilisation d'OpenAlex pour les chercheurs, sur demande

En complément, atelier en ligne mensuel à l'utilisation de VOSviewer à partir d'OpenAlex

Prochains rapports bibliométriques, repérage des dépenses d'APC du site lorrain et Baromètre lorrain de la science ouverte basés sur OpenAlex

# Formation et sensibilisation de la communauté

Depuis 2023, 4 formations abordant OpenAlex proposées pour un total de 15 participant·es

Présentations à la communauté :

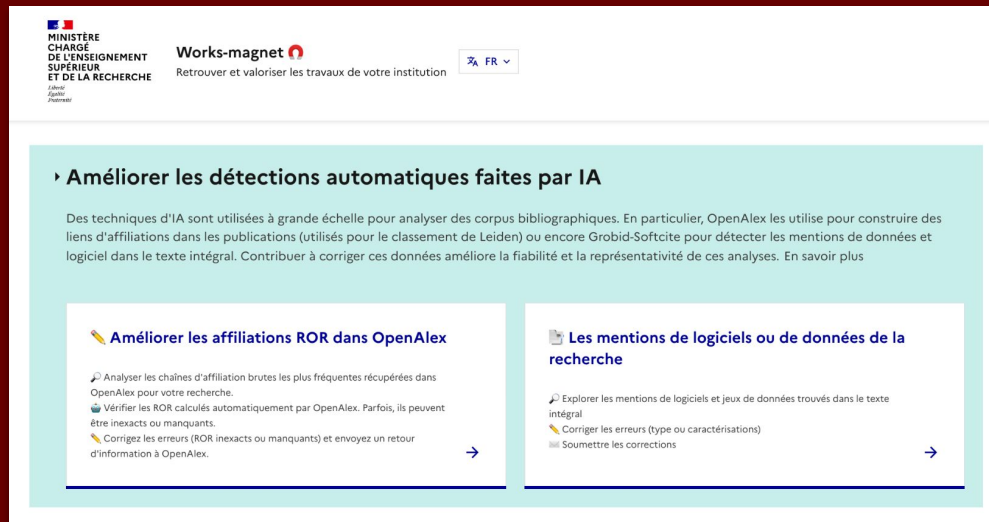
- Conférence de Paris sur les ORI
- Journée Science ouverte CNRS 2024
- Délégation de l'Université de Kyoto
- Conférence Liber 2025
- ScholAgora (association de bibliothécaires - Japon)
- LIS Bibliometrics Annual Conference
- Documentalistes de la Commission européenne

**Sorbonne Université**

---

# Un outil de contrôle qualité : le works-magnet

- Un projet piloté par le SIES (Service Statistique Ministériel de l'ESR) du MESRE
- Objectif : rendre visibles les calculs automatisés par IA et pouvoir les corriger
- 4 grands cas d'usage et en particulier : **améliorer les affiliations ROR dans OpenAlex**
- A partir d'un ROR ou d'un nom, obtenir la liste des affiliations brutes correspondantes et vérifier le lien affiliation / ROR dans OpenAlex
- L'outil permet d'envoyer en masse des corrections à OpenAlex
- +150 000 demandes de correctifs déjà reçues



The screenshot shows the 'Works-magnet' website interface. At the top left is the logo of the 'MINISTÈRE CHARGÉ DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE' with the motto 'Étudier Partager Progresser'. To the right is the 'Works-magnet' logo and the tagline 'Retrouver et valoriser les travaux de votre institution'. A language selector shows 'FR'. The main content area has a light blue background and features a heading 'Améliorer les détections automatiques faites par IA'. Below this heading is a paragraph explaining that AI techniques are used to analyze bibliographic corpora, specifically OpenAlex for affiliation links and Grobid-Softcite for software and data mentions. Two main action cards are displayed: 'Améliorer les affiliations ROR dans OpenAlex' and 'Les mentions de logiciels ou de données de la recherche'. Each card contains a list of steps and a right-pointing arrow.

**MINISTÈRE CHARGÉ DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE**  
Étudier  
Partager  
Progresser

**Works-magnet**  
Retrouver et valoriser les travaux de votre institution

FR

## 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.
- ✍️ Corrigez 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

# La curation, jeu solo ou affaire d'équipe ?

## Université de Lorraine

---

L'équipe Bibliométrie de l'Université de Lorraine est constituée de 3 personnes pour environ 0,5 ETP, 60 laboratoires

Première vague de corrections via le works-magnet répartie sur 6 mois et effectuée par une seule personne, la responsable du service

Près de 12 500 demandes de correction effectuées

Un travail qui porte ses fruits : les laboratoires ont reçu en moyenne 240 publications de plus. Certains sont passés de 0 à plus de 800 !

Mais tout n'a pas été correctement ingéré

Le travail se poursuivra à l'avenir, en étant réparti sur les 15 bibliothécaires d'appui à la recherche qui travaillent dans HAL, à raison d'une à deux vagues de corrections par an

# La curation, jeu solo ou affaire d'équipe ?

GT Openalex créé en 2024, 6 personnes

Comprendre les écarts de référencement en comparant OpenAlex et notre archive ouverte institutionnelle

Abandon de la comparaison pour se consacrer à la curation. D'abord au sein du GT puis sur 1 personne

25 700 demandes de corrections en 3.5 mois

En projet : curation des faux-positifs (affilié à SU de manière erronée, plus difficile à détecter/traiter), curation des auteurs, topics, etc.

**Sorbonne Université**

---

## En conclusion : oui, c'est possible !

- L'amélioration de la qualité des affiliations ne dépend pas que d'OpenAlex, mais aussi : des éditeurs et de la communauté
- Il est crucial de se mettre en ordre de marche pour reprendre le contrôle sur les métadonnées des établissements
- La mise en qualité de ces métadonnées profite à tous
- La couverture élargie des productions scientifiques indexées par OpenAlex contribue à un meilleur repérage de l'activité des chercheurs et donc à une meilleure reconnaissance de la diversité des productions, dans l'esprit de CoARA

Merci de  
votre  
attention

**Nicolas Alarcon**

nicolas.alarcon@sorbonne-universite.fr

**Laetitia Bracco**

laetitia.bracco@univ-lorraine.fr