

Reprendre le contrôle de la donnée scientifique : suivre ses publications avec OpenAlex

c@fé Renatis
07/05/24

Laetitia Bracco laetitia.bracco@univ-lorraine.fr

Comparaison des outils



Clarivate
Web of Science™

- Base de données **bibliographique et bibliométrique** de la société commerciale Clarivate : recherche documentaire et analyse des publications (par les citations, les revues, les pays...)
- Outil propriétaire, **données fermées**
- **Corpus restreint** : les publications non anglophones et les domaines SHS sont sous-représentés
- Couverture des publications scientifiques de l'Université de Lorraine sur la période 2016/2023 : environ **31 700 références**
- Un **référencement qui a un coût humain** : l'équipe biométrie corrige les erreurs d'affiliation de toute la base pour l'UL tous les deux ans
- **Abonnement** : 55 000 € par an
- Outil utilisé actuellement dans quatre grands cas d'usages : **classements internationaux, rapports biométriques, Baromètre lorrain de la Science Ouverte, recherche bibliographique**



-
- Base de données **bibliographique** et **bibliométrique** de l'organisation à but non lucratif OurResearch
 - Outil **libre**, données **ouvertes sous licence CC0**
 - **Corpus élargi** : un effort particulier est porté à l'inclusion des publications non anglophones et dans tous les domaines disciplinaires. Voir ici une [comparaison des couvertures](#) avec d'autres outils, comme Scopus ou le Web of Science
 - Couverture des publications scientifiques de l'Université de Lorraine sur la période 2016/2023 : environ **49 800 références**
 - Des **affiliations qui restent encore à améliorer** au niveau des laboratoires et qui nécessiteront également beaucoup de travail
 - Quelles transpositions possibles des usages du Web of Science sur OpenAlex ?

Contexte



La politique nationale de Science Ouverte : une orientation progressive vers des outils libres

- Un outil de mesure utilisé **dès 2018** dans le Baromètre français de la Science Ouverte, lui-même basé sur des sources libres, pour suivre l'ouverture des publications dotées d'un DOI Crossref avec **Unpaywall** (OurResearch)
- Une mesure forte annoncée **dès 2021** dans le [Deuxième Plan National pour la Science Ouverte](#) : « *Favoriser la création d'un écosystème ouvert de la citation comme alternative aux environnements propriétaires en soutenant l'Initiative for Open Citations et le projet OpenAlex porté par OurResearch* »

La politique nationale de Science Ouverte : une orientation progressive vers des outils libres

- En 2022, le [Conseil de l'Union européenne](#) “*INVITES the Member States, the Commission and stakeholders to promote independence, openness, reproducibility and transparency of the data and criteria necessary for research assessment and for determining research impacts; CONSIDERS that data and bibliographic databases used for research assessment should, in principle, be openly accessible and that tools and technical systems should enable transparency*”
- Un renforcement de la coopération en 2024 avec un soutien financier, [mais pas que](#) : « *la France jouera un rôle essentiel en contribuant à l'amélioration des données générales d'OpenAlex et en aidant à enrichir en particulier les données liées à la recherche française, et au-delà. Cette collaboration dans un écosystème ouvert pourra, à terme, générer des économies d'échelle pour de nombreux acteurs français et internationaux* »

A l'Université de Lorraine

- Une [politique de Science Ouverte](#) fortement développée
- Un [service de bibliométrie](#) intégré dans l'offre de services de Science Ouverte, engagé pour une utilisation vertueuse : « *[L'équipe] n'a pas vocation à calculer des indicateurs de type facteur d'impact ou h-index, ces derniers étant trop peu représentatifs du réel impact de la recherche* »
- La première [déclinaison locale](#) du Baromètre français de la Science Ouverte à l'échelle d'un établissement en 2020
- Une collaboration avec le Ministère de l'Enseignement Supérieur et de la Recherche et Inria pour le portage du projet de [Baromètre pour les données de la recherche et codes logiciels](#)



Cas d'usages à l'Université de Lorraine

Les classements internationaux

Aujourd'hui	Avec OpenAlex
Utilisation des classements médiatiques pour jauger la trajectoire de l'université	Possibilité de créer ses propres classements
Interprétation des critères de qualité d'une université selon les éditeurs des classements	Peut créer ses propres critères d'intérêt
Résultats actualisés une fois par an (sortie des classements)	Résultats actualisés disponibles à tout moment
Difficulté de creuser et mieux comprendre les données utilisées	Toutes les données sont accessibles, meilleure compréhension des résultats

Classement en Mining engineering sur la période 2018-2022 avec OpenAlex

Université	Pays	Nombre de publications (nb)	Citation moyenne par publication (CNCI)	% Publications internationales (PI)	Classement
Duy Tan University	VN	30	41.0	93%	1
China University of Mining and Technology	CN	2,493	11.7	26%	2
Monash University	AU	51	21.1	73%	3
University of Wollongong	AU	77	16.0	77%	4
Hokkaido University	JP	41	16.5	73%	5
UNSW Sydney	AU	127	14.5	73%	6
University of Nottingham	GB	33	15.9	73%	7
TU Bergakademie Freiberg	DE	88	12.1	78%	8
ETH Zurich	CH	42	9.9	83%	9
University of Nevada Reno	US	31	20.1	58%	10
Université de Lorraine	FR	30	5.8	63%	49

Les rapports bibliométriques

Aujourd'hui dans le Web of Science	Avec OpenAlex
Constitution de corpus de références bibliographiques	Usage transposable à l'identique via l'API, quelques fonctionnalités manquantes dans l'interface utilisateur
Extraction de métadonnées : titres, auteurs, affiliations, thématiques, types de document, revues, citations... Le facteur d'impact n'est pas utilisé	Usage transposable à l'identique. Ces métadonnées sont également présentes dans OpenAlex hormis le facteur d'impact
Nettoyage des données d'affiliations et correction des doublons	Ce travail sera également nécessaire, surtout pour les laboratoires
Constitution d'un Baromètre Science Ouverte à l'échelle souhaitée	Usage transposable à l'identique
Création de cartes de réseaux d'auteurs avec VOS Viewer à partir des données du WoS	Usage transposable à l'identique, un export spécifique pour le VOS Viewer est disponible dans OpenAlex

Exemple : les co-publications sur le WoS UL et Université du Luxembourg

The screenshot shows the Web of Science search interface. The left sidebar has a black vertical bar with icons for MENU, FOLDER, CLOCK, PERSON, and BELL. The main search bar at the top right says "Search". Below it, the search query is "Results for Universite de Lorraine (Affiliation) AND University of Luxembour...". A large orange box highlights the text "372 results from Web of Science Core Collection for:". Below this, a search bar contains the query "Universite de Lorraine (Affiliation) and University of Luxembourg (Affiliation)". There are buttons for "Add Keywords" (highlighted in purple), "Quick add keywords:" (with options for CO-ADJOINT ORBIT and H-INFINITY CRITERION), and "Publications". The "Publications" section shows "You may also like..." and "Refine results". At the bottom, there is a search bar for "Search within results", a checkbox for "0/372", a button for "Add To Marked List", and a dropdown for "Export".

Exemple : les co-publications sur OpenAlex

The screenshot shows the OpenAlex search interface. At the top, there is a logo, a search bar with the placeholder "Search OpenAlex", and user account icons. Below the search bar, it says "Saved search". On the right side of the header, there are notification, settings, and share icons.

2 Filters

① Institution is Université de Lorraine X
② and Institution is University of Luxembourg X

Works

Normal Approximations with Malliavin Calculus
2012 · Ivan Nourdin, Giovanni Peccati
Cited by 553

Mineral weathering by bacteria: ecology, actors and mechanisms

Stats

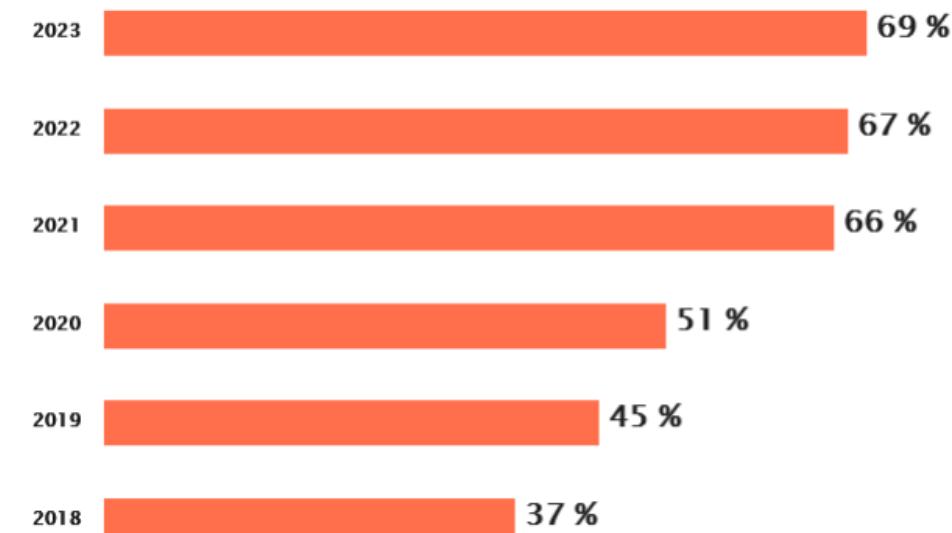
417 results (highlighted with an orange box)

open access 52,5%
219 works

Le Baromètre lorrain de la Science Ouverte

Aujourd'hui dans le Web of Science	Avec OpenAlex
Un export manuel des publications UL dans le Web of Science	Usage transposable à l'identique
Des données bibliographiques non repartagées avec la communauté car sous droits	Des données bibliographiques ouvertes à la communauté et améliorables collectivement
Une limitation technique sur les exports depuis l'interface (5 000 maximum)	Pas de limitation technique

Université de Lorraine (UL) : Taux d'accès ouvert des publications scientifiques de l'université de Lorraine, avec un DOI Crossref, parues durant l'année précédente par année d'observation

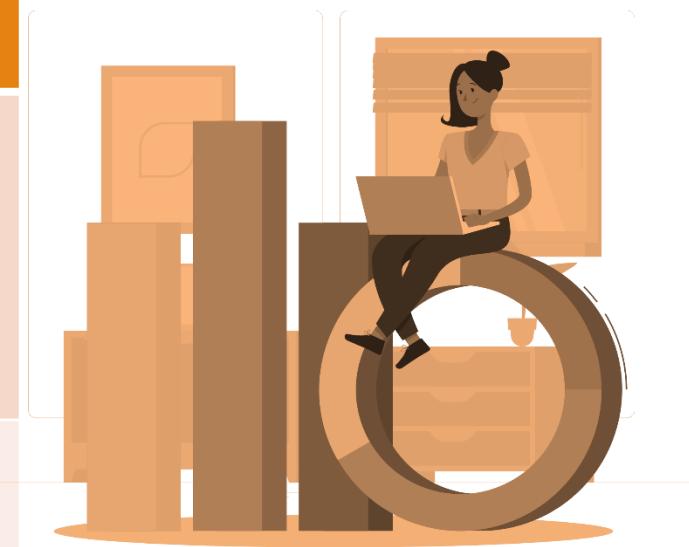


La recherche bibliographique

Aujourd'hui dans le Web of Science	Avec OpenAlex
Une connexion au Web of Science par système d'authentification	Outil accessible avec ou sans connexion
Un moteur de recherche très performant, avec de multiples champs et indicateurs disponibles	Usage transposable quasi à l'identique, quelques fonctionnalités manquantes dans l'interface utilisateur
Des affiliations contrôlées par les établissements	Modalités de correction des affiliations encore à définir
Des regroupements d'auteurs	Usage transposable à l'identique
La possibilité de recevoir des alertes sur des recherches par mail	Usage transposable à l'identique

Comparatif des coûts

	Coût humain	Coût financier
Web of Science	5 jours ETP par an de nettoyage des données d'affiliation	55 000 € / an
OpenAlex	Inconnu à ce jour, probablement similaire	0 € 10 000 € / an pour l'accès à l'API Premium et le soutien à l'infrastructure



Articulation avec le Baromètre français de la Science Ouverte

- Développement du [works-magnet](#) par le MESR
- Croisement d'OpenAlex et du corpus du Baromètre français de la Science Ouverte
- Une vérification fine des affiliations, permettant de vérifier la qualité et la complétude du corpus d'un établissement ou d'un laboratoire
- Trois usages à venir : constituer son corpus de publications / constituer son corpus de jeux de données / corriger les affiliations dans OpenAlex
- **Webinaire prévu le 16/05 à 14h pour**

le Club utilisateurs du BSO

ALPHA
MINISTÈRE DE
L'ENSEIGNEMENT
SUPERIEUR
ET DE LA RECHERCHE
Agence
Publique

Works magnet STAGING
Retrieve the scholarly works of your institution

Start year: 2023 End year: 2023 Search for datasets only

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

Conclusions



Pour l'Université de Lorraine

- Un outil devant encore évoluer, mais en constante amélioration ; mettre à profit notre expérience de la mise en qualité des données
- Une bascule possible et progressive des usages actuels vers OpenAlex
- Premières actions :
 - La formation des doctorants à OpenAlex dans le cadre du séminaire « Approche critique des indicateurs bibliométriques »
 - La sensibilisation des chercheurs par l'animation d'un nouvel atelier en ligne
 - L'utilisation d'OpenAlex préférentiellement au Web of Science pour la réalisation de rapports bibliométriques
 - L'utilisation d'OpenAlex pour constituer le corpus du Baromètre lorrain de la Science Ouverte
 - L'arrêt de la correction des métadonnées d'affiliation dans le Web of Science (travail réalisé actuellement tous les deux ans) pour rediriger ce travail sur la base d'OpenAlex via le works-magnet du MESR

Pour la communauté

1. Faire du principe de la transparence des outils dédiés à l'évaluation une réalité
2. Contribuer à l'amélioration d'un outil bénéficiant à tous
3. Apporter notre expertise de mise en qualité des données, pour construire un workflow ouvert en lieu et place d'un workflow fermé
4. Réaliser des économies d'échelle en travaillant sur un outil ouvert plutôt qu'en alimentant des boîtes noires telles que le Web of Science ou Scopus
5. Développer une stratégie collective pour alimenter ce circuit afin d'atteindre la meilleure qualité au meilleur coût, avec le moins de redondance et de régressions
6. Favoriser le développement de classements transparents basés sur des données ouvertes