Le Roi est nu !

ce qu'OpenAlex change dans notre rapport aux métadonnées

BULAC
MAI 2026

SORBONNE UNIVERSITÉ

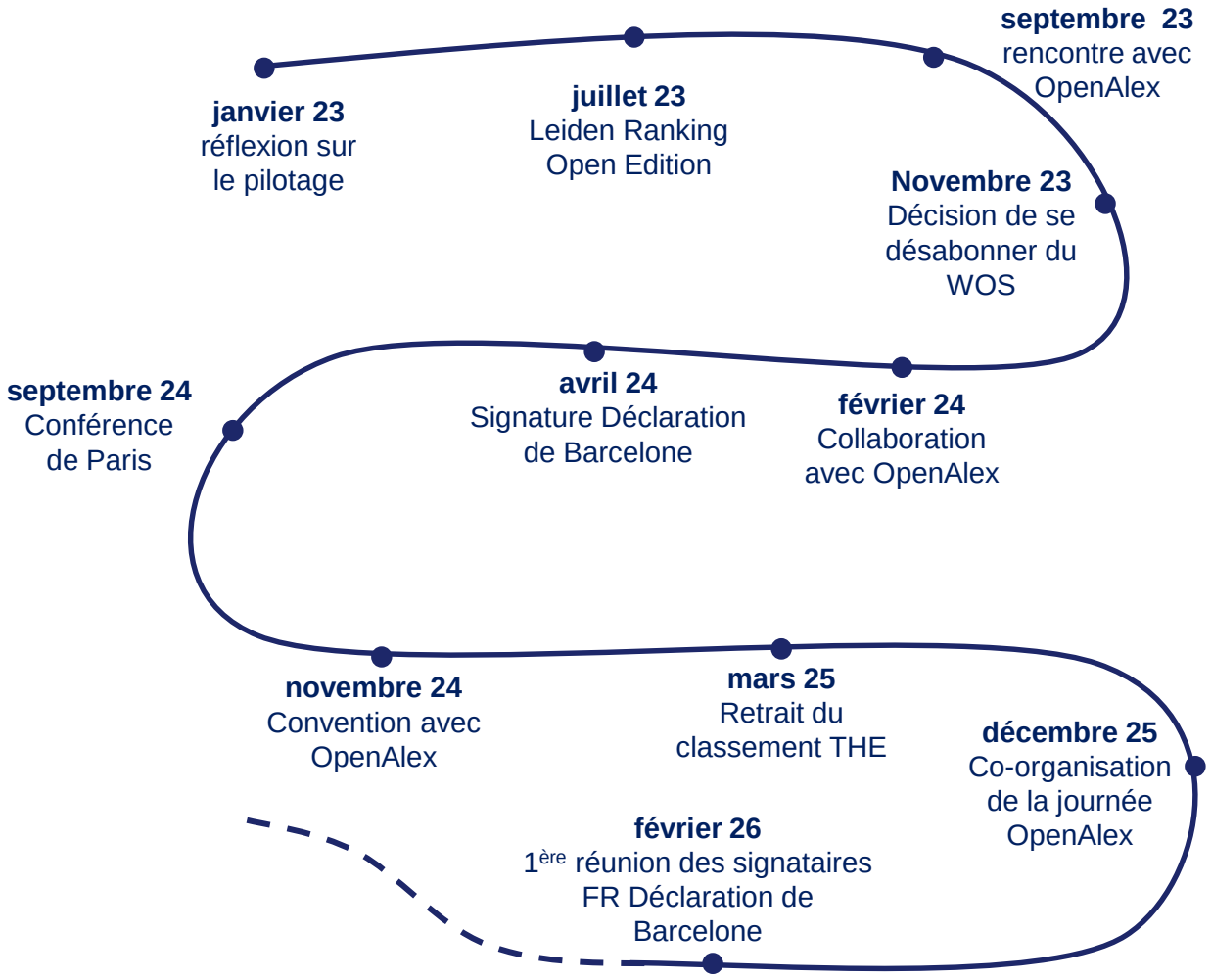
Le long cheminement d'une décision qui fait date

SU se désabonne du WoS

Depuis 2018, SU se désabonne régulièrement de ressources documentaires coûteuses

WOS en cours de renégociation au niveau national

Les biais des bases commerciales ne sont plus questionnés (sous-représentation systémique des publis en SHS et non-anglophones)



Un intérêt national et international



Sorbonne's embrace of free research platform shakes up academic publishing

Paris university's 'radical decision' to join OpenAlex challenges for-profit providers

FINANCIAL TIMES

The Sorbonne says it made the 'radical decision' in order to 'reappropriate the r
control and ownership of what we produce' © Loic Venance/AFP/Getty Images

EL PAÍS

MANUEL ANSEDE

Madrid - 16 ABR 2024 - 13:29 CEST

La universidad francesa de la Sorbona provocó un terremoto en la ciencia mundial el 8 de diciembre, cuando anunció que cancelaba su suscripción a la Web of Science, una omnipresente base de datos de la multinacional londinense Clarivate que se emplea en todo el mundo para evaluar a los científicos y concederles, o no, ascensos y financiación. Cientos de millones de euros de dinero público se reparten en función de estos datos de plataformas privadas, sobre todo de la Web of Science y de Scopus, una herramienta similar de la editorial holandesa Elsevier. La Sorbona y otro medio centenar de organizaciones, incluido el Consejo Superior de Investigaciones Científicas (CSIC) español, han presentado este martes la Declaración de Barcelona, una iniciativa para transformar radicalmente este sistema y hacerlo abierto.

Fev 2025 : Université de Kyoto

Juil 2025 : conférence LIBER Lausanne (conférence + poster)

Oct 2025 : présentation EARMA (association européenne des directions Recherche)

Nov 2025 : Conférence LIS Bibliometrics - Leicester

Nov 2025 : présentation aux bibliothèques universitaires japonaises

Janv 2026 : présentation aux bibliothèques de la Commission européenne

Avr 2026 : Présentation LIBER + EARMA

La qualité des données dans OpenAlex

« OpenAlex va sûrement progresser mais, à ce jour, il est trop risqué d'en faire sa source unique pour une analyse bibliométrique sans prendre le soin de tester la qualité des données en amont » (Bordignon, 2024)

« OpenAlex is a valuable infrastructure, yet its metadata for publication type and language shows notable inconsistencies. Researchers should apply systematic cleaning and validation before using OpenAlex or similar databases. » (Doğan et Sezen 2026)

« A finer-grained palette of document types in OpenAlex, i.e. classifying review articles, as well as an improved type accuracy would be highly beneficial. » (Haupka et al., 2024)

« Results show that OpenAlex exhibits a far more balanced linguistic coverage than WoS. However, language metadata is not always accurate, which leads OpenAlex to overestimate the place of English while underestimating that of other languages. » (Céspedes et al., 2024)

« We currently recommend caution when utilising OpenAlex for scientometric studies due to the volatility and data quality issues discussed earlier in this section. » (Culbert et al. 2024)

« The results reveal numerous records lacking abstracts, affiliations and references; deficiencies in identifying document types and languages; and issues with authority control and versioning » (Borrego et Urbano 2025)

Qui est OpenAlex

Projet porté par une organisation à but non lucratif et une équipe réduite

Financé par Arcadia ($7,5 M/5 ans), Navigation.fund (~ $700K), Wellcome Trust ($3,6M/3 ans) et les contributions d'institutions partenaires

Respect des principes pour une infrastructure académique ouverte (POSI) : transparence, gouvernance, soutenabilité

Gouvernance : Board of Directors et Advisory Group (membres élus par la communauté)

Des partenariats avec le CWTS ou le MESRE

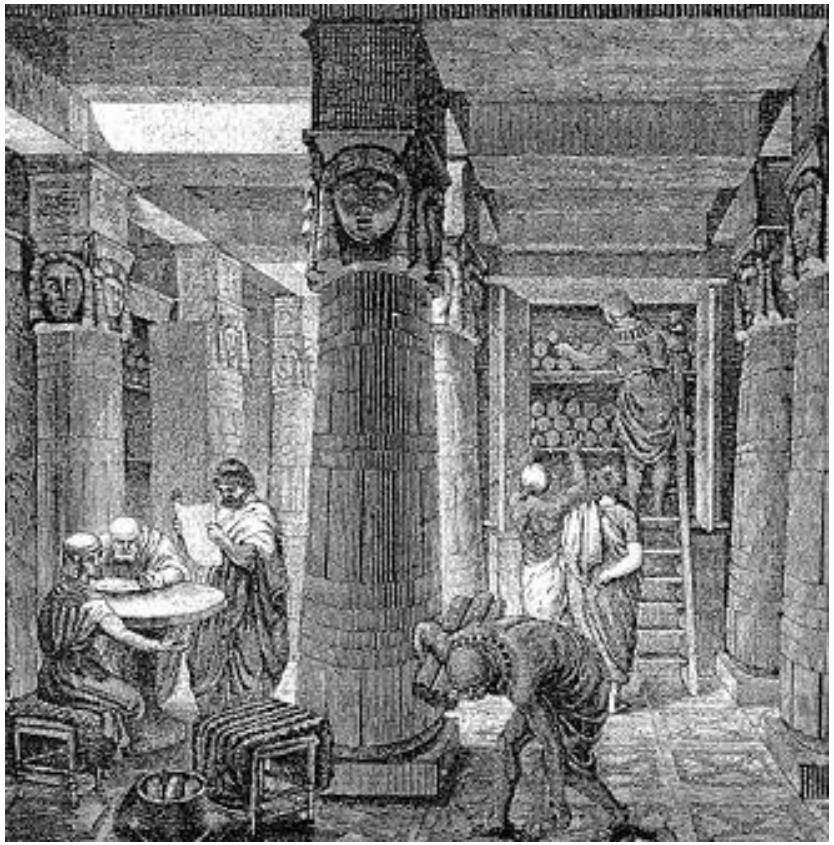
A l’origine

OpenAlex comme projet secondaire

Multilinguisme et couverture géographique

Diversité des sources/types de publication

Accent sur l’ouverture : données et code en CC0



O. Von Corven - Tolzmann, Don Heinrich; Alfred Hessel and Reuben Peiss. *The Memory of Mankind*. New Castle, DE: Oak Knoll Press, 2001

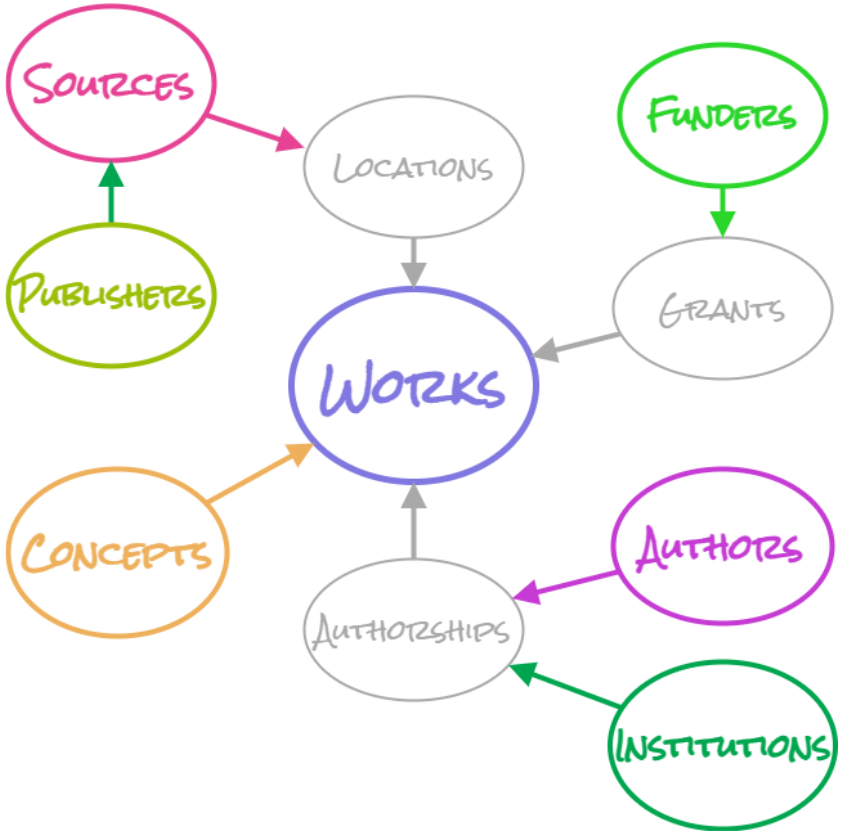
Qu’est-ce qu’OpenAlex

Une base de données des publications scientifiques : articles, livres, actes de colloques, thèses, ...

Et de leurs connexions : auteurs, éditeurs, revues, sujets, institutions, financeurs, citations…

Intérêt bibliographique : ce qui est publié sur un sujet

Et bibliométrique : qui travaille sur quel sujet, quelles collaborations, quel impact



Sources

Crossref : métadonnées des publications

Pubmed : bases bibliographique des publications en santé

Archives ouvertes (HAL, ArXiv.org, DOAJ, Zenodo…) et tous les répertoires de publications identifiés et ouverts

Importance des identifiants pérennes

- DOI pour les publications
- ROR (Research Organization Registry) pour les affiliations institutionnelles
- ORCID pour les identifiants individuels des chercheurs



Cell

CellPress

Volume 186, Issue 2, 19 January 2023, Pages 243-278

Review

Hallmarks of aging: An expanding universe

Carlos López-Otín [1 2 3], Maria A. Blasco [4], Linda Partridge [5 6], Manuel Serrano [7 8 9], Guido Kroemer [10 11 12]

1 Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología (IUOPA), Universidad de Oviedo, Oviedo, Spain
2 Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Spain
3 Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain
4 Telomeres and Telomerase Group, Molecular Oncology Program, Spanish National Cancer Centre (CNIO), Madrid, Spain
5 Department of Genetics, Evolution and Environment, Institute of Healthy Ageing, University College London, London, UK
6 Max Planck Institute for Biology of Ageing, Cologne, Germany
7 Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
8 Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
9 Altos Labs, Cambridge, UK
10 Centre de Recherche des Cordeliers, Equipe labellisée par la Ligue contre le cancer, Université de Paris, Sorbonne Université, INSERM U1138, Institut Universitaire de France, Paris, France
11 Metabolomics and Cell Biology Platforms, Gustave Roussy, Villejuif, France
12 Institut du Cancer Paris CARPEM, Department of Biology, Hôpital Européen Georges Pompidou, AP-HP, Paris, France

Available online 3 January 2023, Version of Record 19 January 2023.

SORBONNE UNIVERSITÉ

Métadonnées Crossref


```
author: [
  {
      given: "Carlos",
      family: "López-Otín",
      sequence: "first",
      affiliation: [ ]
  },
  {
      given: "Maria A.",
      family: "Blasco",
      sequence: "additional",
      affiliation: [ ]
  },
  {
      given: "Linda",
      family: "Partridge",
      sequence: "additional",
      affiliation: [ ]
  },
  {
      given: "Manuel",
      family: "Serrano",
      sequence: "additional",
      affiliation: [ ]
  },
  {
      ORCID: https://orcid.org/0000-0002-9334-4405,
      "authenticated-orcid": false,
      given: "Guido",
      family: "Kroemer",
      sequence: "additional",
      affiliation: [ ]
  }
],
```

Cell

CellPress

Volume 186, Issue 2, 19 January 2023, Pages 243-278

Review

Hallmarks of aging: An expanding universe

Carlos López-Otín [1 2 3], Maria A. Blasco [4], Linda Partridge [5 6], Manuel Serrano [7 8 9], Guido Kroemer [10 11 12]

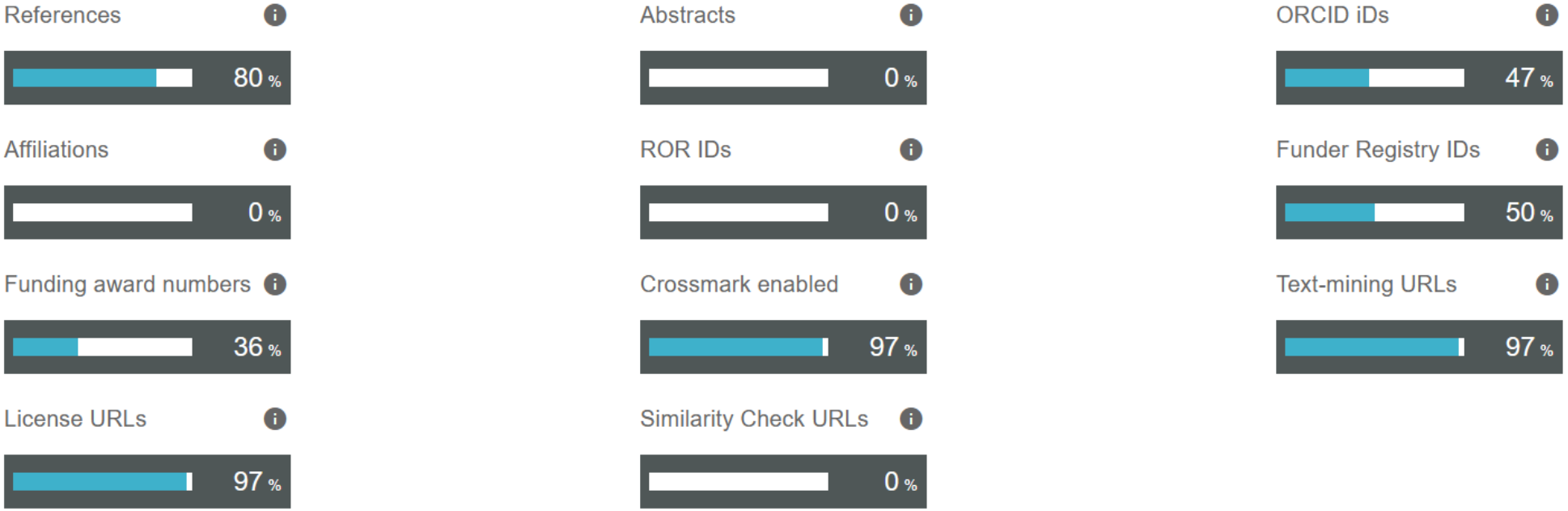
[1] Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología (IUOPA), Universidad de Oviedo, Oviedo, Spain
[2] Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Spain
[3] Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain
[4] Telomeres and Telomerase Group, Molecular Oncology Program, Spanish National Cancer Centre (CNIO), Madrid, Spain
[5] Department of Genetics, Evolution and Environment, Institute of Healthy Ageing, University College London, London, UK
[6] Max Planck Institute for Biology of Ageing, Cologne, Germany
[7] Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
[8] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
[9] Altos Labs, Cambridge, UK
[10] Centre de Recherche des Cordeliers, Equipe labellisée par la Ligue contre le cancer, Université de Paris, Sorbonne Université, INSERM U1138, Institut Universitaire de France, Paris, France
[11] Metabolomics and Cell Biology Platforms, Gustave Roussy, Villejuif, France
[12] Institut du Cancer Paris CARPEM, Department of Biology, Hôpital Européen Georges Pompidou, AP-HP, Paris, France

Available online 3 January 2023, Version of Record 19 January 2023.

Métadonnées Openalex


```
{
  author_position: "last",
  author: {
      id: https://openalex.org/A5075296559,
      display_name: "Guido Kroemer",
      orcid: https://orcid.org/0000-0002-9334-4405
  },
  institutions: [
    {
        id: https://openalex.org/I39804081,
        display_name: "Sorbonne Université",
        ror: https://ror.org/02en5vm52,
        country_code: "FR",
        type: "funder",
        lineage: [
            https://openalex.org/I39804081
        ]
    },
    {
        id: https://openalex.org/I2801843802,
        display_name: "Institut Gustave Roussy",
        ror: https://ror.org/0321g0743,
        country_code: "FR",
        type: "funder",
        lineage: [
            https://openalex.org/I2801843802
        ]
    },
    {
        id: https://openalex.org/I4210092322,
        display_name: "Centre de Recherche des Cordeliers",
        ror: https://ror.org/00dmms154,
        country_code: "FR",
        type: "facility",
        lineage: [
            https://openalex.org/I154526488,
            https://openalex.org/I204730241,
            https://openalex.org/I39804081,
            https://openalex.org/I4210092322
        ]
    },
    {
        id: https://openalex.org/I204730241,
        display_name: "Université Paris Cité",
        ror: https://ror.org/05f82e368,
        country_code: "FR",
        type: "funder",
        lineage: [
            https://openalex.org/I204730241
        ]
```

« Ensemble, nous pouvons créer un monde de la recherche plus inclusif, plus collaboratif et plus transparent. » Elsevier



Source : Crossref Participation Reports

« De nombreuses notices ne comportent ni résumé, ni affiliation, ni références »

La disponibilité des métadonnées progressent de manière inégale

Initiatives internationales et exigences des bailleurs transforment ce paysage

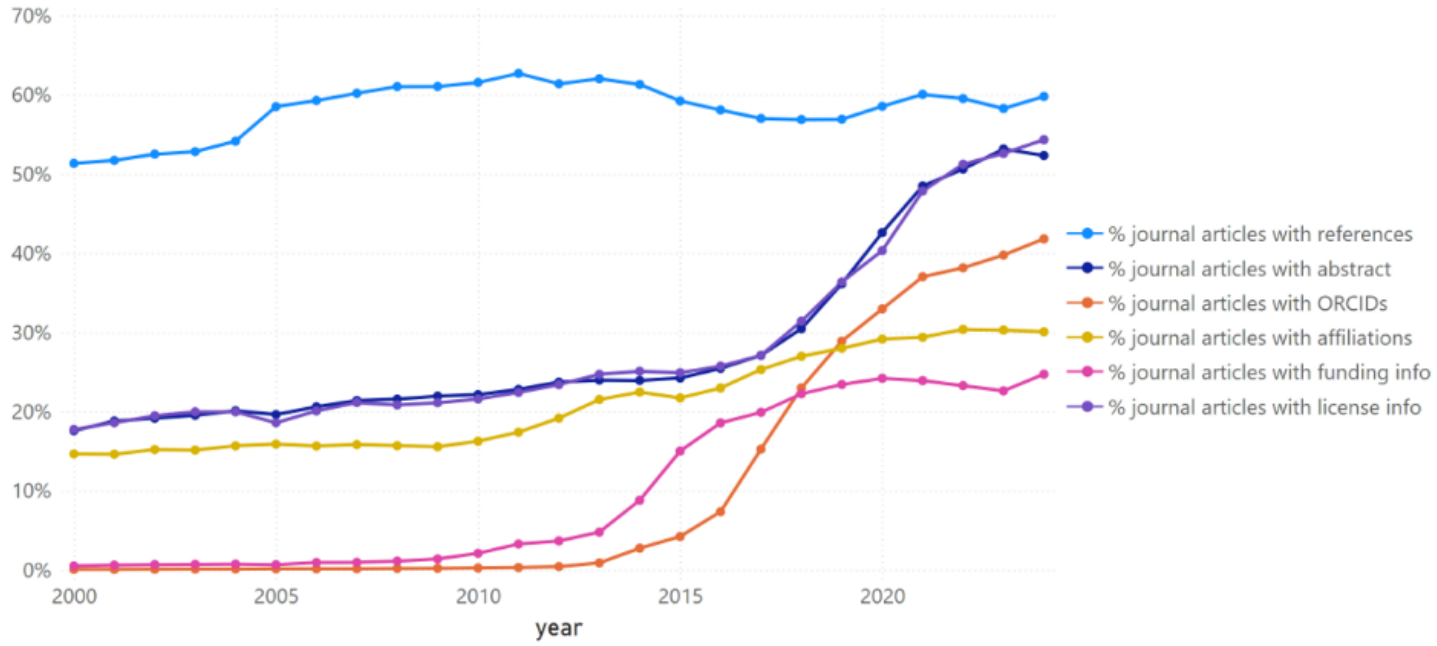


Figure 2. Availability of different metadata elements for journal articles in Crossref.

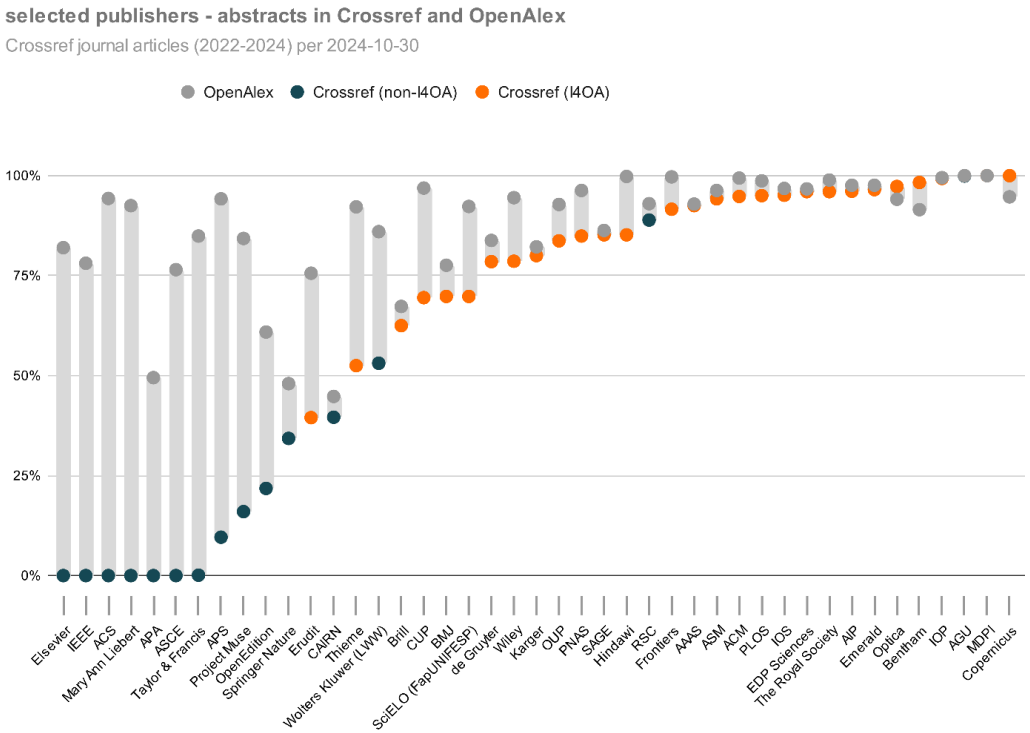
van Eck, N. and Waltman, L. (2025) 'Crossref as a source of open bibliographic metadata'. OSF. Available at: https://doi.org/10.31222/osf.io/smxe5_v2.

« Utiliser OpenAlex avec prudence … »

En fouillant les pages web des articles, OpenAlex récupère des métadonnées qui ne sont pas fournies par les éditeurs

Résumé, affiliations, statut OA

Blocus sur les résumés



Kramer, Bianca. 2024. « More Open Abstracts? » Sesame Open Science, novembre 24. https://bmkramer.github.io/SesameOpenScience_site/thought/202411_open_abstracts/ .

« OpenAlex va sûrement progresser … »

Les données OpenAlex évoluent très vite : par la curation humaine ou l'amélioration du traitement algorithmique

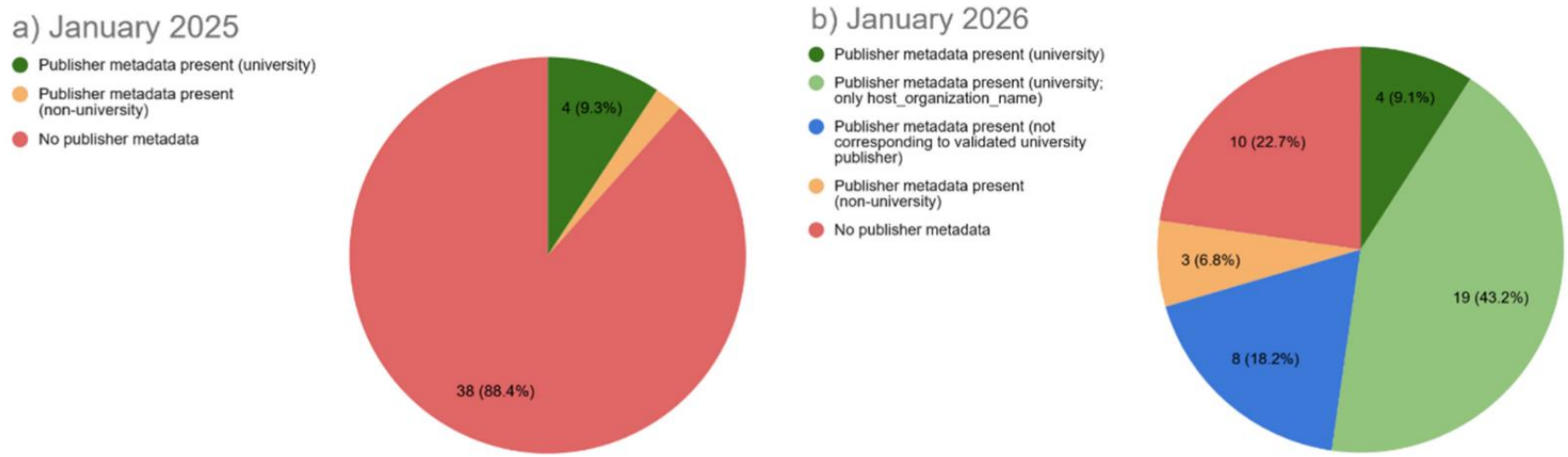


FIGURE 2 | Publisher-affiliation metadata for the selected UJs in OpenAlex: (a) January 2025; (b) January 2026.

Nazarovets, Maryna. 2026. « Visible as Journals, Invisible as Publishers: Limitations of OpenAlex for Analysing University Publishing ». *Learned Publishing* 39 (3): e2061. https://doi.org/10.1002/leap.2061.

Investir dans la mise en qualité des métadonnées

- **Où ?**
  - HAL, Openalex
  - Référentiels de personnes et de structures (ORCID, ROR, RNEST)

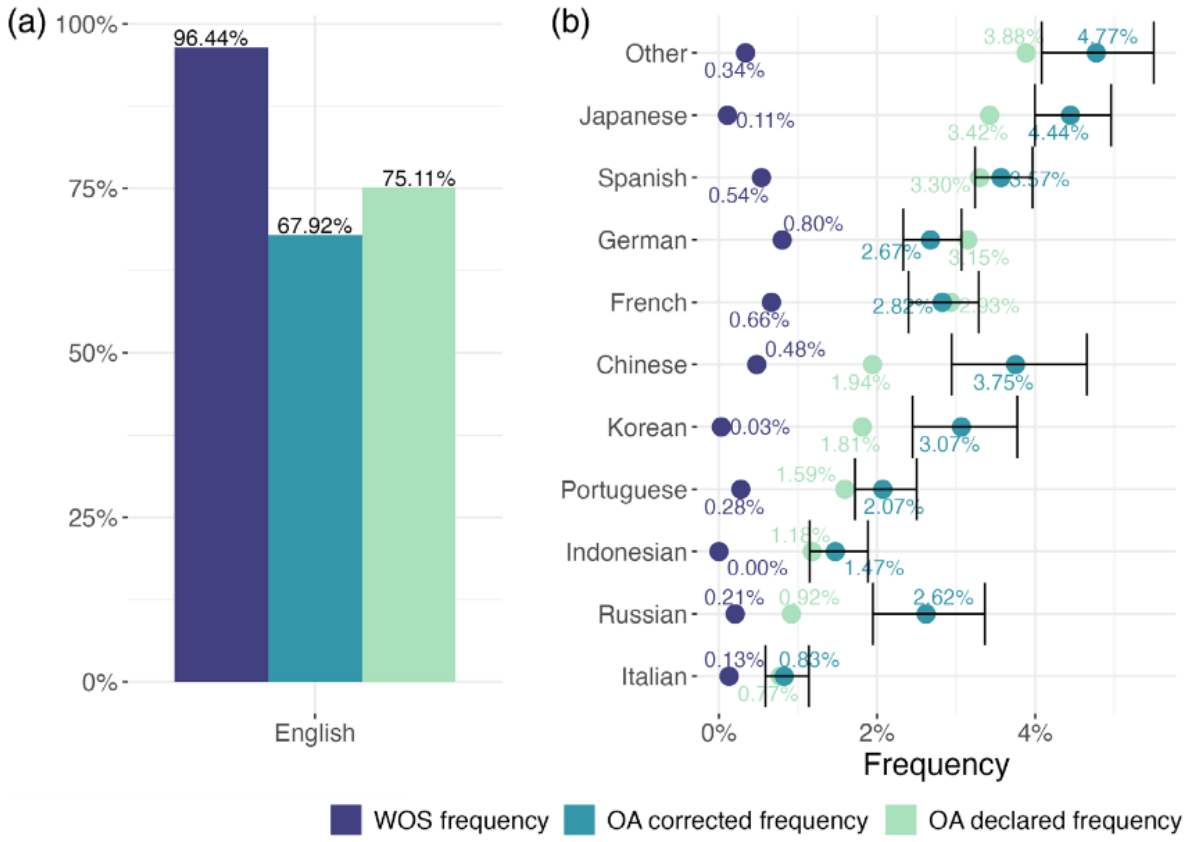
- **Comment ?**
  - Charte de signature des publications scientifiques
  - Identifiants pérennes de la recherche (idref, DOI, ORCID, ROR)
  - Chantier d’alignement
  - Curation des métadonnées
  - Formation & sensibilisation

OpenAlex et le multilinguisme

Deux enjeux :

- Diversifier les sources pour mieux représenter les publications non-anglophones dans le corpus
- Identifier correctement la langue de la publication dans les métadonnées

Améliorer la détection des langues dans OpenAlex grâce aux résultats de ces recherches

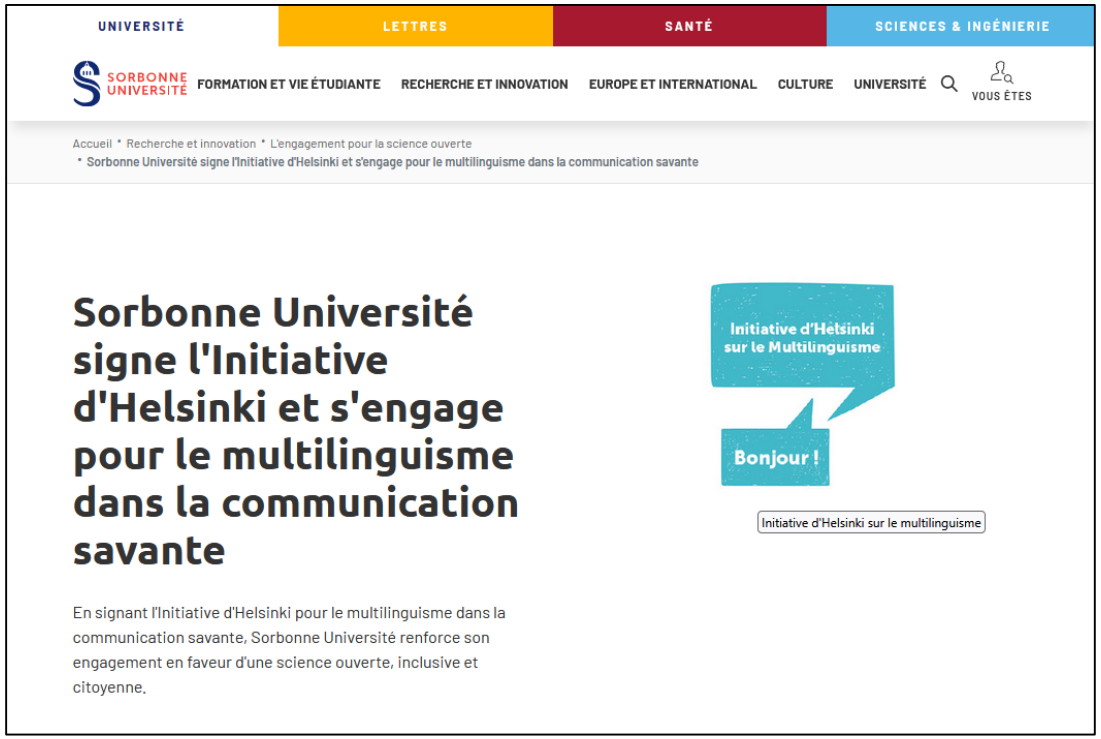


Céspedes, Lucía, Diego Kozlowski, Carolina Pradier, et al. 2025. « Evaluating the Linguistic Coverage of OpenAlex: An Assessment of Metadata Accuracy and Completeness ». *Journal of the Association for Information Science and Technology* 76 (6): 884-95. https://doi.org/10.1002/asi.24979.

Le multilinguisme à SU

Développement d'un « axe francophonie, articulé aux instances nationales et internationales, promouvant, en phase avec l'**Initiative d'Helsinki sur le multilinguisme**, le plurilinguisme et l'usage du français dans les échanges scientifiques. » (Projet d'établissement 2025-2029)

- Soutenir la diffusion des résultats de la recherche dans l'intérêt de la société
- Protéger les infrastructures nationales permettant la publication de recherches pertinentes au niveau local
- 3. Promouvoir la diversité linguistique dans les systèmes d'évaluation et de financement de la recherche



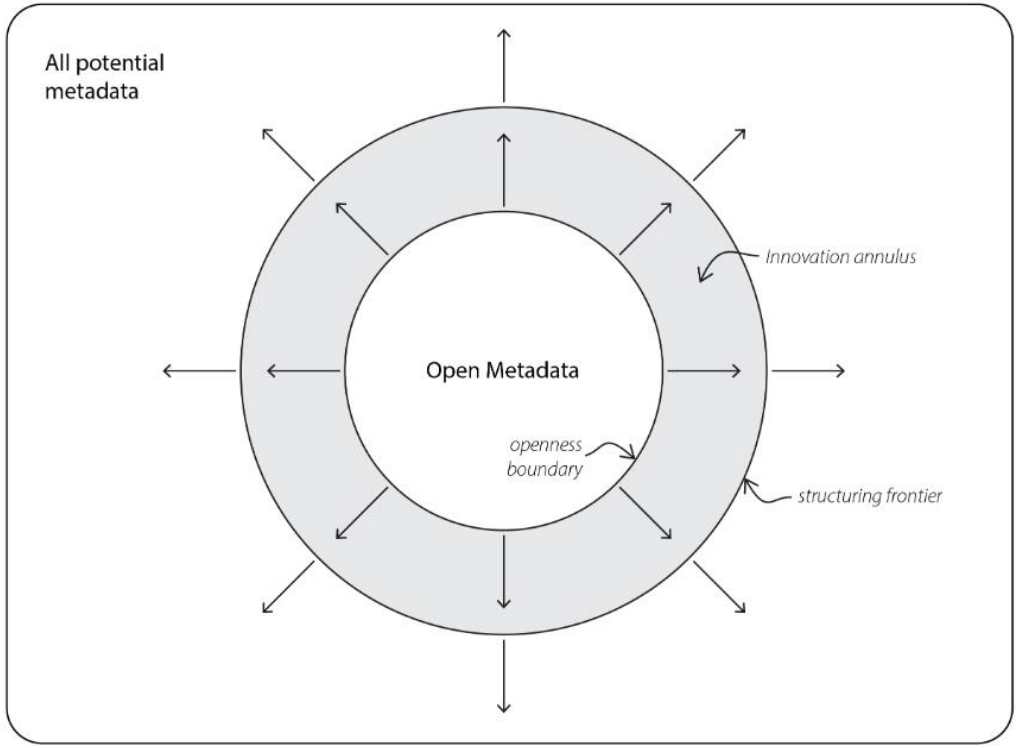
Pourquoi les métadonnées ne sont pas toutes ouvertes ?

Coût de production/structuration VS besoins des communautés

Données produites dans un écosystème hautement hétérogène sans standardisation globale

3 frictions entre offre et demande :

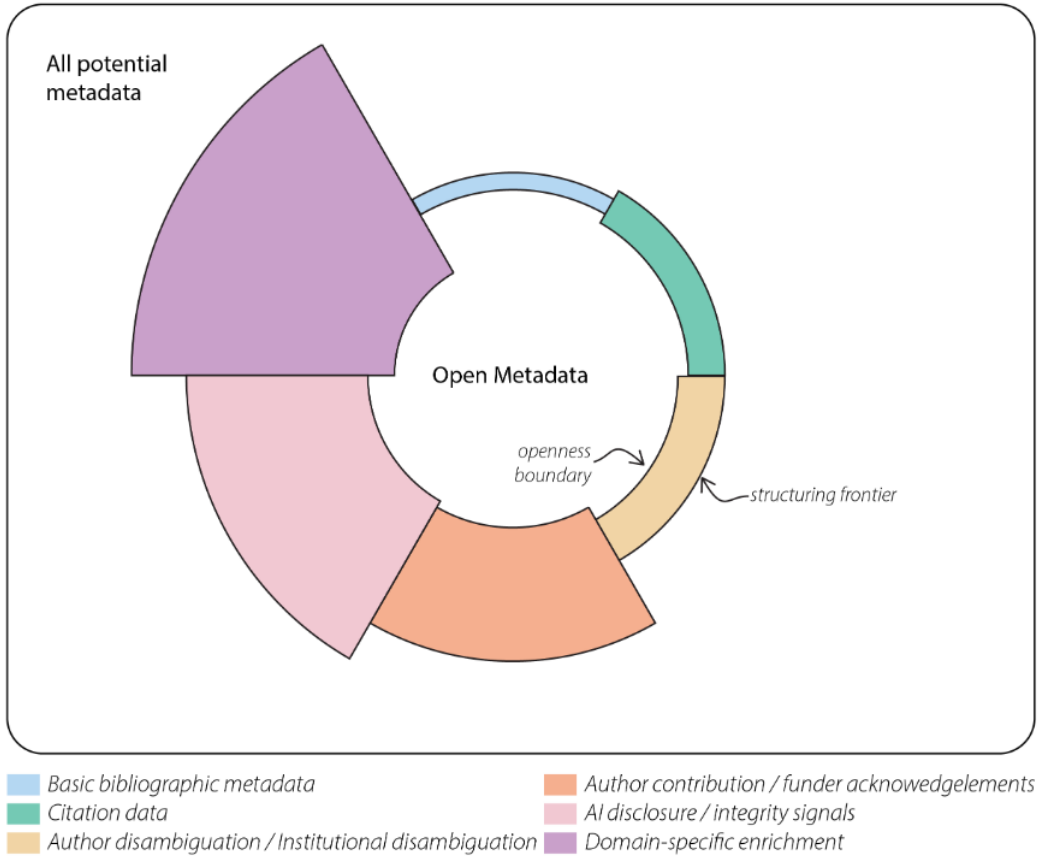
- Hétérogénéité des sources
- Perte en qualité
- Besoins toujours renouvelés



Hook, Daniel W. 2026. « Market Dynamics, Governance and Open Research Metadata in the AI Era ». arXiv:2604.19507. Prépublication, arXiv, avril 21. https://doi.org/10.48550/arXiv.2604.19507.

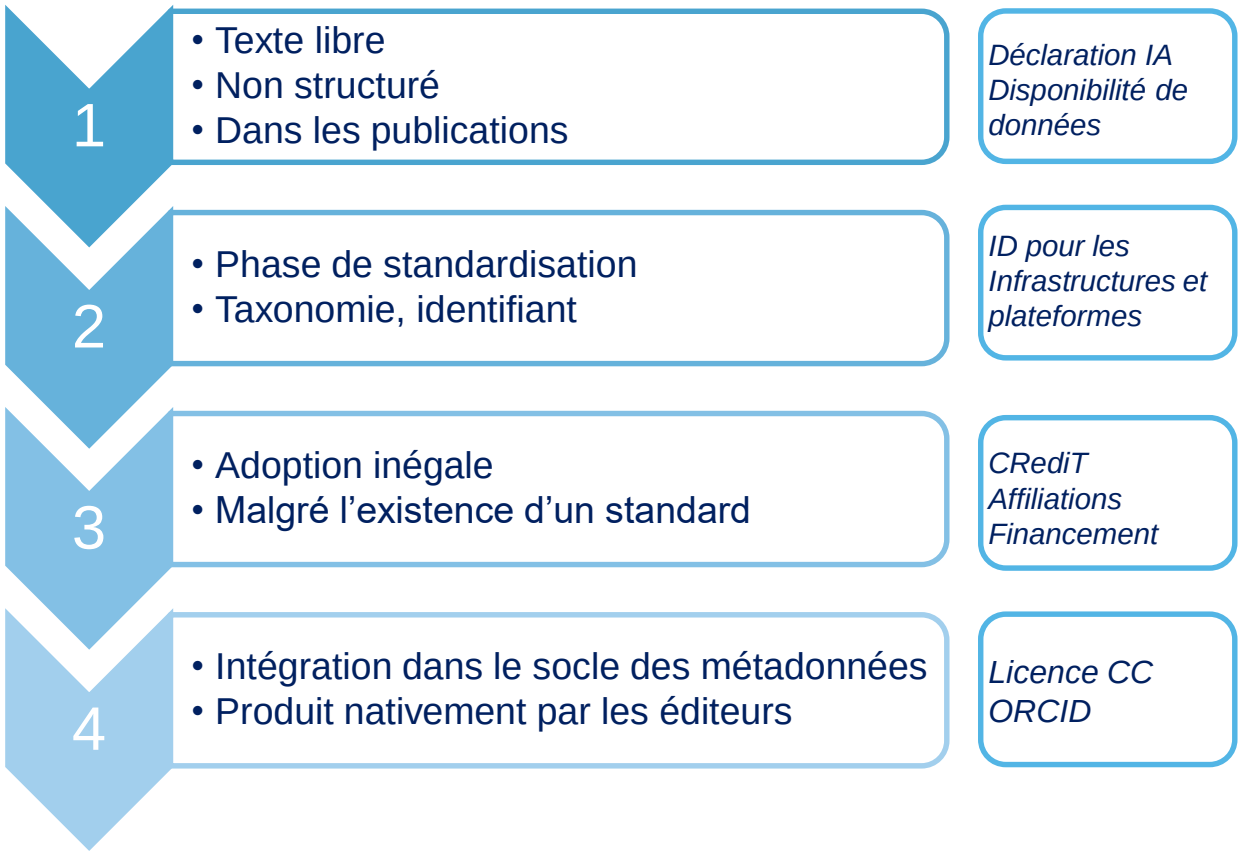
Pourquoi les métadonnées ne sont pas toutes ouvertes ?

Forme de l'anneau	Caractéristiques	Exemple
Fin et proche du centre	Production peu coûteuse, pas d'innovation requise	Peer-review
Fin et éloigné du centre	Système mature, bons standards, ouverture comme norme	Métadonnées basiques DOI
Épais et proche du centre	Structuration établie mais peu d'ouverture ou données coûteuse/difficiles à améliorer	AI Disclosure
Épais et éloigné du centre	Ouverture significative avec des opportunités d'innovation	Liens publis-brevet



Hook, Daniel W. 2026. « Market Dynamics, Governance and Open Research Metadata in the AI Era ». arXiv:2604.19507. Prépublication, arXiv, avril 21. https://doi.org/10.48550/arXiv.2604.19507.

Cycle d’intégration des métadonnées



Barcelona Declaration on Open Research Information (ORI)

1



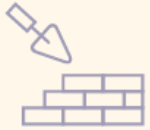
We will make openness the default for the research information we use and produce

2



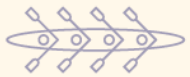
We will work with services and systems that support and enable open research information

3



We will support the sustainability of infrastructures for open research information

4



We will support collective action to accelerate the transition to openness of research information

Signataires et supporters

127 signataires de 28 pays

52 organisations supportrices

Aix-Marseille Université
Sorbonne Université
Udice
Institut Polytechnique UniLaSalle
Nantes Université
Université de Lorraine
Université de Reims
Université de Strasbourg
Grenoble Alpes Université
Université du Havre
Paris-Saclay Université
Université de Toulouse
Université de Poitiers
ANR
CoSO
Région Normandie
COUPERIN

Réseaux linguistiques/nationaux

Germany, Austria, and Switzerland / DACH-Austausch Barcelona Declaration

Netherlands / Dutch Signatories of the Barcelona Declaration

Italy / Italian Node of the Barcelona Declaration

France / Réseau des signataires français de la Déclaration de Barcelone

Déclaration de Barcelone – GT

1 – Métadonnées des articles de revue et des livres (*Inclure les métadonnées ouvertes dans les négociations avec les éditeurs*)

2 – Métadonnées dans les entrepôts ouverts

3 – Métadonnées de financement

4 – Remplacer les bases bibliographiques commerciales

5 – Soutenabilité des infrastructures

6 – Évaluer les sources ouvertes de métadonnées

7 – Preuves des bénéfices des informations de recherche ouvertes

Conclusions

La question n’est pas de substituer une base par une autre

Défendre la cohérence de nos engagements pour la science ouverte

Maîtriser nos données

Penser global, agir local

Le ROI est nu, vive les ORI !



Les Habits neufs de l’empereur.
Illustration de Vilhelm Pedersen.

MERCI DE VOTRE ATTENTION

BIBLIOGRAPHIE SUR OPENALEX

CONTACT : NICOLAS.ALARCON@SORBONNE-UNIVERSITE.FR

SORBONNE UNIVERSITÉ