



Enclosure et commonswashing : Clarivate annonce la base de données « Research Commons »

by Alan Colin Arce | 22 June 2026 | French, Observations



Read in English

Cette observation a été rédigée par Alan Colin-Arce, avec ses remerciements à Brittany Amell, Simon van Bellen et Jack Young pour

leurs commentaires et leur relecture.

Traduction française révisée par Olga Ziminova (MA), Electronic Textual Cultures Lab (ETCL).

At a Glance

Titre	Enclosure et commonswashing : Clarivate annonce la base de données « Research Commons »
Participants clés	Clarivate, OpenAlex, Crossref, Creative Commons, Wikimedia
Période	2025
Mots-clés	Digital Commons / Commun numérique, AI bots / Robots d'indexation IA, bibliodiversity / bibliodiversité, commonswashing

Résumé

Cette observation décrit le lancement par Clarivate de la base de données « Research Commons [Communs de recherche] », qui exploite les données d'OpenAlex et de Crossref et les intègre à une base de données distincte du Web of Science. Cette annonce s'inscrit dans le contexte de tendances récentes observées dans le domaine de la science ouverte, telles que l'appropriation des connaissances communes par les entreprises et le « commonswashing », c'est-à-dire l'appropriation du langage des biens communs sans contribution à leur développement.

Introduction

Le 4 septembre 2025, Clarivate a **annoncé** l'intégration d'une nouvelle base de données, baptisée « Research Commons », à la plateforme Web of Science. «

Research Commons » est une « collection distincte et exhaustive de publications provenant de sources de métadonnées ouvertes » (Goldfinger, 2025), à savoir OpenAlex et Crossref. L'ajout de cette collection s'inscrit dans la stratégie d'expansion du Web of Science visant à accroître la portée mondiale et disciplinaire de la plateforme, tout en la maintenant distincte de la collection principale, une base de données plus restrictive qui n'accepte que les revues répondant aux critères de **sélection éditoriale de Clarivate**. À ce jour, Clarivate a intégré 32 millions de dossiers des dix dernières années dans la base de données « Research Commons ».

La base de données « Research Commons » répond aux critiques reprochant au Web of Science et à son concurrent Scopus d'être biaisés à l'encontre des recherches menées dans des pays non occidentaux, des recherches en langue étrangère et des recherches en arts, lettres et sciences sociales (Tennant, 2020, p. 1). L'intégration de données provenant d'OpenAlex et de Crossref élargit ainsi la couverture géographique, linguistique et disciplinaire du Web of Science, car des études récentes ont démontré qu'OpenAlex offre une indexation plus inclusive et exhaustive dans tous ces domaines (Céspedes et al., 2025 ; Maddi et al., 2025).

Les chercheurs et les intervenants consultés par Clarivate ont salué cette initiative, qui favorise l'inclusion des chercheurs et des revues des pays du Sud, fait la promotion de la diversité bibliographique et enrichit le Web of Science (Goldfinger, 2025). Toutefois, l'accès à la base de données « Research Commons » nécessite un abonnement à la plateforme Web of Science, une plateforme onéreuse dont les frais d'abonnement peuvent atteindre 50 000 euros (environ 81 600 dollars canadiens) (Clavey, 2024). Certaines universités à travers le monde, comme l'**Université de la Sorbonne**, l'**Université d'Utrecht** et l'**Université Dalhousie**, ont résilié leur abonnement au Web of Science en raison de son coût élevé ou pour privilégier des bases de données ouvertes alternatives, comme OpenAlex.

Appropriation des biens communs par les entreprises

Malgré son nom, la base de données « Research Commons » ne correspond pas aux définitions courantes des biens communs numériques, qui sont compris comme « un sous-ensemble des biens communs, où les ressources sont des données, des

informations, des éléments culturels et des connaissances créés et/ou maintenus en ligne. Elles sont partagées de manière à éviter leur privatisation et à permettre à tous d'y accéder et de les enrichir » (Rosnay & Stalder, 2020, p. 2). La base de données « Research Commons », par contre, empêche un accès équitable à ses données en raison du coût élevé de la consultation du Web of Science et de l'accès à grande échelle aux métadonnées associées aux articles indexés.

D'autres initiatives scientifiques ouvertes fonctionnant comme des biens communs numériques, comme [Knowledge Commons](#), [Humanities and Social Sciences Commons](#), [Wikimedia Commons](#) et [Creative Commons](#), répondent à ces critères de gouvernance horizontale et d'accès équitable. L'annonce de la base de données « Research Commons » s'inscrit donc dans une tendance plus large : celle des entreprises qui s'approprient et tirent profit du discours et des travaux des initiatives de science ouverte, et notamment des initiatives de biens communs, sans y contribuer ni les soutenir.

La base de données « Research Commons » contribue à l'accaparement des biens communs en s'appropriant et en privatisant des données initialement partagées au sein d'infrastructures ouvertes telles qu'OpenAlex et Crossref. Cette appropriation accentue la transformation de Clarivate, d'éditeur ou fournisseur de plateforme de recherche, en une entreprise de courtage de données, comme l'explique Sarah Lamdan dans son ouvrage « Data Cartels [Cartels de données] ». Selon Lamdan (2022), les sociétés de courtage de données monétisent le processus de recherche en créant des produits de données qui exploitent des renseignements provenant de données académiques et d'autres types de données personnelles. Dans cette optique, le lancement de la base de données « Research Commons » permet à Clarivate d'accroître son volume de données disponibles afin de les analyser puis de revendre les résultats à ses clients.

D'autres exemples d'appropriation et d'accaparement des biens communs par les entreprises se manifestent dans l'extraction croissante d'œuvres sous licence Creative Commons (ou non) pour l'entraînement des modèles d'intelligence artificielle (un [article récent de l'OSPO](#) abordait notamment l'impact de ce phénomène sur les archives ouvertes). Comme le souligne un [rapport](#) de Creative Commons, « les grands modèles d'IA profitent du travail, des ressources financières et de l'attention investis dans la création et l'entretien des biens communs, tout en menaçant de les épuiser » (Hardinges et al., 2025, p. 4). En réponse, Creative

Commons travaille sur de nouvelles licences, appelées « CC Signals [Signaux CC] », afin d'« offrir aux gestionnaires de vastes collections de contenus un nouveau moyen d'indiquer leurs préférences quant à la manière dont les machines (et les humains qui les contrôlent) doivent contribuer aux biens communs lorsqu'elles réutilisent et utilisent ces contenus » (Hardinges et al., 2025, p. 4). Ces licences n'empêchent pas le développement de l'IA, mais visent à encourager les actions qui préservent et soutiennent les biens communs numériques lors de l'utilisation de ressources sous licence Creative Commons.

Wikimedia a rencontré un problème semblable : des robots d'indexation d'IA ont mis à rude épreuve ses serveurs en extrayant du contenu de projets tels que Wikipédia, Wikimedia Commons et Wikibooks (Edwards, 2025). Pour réduire cette surcharge, Wikipédia a publié un jeu de données basé sur Wikipédia et formaté spécifiquement pour les développeurs d'IA (Maxwell, 2025). Cette initiative a été **saluée** pour sa promotion d'une utilisation éthique des données et sa démocratisation de l'accès aux données d'entraînement pour l'IA. Cependant, des utilisateurs de Wikipédia ont également exprimé des inquiétudes quant à l'exploitation du web ouvert pour créer des modèles d'IA propriétaires (Woodcock, 2023).

Le commonwashing

Ces cas illustrent les préoccupations récurrentes liées à l'accaparement et à la mainmise des entreprises sur les biens communs numériques. Le lancement par Clarivate de la base de données « Research Commons » constitue un autre exemple d'appropriation, par une entreprise, des données d'infrastructures ouvertes à des fins lucratives. Comme le souligne Simon van Bellen (conseiller principal en recherche chez Érudit) dans ses commentaires :

Ce qui est particulièrement problématique dans ce cas, c'est qu'une grande entreprise s'appuie sur les biens communs sans aucune réciprocité. Il n'y a aucune réciprocité avec Crossref ou OpenAlex, et peut-être même aucune attribution correcte, puisque Crossref et OpenAlex rendent leurs données accessibles au public sous licence CC0.

Toutefois, lier cette base de données au concept de biens communs relève également d'un processus de « commonswashing », terme forgé par Mélanie Dulong de Rosnay pour décrire comment des entreprises à but lucratif s'approprient le message des biens communs à des fins commerciales, sans en adhérer aux principes ni en préserver la valeur au sein de la communauté académique (Dulong de Rosnay, 2020).

Dulong de Rosnay (2020) recense certaines caractéristiques communes aux produits ou services relevant des biens communs : la propriété collective des moyens de production, une architecture et une conception techniques gérées par la communauté, une gouvernance sociale des modes de travail et une propriété collective des ressources. Dans le cas de « Research Commons », la base de données appartient à Clarivate, une entreprise à but lucratif ; la conception technique et la gouvernance sont centralisées et gérées par Clarivate (et non par la communauté académique) ; et la base de données appartient à Clarivate. Par conséquent, cette nouvelle base de données ne peut en aucun cas être considérée comme un bien commun.

Le « commonswashing » nuit aux communs de la connaissance, ainsi qu'au mouvement des communs en général, car il s'agit non seulement d'une appropriation des ressources communes, mais aussi des avantages sociaux et des représentations symboliques des biens communs eux-mêmes (Dulong de Rosnay, 2020). Par conséquent, le « commonswashing » rend plus difficile la reconnaissance des initiatives de science ouverte qui s'inspirent réellement des principes des biens communs, comme [Knowledge Commons](#), [Humanities and Social Sciences Commons](#) et certains éditeurs du [Radical Open Access Collective](#).

Questions et considérations clés

Clarivate a annoncé son intention d'introduire de nouveaux indicateurs pour les publications disponibles dans toutes les bases de données du Web of Science, y compris « Research Commons ». Cette initiative pourrait permettre d'évaluer un plus grand nombre de publications, mais elle risque également d'accentuer l'externalisation de cette évaluation vers le Web of Science (Pinfield, 2024), d'accroître la dépendance à une base de données unique et de marchandiser un

nombre croissant de publications qui n'étaient pas quantifiées auparavant dans le Web of Science.

Simon van Bellen a soulevé d'autres préoccupations concernant la commercialisation des données ouvertes et des infrastructures ouvertes :

Il convient de se demander comment Clarivate, en tant qu'entreprise à but lucratif, parvient à lancer cette base de données avec une relative assurance de sa rentabilité. Comme indiqué dans cet article, c'est possible parce que Clarivate vend les données sous forme de paquet : elle intègre une ressource primaire gratuite et la connecte à son produit principal. Cette approche n'est pas nouvelle : les éditeurs commerciaux y ont de plus en plus recours ces dernières années, souvent en s'appuyant sur l'IA appliquée à leurs propres publications, voire sur du contenu protégé par le droit d'auteur. Ces entreprises tirent des tendances des résultats de la recherche et créent des « produits de prédiction » destinés à la vente (Pooley, 2024). Or, dans ce cas précis, on assiste à l'intégration flagrante d'une ressource publique, construite grâce à des fonds publics.

On peut difficilement blâmer une entité commerciale d'adopter une telle approche, mais il convient peut-être, une fois de plus, de s'opposer à la commercialisation de ce qui a été construit par la communauté, avec des fonds publics. Les initiatives ouvertes, de plus en plus nombreuses, devraient être privilégiées. Ces ressources sont déjà offertes gratuitement ; le seul obstacle restant est donc potentiellement d'ordre technique. C'est pourquoi il est nécessaire de renforcer les compétences techniques des communautés pour qu'elles puissent utiliser ces ressources.

Un autre problème réside dans le fait que les métadonnées d'OpenAlex (l'une des bases de données utilisées par « Research Commons ») peuvent être inexactes dans des domaines tels que la langue, l'affiliation et les références (Alperin et al., 2024 ; Céspedes et al., 2025). Clarivate a annoncé son intention d'enrichir les métadonnées disponibles sur « Research Commons ». Toutefois, selon un webinaire de « Research Commons » d'octobre 2025, cette opération se fera en unifiant les noms des organisations à l'aide des noms préférentiels du Web of Science et en classant les dossiers selon les catégories thématiques ou les domaines de recherche du Web of Science. Ce ne sont pas des identificateurs persistants ni des systèmes de classification largement adoptés.

OpenAlex prévoit aussi d'améliorer ses données et métadonnées en ajoutant de nouvelles publications et en peaufinant ses algorithmes de sujets et de mots-clés. De même, Crossref a renouvelé ses partenariats avec le Public Knowledge Project (PKP) et le Directory of Open Access Journals (DOAJ) afin de promouvoir des métadonnées plus inclusives et équitables dans les dossiers Crossref.

Réponses du partenariat INKE

Jack Young (Bibliothécaire en impact de la recherche et bibliométrie, Université McMaster) :

L'adoption du langage de la science ouverte par les éditeurs à but lucratif comme stratégie marketing n'est pas un phénomène nouveau, mais l'exemple le plus récent de cette pratique avec le nouveau produit « Research Commons » de Clarivate est particulièrement frappant.

S'appuyant à la fois sur les données et la terminologie des sources de données véritablement ouvertes dont dépend son existence même (à savoir OpenAlex et CrossRef), le produit « Research Commons » sape les idéaux fondamentaux des biens communs en plaçant ces données derrière un abonnement coûteux au Web of Science en limitant l'accès à ces données au lieu de les rendre ouvertes. Bien que parfaitement légale d'un point de vue juridique, cette pratique de dénomination est incontestablement trompeuse et oblige les bibliothèques à établir des distinctions claires entre les outils qui soutiennent véritablement les idéaux de la science ouverte et ceux qui ne sont « ouverts » que de nom.

Ce travail pourrait avoir un impact significatif, notamment dans la promotion de pratiques d'évaluation de la recherche plus responsables, impulsée par des mouvements tels que la Déclaration de San Francisco sur l'évaluation de la recherche (DORA). L'accent mis par DORA sur la transparence et l'ouverture semble s'opposer à des produits comme « Research Commons » de Clarivate. L'une des principales recommandations de DORA pour une évaluation responsable de la recherche est que les organismes fournissant des indicateurs, tels que Clarivate, « [f]ournissent les données sous licence permettant leur réutilisation sans restriction et offrent un accès informatique aux données, lorsque cela est possible » (Cagan, 2013). S'il est vrai que de nombreux outils d'évaluation de la recherche populaires ne répondent pas actuellement à ce critère, « Research Commons » de Clarivate, en adoptant malheureusement le langage des biens communs sans en incarner les idéaux, se distingue par son inadéquation avec les principes fondamentaux de DORA.

Je reconnais que l'évaluation de la recherche n'est qu'un des nombreux cas d'utilisation potentiels de cet outil et que des changements significatifs vers des processus plus transparents doivent provenir de multiples directions. Néanmoins, à mesure que davantage d'organisations adhèrent à DORA et que les discussions se concentrent sur l'infrastructure pratique nécessaire à une évaluation de la recherche véritablement responsable, j'espère que la communauté exercera une pression accrue sur les éditeurs de bases de données et les développeurs d'outils afin qu'ils ouvrent leurs données et les partagent de la même manière, sous peine d'être exclus des processus d'évaluation de la recherche.

Parallèlement, il me semble essentiel que les bibliothèques soient de plus en plus vigilantes quant à la manière dont ces outils sont évalués, promus et utilisés au sein de leurs propres institutions, en veillant à ce que les affirmations d'ouverture soient

exactes et que les pratiques soient conformes aux valeurs qu'elles défendent.

Références

Alperin, J. P., Portenoy, J., Demes, K., Larivière, V., & Haustein, S. (2024). *An analysis of the suitability of OpenAlex for bibliometric analyses* (No. arXiv:2404.17663). arXiv. <https://doi.org/10.48550/arXiv.2404.17663>

Cagan R. (2013). The San Francisco Declaration on Research Assessment. *Disease models & mechanisms*, 6(4), 869–870. <https://doi.org/10.1242/dmm.012955>

Céspedes, L., Kozłowski, D., Pradier, C., Sainte-Marie, M. H., Shokida, N. S., Benz, P., Poitras, C., Ninkov, A. B., Ebrahimi, S., Ayeni, P., Filali, S., Li, B., & Larivière, V. (2025). Evaluating the linguistic coverage of OpenAlex: An assessment of metadata accuracy and completeness. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.24979>

Clavey, M. (2024, March 4). La recherche française parie sur OpenAlex pour briser l'emprise d'Elsevier et Clarivate. *Next*. <https://next.ink/129485/la-recherche-francaise-parie-sur-openalex-pour-briser-lemprise-delsevier-et-clarivate/>

Dulong de Rosnay, M. (2020). Commonswashing – A Political Communication Struggle. *Global Cooperation Research – A Quarterly Magazine*, Vol. 2, No. 3, pp. 11-13. ISSN 2628-5142 (print). ISSN 2629-3080 (online). <https://hal.science/hal-02986722>

Edwards, B. (2025, April 2). *AI bots strain Wikimedia as bandwidth surges 50%*. Ars Technica. <https://arstechnica.com/information-technology/2025/04/ai-bots-strain-wikimedia-as-bandwidth-surges-50/>

Goldfinger, E. (2025, September 4). *Expanding the Web of Science platform with Research Commons*. <https://clarivate.com/academia-government/blog/expanding-the-web-of-science-platform-with-research-commons/>

Hardinges, J., Pearson, S., & Ross, R. (2025). *From Human Content to Machine Data. Introducing CC Signals*. Creative Commons. https://creativecommons.org/wp-content/uploads/2025/06/Human-Content-to-Machine-Data_Final.pdf

Lamdan, S. (2022). *Data Cartels: The Companies That Control and Monopolize Our Information*. Stanford University Press. <https://doi.org/10.1515/9781503633728>

Maddi, A., Maisonobe, M., & Boukacem-Zeghmouri, C. (2025). Geographical and disciplinary coverage of open access journals: OpenAlex, Scopus, and WoS. *PLOS ONE*, 20(4), e0320347. <https://doi.org/10.1371/journal.pone.0320347>

Maxwell, T. (2025, April 17). Wikipedia Is Making a Dataset for Training AI Because It's Overwhelmed by Bots. *Gizmodo*. <https://gizmodo.com/wikipedia-is-making-a-dataset-for-training-ai-because-its-overwhelmed-by-bots-2000590704>

Pinfield, S. (2024). *Achieving Global Open Access: The Need for Scientific, Epistemic and Participatory Openness* (1st ed.). Routledge. <https://doi.org/10.4324/9781032679259>

Pooley, J. (2024). Large language publishing: The scholarly publishing oligopoly's bet on ai. *KULA: Knowledge Creation, Dissemination, and Preservation Studies*, 7(1), 1–11. <https://doi.org/10.18357/kula.291>

Rosnay, M. D. de, & Stalder, F. (2020). Digital commons. *Internet Policy Review*, 9(4). <https://doi.org/10.14763/2020.4.1530>

Tennant, J. P. (2020). Web of Science and Scopus are not global databases of knowledge. *European Science Editing*, 46, e51987. <https://doi.org/10.3897/ese.2020.e51987>

Woodcock, C. (2023, May 2). AI Is Tearing Wikipedia Apart. *VICE*. <https://www.vice.com/en/article/ai-is-tearing-wikipedia-apart/>

Search



Archives

Select Year



Categories

Community News

English

French

Insights and Signals Reports

Observations

Observations and Responses

Policies

Portuguese

Responses

Spanish

Tags

AI bots / Robots d'indexation IA AI governance / Gouvernance de l'IA

AI safety / Sécurité de l'IA Berlin Declaration / Déclaration de Berlin

Bethesda Statement / Déclaration de Bethesda bibliodiversity / bibliodiversité

Bilingualism / bilinguisme Budapest Statement / Déclaration de Budapest Canada

Canadiana.org Canadian government/le gouvernement du Canada CAPOS

CARL / ABRC collaboration commonswashing

community engagement / engagement communautaire

Compute Canada / calcul Canada copyright / droits d'auteurs

credibility / crédibilité CRKN / RCDR Cybersecurity / Cybersécurité

data management / gestion des données

[diamond open access / le libre accès diamant](#) [Digital Commons / Commun numérique](#)

[digital scholarship / version numérique](#) [en français / French](#) [English / en anglais](#)

[events and gatherings / événements et rassemblements](#)

[Federation for the HSS / Fédération des sciences humaines](#)

[funding agencies / organismes de financement](#)

[generative artificial intelligence / l'intelligence artificielle générative](#)

[identity management / gestion de l'identité](#) [implementation / mise en oeuvre](#) [INKE](#)

[International OA Week / Semaine internationale du libre accès](#)

[international policy / politique internationale](#)

[licensing agreements / accords de licence](#) [Multilingualism / Multilinguisme](#)

[Naylor Report / le rapport Naylor](#) [open access / libre accès](#)

[open data / données ouvertes](#) [open education / éducation ouverte](#)

[open government / gouvernement ouvert](#) [open infrastructure / infrastructure ouverte](#)

[Open Scholarship Press](#) [open science / science ouverte](#)

[open social scholarship / approches sociales des savoirs ouverts](#)

[open source software / les logiciels libres](#) [ORCID](#) [peer review / critique des pairs](#)

[Perpetual Access / Accès perpétuel](#) [PKP](#) [Plan S](#)

[Plan S update / mise à jour du Plan S](#) [policy / politique](#)

[policy guide / guide des politiques](#) [promotion et titularisation](#) [publishing / édition](#)

[RDC / DRC](#) [RDM](#) [RECODE](#) [recommendations / recommandations](#)

[reports / les rapports](#) [repositories / les dépôts](#)

[research creation / recherche-création](#)

[research evaluation / l'évaluation de la recherche](#)

[research libraries / les bibliothèques de recherche](#)

[research output / les résultats de la recherche](#)

[research security / sécurité de la recherche](#) [RPT / révision](#)

[scholarly communication / communication savante](#)

[scholarly communication / la communication savante](#) [SFU Library / Bibliothèque](#)

[social media / les médias sociaux](#) [Tri-Agency / des trois organismes](#) [UK](#)

[UK / Royaume-Uni](#) [UNESCO](#) [UVic Libraries](#) [Érudit](#)



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

