Improving OpenAlex Data: Why Publishers Should Help Authors Improve the Scholarly Record

By PAUL TUINENBURG OCT 23, 2026

AUTHORS | DISCOVERY | INFRASTRUCTURE | INNOVATION | PRESERVATION | RESEARCH | TECHNOLOGY

Editors' note: Today's post is by Paul Tuinenburg, co-founder of Global Campus.

A growing number of the systems researchers and institutions rely on every day are being built on the same underlying map of who published what. The Sorbonne University in France was the first major university moving from using the proprietary Web of Science data to the open infrastructure of OpenAlex in 2024. In that specific case, OpenAlex serves as an assistant layer for the library.

But OpenAlex is also used as an assistant layer for a funder's reporting dashboard, a global ranking like the CWTS Leiden Ranking, and a new AI research assistant. Most people using these tools never see it directly, which is exactly why it matters: an error in that underlying map doesn't stay contained to one platform; it propagates silently into every tool built on top of it, including the ones publishers and researchers use to track their own impact.



OpenAlex: The Data Infrastructure Organizations Quietly Build On

Named after the ancient Library of Alexandria, OpenAlex launched in 2022 as an open successor to Microsoft Academic Graph, maintained as a public good by the non-profit OurResearch. It indexes hundreds of millions of works, authors, institutions, sources, and funders, along with the links between them, and it makes all of it available under a fully open license (CC0), free for anyone to read, reuse or build on (within certain use limits).

As a piece of infrastructure, what OpenAlex has built at this scale, with no subscription wall, is genuinely impressive. It has become one of the largest connected records of scholarship in existence. But scale is also where its hardest problem lives. Matching tens of millions of papers to the correct person,is one of the most persistent unsolved problems in bibliometrics, and even a very good algorithm produces a meaningful number of errors when applied across hundreds of millions of works.

477M	CC0	2022
works indexed — the largest connected record of scholarship published	100% open data, free to read, reuse and build on, with an open API	open successor to Microsoft Academic Graph; an alternative to Scopus & Web of Science

Why Disambiguation Breaks, and How it Gets Fixed

The failure modes are well understood. Common names create ambiguity that only grows with scale. Researchers publish under inconsistent variants of their own name over a career. Whether it's maiden names, middle initials present in some records but omitted in others or simple typing mistakes. Institutional affiliations change, and two researchers at the same institution with similar names can be nearly indistinguishable from citation data alone.

The result is either split identities, where one person's work is scattered across several profiles, or merged ones, where the work of two academics is merged in one profile. Some researchers have a linked ORCID in OpenAlex, which is the strongest signal the algorithm has to work with, but not everyone has an ORCID, or uses it, so this only solves part of the problem. This leads to wrong data; not wrong on purpose, simply unconfirmed until a human checks them. And this isn't only a discovery problem. Research-integrity tools built to catch paper mills and citation manipulation depend on the same authorship data. When a record is thin or misattributed, those tools lose precision precisely when it matters most.

An unclaimed profile isn't wrong on purpose — it's simply unconfirmed. And unconfirmed data propagates into every tool built on top of it.

The fix is almost entirely a human one: an author claiming their own profile, merging their name variants, and linking their ORCID resolves the ambiguity an algorithm alone cannot. Publishers are unusually well placed to prompt this, because they're in direct, trusted contact with authors at the moments they're most engaged with their own scholarly identity: upon acceptance and at publication. A short, well-timed prompt from a publisher will do more than a generic invitation for an administrative chore ever could. In addition: publishers have an additional source of unification in accounts of submission systems: An academic with a common name that has published without specifying an ORCID, might still have used a single account in a manuscript submission system, proving that the output belongs to a single academic author.

Our ask: help your authors claim and maintain their profiles

Encourage every corresponding author to take a few minutes to confirm their OpenAlex record. Concretely, publishers can:

→ **Add a line to acceptance and publication emails** inviting authors to claim their OpenAlex profile and check it is accurate.

→ **Encourage authors to connect their ORCID**, so future work is attributed automatically.

→ **Remind authors to merge name variants and remove mis-attributed works,** so their metrics reflect reality.

→ **Reinforce the message that accurate author data benefits the author** — in discoverability, evaluation and reputation — not just the database.

Can We Ask Researchers to Update Another Profile?

Researchers already update their presence on Google Scholar, ResearchGate, ORCID, institutional repositories such as Pure, and profile sites like Loop. So is it reasonable to ask them to maintain another?

We believe researchers should because OpenAlex is combining taking the responsibility to build and maintain this ledger of academic content with the commitment to doing that in the open, thereby returning its value to the academic community. This combination of ambition in setting the record straight, combined with opening up both the source code and the data, is why OpenAlex deserves the time and support of the academic researchers.

Conclusion

So why should publishers care about authors claiming their OpenAlex profile? For those publishers that use OpenAlex data directly, the answer is very simple, it directly improves their own analyses on the data. But besides that, a lot of vendors that publishers rely on use the OpenAlex source data to deliver their value. Improvements in that data directly impact the quality of reviewer suggestions, citation checks, integrity checks, potential conflicts of interest, etc…The same records also affect how a journal's own output is attributed, discovered and counted in rankings and funder reports. Every author claiming and curating their own profile makes the data feeding these reports and rankings more reliable.

The cost of acting is very small when compared to the benefit. A single sentence in an acceptance email reaches authors when they are most invested in their work. It requires no new systems, budget or time from staff. Because OpenAlex is open, a correction made once improves every platform that builds on it. And every profile an author confirms makes the shared map a little more trustworthy — for that author, for the publisher whose metadata depends on it, and for every reader and tool downstream.

Open knowledge is maintained not by its owners, but by all of us. Publishers can help more authors play their part. Publishers have long been responsible for the quality of the scholarly record within their own journals. Helping authors claim their OpenAlex profiles extends that responsibility to the shared infrastructure a lot of people in academia rely upon, and very few organizations are as well placed to do it.

Author note: AI-assisted editing tools were used to create images and to help revise wording and structure for clarity.

Paul Tuinenburg

Paul Tuinenburg is a co-founder of Global Campus, an AI-powered platform that helps publishers, institutions, and funders navigate the scholarly landscape. Mostly to identify (peer) reviewers and collaboration partners. Global Campus builds on OpenAlex data, an open repository of scholarly infrastructure. The fact that this data is openly accessible, doesn’t mean it’s free to maintain. That’s why Global Campus is a paying supporter of OpenAlex. To contribute to a flourishing academic ecosystem in the future!