Unlocking references from the literature: The Initiative for Open Citations

Dario Taraborelli (1) and Mark Patterson (2)

(1) Wikimedia Foundation, San Francisco, USA. ORCID: 0000-0002-0082-8508, (2) eLife, Cambridge, UK. ORCID: 0000-0001-7237-0797

The idea of creating a freely accessible repository of citation data—representing how scholarly works cite each other—has been hampered until recently by restrictive licenses and by the lack of comprehensive, machine-readable data sources: for decades, references have been locked inside PDFs or proprietary databases. Launched in April 2017, the Initiative for Open Citations (I4OC) has made nearly half of all indexed scholarly references freely available to everyone, with no copyright restrictions. The percentage of indexed scholarly works with open reference data was 1% before the launch of the I4OC: as of July 2017, over 16 million scholarly works have open references available as machine-readable public domain data. There's now momentum and a growing number of organizations, scholarly societies, funders, and publishers in support of the unconstrained availability of scholarly citation data. However, this is just the beginning of a journey to grow the scholarly commons with high-quality citation metadata. In this presentation, we'll talk about how the I4OC was created, and where it's going next, its current vision and challenges. We'll showcase examples of real-world applications demonstrating how data unlocked by the initiative is being reused to accelerate scientific discovery and to strengthen open scholarly infrastructure.