

# Insights into the Economy of Open Scholarship: A look into ScienceOpen with Stephanie Dawson, CEO



**DFG**



# About ScienceOpen

ScienceOpen is an interactive discovery environment for scholarly research across all disciplines. It is freely accessible for all and offers hosting and promotional services within the platform for publishers and institutes. ScienceOpen is a privately funded startup company, owned by Alexander Grossmann and Tibor Tscheke. The main office is in Berlin, Germany with technical offices in Boston, USA and Budapest, Hungary. While the service is free for end users, publishers, institutes, and scholarly societies are charged either for full content hosting or for promotional services, which consist of the creation of 'collections', that is, curated landing pages collecting articles on a certain topic.

[scienceopen.com](https://scienceopen.com)

# ScienceOpen: Business model

## Key activities

- ▶ Discovery platform for open access articles
- ▶ Option to form cross-journal topical collections with advanced indexing and altmetrics options
- ▶ Conference poster publication platform



## Organisation type

- ▶ Commercial startup
- ▶ Privately funded, owned by Alexander Grossmann and Tibor Tscheke



## Key partners

- ▶ Publishers, scholarly societies, institutions
- ▶ UCL Press
- ▶ Member of CrossRef, ORCID, the Open Access Scholarly Publishers Association (OASPA), The International Association of Scientific, Technical and Medical Publishers (STM), and the Directory of Open Access Journals (DOAJ)



## Revenue streams

- ▶ Private funding
- ▶ Fee paid by publishers, institutions, and scholarly societies
- ▶ Contract with UCL Press for mega journal



## IP/Copyright

- ▶ All articles aggregated are open access (all licences allowed) – preprints/conference posters: CC BY 4.0
- ▶ Promotes open metadata and abstracts
- ▶ Own outputs: CC BY 4.0



## Customers/users

- ▶ Publishers
- ▶ Scholarly societies
- ▶ Research institutions
- ▶ Individual researchers



# Interview with Stephanie Dawson

The original idea behind ScienceOpen was to speed up science by removing the constraints of the traditional journal, and to experiment with digital tools to achieve this.

“ScienceOpen started in 2013, the year of the ‘mega journal’. **PLOS ONE** ([journals.plos.org/plosone](https://journals.plos.org/plosone)) appeared to be successful and changed the perception of what publishers can and should do in order to speed up the process of scientific communication. Post-publication collections suddenly became possible, for example, around Zika,” says **Stephanie Dawson**, CEO of ScienceOpen.

Many ScienceOpen employees had previous experience at big traditional publishing houses and experienced the digitisation process from the start, when it entailed nothing more than the digitisation of articles into pdfs. “The baggage of these legacy publishers was huge,” says Dawson, “and this was very frustrating. We couldn’t make colour images available online for technical reasons, there were always extra costs charged to the author, and the print version always remained the reference. We all felt that the full potential of going digital was not being used at all by the big legacy publishers.”

“ We all felt that the full potential of going digital was not being used at all by the big legacy publishers. ”

ScienceOpen set out to rethink this old-fashioned process. The idea was to get the research out as soon as possible with the community crowdsourced peer review happening afterwards in a mega journal platform environment. This mega journal would then be embedded in a larger discovery platform that would aggregate other information from **ArXiv** ([arxiv.org](https://arxiv.org)), **PubMed** ([ncbi.nlm.nih.gov/pubmed](https://ncbi.nlm.nih.gov/pubmed)) etc so that every single article would be embedded in a much wider scientific context. “We hoped that our network would be able to communicate in a faster, digital and very transparent way about their research results because we also wanted to open up the peer review system.”

Although many researchers were very enthusiastic about the idea, ScienceOpen soon reached the level of innovation that authors were open to. “When it came down to publishing directly into our mega journal, they were suddenly worried about impact and reputation,” says Dawson. “In the end, it turned out that we had underestimated the implicit contract between publisher and researchers, which was still largely about reputation. This is a currency that you don’t get to offer as a brand new product.”

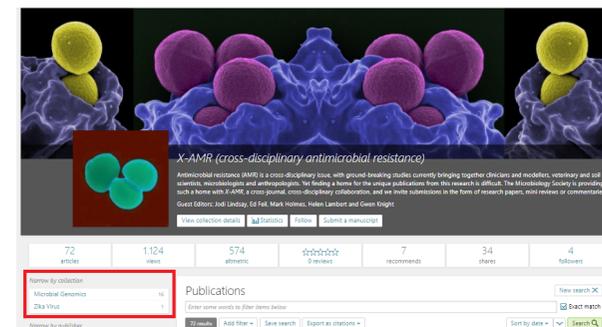
ScienceOpen decided to set its publishing ambitions aside, together with the idea of sustaining the platform by charging article processing fees. The decision was made to focus instead on offering services to publishers and societies that wanted to enhance the visibility of their niche journals by putting them in a broader context, including a search engine. The discovery platform remained in place, but an interactive overlay for user interaction and peer review was added. Dawson: “We focused on improving our overlay framework and services for researchers to convince publishers that our system could improve their visibility. The curation/collection

aspect is our unique selling proposition: we’ve made it possible to visualise different ways to curate content in a way that is not threatening to publishers or journal titles.”

For now, the main activities of ScienceOpen focus on the hosting of external content in a curated collection structure. Dawson: “We take a topical selection of articles. When working with publishers we can make a collection based on a single journal or, if we’re working with other entities such as scholarly societies, we can spread the net wider by making a selection from multiple journals.”

“For example, the **Microbiology Society** ([microbiologysociety.org](http://microbiologysociety.org)) created a collection on antimicrobial resistance that pulled from all of their journals, thus, creating a sort of ‘virtual’ journal. We promote these collections on our platform using banners and integrate it in our search function. The content remains hosted elsewhere though; we simply promote it via the metadata, linking back to the version of record on the publishers’ servers. It’s important to note that we only host and link to open

access articles, we are not interested in creating tools that will stop or even hinder the free flow of information.”



Example of a collection curated by the Microbiology Society

“Our plan to implement post-publication peer review has been pushed back for a while. I really wanted this to be an integral part of the contribution of ScienceOpen to ‘open’ but we haven’t yet been able to get researchers excited about this. We were the first ones to set up the **Crossref** ([crossref.org](https://crossref.org)) XML schema for post-publication peer review but it has not taken off yet, although it could be the next big opportunity.”

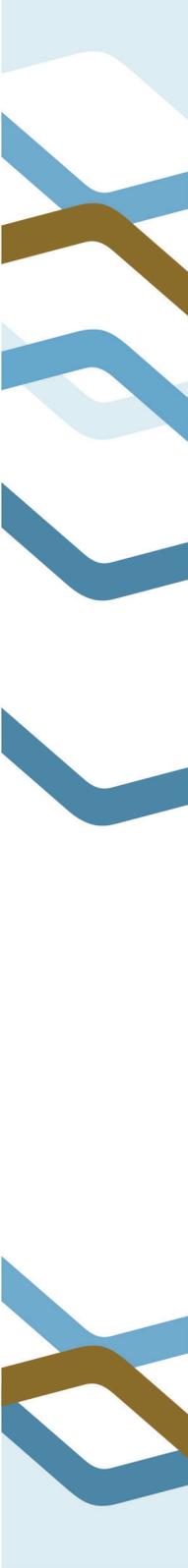
However, by engaging with **UCL Press** ([ucl.ac.uk/ucl-press](https://ucl.ac.uk/ucl-press)) in London, for which ScienceOpen is building a mega journal platform, ScienceOpen wants to revive this old ambition. “They really want authors to publish the preprint with the peer review happening in the open, and then publish the final peer reviewed versions. UCL is probably in a better position to make this work because they have enthusiastic researchers who will engage their peers and talk about open peer review. I think this grassroots approach can be more effective than if we, as a for-profit company, impose it on them.”

“ I think a grassroots approach towards post-publication peer review can be more effective than if we, as a for-profit company, impose it on researchers. ”

“It could be a great model for other institutions that want to start up a mega journal. This can be more impactful than just publishing in a repository or creating more small niche journals. Our set-up is perfect for overlay journals based on this peer review system. But we can’t do it on our own, so that’s why we consider our collaboration with UCL as a great demo project.”

Because the majority of its content is externally hosted, ScienceOpen relies heavily on the quality of the metadata it receives. The company is part of the **Metadata2020** ([metadata2020.org](https://metadata2020.org)) group, an organisation pushing for better quality metadata.

Dawson: “ScienceOpen offers an important indirect service to libraries, institutions, and individual researchers. We enable them to evaluate which publishers do a good job and which do a bad job in terms of discoverability of research, by providing good quality metadata, openness of references, and abstracts. These are all part of the metadata and should never be copyrighted, not only for the benefit of ScienceOpen, but for everybody’s benefit.”



“I feel it’s very much part of our mission to try to get the most out of the available digital tools,” says Dawson. “We really push all publishers we work with to care about the importance of metadata. Having your content tagged as open access, is something you need to do as a publisher, or you are doing a bad job at distributing the content. Authors are paying often high article processing charges (APCs) for this, and as a publisher you need to make this visible digitally. If you don’t do this, you deserve criticism! Publishers need to pay more attention to it. Algorithms such as that of **Unpaywall (unpaywall.org)**, which we use, are very good at detecting false positives but I am convinced that there are a lot of open access materials out there which are just not being picked up, simply because the metadata is inadequate. I fear this leads to a serious underestimation of the total amount of open access material that is actually available.”

“Inadequate metadata leads to a serious underestimation of the total amount of open access material that is actually available.”

Aside from its own hosting services, ScienceOpen is working with publishers to drive traffic to its version of record via their metadata. They get help to clean up and enrich their metadata and to add it to Crossref. ScienceOpen often works with publishers directly to convince them to improve their metadata.

Dawson: “Often it’s only a matter of changing or adding one metadata field! The smaller publishers who are struggling technically to add the right information to Crossref are more frustrating, either because they don’t have the technical knowhow or because they lack manpower. Then again, it’s not only the small companies. I know of some larger publishers who also just don’t care about getting the right metadata. For me this is all part of the same mission: to get better article and author information so that the entire research community can work more smoothly and more in concert with each other.

“We encounter the same reluctance when we discuss making abstracts openly available. We know that content with an abstract gets accessed and used up to ten times more than content without

one but not all publishers are actually taking the trouble to deposit abstracts with Crossref! If you don’t do this, which is the case for some major publishers, you are doing a disservice to your authors. Some are trying to monetise this essential part of research dissemination. I believe that making some research items open as default is a basic courtesy. Also, there are some cool artificial intelligence (AI) initiatives that can do such interesting stuff by mining the abstracts and references! I would like to see publishers take this more seriously.”

“If you don’t make abstracts openly available you are doing a disservice to your authors. Some publishers are trying to monetise this essential part of research dissemination.”

ScienceOpen offers open access hosting within its interactive discovery environment, including an individualised collection structure for a list price of 2,000 US dollars (USD) per year and a per article fee of 25 USD. Promotional collections can be booked for the same yearly fee with an article aggregation fee of five USD per article.

Dawson: “We always crosscheck with Crossref and the **Directory of Open Access Journals (DOAJ)** ([doaj.org](http://doaj.org)) for ‘new’ content, combining it with available metadata. If a publisher wants to go beyond the articles that we pick up and the metadata we can collect, they can also give us full metadata access to their catalogue. We can really work with them to improve visibility for individual articles, journals and so on. There are so many out there, it can be really hard if the publisher cannot differentiate its own brand from the competition.”

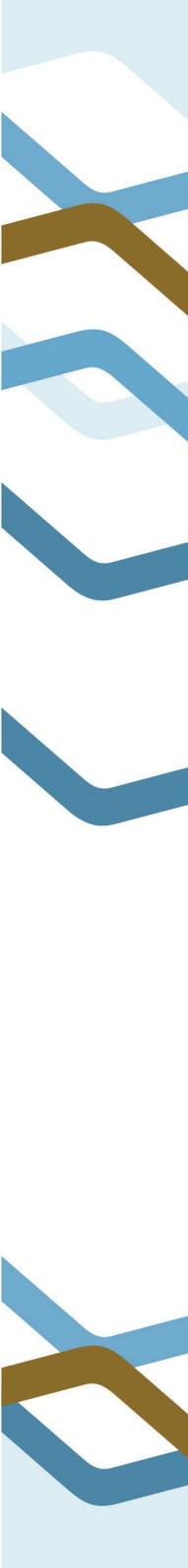
However, the service remains free for end users. In addition, researchers who want to create and edit their own topical collections can use the collection functionality. ScienceOpen even advises them about

curation and enhancement of these collections, for example, by suggesting additional articles or illustrations. Currently, 35,000 researchers use the service to promote their own work or to create a new collection for their field of interest. Dawson: “Publishers are also showing an interest in this, because for them it’s a very direct way to get feedback from researchers, both in terms of which collections are being created and which articles are being added to them.”

Above all, to become sustainable, ScienceOpen needs to ensure enough buy-in from paying customers. Right now, 150 paid-for collections are hosted on the platform, but in order for ScienceOpen to be both sustainable and to be able to invest more in development this number needs to increase. The company is trying to diversify revenue streams by licensing features to other projects, but thus far this has not been proven to be scalable.

For its end users ScienceOpen fosters a very open philosophy: it does not even require registration to use the discovery tools. Users only have to register if

they want to do something on the platform such as collection curation or peer review. From the beginning, ScienceOpen has been keen to promote the use of **ORCID** ([orcid.org](http://orcid.org)) (the persistent identifier for researchers). ScienceOpen asks people to edit and augment their ORCID profiles, instead of editing user profiles on the platform. Everything is imported from the ORCID servers.



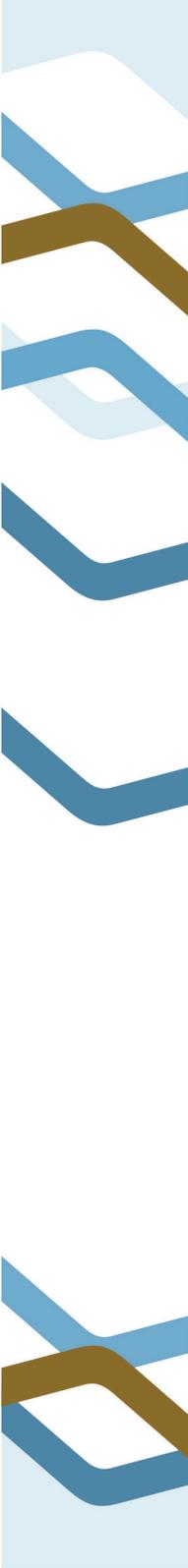
Dawson: “I think this way of working illustrates our philosophy nicely! ORCID is an excellent digital tool and it makes research dissemination more efficient and interoperable between systems because there is less confusion about who wrote what. The extra step might have deterred some people in the beginning, but we still think it’s worth it. Also, it’s more efficient for us as we don’t have to manage all this personal information.”

“ScienceOpen is a for-profit company. We do occasionally get criticised for that. However, I think for-profits have a place in the service provider infrastructure. Universities work with commercial companies in so many different areas. I don’t see it as a big issue. But it has made it difficult for us to work with libraries, for example. They work with commercial publishers all the time, but when they start thinking about open access infrastructure they project their bad experiences with those publishers onto us.

“There’s certainly wariness about potential data or vendor lock-in. But programming for yourself is usually very expensive, so I don’t think it’s possible to have all your services built in-house based on not-for-profit open source infrastructure.”

“We try to be very cautious about not creating any data lock-in ourselves, but we need to get even better at respecting the work that goes into hand-curated collections. If an editor wants to move their collection they can export the metadata, but only for 200 items at a time. We are working on a system so that researchers will always have the assurance that they can download the metadata of their collections in its entirety, hassle-free. We don’t have contracts with researchers but that’s one of the places where we have to come up with a good technical solution so that researchers know they’ll always be able to take their data with them.”

Although services similar to ScienceOpen, such as **Scopus** ([scopus.com](https://scopus.com)) and **Web of Science** ([wok.mimas.ac.uk](https://wok.mimas.ac.uk)), cover a large part of the market, Dawson doesn’t think they represent direct competition because these are paywalled services. However, she believes **Google Scholar** ([scholar.google.co.uk](https://scholar.google.co.uk)) is much more of a threat because of its powerful algorithm: “On the other hand it can be argued there are weaknesses in their model – because they are essentially a ‘black box’.



“I really see open scholarship as being committed to shifting the research paradigm to a more collaborative mode in the digital space, with the brand new digital tools that we now have available to us. We know how to solve problems faster with open science. For me, a great example is the H1N1 swine flu outbreak in 2009: the **American Centers for Disease Control and Prevention (CDC)** ([cdc.gov](http://cdc.gov)) coordinated efforts, uploading the complete gene sequences to a public database to enable global cooperation. The goal was not to get credit but to get their results out there as fast as possible, to get a vaccine as soon as possible. For me, that is really what open science is about. Why can't we all work more collaboratively, globally, with the digital tools that we have? How do we need to change the reward system, change scholarly structures, so that scholars can solve the big issues?

“We have the tools, but we haven't figured out how to get everybody on board yet. It will require a cultural shift that will take some time – but with ScienceOpen, we are trying to make a contribution,” concludes Dawson.

### References and relevant links

- ▶ ScienceOpen website: [scienceopen.com](http://scienceopen.com)
- ▶ PLOS ONE: [journals.plos.org/plosone](http://journals.plos.org/plosone)
- ▶ ArXiv: [arxiv.org](http://arxiv.org)
- ▶ PubMed: [ncbi.nlm.nih.gov/pubmed](http://ncbi.nlm.nih.gov/pubmed)
- ▶ Microbiology Society: [microbiologysociety.org](http://microbiologysociety.org)
- ▶ Crossref: [crossref.org](http://crossref.org)
- ▶ UCL Press, London, UK: [ucl.ac.uk/ucl-press](http://ucl.ac.uk/ucl-press)
- ▶ Metadata 2020: [metadata2020.org](http://metadata2020.org)
- ▶ Unpaywall: [unpaywall.org](http://unpaywall.org)
- ▶ DOAJ: [doaj.org](http://doaj.org)
- ▶ ORCID: [orcid.org](http://orcid.org)
- ▶ Scopus: [scopus.com](http://scopus.com)
- ▶ Web of Science: [wok.mimas.ac.uk](http://wok.mimas.ac.uk)
- ▶ Google Scholar: [scholar.google.co.uk](http://scholar.google.co.uk)
- ▶ Centers for Disease Control and Prevention: [cdc.gov](http://cdc.gov)

# About Stephanie Dawson CEO



Stephanie Dawson grew up in northern California and studied biology at Yale University. She then worked at the labs of Susan Parkhurst at the Fred Hutchinson Cancer Research Center in Seattle, WA and Ralph Rupp, at the MPG Friedrich Miescher Laboratory, Tübingen, Germany before changing fields and getting a PhD in German literature from the University of Washington under Jane Brown. From 2001-2012 she worked in various positions at the academic publisher De Gruyter in Berlin in the fields of biology and chemistry in both journals and book publishing. In 2013 she joined the ScienceOpen management team as CEO.

© Knowledge Exchange 2019

Title: Insights into the Economy of Open  
Scholarship: A look into ScienceOpen  
with Stephanie Dawson, CEO

Authored by: Gwen Franck

Contact: [gwenfranckgcv@gmail.com](mailto:gwenfranckgcv@gmail.com)

Interview date: November 15, 2018

DOI: [10.5281/zenodo.3084650](https://doi.org/10.5281/zenodo.3084650)

This report is available under a CC BY 4.0  
licence [creativecommons.org/licenses/by/4.0](https://creativecommons.org/licenses/by/4.0)



Knowledge Exchange Office  
C/ O Jisc,  
One Castlepark,  
Tower Hill,  
Bristol, BS2 0JA

t: 0203 697 5804

e: [office@knowledge-exchange.info](mailto:office@knowledge-exchange.info)