# ImpactStory.

Insights into the Economy of Open Scholarship:
A look into Impactstory with
Heather Piwowar, co-founder
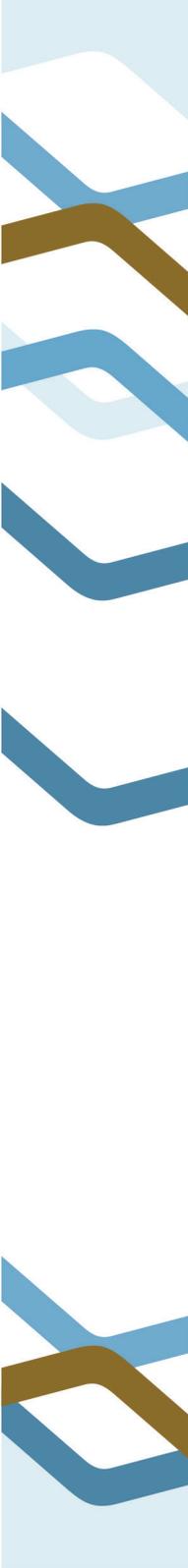
‹ KE ›
*Knowledge Exchange*

DFG    CSC    Jisc    Cnrs    SURF    Deff
Denmark's Electronic
Research Library

# About Impactstory

Impactstory is a non-profit dedicated to making scholarly research more open, accessible, and reusable. It offers free services including Unpaywall, Impactstory Profiles, and Depsy. The company is based in Vancouver, Canada. Its mission is to change the focus of the scholarly reward system to value and encourage web-native scholarship. The organisation receives or has received grant funding from The Open Society Foundation, The Alfred P. Sloan Foundation, The National Science Foundation (NSF), Clarivate Analytics, and Arcadia. Aside from this, Impactstory earns income from selling the Unpaywall Data Feed to 12 high-profile customers and from a smaller subcontracting grant with the University of Texas at Austin.

impactstory.org

# Impactstory: Business model

### Key activities

▶ Unpaywall: free version and data feed (paying service)
▶ Depsy
▶ Impactstory Profiles
▶ Research/consultancy

### Organisation type

▶ Non-profit
▶ Two full-time equivalent (FTE) staff (+ One additional hire in December 2018)

### Key partners

▶ Funding agencies
▶ University of Texas at Austin
▶ Data providers

### Revenue streams

▶ Earned income from Unpaywall data feed
▶ Grant funding
▶ Subcontracting/ consulting for University of Texas at Austin

### IP/Copyright

▶ Code base: open source
▶ Data: depends on status of the data provided

### Customers/users

▶ Publishers, funders, institutions, companies
▶ Individual researchers

Partially based on the Business Model Canvas designed by: **Strategyzer AG (strategyzer.com)** (available under CC BY-SA 3.0
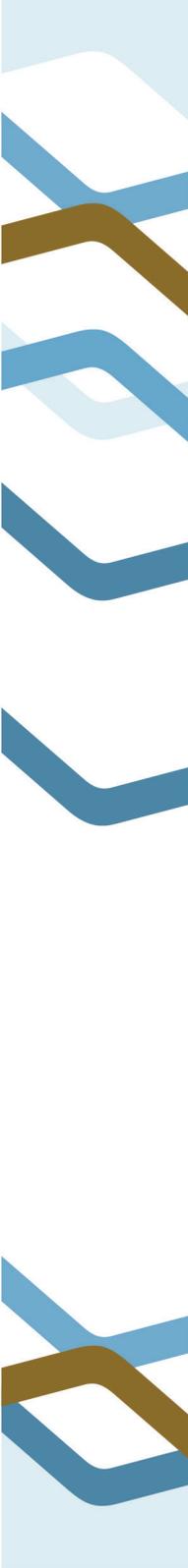
# Interview with Heather Piwowar

Since its foundation in 2011, Impactstory has released a range of services to help researchers measure the impacts of their research outputs. What started out with an attempt to build an author profile system intended to encourage altmetrics and offer incentives for sharing research, has now become a well-established organisation with Unpaywall as its most well-known product.

"Essentially, Impactstory is a two-person non-profit," says co-founder **Heather Piwowar.** "My co-founder Jason Priem and I met at a hackathon during our PhD years. Our initial idea was to build a profile system for researchers that offered metrics and incentives for sharing. A small grant from **Sloan (sloan.org)** and a collaboration with **Duke University (duke.edu)** allowed us to work full time on this project. However, despite people loving our profile project on social media, we only got about 20,000 people to actively use it, which was not enough for it to be sustainable. Even after we experimented with asking a fee for it, it became clear that we would not be able to continue to run the service."

Although the research profile service did not take off, Impactstory procured some other grants to fund **Depsy (depsy.org)**, a tool for researchers to showcase open source software they have created and to add value to it, promoting credit for software as a fundamental building block of science.

The organisation is also acting as a subcontractor for the **University of Texas at Austin on a project to improve research software (ischool.utexas. edu/tags/impactstory)** by helping its creators get proper credit for their work. Another project Impactstory worked on was 'Open Heroes', a badge system to reward researchers who made their work openly available.

Piwowar: "While developing these services, we realised there was no good all-encompassing way to determine whether papers were actually open access. A lot of services we relied on (such as **CORE** (core.ac.uk), **Open Access Button** (openaccessbutton.org), and **OpenAIRE** (openaire.eu)) did not, at that time, deliver the kind of results we required, at the very high volume we needed – namely a clear 'yes' or 'no' that was actually accurate. That's why we decided to build our own code based on the **BASE** (base-search. net), **PubMed** (ncbi.nlm.nih.gov/pubmed), and **ArXiv** (arxiv.org) APIs [currently, Impactstory is no longer using the BASE API, Gwen Franck]. For us, precision is key: if we say something is open access, it definitely is. We don't want surprise embargoes where a paper that is labelled as 'open' turns out to be embargoed after all, as happens with many repositories. With us, the link will always lead to the open access version of the paper.

> " For us, precision is key: if we say something is open access, it definitely is. "

"But as it turned out, the **Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH)** (openarchives.org/pmh) standard does not have a single way to ensure that the version directed to is open. That's why many repositories do not show this data accurately and that's why some aggregators don't display the correct information."

However, even after building their own code, Impactstory still had issues, either because the data was not accurate enough, it was not delivered fast enough, or because they could not call the API often enough. In one case, it was difficult to get permission for commercial reuse and this was problematic, because they intended to sell the service. Starting from this initial product, Impactstory decided to spin it off to create its own API.

The Unpaywall API finds open access content in many places, including using data from open indexes like **Crossref** (crossref.org) and the **Directory of Open Access Journals (DOAJ(** (doaj.org)**)**, but the majority of the open access content comes from independently monitoring over 50,000 unique online content hosting locations. This new service, the 'Unpaywall Data Feed', quickly took off as libraries started to integrate it. They built a browser extension and started doing regular data snapshots. This is now Impactstory's best-known service.

**Data Feed: Changefiles**

The Data Feed changefiles show all the changes in the Unpaywall database over time. They are provided for subscribers to the Unpaywall Data Feed. A new file is added every Thursday. Files use the same schema as the REST API and database snapshot.

This list is also available via a JSON endpoint for programmatic access.

API key required

Paste your API key here

ACCESS CHANGEFILES

*Unpaywall Data Feed*

Apart from the API, which gets about three million calls a day, Impactstory makes a data dump available for commercial and non-commercial use every six months. It also has a weekly dump showing access status changes of individual papers, based on their **Digital Object Identifier (DOI) (**doi.org**)**. This weekly dump is a paying service, for which Impactstory currently has 16 high-profile paying customers, including Elsevier, Clarivate Analytics, and Digital Science. Of the around 100 million DOIs available worldwide, approximately 0.5% have an access level change every week. "With half a million publications with a DOI having a status change every week, our product has a significant impact on the accuracy of the services delivered by our customers," says Piwowar.

"This service brings in about half a million dollars a year and, as we don't really have any competition there, this is a pretty steady stream of income for us which will allow us to hire a programmer to finally expand our team of two."

Impactstory's income sources are diverse: earned income from Unpaywall, the subcontract for research at University of Texas, and a significant grant funding stream. "Despite the success of the Unpaywall paid service we don't plan to move away from grants, because they allow us to solve problems that the market isn't solving," says Piwowar. "The grants are still important for us to remain innovative, for example, right now we are building a search engine on top of Unpaywall with an **Arcadia Fund (**arcadiafund.org.uk**)** grant."

> **Despite the success of our paying service, we don't plan to move away from grants, because they allow us to solve problems that the market isn't solving.**

With only the two co-founders on the payroll at the time of the interview [a third hire was made in December 2018, Gwen Franck], Impactstory is a very nimble organisation. They manage their work according to the 80/20 principle: finding out which 20 percent of the effort will solve 80 percent of the problem. Because of Impactstory's small size it is good at changing course if something is not a good fit. These are characteristics that are more often associated with startup businesses, not with non-profit organisations, and that is not a coincidence.

> **We thought hard about whether we wanted to be a non-profit or a commercial company. I think the fact that we are so much in doubt about this, makes us different from other players in the field.**

"We thought hard about whether we wanted to be a non-profit or a commercial company," says Piwowar, "and I think the fact that we are so much in doubt about this makes us different from other players in the field. I don't necessarily think that non-profits are the only ones that can be mission-driven, but we realise that our non-profit status allows us to reassure some minds, especially in academia, and it also allows us to pursue certain types of grants. And because we are not seed funded, unlike many commercial startups, we don't need to pay back our funding."

" Non-profits often have a reputation for being bloated, slow, unsustainable, dependent on grant money [...] we try to incorporate the start-up mindset in our own business. "

"Non-profits often have a reputation for being bloated, slow, unsustainable, and dependent on grant money. You become like the people you hang out most with, so that's why we like to hang out with nimble startups in incubators! We try to incorporate the startup mindset in our own business."

Impactstory's small size is not the only thing that makes the organisation unique. Both Piwowar and Priem are academics, so they have in-depth knowledge of academic sensitivities, but they are also very technically oriented with a programmer mindset.

They can solve technical and coding issues themselves instead of having to hire extra staff (although, as mentioned before, they are now hiring a dedicated programmer to help out with the new project).

Both are also excellent communicators who like to do their own marketing and advocacy, and Piwowar states that they intend to continue that way: "We've actually never considered hiring a dedicated person for communications and marketing. Right now, we approach our communications very intuitively and I

am not sure what would be considered a good job, objectively. Despite Jason having coined the term 'altmetrics', we've learned the hard way that tweets are not a substitute for uptake. Twitter can provide a skewed perspective – it does not always reflect growth rate. Our first attempt to create a research profile product generated a lot of buzz online but this wasn't converted into growing usage numbers."

A downside of being a small organisation is that there's no in-house support for administrative and legal matters and for pursuing grants, so a lot of the already limited staff time goes into working on those, instead of into developing services.

Another drawback of Impactstory's small size is that, despite being in a very good place to build innovative tools, its progress is not always as fast as desired.

"Ironically, the best-case scenario for open scholarship in general could be potentially harmful for Impactstory's business model. If every DOI were to resolve into an open access version, the Unpaywall Data Feed would be superfluous for new research. But the services would still be needed for clearing back catalogues and, as long as there are policies requiring green open access (via self-archiving), our services will definitely remain in demand."

With the number of active users continually increasing (about 1,000 new users a week), a high scoring browser extension, and a lot of positive feedback via e-mail and on social media, Impactstory is proving that the services it provides offer real added value to researchers. Because of their highly personal marketing approach, Piwowar and Priem follow up closely on interesting uses – and are always happy to get detailed feedback on how their services are being implemented.

A bigger worry is that the competition might reuse the Unpaywall code base to create their own product – which is technically possible because the code licence allows for commercial reuse: "It is a risk we run because of our liberal licensing policy. A big company might consider running it in-house, but I don't think they can do it as efficiently and cheaply as us, at least not in the US. A small company might try to offer the same service as we do. However, a competitor would still need to convince our big clients that they'd do a better job of it than us and we really have a headstart on our competitors. Changing our licence to remove this risk entirely could be an option but for now we've decided against it, mainly because non-commercial licences are notoriously difficult to enforce and non-compliance is difficult to detect."

"Non-commercial licences are notoriously difficult to enforce and non-compliance is difficult to detect."

## References and relevant links

- Impactstory : **impactstory.org**
- Sloan Foundation: **sloan.org**
- Duke University: **duke.edu**
- Depsy: **depsy.org**
- University of Texas at Austin collaboration: **ischool.utexas.edu/tags/impactstory**
- CORE: **core.ac.uk**
- Open Access Button: **openaccessbutton.org**
- OpenAIRE: **openaire.eu**
- BASE: **base-search.net**
- PubMed: **ncbi.nlm.nih.gov/pubmed**
- ArXiv: **arxiv.org**
- OAI-PMH: **openarchives.org/pmh**
- Crossref: **crossref.org**
- DOAJ: **doaj.org**
- DOI: **doi.org**
- Arcadia Fund: **arcadiafund.org.uk**

# About Heather Piwowar
## Co-founder

Dr Heather Piwowar is a co-founder of Impactstory, the non-profit company behind the Unpaywall database for open access discovery. A longtime advocate for open science, Dr Piwowar is also a leading researcher in research data availability and reuse, including a seminal paper measuring the citation benefit of publicly available research data. Dr Piwowar has a bachelor's and a master's degree from Massachusetts Institute of Technology (MIT) in electrical engineering, ten years of experience as a software engineer, and a PhD in biomedical informatics.