SPARC
Europe

10 Key Interviews

INSIGHTS INTO THE SUSTAINABILITY OF OPEN INFRASTRUCTURE SERVICES

An Invest in Open Infrastructure project
More info: www.sparceurope.org/ioiinterviews

ABOUT THE PROJECT

Gleaning Insights

Research and scholarship is underpinned by a variety of tools, technologies and services ranging from for-profit commercial solutions and offerings from vendors to community-owned, open technologies and infrastructure. We often hear about the challenges for open infrastructure tools and services to scale, maintain, and compete in the broader market.

The 10 interviews comprised in this project highlight some of the key decision-making points, funding mechanisms and models, and other learnings from a series of commonly used services and technologies used to support research and scholarship. These include both for-profit and not-for-profit services, highlighting perspectives on sustainability across the sector.

This work is supported by Open Society Foundations and SPARC Europe, in collaboration with Invest in Open Infrastructure.

An Invest in Open Infrastructure project
More info: www.sparceurope.org/ioiinterviews

WHO WAS INTERVIEWED?

Overview

Figshare *Featured in this document.*
Code Ocean
Dryad
EDP Sciences
F1000 Research
Mendeley
Our Research
arXiv
Redalyc
4TU.Research.Data

"Sustainability is the most important goal for academic infrastructure"

Mark Hahnel, CEO and founder of Figshare, UK

FIGSHARE

Mark Hahnel, CEO and founder of Figshare, UK

At a Glance

Type of activities:
Repository services for funders, institutions and publishers, exploring the next phase of data publishing

Life-cycle stage:
Founded in 2011, now a small-medium enterprise, sustainable in financial sense and customer retention, growth/scaling phase and exploring new services

Current legal structure and funding model:
Owned by Digital Science, charging customers for the services provided

Technology:
Proprietary, considering the pros and cons of now going open source

Sustainability

'There are two levels of sustainability. First, there is the technical sustainability in the field that we work in, including policy and compliance. The second level of sustainability is financial: do you bring in more money than you spend? There are some fantastic examples of open infrastructure being hugely viable and successful. But everybody has to make sustainability trade-offs. I feel there are various sustainable models, whether open source or in other ways.'

Piece of Advice

'Financial support at the beginning of an initiative is not enough to become sustainable. If Figshare would have received a 100k grant at the beginning and was told, "Off you go, make yourself sustainable," I don't think we'd have got there. Understanding what you're comfortable with, and plotting out what the future will hold, helps you make sensible decisions..'

An Invest in Open Infrastructure project
More info: www.sparceurope.org/ioiinterviews

FIGSHARE

Figshare was founded by Mark Hahnel, a former life sciences PhD, to address his wish to publish his research results in formats other than PDF. It received investment from Digital Science and is now a Digital Science portfolio company.

Original vision

Mark Hahnel was doing his PhD in life sciences when he ran into a problem: he wanted to publish his research results in multiple formats, but found out that the files were too big to be accepted by traditional journals. 'This made it very hard for me to tell the story of the research I was doing. I wanted to be able to publish my videos of stem cells moving from one side of the screen to the other. So, I started to publish my own data, giving things persistent identifiers and using Creative Commons licensing and technology that was available at the time, like cloud computing.' Hahnel created a regular limited company in the UK in his own name, because 'I thought it was the right thing to do.'

At the time, he envisioned a Github 'freemium' model for the initiative, where its main features would be free and subscribers would pay a monthly fee for extras. That's not how it happened, however.

Growth and sustainability challenges

When the Figshare idea took shape Hahnel was approached by the Sloan Foundation and the Wellcome Trust. These conversations did not lead to funding because of location and academic seniority requirements for grants, that he did not have as a PhD student. 'Digital Science had just opened up down the road as a traditional incubator/investor, and I was invited to pitch them. That was actually my only option at the time.'

In September 2011, Digital Science stepped in and allowed Hahnel to develop his idea by providing him with a grant that resembles today's Digital Science catalyst grants: 'They said: we'll give you some money for Figshare, and we'll give you a full-time job for a year to work on Figshare – but we'll take no equity. We'll see where we are twelve months from now.' It was in all essence a grant.

After the first year was over, Digital Science invested in Figshare and took an equity share. The company turned into an LLP (Limited Liability Partnership), which then became fully owned as a Digital Science portfolio company. Hahnel has always been CEO.

Financial support at the beginning of an initiative is not enough to become sustainable. More help is needed for founders in that respect. 'If Figshare would have received a 100k grant at the beginning and was told, "Off you go, make yourself sustainable," I don't think we'd have got there.' Hahnel recommends that support for newcomers take the entire picture into account, like complete start-up support, focusing on what can help organisations as opposed to what should be the goal of organisations.

Opportunities, considerations and choices

Hahnel realised later that the start-up support Digital Science provided was crucial. 'I was a life sciences PhD student; I had never gone to a business class in my life. I experienced a very steep learning curve in understanding what 'EBITDA' was (Earnings Before Interest, Taxes, Depreciation, and Amortisation), that people's

EVERYBODY HAS TO MAKE SUSTAINABILITY TRADE-OFFS.

salaries are not the full cost of people, and the business of marketing. Basic accounting was also something I didn't know anything about. As an incubator/investor, Digital Science allows you to make use of their central departments (legal, finance, marketing, etc.) for free, until you are big enough to hire your own people.'

As a company, Figshare focused on financial sustainability from the very beginning. 'I feel that there are two levels of sustainability, which a lot of grant-funded projects may not take into consideration in the same way. First, there is the technical sustainability in the field that we work in, including policy and compliance. We preserve data and research outputs in the long term.

In terms of processes, we are ISO27001 certified to make sure we're operating in a way that is sustainable. For example, we had a pandemic recovery plan years ago, and I thought it was a joke, that we'd never need it. And here we are. Because we're a commercial company we've been forced to jump through a lot of hoops around policy compliance set by our clients such as the data needing to live in the country of the client.' The organisations that we work with have requirements that contracts over a certain value have to go to tender, particularly in the EU, the results of

which are made openly available. So institutions are confident they are getting the best solution against their requirements and there is a transparent and fair playing field. Handing out contracts on an ideological basis has inherent risks for the entire ecosystem.

'The second level of sustainability is financial: do you bring in more money than you spend? As a commercial company with a finite amount of money available, we've had to get to a financially sustainable level. If you don't, the people employed will lose their jobs and won't be able to pay their mortgages. So it gets very real, very fast.'

Consequences of current funding model

Hahnel had a strong incentive to reach financial sustainability as soon as possible. 'Like any project that needs funding, I could never guarantee that the next investment would come. It would also be naive to assume that the central services offered by Digital Science would continue to be provided for free. So there was always the push: what are our sustainability models here? We're now at a point with Figshare where we have hundreds of clients, who are probably more resilient than one funder. There is a chance that some of our customers will cancel their annual contract, but

hundreds of them cancelling at once is less likely. So we're less dependent, and the business model we set up provided a good diversity of financial sustainability, which is great.' Digital Science invests in companies with a view to buy, and thus take 100% of the equity. But Hahnel is not worried about losing control of Figshare. 'A lot of people are concerned about changes of ownership and changes of mentality. But it would be strange to change a winning formula. I think, well, I've worked at Digital Science for ten years. I'd be shocked if they'd suddenly change their plan that has been successful so far.' Moreover, he counts on the contracts he is signing. 'I always tell customers: "Don't trust me; trust the contract that you're signing."'

'In terms of negative consequences of the commercial funding model with Digital Science, we get two kinds of criticisms: first, that the business is not community-owned, and second and relatedly, that we are not open source.'

Hahnel answers the first concern by stating that Figshare has always aimed to add more value than it extracts. He also defends the company's status. 'We get asked: "How do we know you're not going to lock away content, or charge us for it?" I then ask: "What are the pros

of community-owned, versus the cons?" The fact that we're built on open principles helps us here, anyone can pull any and all of the content via our open APIs. Published data are licensed CC-0 or CC-BY on the free version of Figshare. We could have given people more restrictive licenses (like others do), but we made the decision not to do it, because we don't want to give academics the option to restrict access to their research more than they need to.'

Hahnel invites us to take a nuanced perspective instead of seeing things as mutually exclusive or as clearly defined opposites. 'We are commercial and have an open science platform that is sustainable, meaning it is free to use, anyone can come along anytime and use it. But we are not an open source platform. It is good that this investigation around sustainability models does not involve just open source software. I feel you can have sustainable models around open infrastructure, whether open source or open in other ways.'

He addresses the second concern by pointing to the risks for a young and small company to go open source too early. 'In the early years of our ten-year existence, I was concerned that a large commercial publisher or technology company would build a clone of us. If we would have been open source then, I believe

bigger organisations with large, organised sales teams may have put us out of business. That was one of the early reasons why we didn't open source the platform at the time.'

Future vision for sustainability

'Before COVID, I would have put us in a growth and scaling phase: work with more organisations, bring in more revenue, and hire more developers to build better products and features to provide policy-compliant infrastructure for organisations. We are now exploring the next phase of data publishing, working with organisations with data-publishing needs such as the National Institute of Health to check the metadata for each published dataset; for more, see (https://figshare.com/blog/Why_fast_but_good_publishing_matters/587). It is very relevant that we're moving from how to publish easily, no matter what content, to how to publish fast but with good quality. COVID has shone more attention on this issue. We're looking at new models around that too.'

The idea is just to keep building, so our clients have the best repository, making use of the best modern technologies to allow universities and funders to have some enterprise-level software taking care of some of their core functionality.'

Hahnel says that there are good reasons to go open source, but that there are some challenges as well. 'We'd love to open source the platform. We've built a lot of cool stuff that could help sustain a lot of other tools. We also see benefits in other people building on top of our stack.' In terms of affecting our sustainability model, we know that our infrastructure stack is so big that if a university wanted to implement their full stack, a fully working Figshare based repository themselves, it would cost them three times as much as what we sell them a license for. It is a bad business decision to take our code base and implement it yourself. I'm sure some people would do it, and there would be some effect, but I don't think it would affect our sustainability.'

'The question is, why don't we do it now? Because it would take six months of development work with our current team, so if someone would give us a million dollars to do it, we would. But in those six months, all our clients who got existing feature requests or bugs will get ignored.

Advice for peers

Finally, Hahnel advises peers to be diverse in their financial stability and not rely on one grant that may not be renewed. He also encourages everyone to not have a strong binary view on the commercial/non-commercial divide. 'There are some fantastic examples of open infrastructure

being hugely viable and successful. But everybody has to make sustainability trade-offs.'

'For instance, Firefox Mozilla is basically run by the amount of money they get from Google or Bing to be on their homepage. I think understanding what you're comfortable with, and plotting out what the future will hold, helps you make sensible decisions. In the Figshare case, I was adamant that we offer a free at point of use and free at point of consumption – platform, for all academics. Lots of people think that every decision at a commercial company is made of the basis of 'is this profitable' this simply isn't true, and you can look at the free Figshare.com as an example of that. Whether commercial or open source you need to make sure you are sustainable and being profitable in some areas gives you the freedom to experiment and give back to the community in others.

'All academic infrastructure should be inter-changeable. The core problem for a lot of academic projects is sustainability, membership fatigue, and these types of things. So, if we can focus on sustainability models that work, the question then for me is: how do you get all these systems to be good actors, for instance, that people will not limit their APIs?'