UKSG

Social infrastructures in research communication: a personal view of the FORCE11 story

There are a wide range of community organizations and projects that aim to support scholarly communications in one way or another. Scholarly societies are some of the longest standing and, as research communication has changed, new groups, organizations and projects appear – to solve problems and fill gaps. Amongst these organizations, FORCE11 has acted as a nucleus, support platform and convening point for community efforts on issues as wide-ranging as identifiers for research materials, implementing data and software citation systems, and on what we might mean when we talk about building a 'scholarly commons'. I have personally been involved in FORCE11, attending the meetings that led to its formation, as a founding board member of the formal organization and, most recently, serving as President of the Board up until the end of 2017. In this article I give an entirely personal view, drawing on my perspective of those experiences and what it can do to illuminate the roles and interactions of the many organizations seeking to support change in research communications.

Background: a proliferation of efforts and organizations

It is a lazy truism to say that scholarly communications is in flux. How historically accurate the view that recent change is more profound or more rapid is perhaps an open question, but the perception that we are living through a period of unprecedented rapid change is widely held. The nature of this change means different things for different people. For some, it is a threat to cherished ways of working and communication, for others, an opportunity to do things radically differently, or perhaps just a bit better, or fix one specific problem.

CAMERON NEYLON

Immediate Past President, FORCE11
Curtin University
Perth

Photo: Elena Giglia
CC BY 2.0 via
Wikimedia Commons

For innovators, there seems to be a moment in which it is possible to build entirely new technical systems, small or large, that use the web to reconfigure how we work. For publishers, there are threats to existing business assumptions and also opportunities to branch out; for libraries, a similar set of threats and opportunities. For those with a concern about a specific problem, it is a moment when that specific thing can be fixed. For those with dedication to ensuring some part of our communications process is done properly and to high standards, it is a moment when those core values must be defended.

Out of this diversity, a wide range of different priorities naturally emerge. Radicals will differ from conservatives; different stakeholders from each other; those concerned with maintaining the qualities of peer review from those concerned with improving data management. From the ferment, it is natural that different groups emerge to tackle the particular problems that they each agree internally are important. What is equally natural, but perhaps less obvious, is that the forms that these efforts take will mirror the ways of working and assumptions that these different groups bring from their home turf.

> 'Out of this diversity, a wide range of different priorities naturally emerge'

For example, scholarly societies, as the longest-standing players with a disciplinary focus, work in particular ways. They are by nature conservative, so efforts for change usually arise

from the radicals within the disciplinary community. They form internal and ad hoc groups that advocate for change, often with a focus on the publishing operations of the society. Sometimes, society efforts arise from an external threat or a crisis in quality control, but they are generally introspective and rarely look outside to parallel efforts in other disciplines. This is the natural result of a focus on the specific disciplinary community. Conservatism is a natural position to take for communities that have significant history.

The Research Data Alliance is an international collaboration and an organization with a mission to 'build … the social and technical bridges that enable open sharing of data'.[1] Its structure and efforts, with working groups being co-ordinated through a central secretariat and regular plenary meetings that bring these groups together, show the traces of its origins in standards organizations. This in turn can be traced to its founding community, who were (and are) powerful players in the building and provision of technical systems for the research community.

> 'Conservatism is a natural position to take for communities that have significant history'

By contrast, the Open Scholarship Initiative,[2] which aims to chart a route towards similar goals, was founded with support from UNESCO and has at its centre staff with experience of intergovernmental treaty negotiations. It therefore takes a form that seeks a kind of Kyoto Agreement for scholarly communications, aiming to bring leadership from stakeholder groups together to negotiate a settlement on how to move forward to a networked, open future for research communications.

By nature, more radical voices tend not to find an easy place to sit in these kinds of formal structures. For a range of reasons, some structural, and some rooted in the history and politics of the technological developments that underpin the web and 'open' agendas, radicals tend to come together in groups that are more ad hoc and experimental. Such groups may form to 'try something out' – though rarely building anything that might be described as a viable product. They will often be small groups focused on very particular efforts targeting fundamental or radical shifts in practice or technology. By both nature and inclination, these groups are often less organized and less plugged into more traditional and organized community-wide efforts.

A potted (and very partial) history of FORCE11

2010–12: roots in radical advocacy

One such group of radicals, with a mission to topple the PDF from its central role in online research communication, came together in 2011 in San Diego. The 'Beyond the PDF' meeting,[3] funded by the Gordon and Betty Moore Foundation, was convened by Phil Bourne with the aim of 'moving digital research communications beyond putting digital paper online'. In parallel, Anita de Waard, with Ed Hovy, Tim Clark, Paul Groth, David Shotton and Ivan Herman, had been working on convening a Dagstuhl symposium under the modest title of 'The Future of Research Communication'.[4]

The agenda for both meetings was radical, ripping up the current format and form of online research communications (at least in the sciences) to move it forward into the sunny uplands of online networked knowledge. The FORCE11 Manifesto that came out of the Dagstuhl meeting[5] in particular charts a vision of how the web could offer both small-scale and radical change to improve the systems through which we communicate research. At the critical distance of half a decade, it is worth noting that the Manifesto did not grapple with issues of structural power, diversity or geographical inclusion, assuming as many of us did at that time, that the utopian vision it outlined would solve all of those problems.

> 'a group of radicals with a radical vision for change'

This, then, was a group of radicals with a radical vision for change. The serendipitous collision of the two separate but parallel meeting organization processes made it an unusually eclectic group. Phil Bourne was at that time Associate Vice Chancellor for Industry and Innovation at the University of California, San Diego, had been deeply involved in the

founding and development of PLOS, and was a key player in the history of bioinformatics and biomedical computing. He later moved on to the role of Associate Director for Data Science at the National Institutes of Health (NIH) and is now Director of the Data Science Institute at the University of Virginia. Anita de Waard was then Director of Disruptive Innovation (now Vice President, Research Data Collaborations) at Elsevier. Others were academics, technologists, publishers and librarians. It was an unusually disparate group of people who were well connected in their own spaces. Overall, they were perhaps *influential* without necessarily being particularly *powerful* (at least at that stage).

A core group of people continued to meet by teleconference after the meetings. The first focus was to complete the drafting of the Manifesto, and then later to consider how to take it forward. Looking back, many aspects of the Manifesto have seen adoption in one form or another, but this seems to be less due to specific projects to deliver them and more due to a slow process of reinforcement and repetition. In addition, occasional (sometimes seemingly regular) systemic scandals or crises have raised the profile of data sharing, peer-review practices, more dynamic and networked research objects and a growing concern with credit being received where due. While the Manifesto did have an important role to play, it was not a simple process of implementing a specific vision that ended up being the way that FORCE11 led to change.

Community building through meetings

I cannot recall the details of the discussion that led to a decision to hold a 'Beyond the PDF2' meeting but I am pretty sure it went something along the lines of '… Wouldn't it be fun to get everyone back together …'. Paul Groth, then at Vrije Universiteit Amsterdam and now working (fittingly with Anita de Waard's old job title of Director of Disruptive Innovation) at Elsevier, stepped up to organize the conference in Amsterdam.

With some – somewhat confusing – name changes along the way, the Amsterdam meeting set a pattern to be continued first in Oxford with FORCE2015, Portland for FORCE2016 and Berlin in 2017. Conscious efforts to expand the community reach started with a focus on engaging researchers from the humanities in Amsterdam. This was followed by the development of a travel fellows programme championed by Melissa Haendel (Oregon Health Sciences University, who also led the Portland FORCE2016 conference) that has brought over 100 fellows from more than 40 countries to FORCE meetings, with core funding from the Alfred P Sloan Foundation and additional support from the National Science Foundations and NIH.

The FORCE conference has become an important part of the calendar for many different communities, and is unusual in the diversity of different groups that it brings together. That diversity is always a work in progress, of course, and it remains an issue that brings criticism of what still has not been done, or has not been done well enough. However, slow progress is being made. As a result, the conferences are different for many attendees from their 'home' meetings, whether those are stakeholder, technology, or disciplinary in nature. As the chairs of those conferences will attest, this makes them very challenging to organize. Very different expectations of what a conference should do, of what justifies attendance, of when programmes are fixed and who will be there are a challenge to manage. The magic that happens when that mix of expectations is pulled together makes it worth it, however.

Consensus principles and implementation

One of the most important outcomes, for FORCE11, of the Beyond the PDF2 in Amsterdam came out of exactly that ferment. An ad hoc session was convened to discuss data citation, and out of that came a draft set of principles for data citation. The response to this, both positive and negative, was incredibly important for the future development of FORCE11 activities. First, the positive responses showed that this was an issue that people cared about, but in some ways more important were the responses that could be summarized as, 'but we're already working on that', or, 'who are these people to articulate principles'.

As I noted at the beginning, this is not exactly unexpected. Different groups, who are not necessarily in contact, may see similar problems and seek to solve them in their own ways. What often happens is confusion, sometimes bogging down progress, sometimes with one effort 'winning' against others. But what happened here was different. The initial stimulus provoked many different groups to come together, and an effort was made to develop a set of consensus principles.[6,7] These were high-level principles, not implementation details, which helped a very wide-ranging group agree them. Again, the theme emerges of bringing together different perspectives to find a common ground. It was not necessarily planned to work in the way that it did but, having seen the pattern, efforts to bring broader groups together became part of the way that FORCE11 operates.

With the principles agreed, NIH funding was successfully obtained to move the work from consensus principles towards implementation. The Data Citation Implementation Pilot (DCIP) brought together repositories, publishers, data providers, standards developers and technologists to work towards developing common data citation protocols that are now being implemented. Convening as broad a group as possible, reaching consensus on principles, piloting implementation plans and then passing that on to the right stakeholders to roll out, became a pattern.

> 'efforts to bring broader groups together became part of the way that FORCE11 operates'

Following on from data citation, software citation is now being worked on. FORCE11 provided a platform for working on and then disseminating the FAIR Data Principles. Working groups that are focused on research material identifiers and, more recently, on trying to wrangle the ideas behind what it means to have a 'scholarly commons', have moved their own agendas forward. Many efforts fizzle, and it is not always the case that agreement is reached, but a pattern has developed that seems to work well.

2017 and onwards: training and platform building?

In 2017 the first FORCE11 Scholarly Communications Institute (FSCI) was convened at UC San Diego. Conceived as a five-day intensive training 'summer camp', this represents a new activity and evolution of what FORCE11 does. Drawing inspiration from the Digital Humanities Summer Institute run in Victoria, British Columbia each year, the concept behind FSCI can be summed up in the hypothesis that learning with others from different backgrounds will radically enhance the value of training and skills development in scholarly communications. That hypothesis seemed to be strongly validated, although again, there is much that can be done to improve the experience. Courses were submitted from around the world and on topics from reproducibility to network analysis and from the humanities to the sciences.

FSCI for me represents a natural evolution beyond the implementation phase of work we have seen towards dissemination. Our space is littered with white papers and standards, the means of formal dissemination. There is a huge amount of expertise but there appeared to be a gap in opportunities for training librarians, researchers, publishers and technologists in applying and using the tools and systems available. More than that, training these disparate groups *together* so they begin to appreciate the different experiences and perspectives is incredibly valuable. Whether at the conference, at FSCI or in the working groups, the most important progress often seems to come when people who would naturally assume themselves to be enemies in one of our many political battles find common ground.

> 'people who would naturally assume themselves to be enemies … find common ground'

The 'FORCE11 way': what is uniquely valuable about this organization?

Many years ago John Wilbanks posed a question that has stuck in my head ever since: how does a mission-driven organization know when it has finished? Because a mission should be achievable. Given the issues I raised at the beginning about the proliferation of organizations, there is a related question: do we really need another organization? What is unique about FORCE11, and what is therefore worth working to sustain?

It was an observation by Maryann Martone, a founding board member, Executive Director and later President of the Board, that led me to my view. She noted that, while FORCE11 started as a group with an agenda for change, many of the most valuable impacts had arisen from activities that bubbled up from within groups of attendees, such as the Data Citation Implementation Principles from an ad hoc session, and the Scholarly Commons work from a question posed on Twitter by a conference attendee. Beyond that were all the small-scale interactions and efforts that arose from chance meetings.

For me, there are two things that mark FORCE11 out as different. First, that it is not a community, but a place where different communities come together. Second, that by design it aims to use that meeting place to support convergence and bottom-up identification of what is worth working on together. At its best, it is a platform where those different perspectives can come together and find common ground to work towards common problems. And when it is not at its best, it continues to surface the fact that we have not built a good enough platform – yet – to allow the full diversity of voices to be heard and appreciated. The pattern of broad convening, seeking consensus principles, and only then moving to implementation, was never designed but in its outlines seems a valuable way to identify where useful work can be done.

There are other important, but less visible, aspects of the way that FORCE11 operates that reinforce this. It provides a platform for people to come together and work, but does not seek ownership of the outputs of that work. They are the 'Data Citation Implementation Principles' or the 'FAIR Data Principles', without the FORCE11 name appended at the front. The Data Citation effort is as much connected with work by the RDA, but all of these outputs belong to the community that created them, not the organization.

FORCE11 never endorses statements or positions of others, including that of its own working groups. This piece is my personal vision and opinion, not that of the organization. FORCE11 seeks to be a neutral platform (as far as that is possible) for the communities of interest to come together. Some stakeholders may disagree with the statements of the Scholarly Commons Working Group, or of the implementation details proposed by the Software Citation Working Group, but those are issues to address to those groups. FORCE11 aims to provide a platform where those criticisms can be brought to the table, but not to adjudicate.

> 'FORCE11 seeks to be a neutral platform'

Many other organizations, such as scholarly societies, have a disciplinary focus. FORCE11 seeks to work across disciplines. Other organizations have a particular sector focus or target, for example, RDA works to an agenda of advancing data sharing and OSI works towards negotiation of a settlement between parties to deliver open access. The Committee on Data of the International Council for Science (CODATA), the International Council for Scientific and Technical Information (ICSTI), the Confederation of Open Access Repositories (COAR), the Scholarly Publishing and Academic Resources Coalition (SPARC), the Open Research Funders Group and others have particular roles to play in governance or community co-ordination within specific spaces, or towards particular, well-defined goals. Beyond a shared, but quite vague, notion of using digital and networked technologies better, FORCE11 seeks ways to support the ideas and projects that can bubble up when disparate groups of people come together. My view of what makes FORCE11 unique and valuable is that it is a kind of social infrastructure that can support this.

The challenge of sustainability – social infrastructures and funding

If you accept my notion of FORCE11 as a form of infrastructure and know anything about my own work over the past decade, it will be clear that sustainability is a central issue. In a world with limited (and currently seemingly diminishing) financial resources, but also one in which we expect many different efforts with different emphases to arise, how do we first find resources that can support sustainability in the long term, but also find the right distribution between different efforts with different aims?

FORCE11 has done well at funding particular activities and particular projects. The DCIP was supported by the NIH, Scholarly Commons by the Helmsley Foundation. The Sloan and Moore Foundations have supported many meetings and projects. If we look to the other comparator organizations that I have mentioned, the RDA and OSI have both been largely supported by grants for projects and some sponsorship. More generally, funders are starting to take seriously the different qualities of funding required to support technical infrastructures,[8,9] but sustaining shared social platforms, social institutions that mediate difficult conversations, lags far behind.

This idea, that we need platforms that first bring people together and then create the right conditions for conversation, developing consensus and then working towards implementation, seems to be difficult to explain. Everyone assumes that just happens, or perhaps that particular individuals make it happen. We look to success stories where through luck and the hard work of specific individuals a community has come together organically to solve a problem. However, we do not examine how many efforts have failed, or never even started. Each new project has to replicate this social infrastructure for itself. How much extra work was required to put together a shared system like ORCID because the platform to bring people together had to be invented over and over again? One of my frustrations has been that funding to collate what we have learned and turn it into systems, guidelines and templates that could be used by others has been difficult or impossible to find.

An ecosystem of platforms

It should by this point be evident that I believe that there is an important role for an organization or organizations like FORCE11 that can provide the glue and support for communities to come together and find common ground for experimentation and innovation. My perspectives and interests lead me to those intersections and the question of how to support them, just as those with different histories and experiences might naturally prefer to work in the different ways that other organizations in the space have developed. However, I do not have any simple answers for how we should choose to allocate scarce resources amongst different community efforts and support platforms. But what I will observe is that it is the underpinning support work that is most often invisible and hardest to sustain. When there is under-resourcing, things do not necessarily go badly wrong, rather they just tend not to happen: voices are not heard, beyond those who already have a platform and potential innovations are missed because the ideas are not obviously fundable as projects or start-ups.

> 'the glue and support for communities to come together and find common ground'

I would make no claim that FORCE11 is perfect at what it does, nor that there are not others in our space making similar efforts. What I will say is that top-down initiatives led by people who are already well known are easier to resource than platforms to support innovation initiated by those we do not know, or are designed to help us to solve problems that we have not yet realized we have.

Thinking about how to best support an ecosystem that includes those platforms is in the best interests of all of us.

References

1. About RDA:
    https://rd-alliance.org/about-rda (accessed 2 February 2018).

2. About OSI:
    http://osinitiative.org/about-osi/ (accessed 2 February 2018).

3. Beyond the PDF:
    https://www.force11.org/meetings/beyond-pdf-1 (accessed 2 February 2018).

4. FORC Workshop 2011 Dagstuhl:
    https://www.force11.org/meetings/forc-workshop-2011-dagstuhl (accessed 2 February 2018).

5. Bourne P E et al., *Improving Future Research Communication and e-Scholarship (The FORCE11 Manifesto)*, 2011, FORCE11:
    https://www.force11.org/about/manifesto (accessed 2 February 2018).

6. Martone M et al., *Data Citation Synthesis Group: Joint Declaration of Data Citation Principles*, 2014, San Diego CA, FORCE11; DOI:
    https://doi.org/10.25490/a97f-egyk (accessed 2 February 2018).

7. Starr J, Castro E, Crosas M, Dumontier M, Downs R R, Duerr R, Haak L L, Haendel M, Herman I, Hodson S, Hourclé J, Kratz J E, Lin J, Nielsen L H, Nurnberger A, Proell S, Rauber A, Sacchi S, Smith A, Taylor M and Clark T, Achieving human and machine accessibility of cited data in scholarly publications, *PeerJ Computer Science*, 2015, 1: e1; DOI:
    https://doi.org/10.7717/peerj-cs.1 (accessed 2 February 2018).

8. Anderson W et al., Towards coordinated international support of core data resources for the life sciences, *bioRxiv*, 2017; DOI:
    https://doi.org/10.1101/110825 (accessed 2 February 2018).

9. OECD, *Business models for sustainable research data repositories*, Paris, OECD; DOI:
    https://doi.org/10.1787/23074957 (accessed 30 December 2017).

Cameron Neylon
Professor of Research and Communications – President, FORCE11
Centre for Culture and Technology
Curtin University Perth
Western Australia
E-mail: cn@cameronneylon.net

ORCID ID: http://orcid.org/0000-0002-0068-716X