Building community: the first 5 years of DataCite

Frauke Ziedorn

Blog post published April 9, 2019 in *DataCite Blog - DataCite*
https://doi.org/10.5438/edb7-0r15

Abstract

As part of our 10-year anniversary, we want to tell you the story of how DataCite was founded 10 years ago. Therefore, we approached several people 'who were there' to tell you their part of the story. This guest blog post is by Frauke Ziedorn, who was TIB's DOI Service Manager from 2010 until 2015 and the first co-chair of DataCite's Metadata Working Group. I came to the TIB and DataCite in 2010.

Keywords

Information Systems (Subfield), Anniversary

Recommended citation

Ziedorn, F. (2019, April 9). Building community: the first 5 years of DataCite. *DataCite Blog - DataCite*. https://doi.org/10.5438/edb7-0r15

Copyright

Copyright © Frauke Ziedorn 2019. Distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

As part of our 10-year anniversary, we want to tell you the story of how DataCite was founded 10 years ago. Therefore, we approached several people 'who were there' to tell you their part of the story. This guest blog post is by Frauke Ziedorn, who was TIB's DOI Service Manager from 2010 until 2015 and the first co-chair of DataCite's Metadata Working Group.

I came to the TIB and DataCite in 2010. Being the managing agency of DataCite, the TIB was the organizational hub of DataCite, providing not only a lot of man-power but also technical and organizational resources. Working for DataCite in the early years was a great adventure. We were a close and growing community expanding fast from the initial seven members to 21 members by the end of 2015. In this blog post, I will describe the work we did during the first five years and the working groups we started that allowed us to build the community.

Image

At the beginning, every member committed to contributing to build DataCite's services – a commitment without which it would have been impossible to grow DataCite. Although members were situated all around the world, everybody knew everyone and we worked as a motivated and close team. Our annual General Assembly and the Summer Meetings – yearly conferences to connect the research data community – were welcome opportunities to meet each other in person. We created working groups dedicated to different areas of DataCite's business and services. The most important and persistent ones being the Metadata Working Group, the Policy and Best Practices Working Group, and, emerging from the Technical Infrastructure Working Group, the TechTeam. From the beginning, the feedback of users of DataCite's services gave valuable input in all areas. Some organizations, such as PANGAEA, also set best practices for data citation that are still being used today.

The Policy and Best Practices Working Group

This working group started as "Business Practices WG" with the goal to develop best business practices and articulate DataCite's policy. The WG published a "Business Models Principles" document in 2012 with requirements and expectations for DataCite members and their clients, the data centers.

The scope of the WG changed afterwards and it was renamed "Policy and Best Practices Working Group", which investigated and documented best practices related to the business operations of DataCite's members. The WG developed and implemented policies and best practices, monitored new developments, initiatives and norms, supported evaluation and implementation of new strategies, and proposed new initiatives for further development by DataCite.

The Metadata Working Group

The Metadata Working Group is still active today. The Metadata Schema is the backbone of most of DataCite's services and has become a widely known standard. We started the group in 2010 with the objective to create a metadata schema that was as compact as possible, could be used by all kinds of scientific disciplines and still covered the most important elements to describe data. We based the schema on Dublin Core but expanded it to be a better fit for research data. The group

held monthly virtual meetings discussing what elements should be included in the schema and how to define them. In between meetings, specific questions were investigated by sub-groups or individuals, until the first version of the DataCite Metadata Schema was released in January 2011. It consisted of 17 elements, of which only 5 were mandatory. We reached our goal to keep it lean and inter-disciplinary, and it was well-received by the metadata community. The schema was adopted by the data centers that registered DOIs with DataCite, and we got a lot of feedback from them – a valuable source of input to this day.

The TechTeam

The Tech Team consisted mostly of two developers, Sebastian Peters at the TIB and Ed Zukowsky at the British Library. They developed the first version of the Metadata Store (MDS) which was in use until last year. They added a lot of basic and useful services built upon the MDS and the metadata associated with DOIs, such as the search portal, OAI-PMH metadata harvesting, statistics, and link redirection to allow computers to directly access the data behind the DOIs instead of the metadata. They also started the collaboration with Crossref. One of the first services that came out of this collaboration was the DOI Citation Formatter, which allows users to extract metadata automatically from a DOI and build a full citation. Many of their services are still in use today or have become the basis for more modernized versions.

Over time, DataCite became more professional and started to employ its own staff, which helped it grow and incorporate new services like re3data.org or Make Data Count. It was great to help build this amazing organization during its first five years and I wish DataCite and its staff a lot of success for the future!

The post Building community: the first 5 years of DataCite appeared first on DataCite.