

DataCite: visit their new blog

Martin Fenner, Gobbledygook

January 14, 2011

DataCite is an international consortium for data citation. DataCite originally started as a project at the German National Library of Science and Technology (TIB).

Since 2005 the TIB was providing digital object identifiers (DOIs) to research datasets. In December 2009 research libraries and technical information centres from 6 countries founded the DataCite initiative. In December 2010 DataCite reach an important milestone by registering the 1,000,000th DOI name.

Many people are familiar with CrossRef which is providing DOIs for research papers. CrossRef is operated by Publishers International Linking Association (PILA), a non-profit organization formed by scholarly publishers in 2000. Think of DataCite as the counterpart that is providing DOIs for research data. And as research data are typically stored in data centers associated with technical information centers - e.g. the British Library or the Canada Institute for Scientific and Technical Information (CISTI) - and large research libraries (e.g. California Digital Library), DataCite members come from the academic community.

The TIB is located in Hannover, just a few miles away from Hannover Medical School where I work. In December I sat down with Jan Brase from the TIB, one of the driving forces behind DataCite and on its Board of Directors. We talked not only about DataCite, but also how DataCite could interoperate with ORCID, the unique author identifier initiative (where I am a member of the Board). One of the aims of ORCID is to create a permanent, clear and unambiguous record of scholarly communication, and this of course includes not only publications, but also the contribution of research datasets. Gudmundur Thorisson gave a nice presentation about the possible integration of DataCite and ORCID in September.

Earlier this week, Jan presented about DataCite at the Academic Publishing in Europe conference in Berlin. His slides are available on SlideShare, the video of his talk (like all the other presentations) will soon be available here.

Also this week Jan started the DataCite blog. In his first post (after the obligatory welcome post) Jan talks about the plans DataCite has for 2011. I am particularly

excited about the central metadata repository that should be up and running by June. DataCite has allowed Thomson Reuters to crawl the repository so that the metadata will also appear in the Web of Science. A lot of interesting stuff is going to happen in 2011 around unique identifiers for researchers and their scholarly works.