

PIDapalooza Greetings! from Maria Johnsson

About the conference

On 23-24 January I participated in the conference PIDapalooza in Girona, Spain. This conference is solely dedicated to the technologies and services around persistent identifiers. The main organizations behind it were: ORCID, CrossRef, DataCite and Digital Library of California, and they also contributed a lot to the content of the program. This year the conference was characterized by short speeches of 30 minutes each, and which were supposed to invite to discussions and feedback from the audience. It was a small event with about 60-70 participants in total and with a very relaxed and good ambiance during the whole conference.

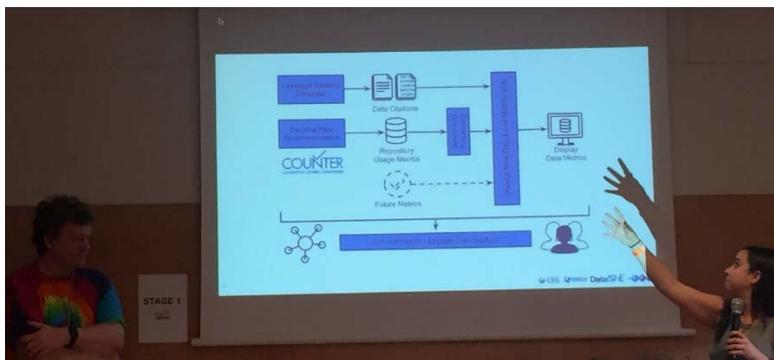
The conference was organized in parallel sessions with many exciting presentations, it was difficult to choose between. I hereby report on my impressions from the presentations I visited. For an overview of all presentations see: <https://pidapalooza.org/>.



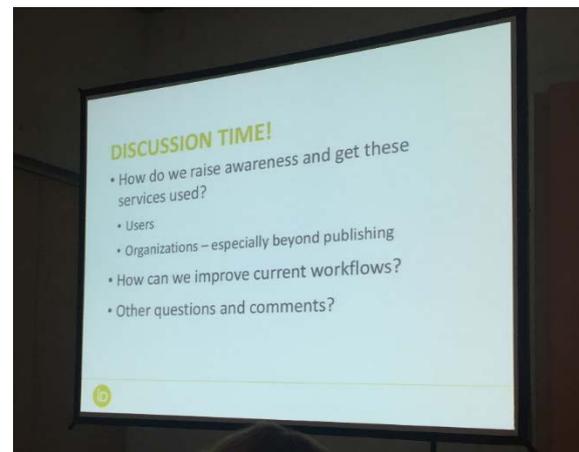
River view in Girona



The venue: Palau de Congressos de Girona



Martin Fenner and Daniella Daniella Lowenberg shows Make Data Count.



There was always time for discussions...



The breaks gave opportunities to meet and chat...

Some of my impressions

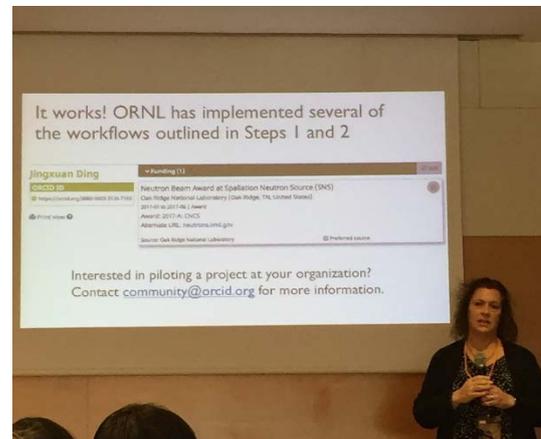
- Chris Erdmann (NCSU) and Lars Nielsen (CERN) showed examples on how to use Jupyter Notebooks to share scripts and solutions for doing specific searches for data in their talk

“Jupyter and PIDs”.

- Angelina Kraft (TIB Hannover) and Britta Dreyer (TIB Hannover) reported from a survey directed to researchers about their views on PIDs. The survey disclosed interesting results about the basic knowledge and experiences among researchers when it comes to PID services. Researchers are becoming more and more surrounded by different types of PIDs which they might not be so aware of. According to Kraft and Dreyer there is no need to involve researchers in too much of details on PIDs. PIDs are part of the infrastructure, and it should be enough for researchers to have a basic knowledge of PIDs.
- Ed Pentz (Crossref) and Alice Meadows (ORCID) shared a presentation on how peer reviews can be registered with PIDs, something that enables to better follow the peer review process.
- Martin Fenner (DataCite) and Daniella Lowenberg (CDL) presented the project on data usage metrics “Make Data Count”, which is an ambitious project between CDL, DataCite and DataOne. Questions that were discussed: How do we link citations and usage? How do we use PIDs to help filter by organization, funder, institution, etc?
- Jo McEntyre (EMBL) gave a detailed speech on how EMBL/EBL is organizing their working-flow with PIDs. The life science sector is fairly advanced when it comes to registration with PIDs and metadata registration. EMBL/EBL has many collaborations with other services and projects, such as THOR, ORCID.



Liz Krznarich from ORCID.



Crystal Schrof from ORNL presenting their project on PIDs for instruments.

- There were many listening in to Todd Carpenter’s presentation on the importance of standardization of PIDs. Todd Carpenter comes from National Information Standards Organization (NISO) which is working in close collaboration with other standard organizations, for example the consortium Linked Content Coalition (LCC).
- “What questions can PIDs answer?” was the title of Ted Habermann’s (HDF Group) talk (Ted Talk!) in which he brought up the importance of adding metadata to PIDs and the importance of standardized metadata. He made the distinction between three types of metadata:

discovery/access, use (machine), understanding.

- Martin Fenner (DataCite) showed a new solution in which it is possible add PIDs to science blog posts, which may improve the scientific value of blog posts.
- Sebastian Karcher (Quality Data Repository) gave a presentation on how researchers may pre-register their research projects with PIDs. Pre-registration of coming research projects and activities is common in social sciences, and it may help researchers communicate their work, help to measure the impact of the researcher's work. Persistent identifiers and a good metadata core are essential parts to make pre-registration a valuable process.
- In the presentation "Capturing facilities: PID recommendation for identifying scientific equipment and infrastructure" Oak Ridge National Laboratory (research organization), ORCID (PID service provider) and Elsevier (publisher) reported on a common project where they are developing better features for registering researchers' activities and usage of equipment and instruments with PID and to make these activities visible in the publishing flow.
- One of the last presentations that I visited was by Greg Janée (CDL), where he glimpsed into the future by displaying four unsolved problems with PIDs:
 1. Are PIDs being used? How can we tell?
 2. Who owns an identifier? Who may modify it?
 3. How can identifier aliasing be better handled?
 4. When a PID system moves or fails, what needs to persist?According to Greg Janée there is a lot of collaboration needed in the PID community in order to solve these problems and to achieve interoperability across PID types and systems.



The PIDapalooza family 2018

By Chodacki (Own work) [CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>)], via Wikimedia Commons

Final reflection

The PIDapalooza conference was completely packed with good content and it became a very nice and creative event. I came back home with a lot of inspiration and many new contacts. To summarize it all: It is possible to add a PID to nearly everything today – but it is underlying information, call it “metadata”, that makes the difference and may take everything to a new, higher level. There is certainly room for more PIDapalooza to come!

Maria Johnsson,
ICOS/ENVRplus and University Library, Lund University

More reading at:

<https://pidapalooza.org/>

<https://pidapalooza.figshare.com/>

<http://twitter.com> Hastag #pidapalooza