

# PIDapalooza – the open festival for persistent identifiers

In 2016 three persistent identifier (PID) organizations – Crossref, DataCite and ORCID – together with California Digital Library organized PIDapalooza, the first open festival for PIDs. The next PIDapalooza will take place on 23–24 January 2018 in Girona, Spain. This review reports back on PIDapalooza 2016 and looks forward to PIDapalooza 2018.

A persistent identifier (PID) is defined by Wikipedia as ‘a long-lasting reference to a document, file, web page, or other object ... usually used in the context of digital objects that are accessible over the Internet. Typically, such an identifier is not only persistent but actionable: you can plug it into a web browser and be taken to the identified source.’<sup>1</sup>

PIDs have been around for a long time, especially in scholarly communications. Think of the ISBN (International Standard Book Number), first introduced in 1966; or its journal equivalent, the ISSN (International Standard Serial Number), launched five years later in 1971. But they really started to take off when scholarly communications went digital in the late 1990s, and with the launch of Crossref as a provider of digital object identifiers (DOIs) for research articles and other works. Since then, there has been a dramatic increase both in the number of persistent identifiers, with nearly 150 million DOIs assigned at the time of writing (October 2017),<sup>2</sup> as well as close to four million ORCID iDs for researchers.<sup>3</sup>

PIDs enable authoritative, unambiguous, digital connections between people (researchers), places (their organizations), and ‘things’ (their research contributions and outputs). The research infrastructure – from more established tools like manuscript submission and grant application systems, to innovative new services such as Altmetric, Kudos and Publons – is increasingly reliant on these connections.

But, despite the ubiquity of PIDs in scholarly communications, until recently the PID community lacked a dedicated space in which to explore ideas of networked research and scholarly communications infrastructure. To fill this gap, in November 2016 a diverse group of experts from California Digital Library, Crossref, DataCite and ORCID organized the first PIDapalooza.<sup>4</sup>

Described as ‘the first open festival for scholarly research persistent identifiers’, PIDapalooza took its cue from the music festival after which it is named, Lollapalooza.<sup>5</sup> The intention was to bring together PID enthusiasts – those who create and/or use persistent identifiers for scholarly communications – for two days of high-level but informal and interactive discussions.

What kinds of PIDs will we need in future? How should they be used? What are the best ways to get researchers to adopt and use PIDs? What are the theoretical and practical approaches to persistence and interoperability? These and many other questions were addressed in the first PIDapalooza, which was attended by 120 PID experts globally.



ALICE MEADOWS

Director of  
Community  
Engagement &  
Support  
ORCID

‘such an identifier is  
not only persistent but  
actionable’

162 Much of the meeting was spent in short (half hour or less) parallel sessions, but there were also five plenaries. Together with the session on organization identifiers (which was so well attended it was virtually a plenary), these provide a good representation of the PIDapalooza experience.

First up was Jonathan Clark, Executive Director of the International DOI Foundation, whose talk, 'PIDvasive – What's possible when everything has a persistent identifier?',<sup>6</sup> looked at what we should expect from our persistent identifiers (as well as persistence and uniqueness). The answer: provenance, metadata, machine readability, and policies/guarantees. In a broad-ranging talk, Clark then went on to discuss the risk of having too many PIDs, the types of services that might be built on them, and the critical need for both interoperability and a social infrastructure for PIDs.

'critical need for both interoperability and a social infrastructure for PIDs'

Day one ended with the second plenary, by Simon Porter, VP of Research Engagement and Information Architecture at Digital Science, entitled 'Research Information Citizenship'.<sup>7</sup> He called on each scholarly communications sector – universities, publishers, funders, service providers and researchers themselves – to play their part in making the digital research infrastructure work better. Porter also raised the need for collaboration to build shared infrastructure tools and services, especially among service providers.

Clifford Tatum, Project Manager at ACUMEN and researcher at Leiden University, kicked off day two. His talk, 'Towards Governance of PID Portability for Research Evaluation',<sup>8</sup> looked at the use of PIDs in the collection of research information for the purpose of evaluation, and the challenges this creates – in a world where open science and interoperability are increasingly the norm – in terms of privacy, security and commercial concerns. Tatum's proposed solution was to focus on improving the portability of PIDs through better standards and protocols.

The fourth plenary was by Herbert Van de Sompel, team leader of the Prototyping Team at the Research Library of the Los Alamos National Laboratory. His talk, 'Signposting for Persistent Identifiers',<sup>9</sup> demonstrated that many papers cite uniform resource identifiers (URIs) other than the DOI URI, reducing the potential power of PIDs. His solution to this problem was to create a signposting pattern for PIDs to enable the automatic discovery and use of the DOI URI rather than other types of URI associated with the DOI-identified object.

The last official plenary speaker was Carly Strasser, Program Officer for the Data-Driven Discovery Initiative at the Gordon and Betty Moore Foundation. She had the (un?)enviable task of drawing together everything that went on at PIDapalooza and she did so with aplomb – and a little Lollapalooza inspiration. Strasser described her talk, 'Reaching Nirvana: The Future of Persistent Identifiers',<sup>10</sup> as 'a "Greatest Hits" of takeaways, lessons learned, points for discussion, and new directions'.

As mentioned, there was also an unofficial plenary – a very well-attended (and lively!) session on organization identifiers. Led by Patricia Cruse, Laure Haak and Ed Pentz (respectively Executive Directors of DataCite, ORCID and Crossref), it began with an update on the work that the three organizations had undertaken to review the current work on organization identifiers and define use cases. The rest of the time was spent on a wide-ranging discussion about next steps, with a range of (sometimes divergent) views expressed. However, there was general agreement that none of the current providers of organization identifiers meet all scholarly communications use cases – especially in terms of researcher affiliations – and there was support for a community working group to seek a solution to this challenge.

'the goal of PIDapalooza 2018 is to create an open, welcoming atmosphere ... and it's open to anyone who creates or uses PIDs'

The response to PIDapalooza 2016 was enthusiastic, so we are now planning the next one, to be held in Girona, Spain on 23–24 January 2018. Like its predecessor, the goal of PIDapalooza 2018 is to create an open, welcoming atmosphere in which to discuss persistent identifiers, and it's open to anyone who creates or uses PIDs.

163 Content will fall into eight broad themes:

1. ***PID myths***

Are PIDs better in our minds than in reality? PID stands for Persistent IDentifier, but what does that mean and does such a thing exist?

2. ***Achieving persistence***

So many factors affect persistence: mission, oversight, funding, succession, redundancy, governance. Is open infrastructure for scholarly communication the key to achieving persistence?

3. ***PIDs for emerging uses***

Long-term identifiers are no longer just for digital objects. We have use cases for people, organizations, vocabulary terms, and more. What additional use cases are you working on?

4. ***Legacy PIDs***

There are thousands of venerable old identifier systems that people want to continue using and bring into the modern data citation ecosystem. How can we manage this effectively?

5. ***Bridging worlds***

What would make heterogeneous PID systems 'interoperate' optimally? Would standardized metadata and APIs across PID types solve many of the problems, and, if so, how would that be achieved? What about standardized link/relation types?

6. ***PIDagogy***

It's a challenge for those who provide PID services and tools to engage the wider community. How do you teach, learn, persuade, discuss and improve adoption? What does it mean to build a pedagogy for PIDs?

7. ***PID stories***

Which strategies worked? Which strategies failed? Tell us your horror stories! Share your victories!

8. ***Kinds of persistence***

What are the frontiers of 'persistence'? We hear lots about fraud prevention with identifiers for scientific reproducibility, but what about data papers promoting PIDs for long-term access to reliably improving objects (software, preprints, data sets) or live data feeds?

The programme for PIDapalooza 2018 (which of course has its own DOI!)<sup>11</sup> is not finalized at the time of writing, since proposals are still being accepted, but I can guarantee that the content will be just as diverse and thought-provoking as the last one and that the level of audience participation and engagement will be just as high. You can find out more on the pidapalooza.org website, register at <http://pidapalooza2018.eventbrite.com>, and follow @pidapalooza for updates on speakers, sessions, and more. (See Figure 1 for the official logo and a reminder of the dates.)



Figure 1. The programme for the next PIDapalooza is being finalized

#### Abbreviations and Acronyms

A list of the abbreviations and acronyms used in this and other *Insights* articles can be accessed here – click on the URL below and then select the 'Abbreviations and Acronyms' link at the top of the page it directs you to: <http://www.uksg.org/publications#aa>

#### Competing interests

The author helps organize PIDapalooza.

#### References

1. Wikipedia entry for 'Persistent identifier': [https://en.wikipedia.org/wiki/Persistent\\_identifier](https://en.wikipedia.org/wiki/Persistent_identifier) (accessed 10 October 2017).
2. Factsheet: Key Facts on Digital Object Identifier System: <http://www.doi.org/factsheets/DOIKeyFacts.html> (accessed 10 October 2017).
3. ORCID statistics: <http://orcid.org/statistics> (accessed 10 October 2017).
4. PIDapalooza: <https://pidapalooza.org/> (accessed 10 October 2017).
5. Lollapalooza: <https://es.wikipedia.org/wiki/Lollapalooza> (accessed 11 October 2017).
6. Clark J, PIDvasive – What's possible when everything has a persistent identifier? PIDapalooza Keynote.pdf. figshare, 2016: DOI: <https://doi.org/10.6084/m9.figshare.4233839.v1> (accessed 10 October 2017).
7. Porter S, Exploring the relationship between persistent identifiers and research information citizenship. figshare, 2016: DOI: <https://doi.org/10.6084/m9.figshare.4220454.v1> (accessed 10 October 2017).
8. Tatum C, Towards governance of PID portability for research evaluation, figshare, 2016: DOI: <https://doi.org/10.6084/m9.figshare.4212732.v1> (accessed 10 October 2017).
9. Van de Sompel H, A Signposting Pattern for PIDs, figshare, 2016: DOI: <https://doi.org/10.6084/m9.figshare.4249739.v1> (accessed 10 October 2017).
10. Strasser C, Reaching Nirvana: The Future of Persistent Identifiers, figshare, 2016: DOI: <https://doi.org/10.6084/m9.figshare.4220520.v1> (accessed 10 October 2017).
11. PIDapalooza's DOI: <https://doi.org/10.5438/11.0002> (accessed 11 October 2017).

**Article copyright: © 2017 Alice Meadows. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use and distribution provided the original author and source are credited.**



Alice Meadows  
Director of Community Engagement & Support  
ORCID, US  
E-mail: [a.meadows@orcid.org](mailto:a.meadows@orcid.org)

ORCID ID: <http://orcid.org/0000-0003-2161-3781>

To cite this article:

Meadows A, PIDapalooza – the open festival for persistent identifiers, *Insights*, 2017, 30(3), 161–164; DOI: <https://doi.org/10.1629/uksg.393>

Published by UKSG in association with Ubiquity Press on 08 November 2017