

Stacy Konkiel & Nickoal Eichmann  
Please send correspondence to: stacy.konkiel@gmail.com  
OR 2015 Poster proposal

### **Better metrics through metadata: a study of the use of persistent identifiers in IRs**

With the rise of researcher awareness of altmetrics, which track the usage and sharing of scholarly outputs on the social web (Priem, Groth, Taraborelli & Neylon, 2010), there has been an increased interest in tracking the impact of institutional repository content using these metrics--both as a value added service for depositors, and as a means of learning more about the usefulness of the repository as a service overall (Konkiel & Scherer, 2013).

Yet institutional repository content often includes preprint versions of articles later published in scholarly journals or available in subject repositories such as PubMed Central and ArXiv.org. These disparate versions of content can each be cited, shared, bookmarked and discussed on the social web independently of each other, making it difficult to keep track of the many metrics that a single study might accumulate overall.

One way to aggregate the metrics for a single study is to, in the metadata for various versions of the study, cross-reference other versions of the same study using permanent identifiers (PIDs) like Handles, DOIs, PubMed IDs, and ArXiv. Any repository interested in adding altmetrics reports for repository content should be concerned with having clean and complete PID metadata included in their item records, to help altmetrics aggregators take into account multiple versions of a single publication.

Assuming that such metadata will be incomplete, we set out to calculate the benchmarks for staff hours needed to compile permanent identifier metadata of research outputs held by institutional repositories, where those publications are alternative versions of scholarship published in scholarly journals or subject repositories.

This small study sought to determine:

1. The number of IR holdings where the item is an alternative version of content previously published elsewhere.
2. The number of IR holdings from [1] that include PID information in the dc.identifier metadata field.
3. The total number of publications by our institution's faculty published since the advent of DOIs, PubMed IDs, and ArXiv IDs. This information will be compared with [1] and [2] to establish a benchmark of how much IR content includes proper PID metadata, where the option to do so exists.

We found that--where permanent identifiers exist--only 28.6% of deposits to our IR included DOIs in the item record metadata, and no deposits that had ArXiv IDs or PubMed IDs included this information in the IR item record metadata.

This poster reports these tallies, and also our benchmarks for staff hours needed to complete metadata (analyses of which are still in progress).

Stacy Konkiel & Nickoal Eichmann  
Please send correspondence to: stacy.konkiel@gmail.com  
OR 2015 Poster proposal

## References

Konkiel, Stacy, and Dave Scherer. 2013. "New Opportunities for Repositories in the Age of Altmetrics." *Bulletin of the American Society for Information Science and Technology* 39 (4) (April 15): 22–26. doi:10.1002/bult.2013.1720390408

Priem, Jason, Dario Taraborelli, Paul Groth, and Cameron Neylon. 2010. "Alt-Metrics: A Manifesto." <http://altmetrics.org/manifesto/>