ORCID: Using API Calls to Assess Metadata Completeness

Naomi Eichenlaub
Ryerson University, Canada
neichenl@ryerson.ca
https://orcid.org/0000-0002-4348-1700

RULA
Ryerson University Library & Archives

FLORIDA SOUTHERN COLLEGE
Roux Library

Marina Morgan
Florida Southern College, United States
mmorgan@flsouthern.edu
https://orcid.org/0000-0002-1829-1016

ABSTRACT

The aim of this poster is to demonstrate the importance of adequate metadata in ORCID profiles to ensure name disambiguation. It is only through more complete metadata that ORCID will ensure success in terms of interoperability with institutional scholarly, publishing and funding bodies. We ran queries against the public ORCID API to get a better sense of how many ORCID records have only minimal information.
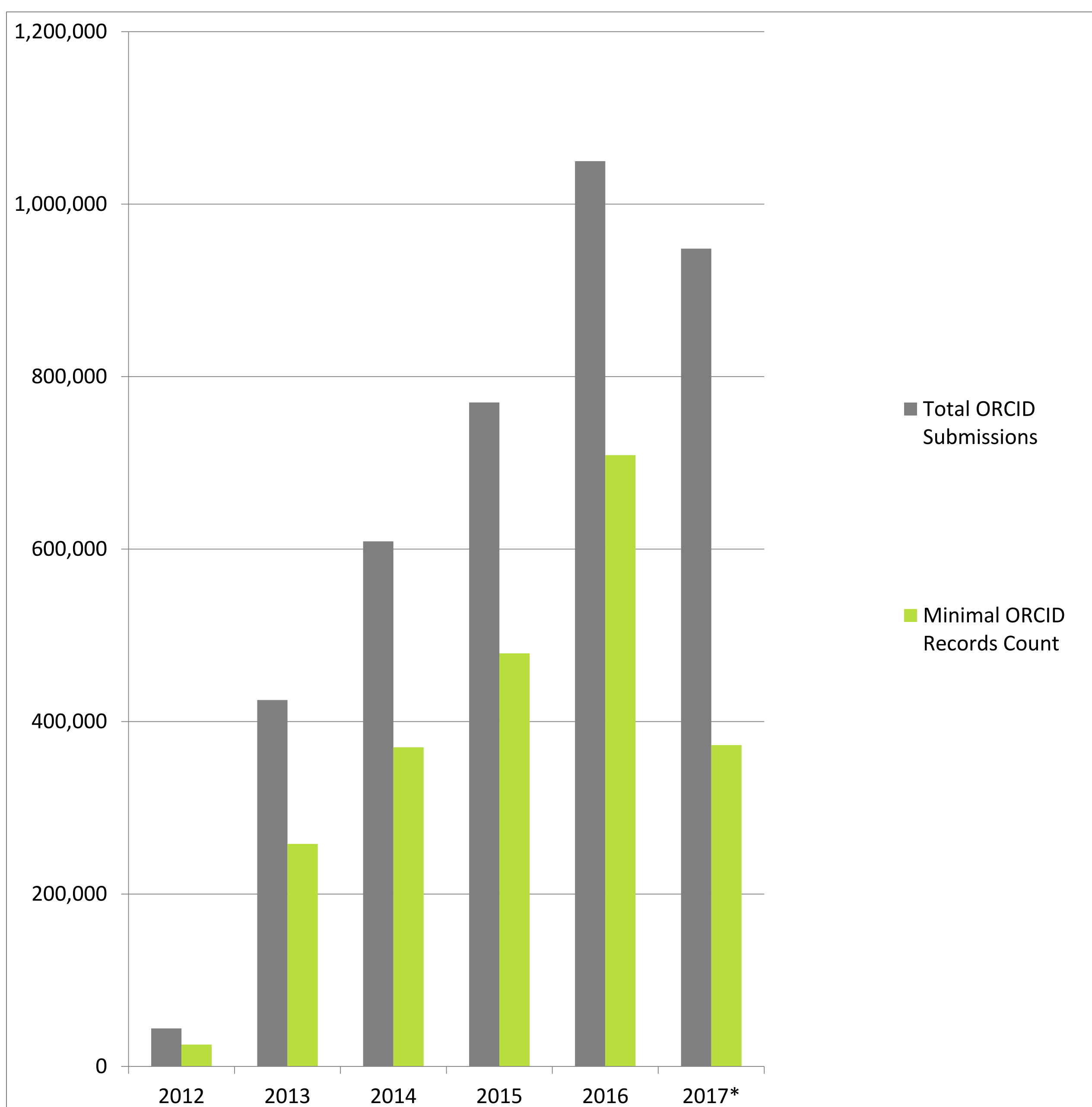
ORCID

ORCID (Open Researcher and Contributor ID) is a non-profit persistent digital identifier open registry offered to researchers across disciplines. Their mission is to provide an identifier in the form of a unique alphanumeric code to provide persistent identity for researchers. There are three easy steps to receive a persistent ID and distinguish oneself from other researchers: register, add your info, and use your ORCID ID.

METHODOLOGIES

To get a sense of the big picture in terms of records with only the minimal required information, i.e., name and email, we searched records created between 2012-2017 that did not include affiliation / organization name, Ringgold ID, and any work titles. Results were gathered by year as well as an overall API call to search the entire ORCID database.

Year	API Calls	Count
2012	**ORCID records with any given names, without any affiliation, Ringgold ID, work titles, and submitted between January 1st to December 31st, 2012:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+ AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *] +AND+profile-submission-date:%5B 2012-01-01T00:00:00Z%20TO%202012-12-31T00:00:00Z%5D	**25,351**
2013	**ORCID records with any given names, without any affiliation, Ringgold ID, work titles, and submitted between January 1st to December 31st, 2013:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *] +AND+profile-submission-date:%5B2013-01-01T00:00:00Z%20TO%202013-12-31T00:00:00Z%5D	**258,182**
2014	**ORCID records with any given names, without any affiliation, Ringgold ID, work titles, and submitted between January 1st to December 31st, 2014:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *] +AND+profile-submission-date:%5B2014-01-01T00:00:00Z%20TO%202014-12-31T00:00:00Z%5D	**370,074**
2015	**ORCID records with any given names, without any affiliation, Ringgold ID, work titles, and submitted between January 1st to December 31st, 2015:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *] +AND+profile-submission-date:%5B2015-01-01T00:00:00Z%20TO%202015-12-31T00:00:00Z%5D	**479,144**
2016	**ORCID records with any given names, without any affiliation, Ringgold ID, work titles, and submitted between January 1st to December 31st, 2016:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *] +AND+profile-submission-date:%5B2016-01-01T00:00:00Z%20TO%202016-12-31T00:00:00Z%5D	**709,046**
2017 to 05/17/17	**ORCID records with any given names, without any affiliation, Ringgold ID, work titles, and submitted between January 1st, 2017 to May 17, 2017:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *] +AND+profile-submission-date:%5B2017-01-01T00:00:00Z%20TO%202017-12-31T00:00:00Z%5D	**372,709**
2012 to 05/17/17	**All ORCID records with any given names, without any affiliation, Ringgold ID, and work titles:** https://pub.orcid.org/v2.0/search/?q=given-names:[* TO *]+AND+-affiliation-org-name:[* TO *]+AND+-ringgold-org-id:[* TO *]+AND+-work-titles:[* TO *]	**2,216,944**

Examples of ORCID dummy accounts

ORCID iD	First/given name	Last/family name	Other names
0000-0001-5429-1577	John	Doe	Jonathan Doe
0000-0001-6599-8818	John	Doe	
0000-0001-7295-1671	John	Doe	
0000-0001-7786-8180	john	doe	
0000-0002-7155-7311	John	Doe	
0000-0002-0365-8743	John	Doe	
0000-0001-7580-2137	John	Doe	
0000-0001-5393-1421	John	Doe	
0000-0003-4488-8251	John	Doe	
0000-0002-9187-455X	John	Doe	
0000-0003-2118-5887	John	Doe	
0000-0002-9509-7620	Jane	Doe	
0000-0001-5442-5838	Jane	Doe	

CHALLENGES

Missing data such as affiliations or work titles poses a major challenge on one hand to gather appropriate data and on the other to consider PID service adoption. From the preliminary results we noticed that the problem of orphan (empty) records is very common. Many records used random placeholder names such as John/Jane Doe to keep their true identity unknown, filler text such as lorem ipsum, fictitious funding, works, and institution names for employment and education.

FUTURE WORK

In order to address a large number of ORCID records missing critical metadata fields the following solutions are proposed:

1. Advocacy and education
2. Continued ORCID outreach
3. Increased system interoperability

CONCLUSIONS

ORCID IDs are very useful, specifically when most names are not unique. However, many records lack critical elements required to perform the main function of a personal identifier, namely the name disambiguation. It is hoped that this poster draws attention to the problems associated with a lack of identifying metadata in ORCID records and highlights the value of more complete metadata in disambiguating researcher names and identifiers.

REFERENCES

Meadows, Alice. (2016). Everything you ever wanted to know about ORCID … but were afraid to ask. College & Research Library News, 77(1), 23-30.
ORCID. (2017a). ORCID open letter: One year on. Retrieved, May 15, 2017, from https://figshare.com/articles/ORCID_Open_Letter_-_One_Year_On_Report/4828312.
ORCID. (2017b). Challenges of measuring PID adoption. Retrieved, April 17, 2017, from http://orcid.org/blog/2017/03/24/challenges-measuring-pid-adoption.
ORCID. (2017c). Collect & Connect. Retrieved, March 15, 2017, from https://orcid.org/content/collect-connect.
ORCID. (2017d). ORCID API. Retrieved, March 15, 2017, from https://orcid.org/organizations/integrators/API.
ORCID. (2017e). ORCID Public File. Retrieved, March 15, 2017, from https://orcid.org/content/orcid-public-data-file.

Reed, Kathleen, Dana McFarland, and Rosie Croft. (2016). Laying the groundwork for a new library service: Scholar-practitioner & graduate student attitudes toward altmetrics and curation of online profiles. Evidence Based Library and Information Practice, 11(2), 87-96. doi:10.18438/B8J047.
Ringgold. (2017). Case Studies ORCID. Retrieved, March 15, 2017, from http://www.ringgold.com/case-studies-orcid.
Thompson, Elleb and Sally French. (2016). Pimp my Profile and the researcher profile health check: Practical, individualised researcher support initiatives co-created by library and faculty. In ALIA National Conference 2016, 29 August - 2 September 2016, Adelaide, SA. Retrieved, March 15, 2017, from http://eprints.qut.edu.au/98649/.
THOR. (2017). Researcher metrics: ORCID IDs. Retrieved, March 15, 2017, from http://dashboard.project-thor.eu/dashboard/researcher/.
Tran, ClaraY. and Jennifer A. Lyon. (2017). Faculty use of author identifiers and researcher networking tools. College & Research Libraries, 78(2), 171-182. doi:10.5860/crl.78.2.171.