

# Open Citations

Christina Winter, Copyright and Scholarly  
Communications Librarian &  
Dale Storie, Associate University Librarian, Content  
Management and Discovery, University of Regina

**April 26, 2019**



University  
of Regina

# Agenda

- Background
- Open Citations Movement
- Comparison
- Conclusion and Further Discussion

# Citations

- Reveal research and publication patterns
- Identify new research questions
- Determine research impact
- Bibliometrics, Citation Analysis, Scientometrics, Informetrics is the study of citations



# Sources of Citation Data

- Web of Science (Clarivate Analytics)
- Scopus (Elsevier)
- Google Scholar



# Open Citation Data

“citation data now needs to be recognized as part of the Commons”

(Shotton, 2013)

# Initiative for Open Citations

“Citations are the links that knit together our scientific and cultural knowledge. They are primary data that provide both provenance and an explanation for how we know facts. They allow us to attribute and credit scientific contributions, and they enable the evaluation of research and its impacts. In sum, citations are the most important vehicle for the discovery, dissemination, and evaluation of all scholarly knowledge.”



# Initiative for Open Citations

“The present scholarly communication system inadequately exposes the knowledge networks that already exist within our literature. **Citation data are not usually freely available to access, they are often subject to inconsistent, hard-to-parse licenses, and they are usually not machine-readable.**”

(I4OC, n.d., is available under a [CC0 Public Domain Dedication](#).)



# Public Domain Dedication





# Open Metadata

“10. No closed metadata: Publishers shall make bibliographic records, usage metrics, and citation data for our authors freely available, easy to parse, and machine readable.”

(2018, University Committee on Library and Scholarly Communication, Assembly of the Academic Senate, University of California)



# How to participate

## **Crossref Access to Citation Data:**

Open – Journal makes the references open

Limited – Available through a Crossref Metadata Plus Account

Closed – Journal does not make them available to Crossref

I4OC is encouraging open whenever possible



# How to participate

If you register journal articles with Crossref, make sure you set your references to open

Requested to make reference metadata “open” allows access through Crossref REST API (no license restrictions) and will get harvested by other citation data projects (OpenCitations.net)



# Participants

<https://i4oc.org/#publishers>

# Open Citations.net

- Open and Linked Data for Citations
- Three services
  - Open Citation Corpus Linked Data Index
    - Harvested from Crossref data
  - CRowdsourced Open Citation Index (CROCI)
    - Crowdsourced citations
- <https://opencitations.net/browser/br/428357>

# Comparison:

- Web of Science ((WoS) vs. the Open Citation Corpus.
- Sample set of 4515 papers by URegina authors published between 2008-2017 (identified through WoS)
- OCC dataset from Nov. 2018
  - 46,534,705 bibliographic records
  - 445,826,118 citation links
- Only compared records with a DOI

# WoS vs. OCC – Times Cited

Total # of times the URegina papers were cited

- OCC: 37,356
- WoS: 88,655

OCC coverage for citing articles was only about 42% of WoS coverage.

# Most Cited Articles

		PubYear	Google Scholar	Web of Science ▾	OCC
1.	Observation of a new parti... <small>Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC</small>	2012	13,459	5,278	3,614
2.	The ATLAS Experiment at ...	2008	8,881	1,890	0
3.	The Physcomitrella geno...	2008	1,436	1,030	624
4.	Lakes and reservoirs as re...	2009	1,402	982	646
5.	Indication of Electron Neu...	2011	1,694	969	0
6.	The ATLAS Simulation Infr...	2010	4,135	803	497
7.	Observation of a Centralit...	2010	960	488	0
8.	Combined search for the ...	2012	803	412	243
9.	The Selaginella Genome I...	2011	532	388	267
10.	Three-way decisions with ...	2010	613	378	182



# Network Graph of Citations

- [30 Psychology publications by U of R authors published between 2008-2017](#)



# Comparison

- Using OCC for informetric studies and research evaluation is a work in progress at this point in time.
- For informetric research, the more complete the dataset the better.
- Although the citations are open, you still need to turn to other tools for author disambiguation, subject categories, and institutional links.



# Why?

- Web of Science was collecting citation data long before Crossref started collecting it, and likely has better data.
- Unsure of retroactive coverage of Crossref data
- Elsevier does not participate in the CrossRef Open program

# In the News

- Journal of Informetrics (JOI)
- Quantitative Science Studies (QSS)

# Questions:

- How do we promote open citations?
- How will this movement affect research impact practices?
- Will having open citations help better represent the work done in humanities and social sciences?



# References

- Enis, Matt. (2019). Informetrics Journal Quits Elsevier, Relaunches as Open Access. *Library Journal*. Retrieved from <https://www.libraryjournal.com/?detailStory=/190117qssgoesoa>
- Initiative for Open Citations. (n.d.). About. Retrieved from <https://i4oc.org/>
- Shotton, David. (2013). Open citations. *Nature*. 502 (7471): 295-297. Retrieved from <https://www.nature.com/news/publishing-open-citations-1.13937>

# References

- University Committee on Library and Scholarly Communication, Assembly of the Academic Senate, University of California. (2018). Declaration of rights and principles to transform scholarly communications. Retrieved from <https://senate.universityofcalifornia.edu/files/committees/ucolasc/scholcommprinciples-20180425.pdf>

# Image Credits

- “[Zero](#)” by [Creative Commons](#) is licensed under [CC BY 4.0](#)