

Open Data & Open Access: Connections in the Life Sciences

Sarah C. Williams, University of Illinois at Urbana-Champaign
Head, Funk ACES Library | scwillms@illinois.edu

Abstract

In order to expand the limited research on connections between open data and open access (OA) publications, this poster presents the findings of a small-scale study exploring the current situation in the life sciences. Starting with the institutional data repository of a land-grant university, this study included 44 openly available life sciences datasets that had 45 related research articles published. The results reveal the percentage of articles in OA journals, the use of OA in hybrid journals, and the availability of articles through other sources. This study also examined how and where the open datasets were included in the related articles, especially considering whether OA articles handled data differently. These findings provide a glimpse of the current state of open data and OA publication connections in the life sciences, and suggest future research studying researchers' decisions to make data and publications openly available.

Methods

Started with:
Illinois Data Bank || <https://databank.illinois.edu>

- Life sciences datasets: 70
- Life sciences datasets with related articles: 44
- Articles connected to life sciences datasets: 45

Numbers as of April 30, 2019

Checked:
Journal OA status

- Searched Directory of Open Access Journals (DOAJ), and reviewed the journal website

Article OA status on journal website

- Tried to access article off-campus without a VPN connection

Unpaywall (database of legally OA scholarly articles)

- Searched article DOI or title in Easy Search, Illinois' homegrown system that includes Unpaywall

ResearchGate (academic social media platform)

- Searched article title and only counted access if article was immediately available to download

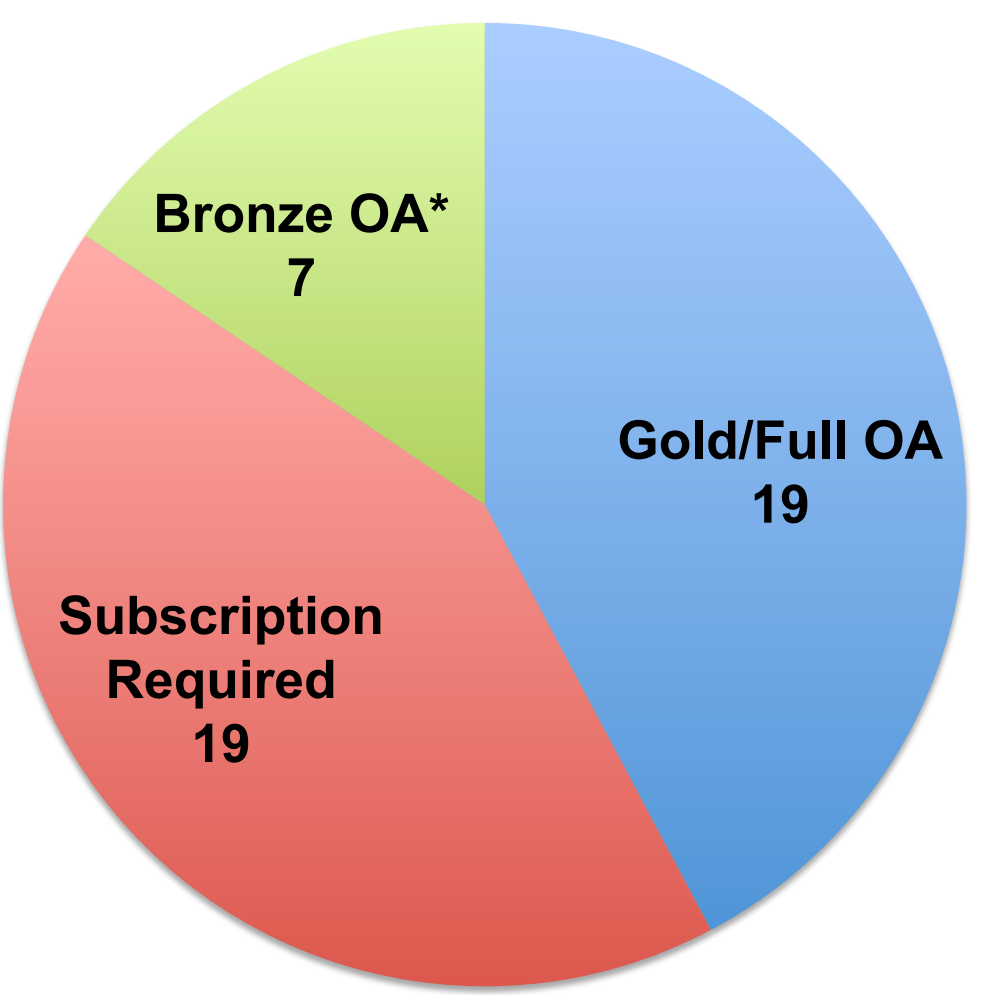
Also examined:

- How and where Illinois Data Bank data was included in the articles

Results

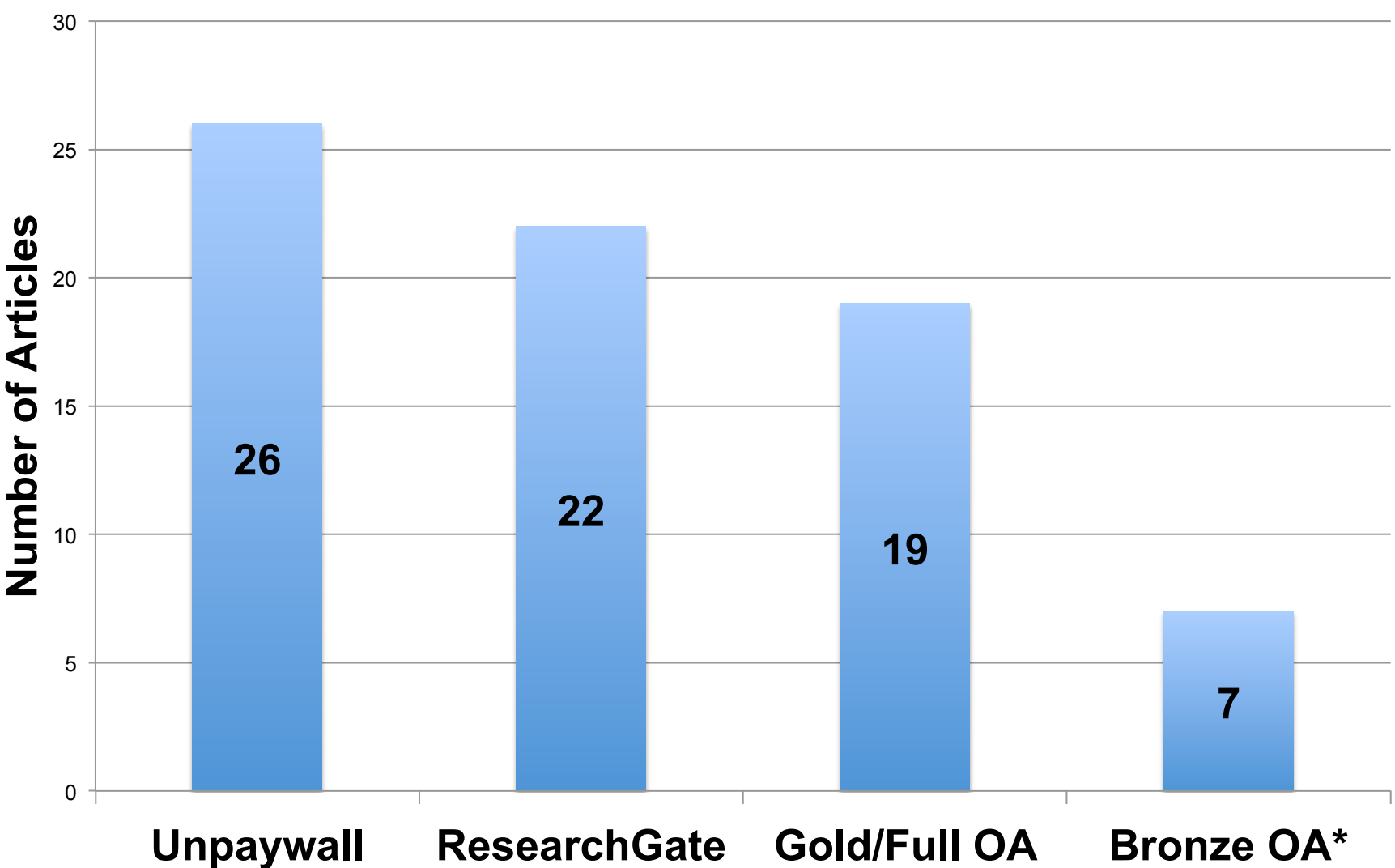
Article Access

Article Status on Journal Site
n=45



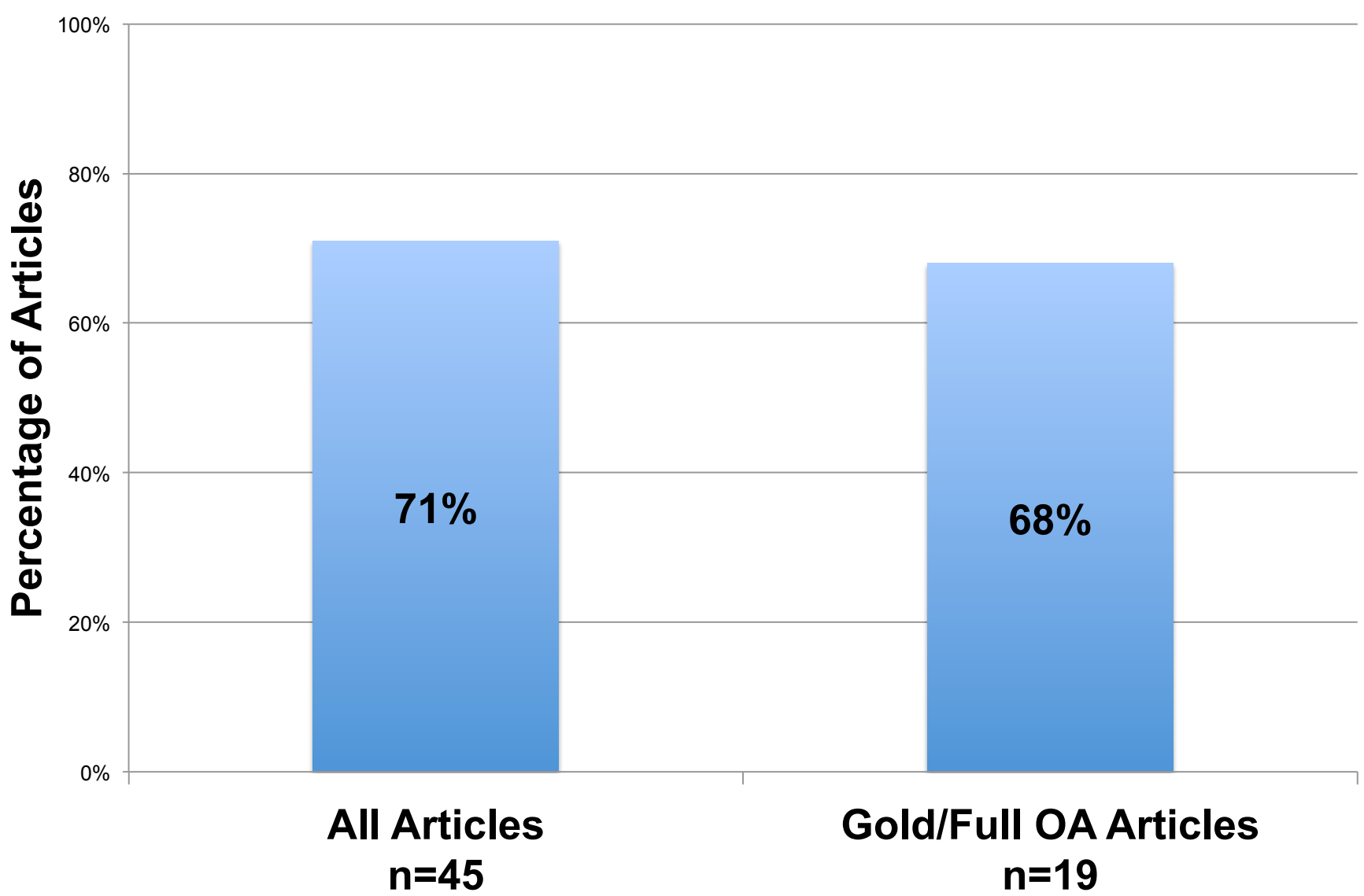
* **Bronze OA** – Free to read on the publisher page, but without a clearly identifiable license. (Piwowar et al. 2018)

Free Access Availability
n=45

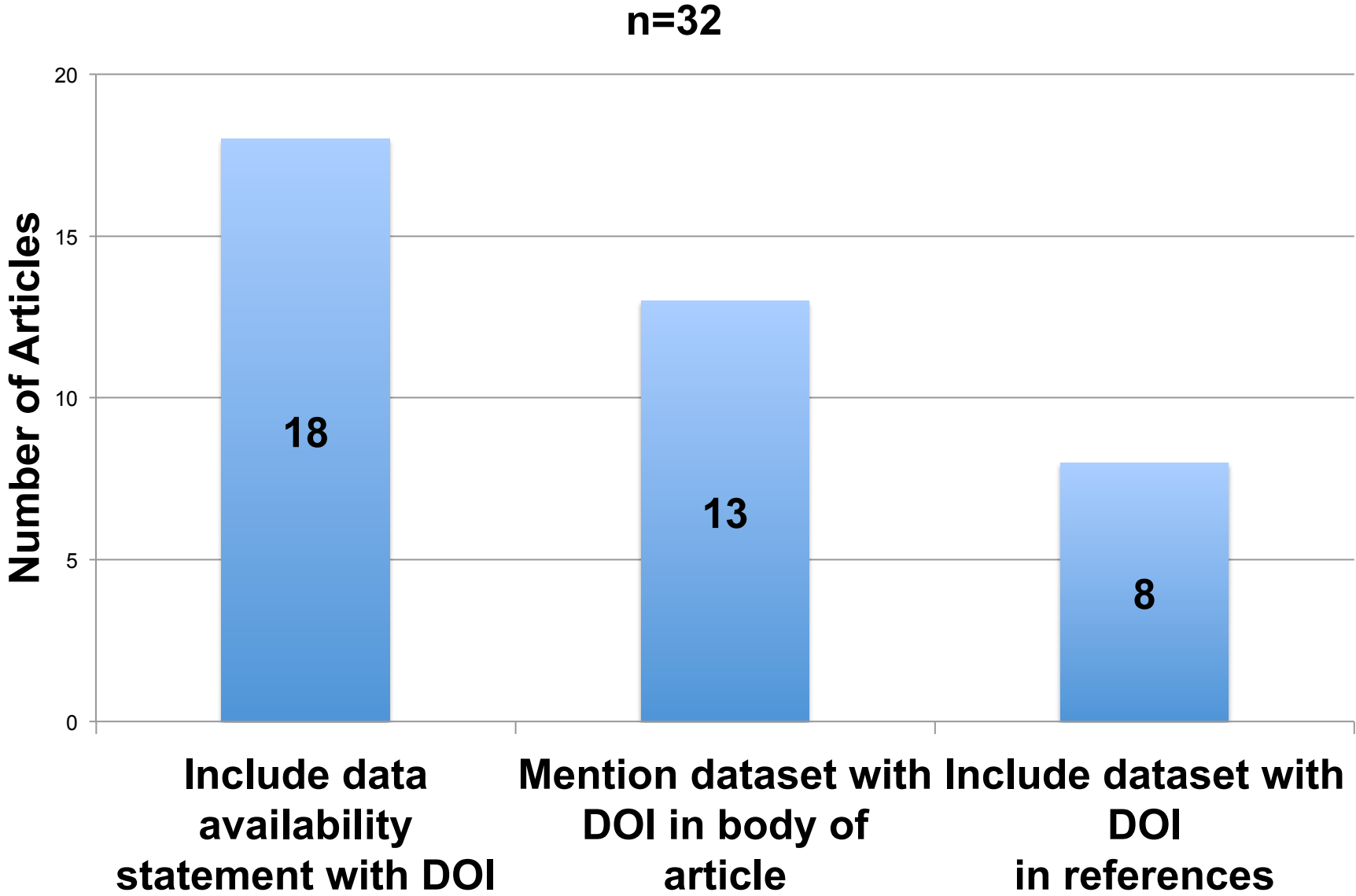


Connections between Articles and Data

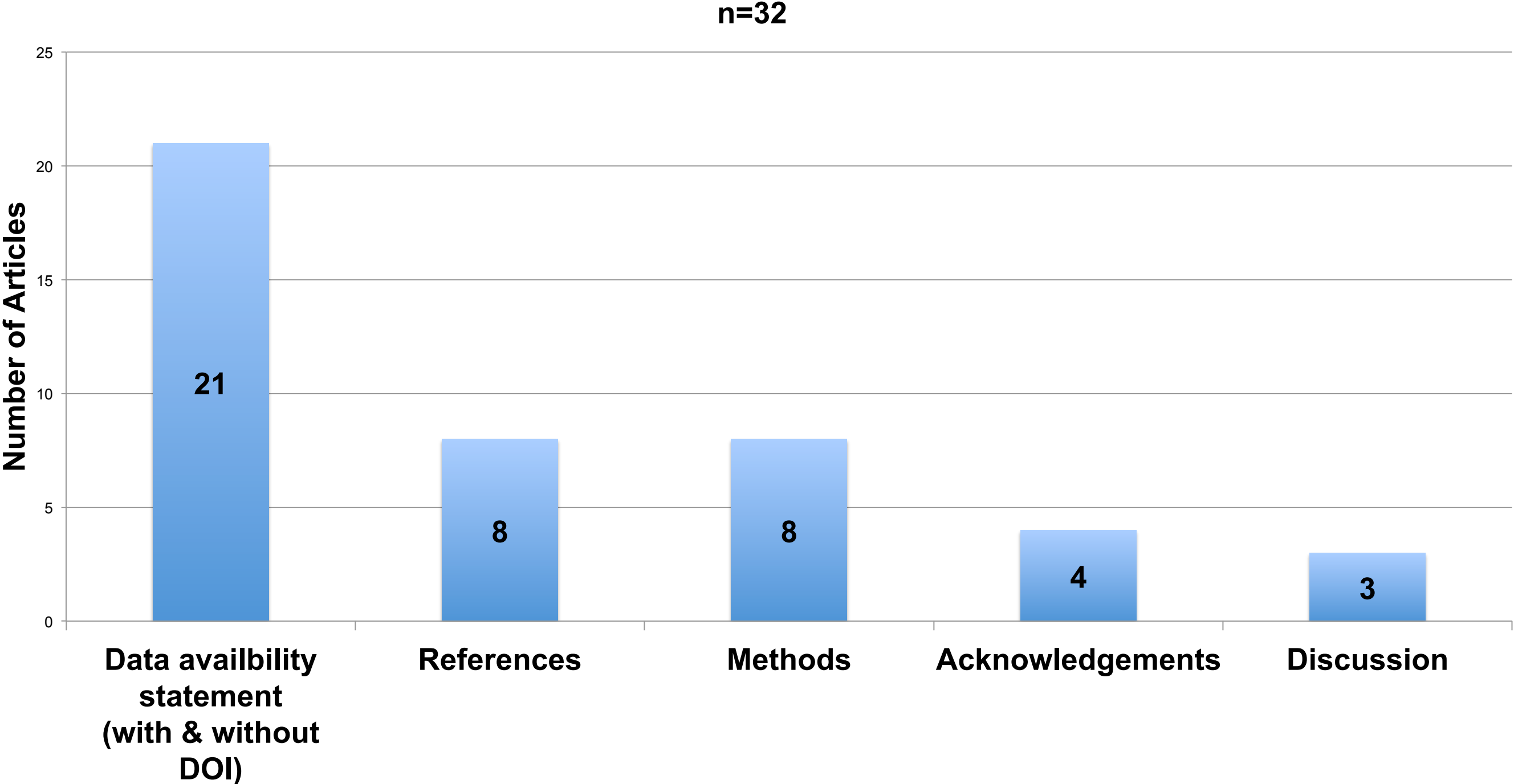
Articles Pointing to Open Data



Common Ways to Point to Data
n=32



Common Places to Point to Data
n=32



Key Takeaways

Bronze OA
This term was coined in Piwowar et al. (2018), and it refers to articles free to read on the publishers page, but without a clearly identifiable license. Bronze OA access is precarious. Of the 7 Bronze OA articles in this study, 2 were no longer freely available just 3 months later. Bronze OA will be an important topic for future research and discussion.

Unpaywall
In this study, Unpaywall linked to 26 articles: 19 Gold/Full OA, 5 Bronze OA, and 2 subscription articles that were available via PubMed Central.

OA adoption in hybrid journals
None of the 26 articles published in hybrid journals were available as licensed OA articles. Seven of these articles were freely available as Bronze OA.

OA articles & open data
Gold/Full OA articles included open data at a similar rate and in similar ways to other articles. Of the 19 Gold/Full OA articles, 68% pointed to the Illinois Data Bank data, while 71% of all articles did. For both Gold/Full OA articles and other articles, a data availability statement with a DOI was the most common way to point to the Illinois Data Bank data.

Dataset DOIs
Of the 32 articles that pointed to Illinois Data Bank data, 30 of them included the data DOI. Such frequent use of DOIs, whether due to journal requirements or author preferences, reveals the importance of data repositories providing DOIs.

Reference

Piwowar, et al. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* 6:e4375. DOI: 10.7717/peerj.4375

