UNC | UNIVERSITY LIBRARIES

Connecting open data to published articles in the Carolina Digital Repository

Rebekah Kati, Institutional Repository Librarian
Julie Rudder, Head, Repository Services
Rolando Rodriguez, Humanities Data Librarian

Carolina Digital Repository Background

Institutional repository which supports data

Home of the UNC Open Access Policy

- 48,000+ articles deposited (mostly) by library staff

Research data

- 100+ deposits deposited (mostly) via self-deposit

Other types of deposits: student work, student papers, multimedia

UNC Dataverse Background

Data repository

Hosted and managed by UNC Research Data Management Core

26,000+ datasets

- 1,155 met our investigation criteria
  - Created by UNC-CH affiliates
  - Deposited within Open Access Policy timeframe (2016-present)

Examined 307 datasets for this presentation, but intend to continue!

Research Questions

- Can we locate articles connected to datasets deposited in CDR and UNC Dataverse?
  - What method was used to locate the article?
  - How much time did it take to locate the article?

- How many articles connected to datasets are eligible for deposit under UNC-CH's open access policy?
  - How many articles are already available openly via publisher?

Methodology

Locating Articles Associated with Datasets:

1. Review dataset metadata for article URL
2. Search for dataset DOI in Google
3. Search Google for any reference to article
4. Search Google for dataset title

Confirming Article and Dataset Association:

1. Compare article's author and date of publication to dataset metadata
2. Check article text for data availability statement or mention of dataset

CDR Investigation Results

Reviewed 71 datasets

41 datasets had clear association to an article

- 24 depositors mention article in dataset metadata
- Only 6 depositors included URL

81% of articles were already open access, 7 were deposited to CDR

UNC Dataverse Investigation Results

Looked at subset of 307 datasets which included related links

Are articles referenced in related links eligible for deposit? Does the full text link back or reference UNC Dataverse?

26 eligible for deposit

12 link back to dataset

Findings

Bi-directional links between articles and datasets are quickest and easiest to confirm

Links to datasets do not always appear in a standard place

DOI links are favored in Dataverse, but not in CDR

Correlation between Open Data and Open Article

Next Steps

Complete UNC Dataverse investigation

Collaborate with UNC's Research Data Management Core

Determine scale and staffing for future projects

Contact Us!

Rebekah Kati: rkati@email.unc.edu
Julie Rudder: jrudder@email.unc.edu
Rolando Rodriguez: rolando.rodriguez@unc.edu

Work Cited:
- Rudder, J., Kati, R. and Rodriguez, R. (2024). Connecting open data to published articles in the Carolina Digital Repository. Presentation given at IDCC24

UNC | UNIVERSITY LIBRARIES

The University of North Carolina at Chapel Hill