

University of Massachusetts Medical School

eScholarship@UMMS

---

University of Massachusetts and New England  
Area Librarian e-Science Symposium

2016 e-Science Symposium

---

Apr 6th, 12:00 AM

## Classifying Data Deposited by Scientists into a Library's Data Repository


Andrew T. Creamer  
*Brown University*

Ann C. Caldwell  
*Brown University*

Joseph Rhoads  
*Brown University*

Ben Cail

~~Brown University~~  
Find this and additional works at: [https://escholarship.umassmed.edu/escience\\_symposium](https://escholarship.umassmed.edu/escience_symposium)

 Part of the [Life Sciences Commons](#), [Medicine and Health Sciences Commons](#), [Physical Sciences and Mathematics Commons](#), and the [Scholarly Communication Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](#).

---

### Repository Citation

Creamer, A. T., Caldwell, A. C., Rhoads, J., & Cail, B. (2016). Classifying Data Deposited by Scientists into a Library's Data Repository. *University of Massachusetts and New England Area Librarian e-Science Symposium*. <https://doi.org/10.13028/679j-z519>. Retrieved from [https://escholarship.umassmed.edu/escience\\_symposium/2016/posters/19](https://escholarship.umassmed.edu/escience_symposium/2016/posters/19)

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](#). This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts and New England Area Librarian e-Science Symposium by an authorized administrator of eScholarship@UMMS. For more information, please contact [Lisa.Palmer@umassmed.edu](mailto:Lisa.Palmer@umassmed.edu).

# Classifying Data Deposited by Scientists into a Library's Data Repository



AT Creamer, EA Caldwell, J Rhoads, and B Cail

Brown University Library, Providence, Rhode Island USA

## Introduction

Libraries that are developing their own data repositories benefit from data on the nature of objects in established data repositories so that they can prepare their future platforms, services, and policies accordingly. While repository reports usually focus on the number, formats, and sizes of files deposited, we have instead endeavored to describe other useful attributes, such as whether the deposited data are associated with a publication or whether data collections are related to a sponsored project.

This poster reports the results of our review and analysis of the individual data deposits (n=20) as well as collection-level data libraries (n=11) deposited by scientists into Brown University's data repository, the Brown Digital Repository (BDR), from April 2014 to April 2016 using a rubric, developed by the authors, and information provided by depositors.

## About the Brown Digital Repository

Find & Borrow Articles, Journals, & Databases Research Help Hours, Locations, & Events Ask a Question Now Off-Campus Access My Library Account

BROWN UNIVERSITY LIBRARY Search the BDR Search

Brown Digital Repository Feedback Discover... I Want to... Login

Featured collections...

- Brown Science Center: Science Collections
- Herbarium
- Opening the Archives: Documenting U.S. - Brazil Relations, 1960s-80s
- Senate Elections Data Project 1871-1913
- Summer Research Symposium
- The Theater that was Rome

Discover... I want to...

- Upload Content**  
Content added by scholars through self-service upload
- Theses & Dissertations**  
Brown University dissertations since 2008
- Deposited Collections**  
Scholar-created research collections, research data, posters, proceedings, and more.
- Upload something**  
Uploaded items can be kept private, shared at Brown, or shared with the public.
- Digitize something**  
Digital Production Services can assist with digitization of materials in support of scholarship, research, and teaching.
- Use an API**  
The BDR offers a complete set of Application Programming Interfaces (APIs) for using content in your own applications.
- Assign a DOI**  
Digital Object Identifiers (DOIs) help publishers and scholars cite content stored in the BDR.
- Consult on a project**  
The Center for Digital Scholarship works with scholars on digital projects across disciplines.
- Read FAQ**  
Frequently Asked Questions may already answer your BDR-related question.

- Website: [repository.library.brown.edu/](http://repository.library.brown.edu/)
- By depositing their data in open, sharing, and preservation-friendly formats, along with appropriate documentation, in the Brown Digital Repository (BDR), scientists can maximize the potential for their data's discovery, access, use, and reuse.
- BDR platform relies on Fedora Commons software and offers an open API for programmatically retrieving both structured metadata and digital content in public collections. The BDR is searchable using an Apache Solr index and via a public search API.
- Items in the BDR are stored redundantly in off-site storage. Audit trails are maintained for each object to document changes and deletions, and older versions of files are retrievable in the event of unintentional modification. Objects also receive a checksum value to allow for periodic auditing of data integrity.
- Objects receive a permanent link and a digital object identifier (DOI) to maximize their potential for discovery, access, and citation.

## Rubric Criteria for Classifying Deposits

### U: Data underlying a publication

Data results reported in a publication, necessary for the peer review and reproduction of the paper's reported results.

### S: Supplementary files accompanying a publication

Data such as images, tables, or visualizations that were not able to be included in the published paper or were cited as supplementary in the paper.

### DS: Digital scholarship

Digital objects that are equivalent to a publication, of a scholarly nature, and intended for dissemination (e.g., video, animation, database, etc.)

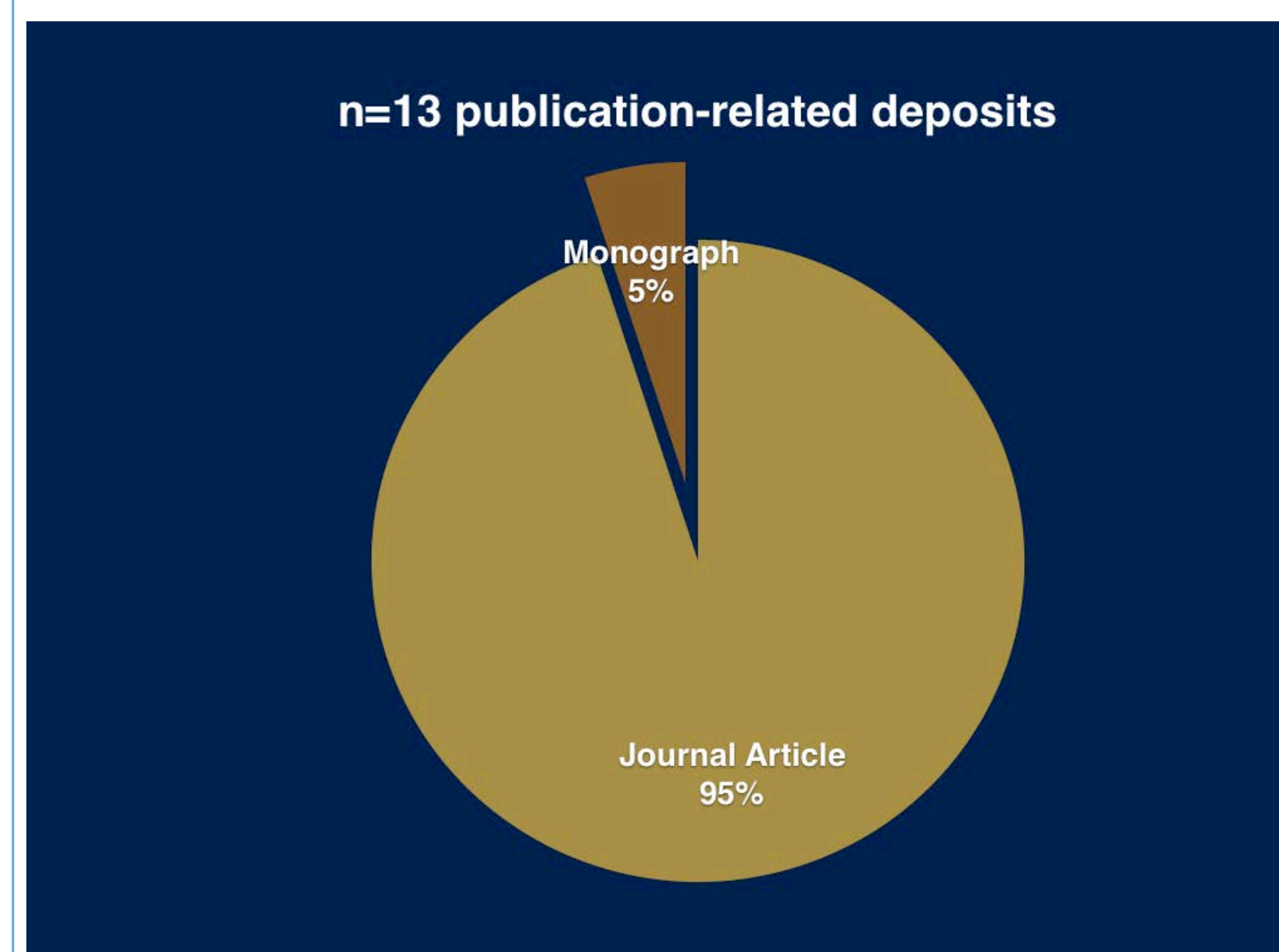
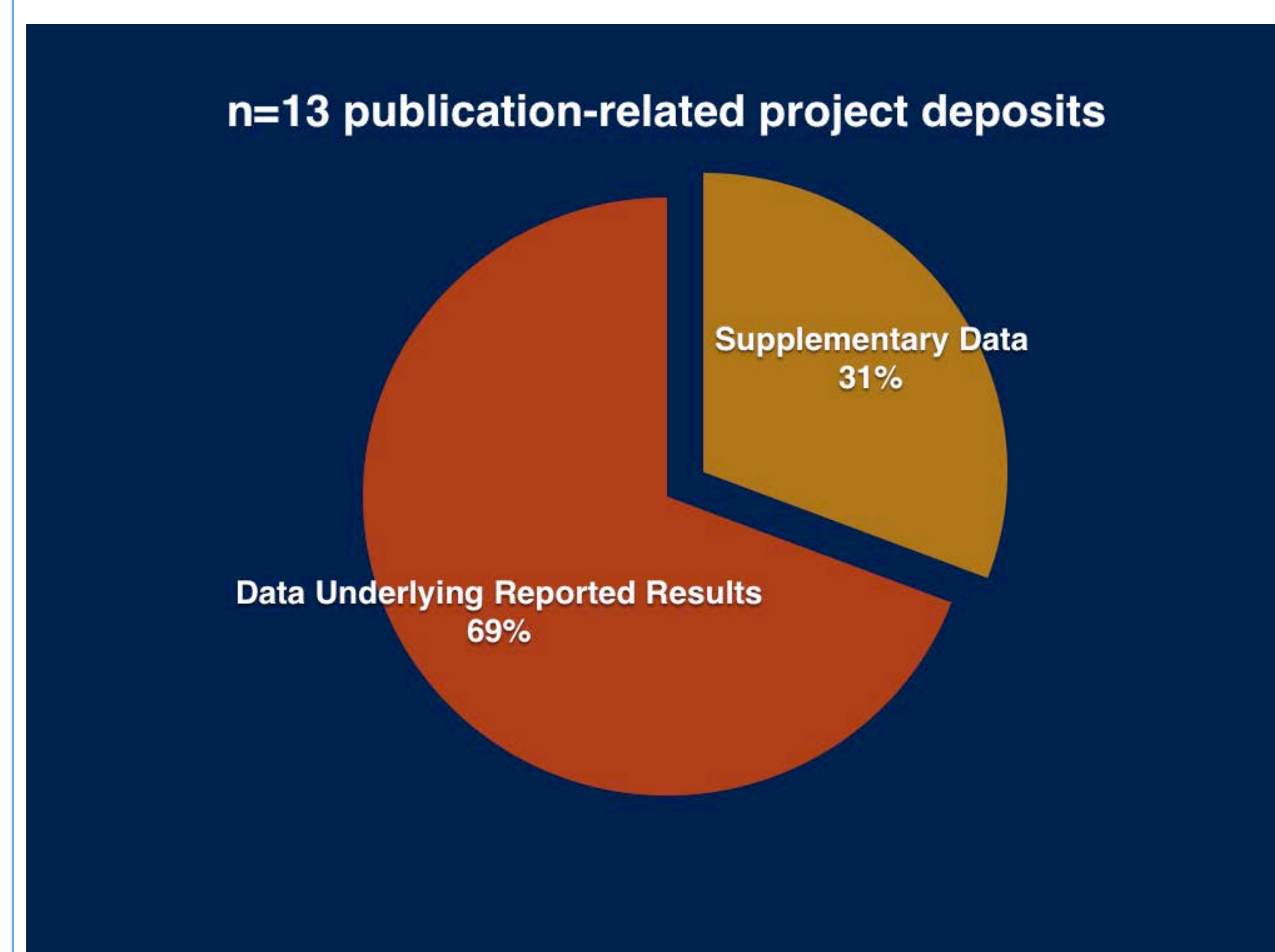
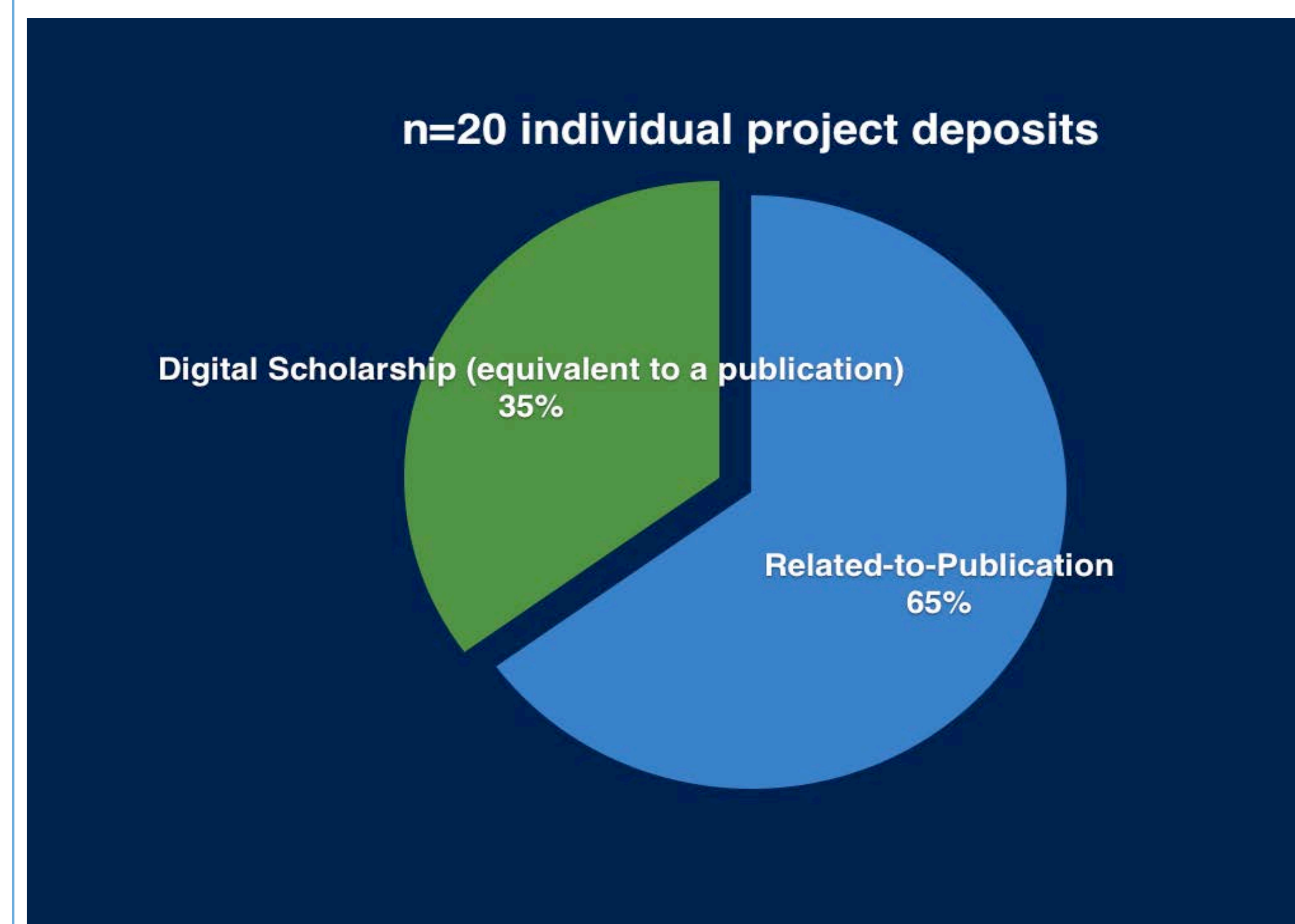
### DC: Data libraries (data collections)

A collection of scientific data, not connected with a single publication or a grant, and is available for other researchers or the public to access and reuse. This category includes legacy data.

### GF: Grant-funded project (data collections)

A collection of data that is associated with a specific funded-project; the purpose is to make these available for other researchers and the public to access and reuse.

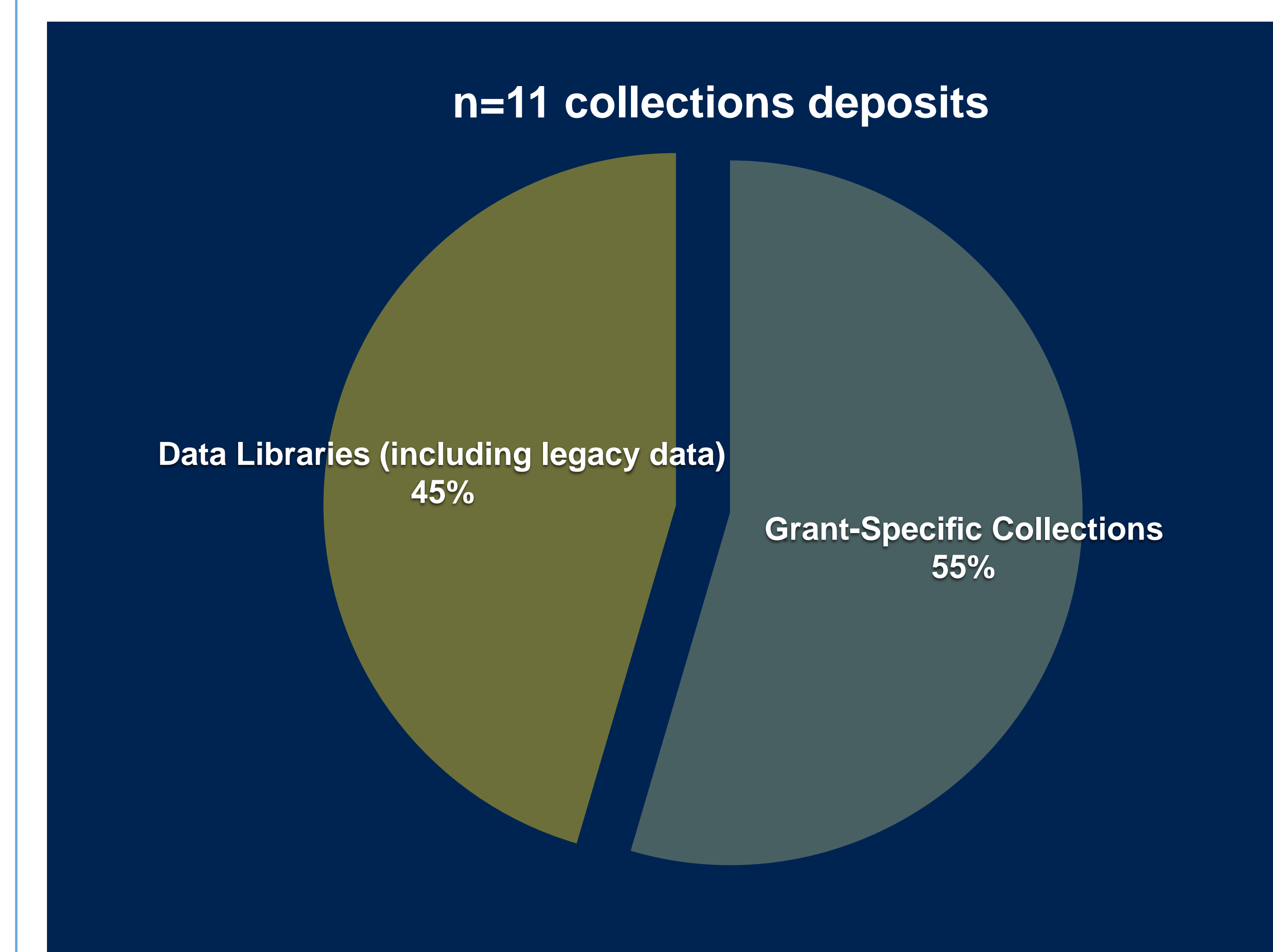
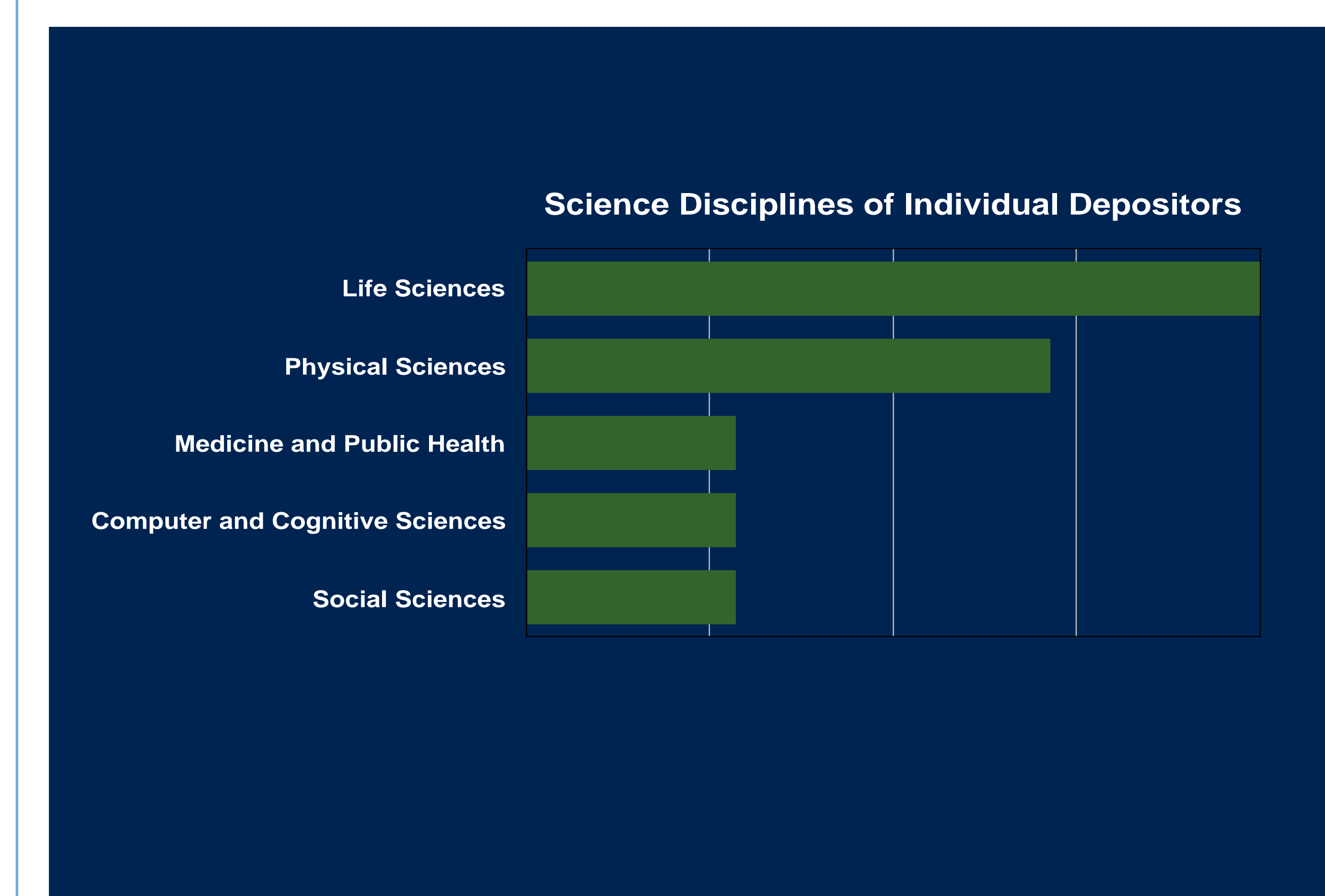
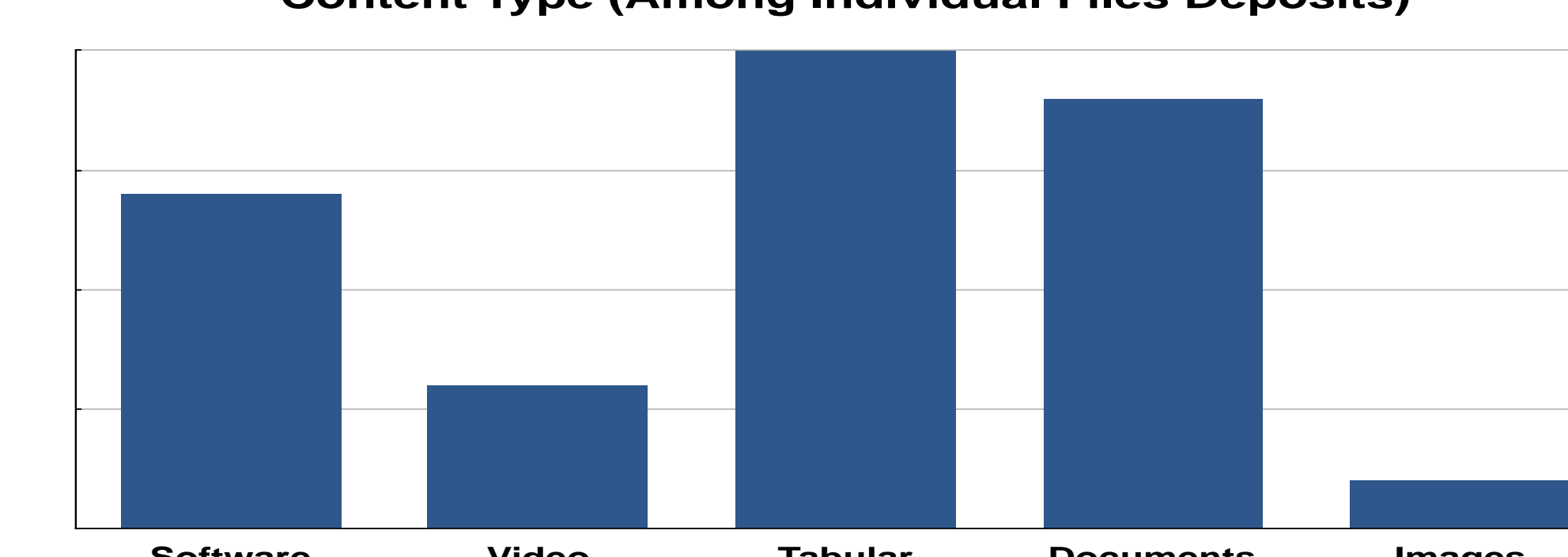
## Results



- The BDR is a means to support reproducibility
- The BDR helps faculty to meet grant-funders' data sharing mandates and publishers' requirements for retaining, citing, and sharing research data
- The BDR offers researchers the means to archive and publish their digital scholarship

## n=20 individual project deposits

Content Type (Among Individual Files Deposits)



## Future Directions

- Begin archiving more digital objects related to undergraduate honors theses, master's-level theses, and doctoral-level electronic dissertations
- Linking and harvesting BDR assets into faculty profile reporting systems, VIVO, and ORCID scholarly profiles
- Enable Google Scholar indexing of BDR assets
- Provide users with analytics, metrics, and evaluation
- Provide users with suggested citations
- Faculty use BDR instead of personal websites and cite BDR assets on their CVs, personal websites, and ResearchGate profiles
- Create digital preservation-related policies