

Shared Repository Infrastructure

CNI Fall 2017 Membership Meeting



- Jimmy Ghaphery, Associate University Librarian for Scholarly Communications and Publishing, VCU Libraries
- jghapher@vcu.edu | [@jimmyghaphery](https://twitter.com/jimmyghaphery)

Measuring our full Elsevier investment

- VCU bepress subscription 2.2% of total Elsevier serial spend
- 26% of VCU collections budget Elsevier and Wiley serial
- VA Elsevier Consortium Total Spend 2017-2021 \$45.47 million
- Does any university know how much? (monographs, textbooks, labor, etc)

2.5% corollary: Vendor subscriptions are an investment in commercial infrastructure

The bepress 115

- 115 “Research Universities” using Digital Commons (DC)
- Publicly available data collected 9/28/2017 (apples::apples)
- 115 of 452 (25%) of all DC institutions
- Est. 60% downloads 75% of items of all DC sites



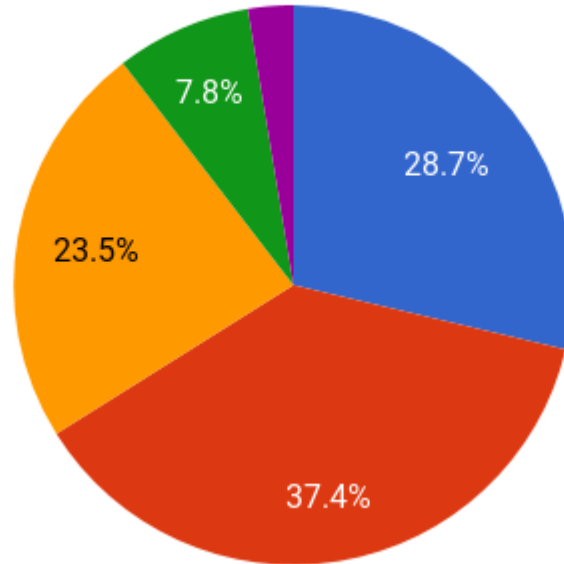
The bepress 115

- Carnegie Classification
- Journal publishing
- ETDs (map data, qualitative notes and quotes)
- Dataset on OSF (DOI: [10.17605/OSF.IO/2KPG6](https://doi.org/10.17605/OSF.IO/2KPG6))



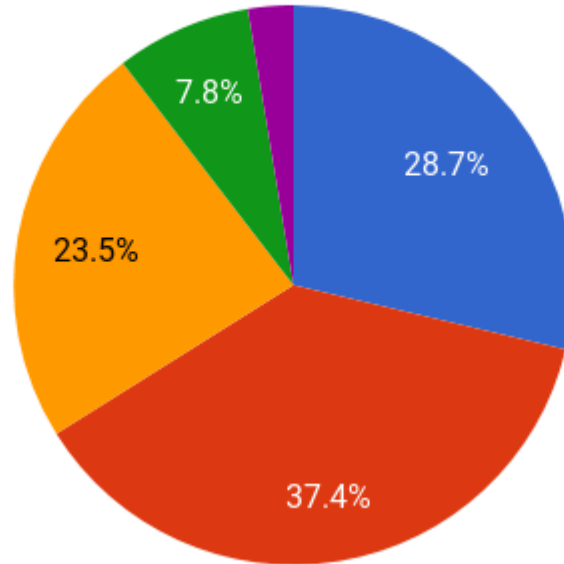
bepress 115 Carnegie Class

● Doctoral Highest ● Doctoral Higher ● Doctoral Moderate
● Master's Colleges & Universities ● Canadian

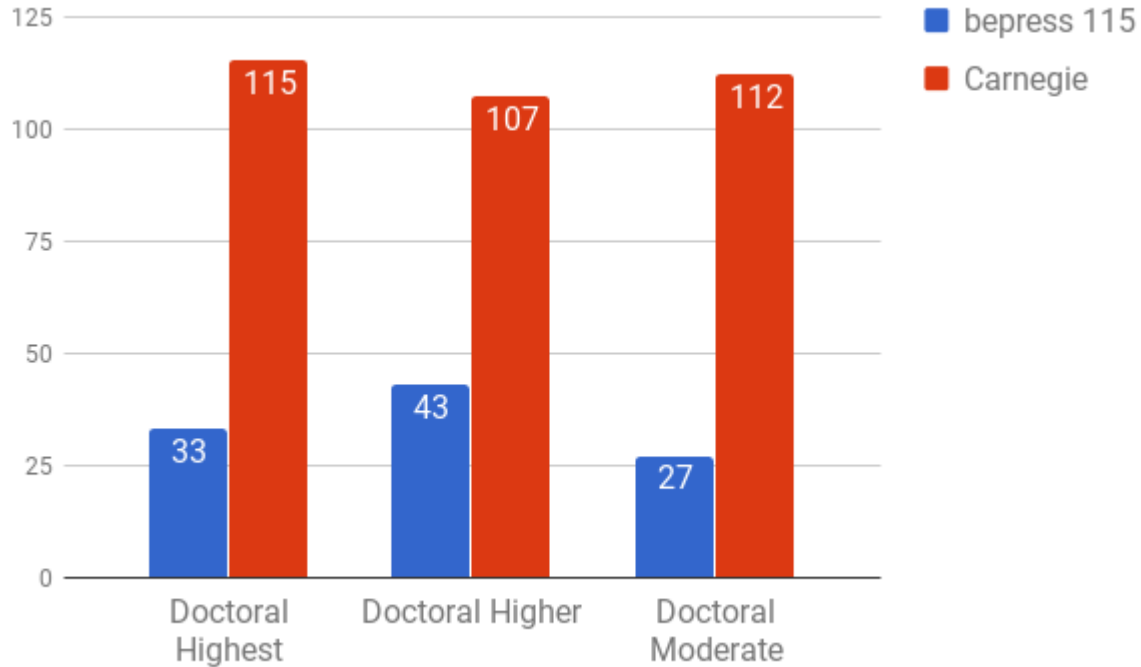


bepress 115 Carnegie Class

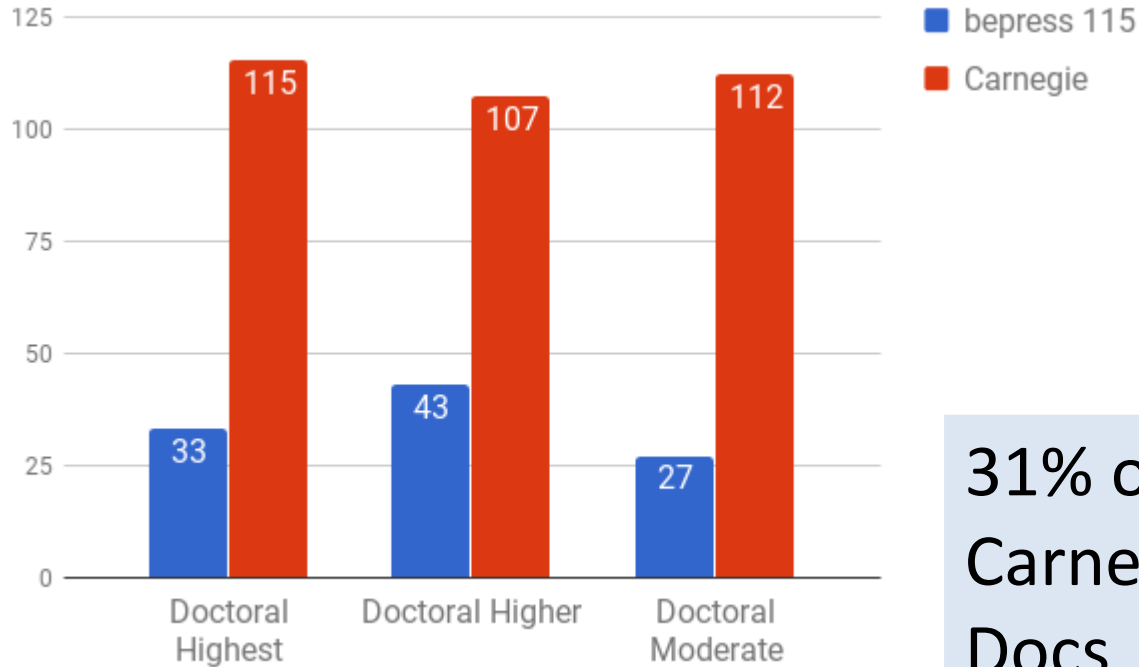
● Doctoral Highest ● Doctoral Higher ● Doctoral Moderate
● Master's Colleges & Universities ● Canadian



bepress 115 and Carnegie Docs

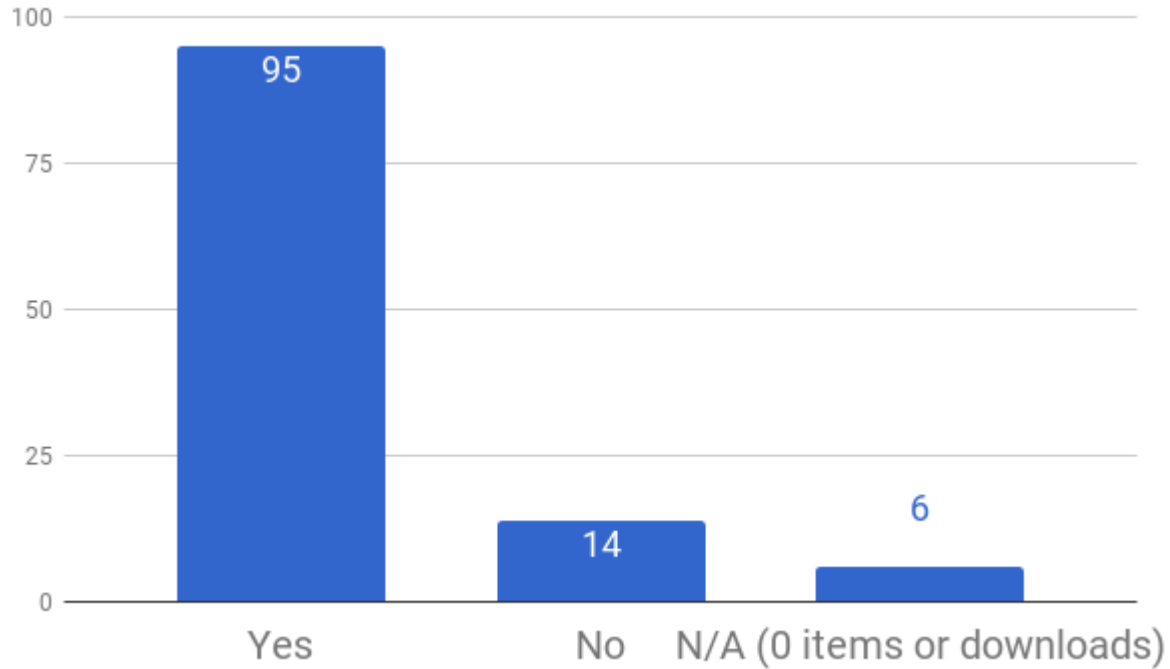


bepress 115 and Carnegie Docs

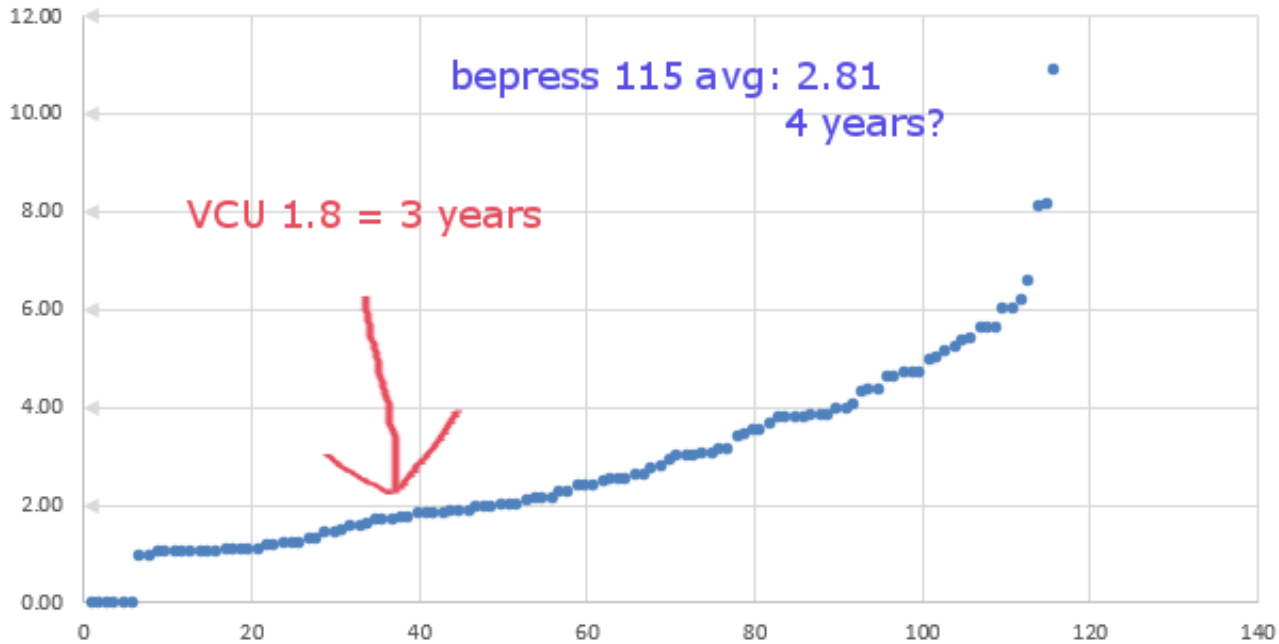


31% of
Carnegie
Docs

bepress 115 Journal Publishing



bepress 115 Age (downloads/downloads past year)



bepress 115 Size and Downloads

- 1.5 million items
- 247 million total downloads
- 62 million downloads in past year
- Top 10: .5 million items 23 million annual downloads

bepress 115 Size and Downloads

- 1.5 million items
- 247 million total downloads
- 62 million downloads in past year
- Top 10: .5 million items 23 million annual downloads

Bepress currently has content containing over 2 million articles and 100 million annual downloads. (Press Release Aug 2, 2017)

bepress 115 Size and Downloads

- 1.5 million items
- 247 million total downloads
- 62 million downloads in past year
- Top 10: .5 million items 23 million annual downloads

Bepress currently has content containing over 2 million articles and 100 million annual downloads. (Press Release Aug 2, 2017)

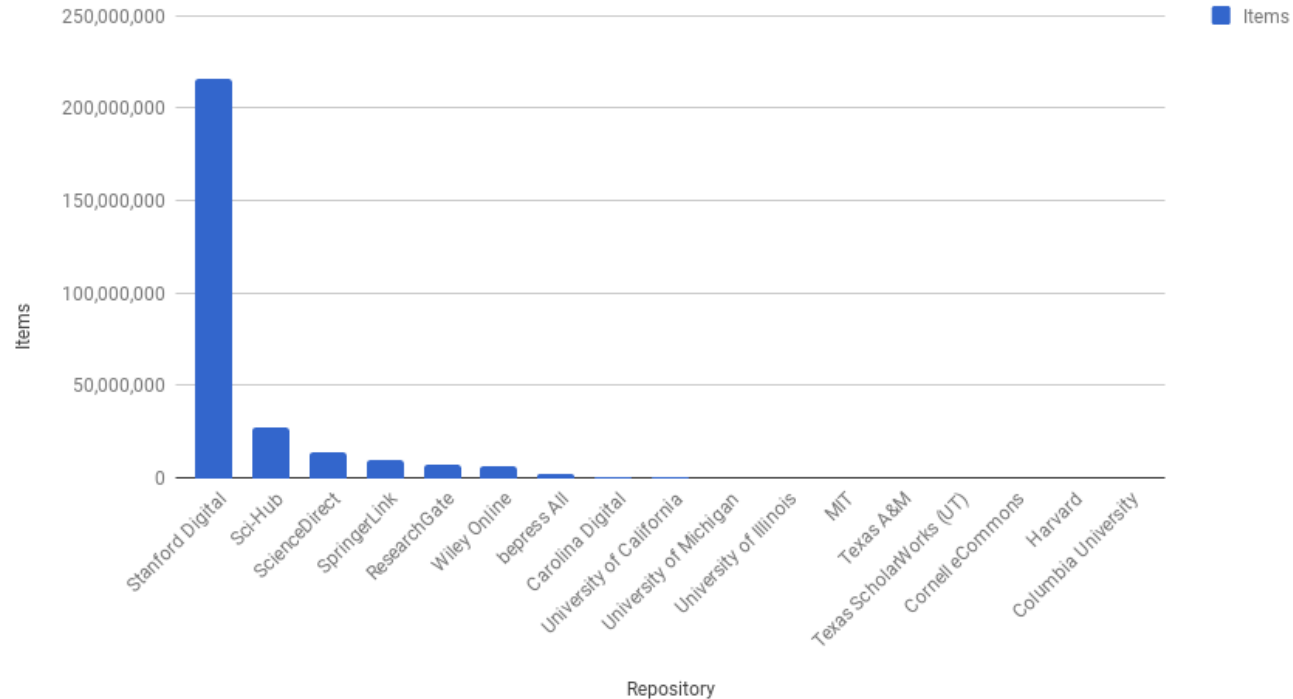
A lot or a little?

Infrastructure: Stats (apples::oranges)

- Bepress items: every published record = 1 item, includes metadata-only record and record with multiple supplemental files.
- Dspace/Samvera: count from browse all titles (similar to bepress definition)
- Stanford/Harvard/MIT: published stats
- Vendors: public stats, many self reported

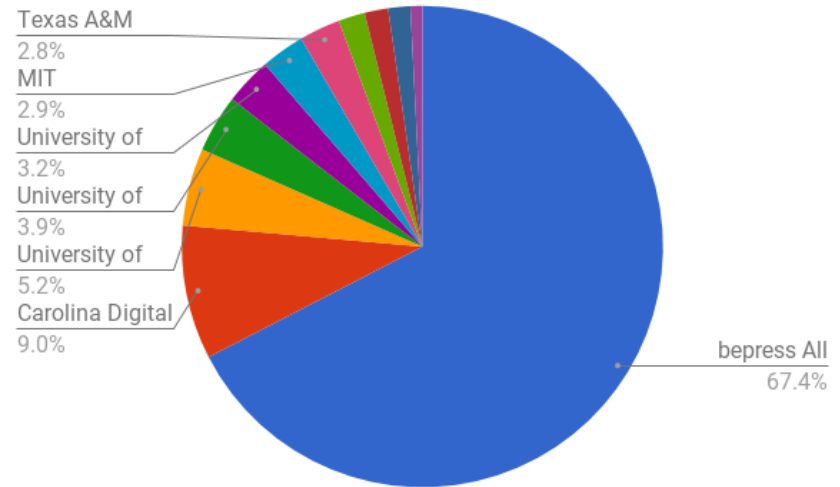
Infrastructure: Size

Items vs. Repository

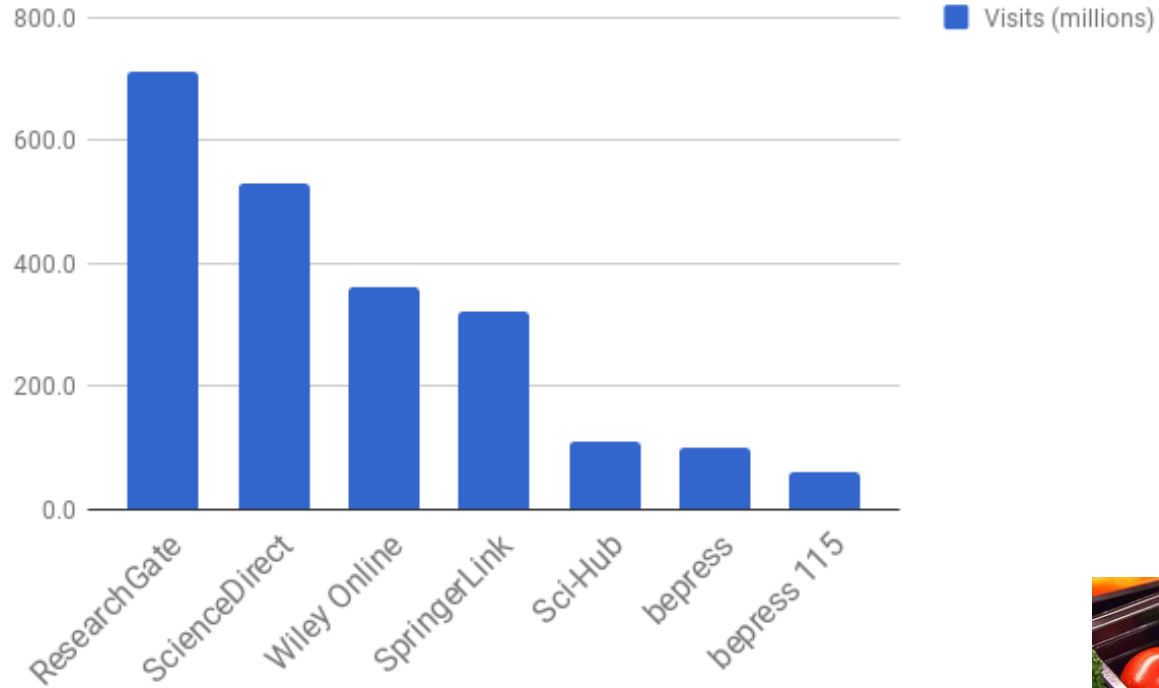


University Infrastructure: Size

| | |
|---------------------------------------|---------|
| Carolina Digital Repository | 266,974 |
| University of California eScholarship | 154,294 |
| University of Michigan DeepBlue Docs | 115,086 |
| University of Illinois IDEALS | 93,973 |
| MIT Dspace@MIT | 87,148 |
| Texas A&M OAKTrust | 81,924 |
| Texas ScholarWorks (UT) | 52,814 |
| Cornell eCommons | 47,032 |
| Harvard Dash | 44,046 |
| Columbia Academic Commons | 23,660 |
| Total | 966,951 |



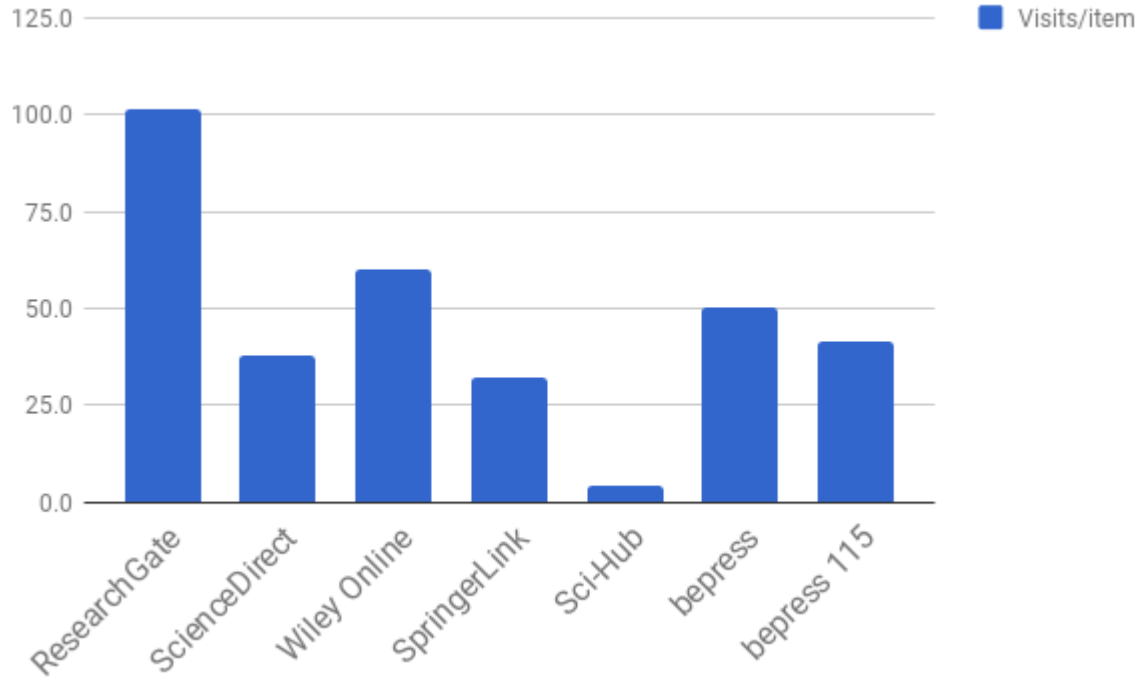
Infrastructure: Annual Visits



bepress=downloads



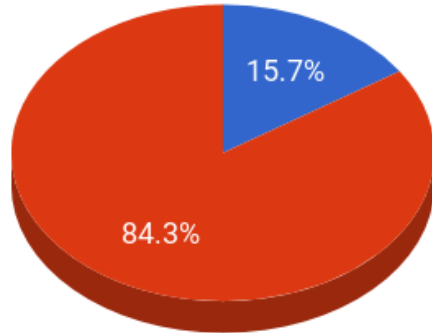
Infrastructure: Visits per item



Bepress 115 ETDs

Items

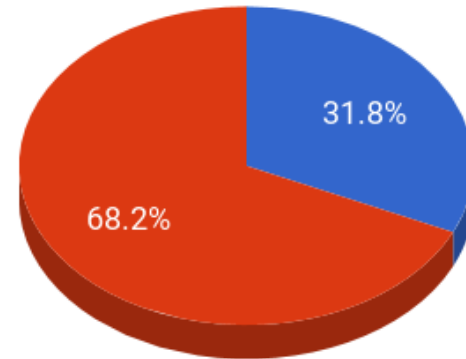
● ETDs ● Not ETDs



72 of 115 bepress institutions

Downloads past year

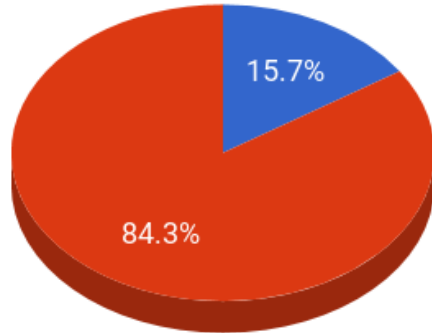
● ETDs ● Not ETDs



Bepress 115: ETDs

Items

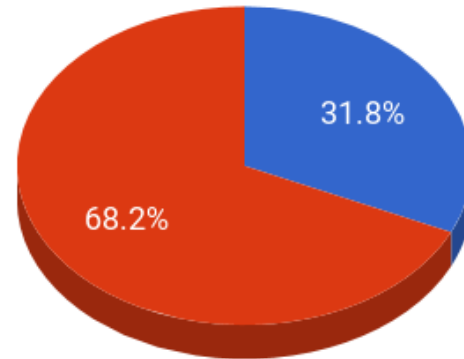
● ETDs ● Not ETDs



Harvard: 9% items | 20% downloads

Downloads past year

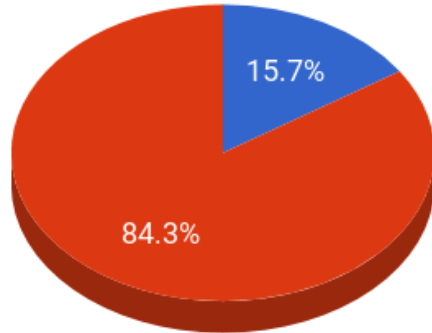
● ETDs ● Not ETDs



ETDs 2:1 rule – Why?

Items

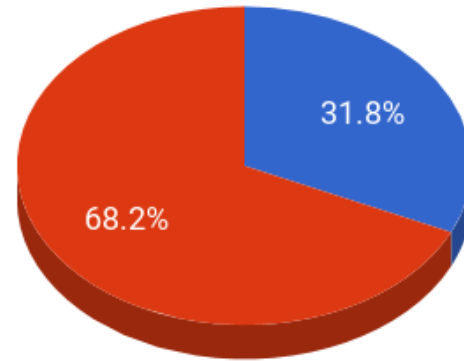
● ETDs ● Not ETDs



Harvard: 9% items | 20% downloads

Downloads past year

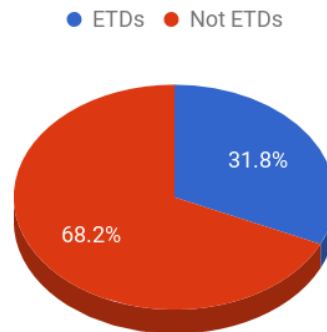
● ETDs ● Not ETDs



ETDs and infrastructure so what

- How Open is it?
- Multiple/Model license considerations
- Education opportunity
- Submission workflow
- National ETD preservation platform “value content transfer”
- Is 15-30% of repository development centered on ETDs?
- OSF for ETDs, Vireo, COAR Next Generation Repositories, ETD community

Downloads past year



The “1 thing” Digital Commons gets right

- Shared publishing/repository
- Readership maps
- Author email alerts
- Faculty profiles
- SaaS
- Others.....

Commercial Goal: Customer Success

- Help and support
- Community of users (platform to policy)
- Backend UI matters
- Broad conception of repository
- Governance

Federated Scale

- Standardization Any color Model T
- Enhancements
- No size or journal limits
- Aggregation
- Benchmarking analytics
- Leverage cloud backups
- Different expectations from in house/subscription services

Credits and References

- Bepress 115 data set. OSF. DOI: [10.17605/OSF.IO/2KPG6](https://doi.org/10.17605/OSF.IO/2KPG6)
- COAR Next Generation Repositories Working Group, [*Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group*](#) (2017.11.28)
- Harrington, Robert . The Scholarly Kitchen. [ResearchGate: Publishers Take Formal Steps to Force Copyright Compliance](#) (2017.10.06)
- [Harvard DASH Stats](#) (2017.10.06)
- Images [Nasa Globe](#) and [USDA Broccoli](#)
- [MIT Open Access Article Statistics](#) (2017.10.02)
- Schonfeld and Ulmschneider. ASERL Webinar. “[Red Light, Green Light: Brainstorming how ASERL members might-could test a new licensing model](#)” (2017.10.13)
- [Stanford Digital Repository Overview](#)
- [Thesis Commons](#) (OSF)
- [Vireo ETD](#)

Thanks and questions

Jimmy Ghaphery

jghapher@vcu.edu