Disciplinary Differences in Citation Aging & its Implications for Librarianship

Sam Hansen
Mathematics & Statistics Librarian
University of Michigan, Ann Arbor

Data

- Web of Science 1900-2017
  - PostgreSQL database created from Clarivate XML files
  - Processed using code from CADRE[a]
- Limited data to publications indexed as articles with at least one citation
- Web of Science Subject Categories
  - 254 Unique Subject Categories[b]
  - Assigned at journal level
  - Not mutually exclusive

Methodology

- Citation age data was first aggregated by Subject Category and then by Year
- Descriptive statistics and metrics were generated from aggregated citation ages using Python and Excel

Metric Definitions

- **Citation Half Life**: Number of years it takes an article to reach half of it total citation count
- **Oldest Citation Age**: The maximum of all the citation ages for an article
- **% of Citations Under X Years**: Number of citations per article younger than X years divided by the citation count

[a] https://cadre.iu.edu/
[b] https://incites.help.clarivate.com/Content/Research-Areas/wos-research-areas.htm

Much of what we believe about **old citation behavior** is **WRONG**.

Citation Aging Metrics

Overall Statistics

Metric	Median	Mean	Median Mean Sum
Half Life	9	14.57	23.57
Oldest Citation	26	30.96	56.96
% Under 5 yrs	40%	46.59%	86.59
% Under 10 yrs	63.33%	65.43%	128.76
% Under 15 yrs	80%	76.94%	156.94
% Under 20 yrs	90.91%	84.12%	175.03
% Under 30 yrs	100%	91.65%	191.65

Oldest Citation Aging Subject Categories

Metric	Median	Mean
Half Life	Law 30.5	Biodiversity Conservation 41.14
Oldest Citation	Biodiversity Conservation 57.5	Biodiversity Conservation 56.62
% Under 5 yrs	Engineering; Geological 20%	Biodiversity Conservation 25%
% Under 10 yrs	Biodiversity Conservation 31%	Biodiversity Conservation 36%
% Under 15 yrs	Mathematics 43%	Biodiversity Conservation 45%
% Under 20 yrs	Mathematics Ecology 50%	Biodiversity Conservation 51%
% Under 30 yrs	Ecology 66%	Biodiversity Conservation 62%

Citation Aging Metrics Data

Youngest Citation Aging Subject Categories

Metric	Median	Mean
Half Life	Physics; Particles Fields 4.57	Physics; Particles Fields 4.34
Oldest Citation	Literary Reviews 11	Literature; African Australian Canadian 10.96
% Under 5 yrs	5 Subjects with 100%	Literature; Romance 99%
% Under 10 yrs	18 Subjects with 100%	Biodiversity Conservation 36%
% Under 15 yrs	37 Subjects with 100%	Literature; German Dutch Scandinavian 100%
% Under 20 yrs	87 Subjects with 100%	9 Subjects with 100%
% Under 30 yrs	191 Subjects with 100%	27 Subjects with 100%

Implications for Librarianship

- Old Research $\not\Rightarrow$ Out of Date Research
- A wide range of subjects regularly cite older articles
  - Many STEM fields regularly cite decades or century old articles
  - Humanities fields often cite the highest percentage of younger articles
- Short citation windows miss lot of research impact
- Selecting materials or publication outlets using metrics like Journal Impact Factor[c] or CiteScore[d] is problematic as their windows miss a majority of citations for most fields
- Materials being old is not a good enough reason to weed them, as researchers across a wide variety of fields use older materials
- There is little researcher behavior to back up stressing currency as a major aspect of evaluating sources when teaching information literacy

Limitations

- Web of Science indexes more STEM content than other subject areas
- Not all Subject Categories are indexed for whole period under investigation

[c] https://www.metrics-toolkit.org/metrics/journal_impact_factor/
[d] https://www.elsevier.com/authors/tools-and-resources/measuring-a-journals-impact