Shifting scope and topics in LIS conferences

A CAIS Study

Geoff Krause
Department of Information Science,
Dalhousie University
gkrause@dal.ca

Presented at BRIC 2025, Montreal, June 5, 2025

Scope of Scholarly Communications

- Scholarly research is disseminated through different venues, notably **journals** and **conferences**
- While these venues may have a **stated scope**, this may differ from their **empirical scope**
- **Lexical similarity** of published works can be used to derive a **quantitative indicator** of scope
- **Conferences** may demonstrate a high **variation in scope** over time, partly due to shifts in their **topical focus** and participating **community**

The CAIS Conference Proceedings

- The Canadian Association for Information Science has held an Annual Conference since 1973
- Proceedings are published with OJS through U. of Alberta (missing 1989, 1991, 1992)
- The conference covers a wide array of work across Library & Information Science
- Early years saw high participation from practitioners in government & industry, focus on information systems & technology
- More recently, shift towards academia & libraries, with high student representation

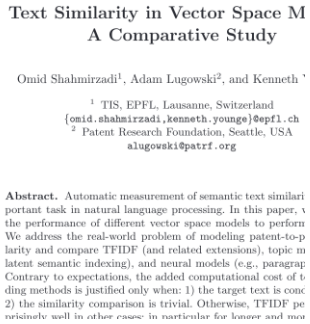
Questions

Can we see a temporal shift in scope within the CAIS Conference Proceedings?

Does this shift correspond to changes in the topical focus of the conference?

Quantifying Scope

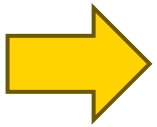
- Metadata and text (PDF) were scraped from the CAIS Proceedings website for all years since 1973 – 1,361 works in total.
- Word tokens were extracted from title + abstract and counted, and cosine similarity calculated for each pair of works
- A work's "fit" is determined by its average similarity to all other works in a given grouping (e.g., year of the conference)
- Scope is determined through the fit of all works


Text Similarity in Vector Space M
A Comparative Study

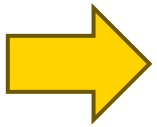
Omid Shahmirzadi[1], Adam Lugowski[2], and Kenneth '

[1] TIS, EPFL, Lausanne, Switzerland
{omid.shahmirzadi,kenneth.younge}@epfl.ch
[2] Patent Research Foundation, Seattle, USA
alugowski@patrf.org

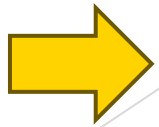
Abstract. Automatic measurement of semantic text similari
portant task in natural language processing. In this paper, v
the performance of different vector space models to perform
We address the real-world problem of modeling patent-to-p
larity and compare TFIDF (and related extensions), topic m
latent semantic indexing), and neural models (e.g., paragrap
Contrary to expectations, the added computational cost of t
ding methods is justified only when: 1) the target text is conc
2) the similarity comparison is trivial. Otherwise, TFIDF pe
riginally well in other cases; in particular for longer and mo

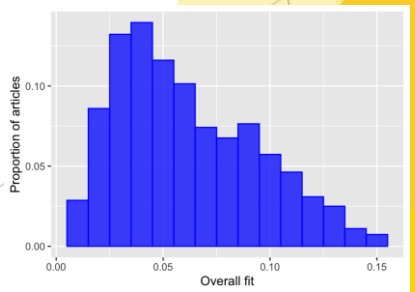


$$[a_1 \quad \dots \quad a_n]$$

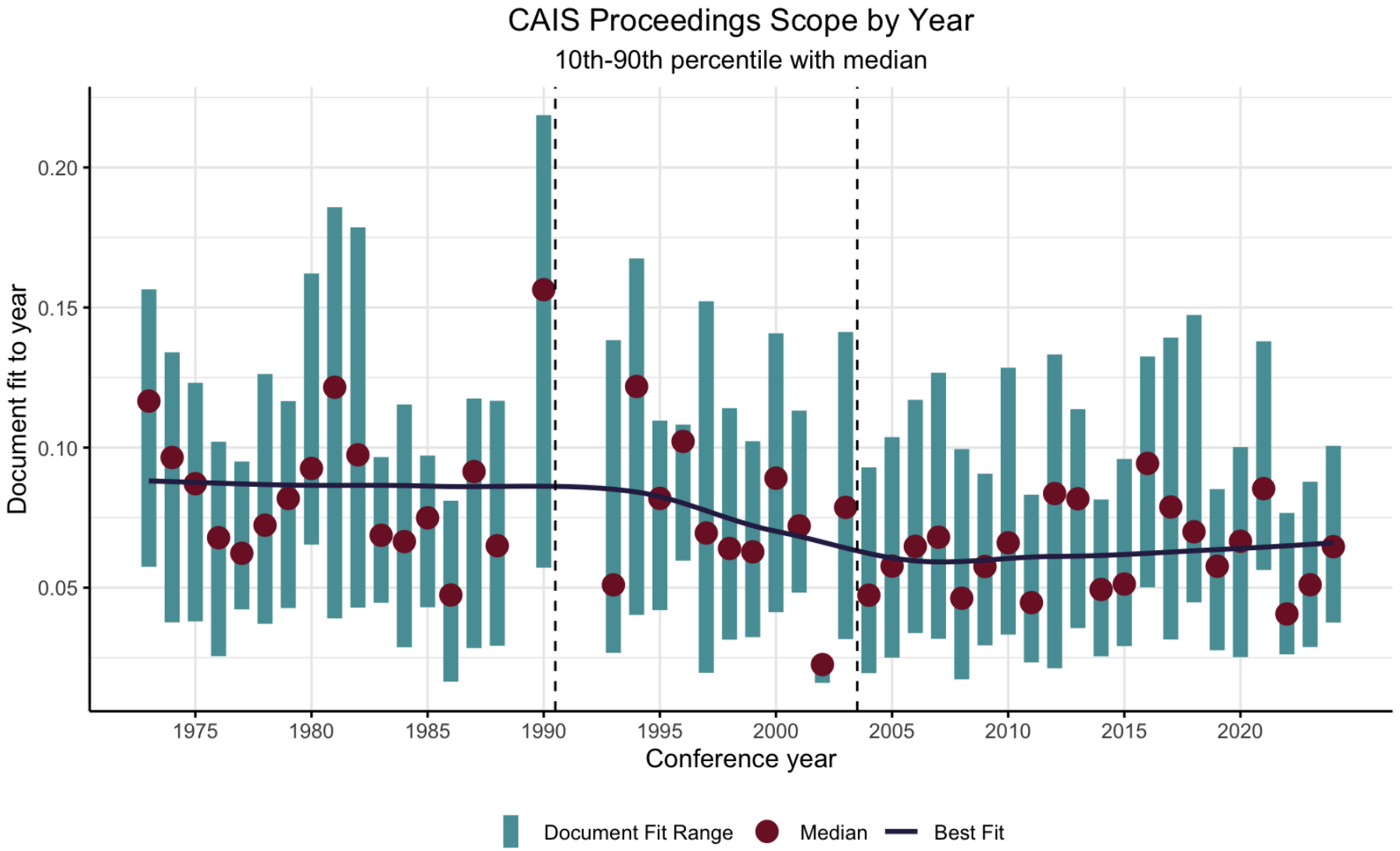


$$\frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \cdot \sqrt{\sum_{i=1}^{n} B_i^2}}$$





Shifting Scope Over Time



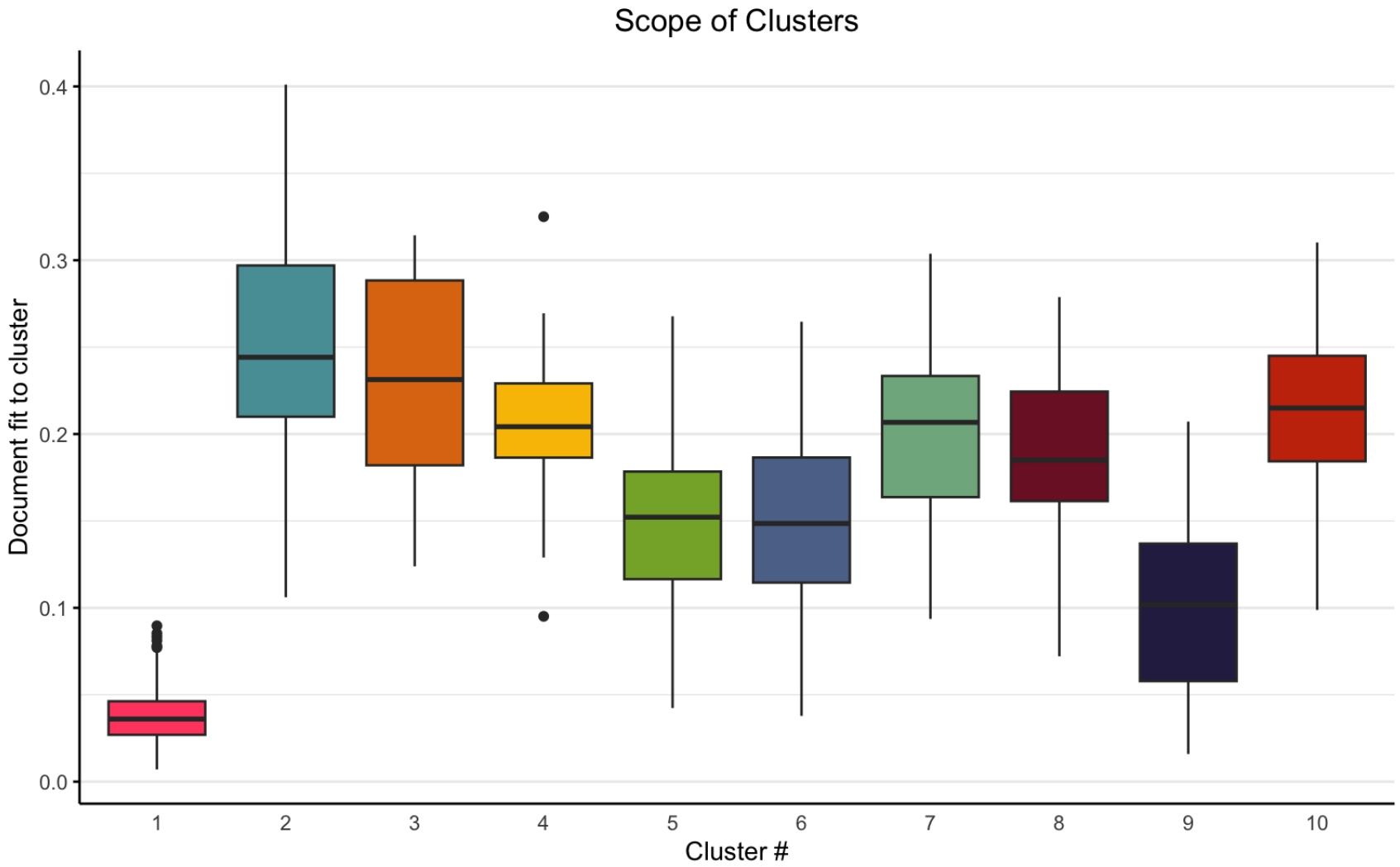
Clustering

Cluster	n	Keywords	Topic
1	629	search, web, study, access, digital, user, analysis, data	*General*
2	193	information, science, health, behaviour, seeking, research, care	Information Seeking/Behaviour
3	39	knowledge, organization, education, cater, constituent, fringe, sider, mix	Knowledge Management
4	29	data, base, bases, census, repositories, access, informetric, sql	Databases
5	78	students, literacy, librarians, school, instruction, reference, academic, skills	Information Literacy & Education
6	104	public, library, libraries, academic, canadian, union, catalogue, services	Academic & Public Libraries
7	40	research, researchers, data, scholarly, collaboration, archeological, ulaanbaatar, conferences	Research & Scholarly Communication
8	57	retrieval, system, line, systems, user, query, language, information	Information Retrieval
9	163	information, services, systems, management, business, lis, practices, technology	Information Systems & Technology
10	29	social, media, npos, profit, organizations, libraries, public, twitter	Social Media

Works were clustered using inverse fit and Hierarchical Agglomerative Clustering

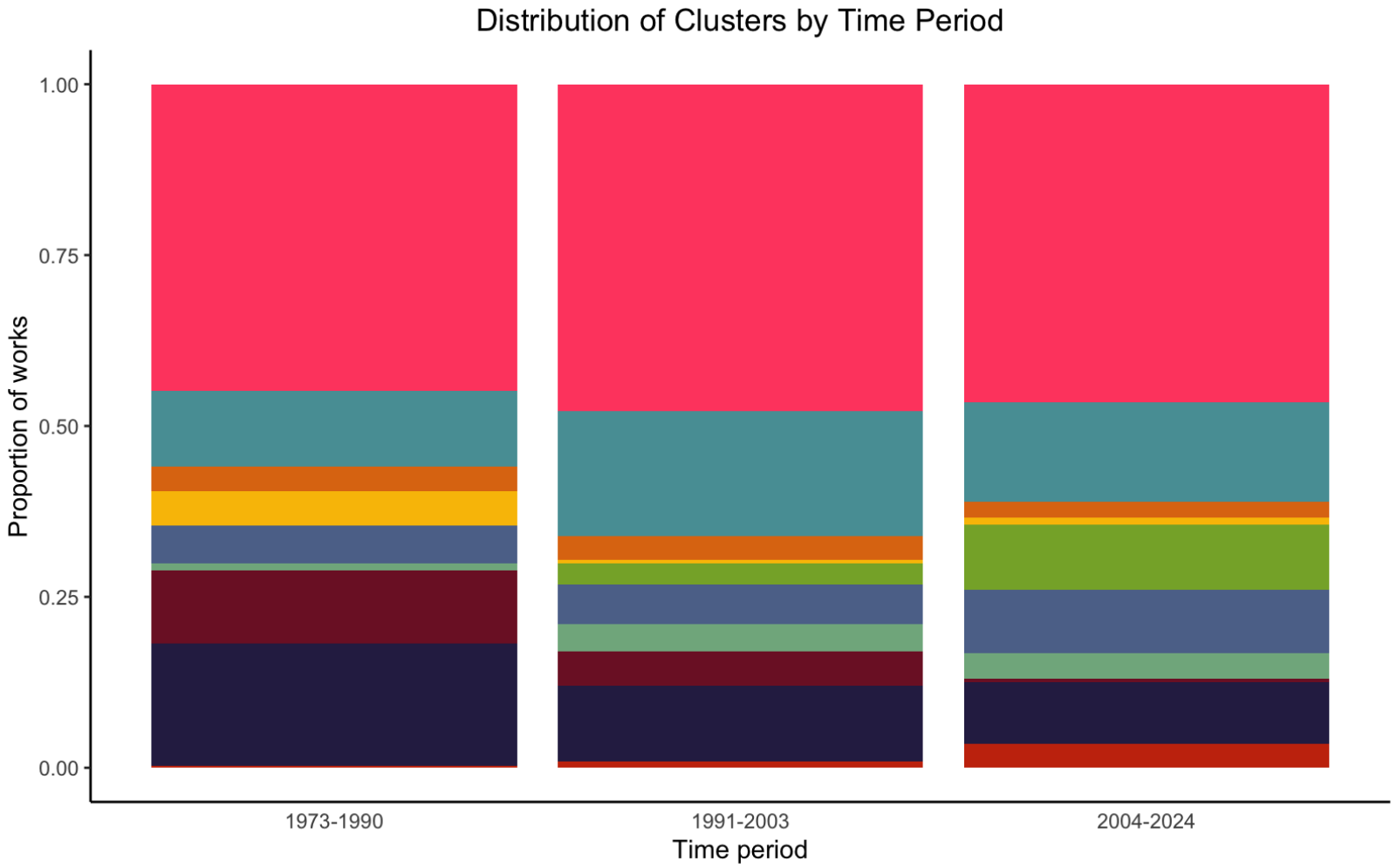
Cluster 1 accounts for nearly half of all works, and has no overarching theme

In-cluster Scope



1 *General*
2 Information Seeking/Behaviour
3 Knowledge Management
4 Databases
5 Information Literacy & Education
6 Academic & Public Libraries
7 Research & Scholarly Communication
8 Information Retrieval
9 Information Systems & Technology
10 Social Media

Cluster Changes Over Time



1 *General*
2 Information Seeking/Behaviour
3 Knowledge Management
4 Databases
5 Information Literacy & Education
6 Academic & Public Libraries
7 Research & Scholarly Communication
8 Information Retrieval
9 Information Systems & Technology
10 Social Media

Shift in Perspectives, Community

- The CAIS Proceedings show a high degree of variation in their scope year-to-year, but can be grouped into periods of higher, changing, and lower similarity
- Clusters are not purely topical, but still reflect certain areas of research & discussion
- Different periods show a different mix of clusters that seem to track changes to the community presenting at CAIS as well as areas of interest

Future Work

- Refining this method of indicating scope
- Application to scholarly journals
- Effects of publication context on a journal's scope

Acknowledgements

This study draws on research supported by the Social Sciences and Humanities Research Council and by Coalition Publica.

References

Canadian Association for Information Science. (n.d.). *Proceedings of the Annual Conference of CAIS*. Retrieved May 9, 2025, from https://journals.library.ualberta.ca/ojs.cais-acsi.ca/index.php/cais-asci/

Julien, H. (2016). Forty-Six Years Young: A Checkup for the Canadian Association for Information Science. *Proceedings of the Annual Conference of CAIS / Actes Du Congrès Annuel de l'ACSI*. https://doi.org/10.29173/cais937

Krause, G., Marjoram, R., & Mongeon, P. (2024, October 18). *Well-tailored words: Comparing the fit of articles within scholarly journals to their citation rates*. 28th International Conference on Science, Technology and Innovation Indicators, Berlin, Germany. https://doi.org/10.5281/ZENODO.13951887

Mongeon, P., Sainte-Marie, M., & Simard, M.-A. (2019). Lexical diversity as an indicator of journal scope. *17th International Conference on Scientometrics and Informetrics, ISSI 2019*, 2547–2548.

Murtagh, F. (1985). Multidimensional Clustering Algorithms. *Compstat Lectures*, *4*.

Nilsen, K. (2007). The Canadian Association for Information Science: A Look at Its Thirty-Five-Year History. *Canadian Journal of Information and Library Science*, *31*(2), 163–177.