Science maps for information retrieval

Bascur Cifuentes, J.P.

Related to the dissertation

1. Cluster-based information retrieval works best when clusters have high granularity, but this creates too many clusters to inspect individually. Visualizing how clusters relate to each other makes it possible to use many clusters without manually checking each one. *(Chapter 2)*

2. Systematic reviews aim to cover all relevant papers on a topic but rely too heavily on Boolean queries, whose performance depends on the choice of search terms and the prior knowledge of experts. Cluster-based information retrieval compensates for these limitations by identifying relevant papers beyond the predefined query. *(Chapter 3)*

3. Citation-based maps dominate scientometric studies but cannot be used when papers have weak citation connectivity. Text-based maps deserve greater attention because they capture similar topics to citation-based maps and rely on more accessible data. *(Chapter 4)*

4. Coauthorship data are usually used to map collaborations between authors, but the same data can also be used to map research topics. Research topics revealed by coauthorship-based maps differ from those identified using citation- and text-based maps. *(Chapter 5)*

Related to the field

5. Universities spend millions on journal access but almost nothing on sustainable science-mapping software, even though such tools would make their investment actually pay off.

6. In science mapping, we tend to treat everything as a network, which is useful for representing discrete relationships. However, for text similarity, it may be more appropriate to use dimensionality reduction of text embeddings.

7. One of the main strengths of science maps is that they let the data speak for itself. Using supervised machine learning to improve topic representation risks silencing that voice by forcing the data into a predetermined paradigm.

8. Cleaning data before network creation, such as removing Twitter bots, can improve clustering quality but may also amplify curator bias, posing the same risk as discussed in proposition 7.

Related to society

9. Public mistrust in science grows partly because people do not have a good overview of what the scientific community is doing. Open and accessible science maps can increase public trust by allowing people to explore scientific literature directly instead of relying on social media gatekeepers.

10. Good bibliometric analyses do not require deep expertise; they require dedication and curiosity. Easy-to-use, open science maps can empower citizen scientists to conduct their own bibliometric analyses without the need for in-depth technical expertise.