Full-text citation pattern analysis based on the enriched cited references: A discipline comparison perspective

Fengxuan Lv (1092386442@qq.com), Wuhan University

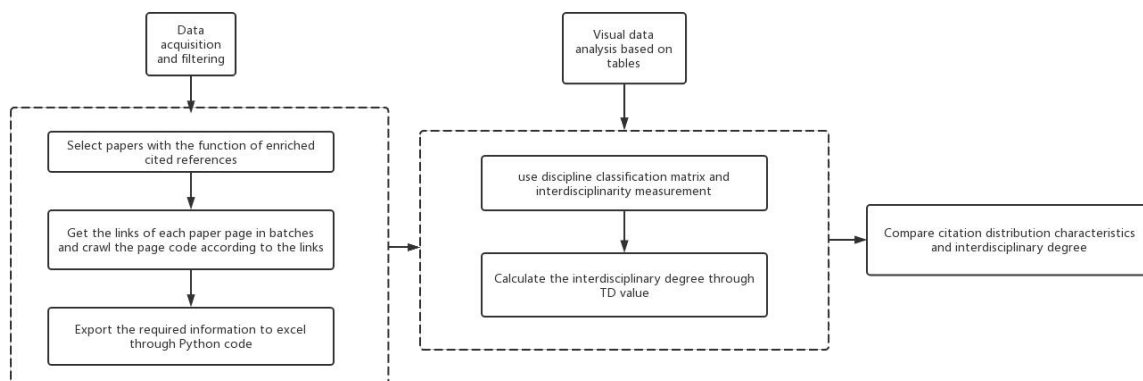Luhe Wang (wangluhe506@163.com), Wuhan University

Yifei Yu (yifeiyu@whu.edu.cn), Wuhan University

Background

In traditional citation analysis, there are some shortcomings in the evaluation index of papers, such as scattered, superficial, one-sided and single-dimension. The citation analysis of the full text is an innovation of the traditional citation analysis, which is a part of the full text analysis[1,2]. It has gradually attracted attention since 2014. Its research content mainly includes the citation position, citation intensity and citation emotion of the cited papers. At present, the difficulties faced by the full-text citation analysis include the difficulty of obtaining data, the lack of comprehensiveness and accuracy of data, the limited availability of the full-text citation and so on. Research gap at this stage is that more standardized and compatible models are needed, as well as more scientific and comprehensive research methods. Web of Science has introduced the enriched cited references, which reduces the obstacles for the citation analysis of the full text.

Methodology

In Humanities, Sociology, Basic Sciences and Life Sciences, four appropriate disciplines are selected as data sampling and research objects. Based on the in-depth analysis function data of cited references provided by Web of Science Core Collection, we take the papers in the categories of library and information sciences, bioinformatics, mathematics and history published in 2021 as research samples. Then we select the papers that have the function of enriched cited references. For those papers with enriched cited references module, we obtain the relevant chart information and corresponding paper description information cited in the full text in batches. With the help of text analysis and visualization tools, we analyze citation patterns from four aspects: citation location, citation intensity, citation context and citation emotion. On this basis, with the help of discipline classification matrix and interdisciplinarity measurement, the degree of intersection on different paper chapters and functional modules is measured respectively, so as to reveal the interdisciplinary characteristics of the targeted paper. Calculate the interdisciplinary degree through TD value[3]. Comparing citation distribution characteristics and interdisciplinary degree, we can profile the interdisciplinary characteristics of different research fields.
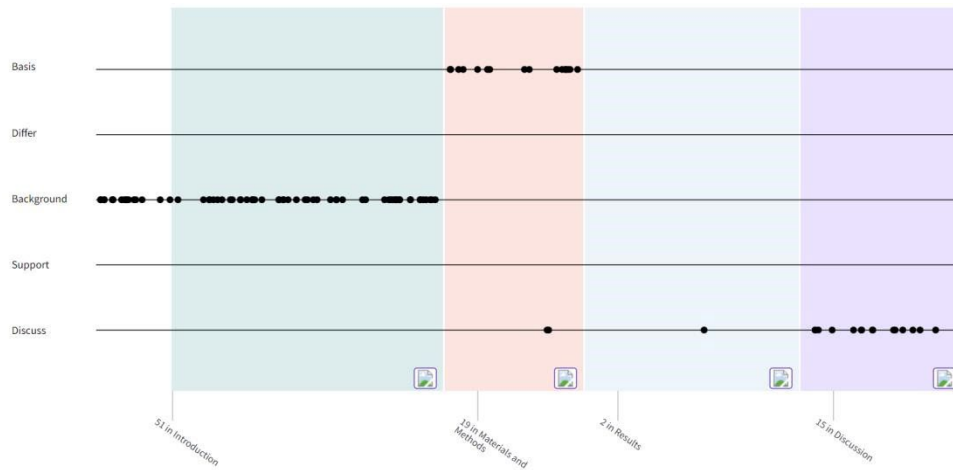
(Flow chart of the data acquisition method)

Results

We believe that different types of disciplines have different citation characteristics. Papers in Humanities and Social Sciences involve more interdisciplinary research methods and comprehensive research and are more dispersed in citation locations. For example, citation patterns in the field of Library and information technology have distinct thematic differences, and the intersection degree of different functional modules shows obvious discrete characteristics. Papers in Basic Sciences will focus more on citing basic formulas and basic theories as background. Papers in Emerging Sciences such as in the field of Life Sciences show the diversity and obvious intersection of citation disciplines.

Conclusion and Discussion

The function of enriched cited references in Web of Science is a further classification and visual presentation of the cited papers. It creatively realizes the standardization of citation analysis in full text, which is of great value to reveal the characteristics of citation patterns and explore the intersection of disciplines. It can deepen the understanding of the connotation of citation, then better assist scientific evaluation. Citation patterns in different fields have distinct thematic differences. Web of Science provides a two-dimensional full-text analysis chart, which makes the relationship between the original paper and the cited papers more clear and obvious. It is conducive to the understanding of the paper and the scientific research in specific fields. At the same time, it shows the citation status and characteristics of the paper. Multiple citations of the same paper are also reflected in the chart. The disadvantage is that there is no clear distinction between self citation and other citation, and the standardization of tables needs to be further improved.

(An example of enriched cited references in Web of Science)

References

[1]Yang Siluo, Nie Ying.:Research on Evaluation Model of Papers' Influence Combined with Full-text Analysis. Journal of Modern Information. 42(03), 133-146(2022).

[2]Patrick Glenisson, Wolfgang Glänzel, Olle Persson.:Combining full-text analysis and bibliometric indicators. A pilot study, Scientometrics 63, 163-180(2005).

[3]Zhang Lin, Liu Dongdong, Lyu Qi, Sun Beibei, Huang Ying.:Interdisciplinarity Measurement in Publications: From Full Reference Analysis to Sectional Reference Analysis. Journal of the China Society for Scientific and Technical Information, 39(5), 492-499(2020).