Research in progress

STI 2022 Conference Proceedings

Proceedings of the 26th International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Proceeding Editors

Nicolas Robinson-Garcia
Daniel Torres-Salinas
Wenceslao Arroyo-Machado

Citation: Sun, R., Zhang, T., Fensel, J., Yu, A., Bu, Y., & Xu, J. (2022). Disciplinary Boundary for Knowledge Diffusion: A Case Study in Psychology and Computer Science. In N. Robinson-Garcia, D. Torres-Salinas, & W. Arroyo-Machado (Eds.), *26th International Conference on Science and Technology Indicators*, STI 2022 (sti22104). https://doi.org/10.5281/zenodo.6948502

Collection: https://zenodo.org/communities/sti2022grx/

Disciplinary Boundary for Knowledge Diffusion: A Case Study in Psychology and Computer Science[1]

Ran Sun[*], Tongyang Zhang[**], Julia Fensel [***], Andrew Yu [***], Yi Bu[****] and Jian Xu[**]

[*] *s_ran131497@whu.edu.cn*
School of Information Management, Wuhan University, Wuhan, Hubei, 430072 (China)

[**] *zhangty65@mail2.sysu.edu.cn; issxj@mail.sysu.edu.cn*
School of Information Management, Sun Yat-sen University, Guangzhou, Guangdong, 510006 (China)

[***] *Julia.smag1@gmail.com; andrewyu1080@gmail.com*
Westlake High School, Austin, Texas, 78746 (America)

[****] *buyi@pku.edu.cn*
Department of Information Management, Peking University, Beijing, 100871 (China)

Background

Due to differences in epistemologies, investigated area and research methods distinguishing disciplines from one another (Smelser & Baltes, 2001), scientific discipline can be seen as a primary unit to differentiate science which forms the fundamental boundaries crossing distinct systems of knowledge called disciplinary boundary (Dirks, 1996). Disciplinary boundary can reshape the communication of knowledge between disciplines, thus leading to differing impacts of a same knowledge in different areas. With the continuous and in-depth development of disciplines, the hindering effect caused by discipline barriers is strengthened.

To further understand the impact of disciplinary boundary on knowledge diffusion, this study focuses on a case study that demonstrates the phenomenon that the same knowledge published in two versions can keep on exerting vastly different effects in different disciplinary channels. We found that Matthew Turk and Alex Pentland, scholars at the Massachusetts Institute of Technology, published two papers on a computational eigenfaces approach. Among them, "Eigenfaces for Recognition" (Turk & Pentland, 1991a) (Article_PSY) is published on January 1, 1991 in a psychological journal, while "Face Recognition Using Eigenfaces" (Turk & Pentland, 1991b) (Article_CS) is another version presented at a computer science conference on June 3-6, 1991. Article_PSY was once quoted by Article_CS. The two articles have almost the same content except for a small portion of psychological content for the theoretical validation of the results only covered in the Article_PSY. The two articles have aroused a profound impact and have been cited 15,859 times until 17th April 2019.

Through tracing the path of how Article_PSY has been diffused in Psychology and how its companion Article_CS has been diffused in Computer Science, we demonstrate different

[1] This work was supported by National Social Science Fund of China [18BTQ076].

influences of the study in these cross-disciplinary knowledge flows. Findings of this case study have important implications for exposing the functionality of disciplinary boundary and how it reshapes the diffusion process of knowledge.

Data and Methodology

11,792 citations for Article_PSY and 4,067 citations for Article_CS were collected in the Microsoft Academic Graph (MAG) dataset. The year span of collected citations is from the publication date (January 1991) of the target article to 2020. Useful metadata includes title, year, date, abstract, the normalized discipline name assigned to the paper, etc. are also extracted for the following investigation.

To analyze the diffusion process of an article across Psychology and Computer Science, we apply LDA (Blei et al., 2003) to understand which topics were brought up, the chronological order of the most attracted topics in the process of adopting the article, and the most relative topics corresponding to each discipline in the two different diffusion channels. The parameters of LDA are set as follow: alpha = 0.3, beta =0.1, iterative = 200. It was observed that the 30-topic model is the best in its fitting between interpretability and the size, processing duration, and diversity of the dataset.

The popularity of a topic g can be calculated by aggregating the popularity of each topic per document obtained in LDA (Zhai et al., 2018). The topic popularity *Pop*(g) of the topic g in year t can be calculated as:

$$Pop(g|t) = \sum_{d|py(d)=t} pop(d) \quad (1)$$

In the formula, *py*(d) refers to the publication year of document d, *pop*(d) denotes the topic proportion of document d.

Empirical Study and Results

We conduct a topic modeling analysis of citations of the two articles to illustrate the differences of their impact on different discipline channels and reveal the impact of the disciplinary boundary.

Interdisciplinary Influence of a Study

Figure 1 shows the annual distribution of the citations of the two articles from 1990 to 2020. It also illustrates the annual percentage of these citations in the total number of papers published in each discipline in the MAG. Differences in citation patterns, influence scale, and persistence of impact across disciplines are also indirect evidence of disciplinary boundaries.

Figure 1: Annual distribution of citations in the top 10 most productive disciplines in MAG.

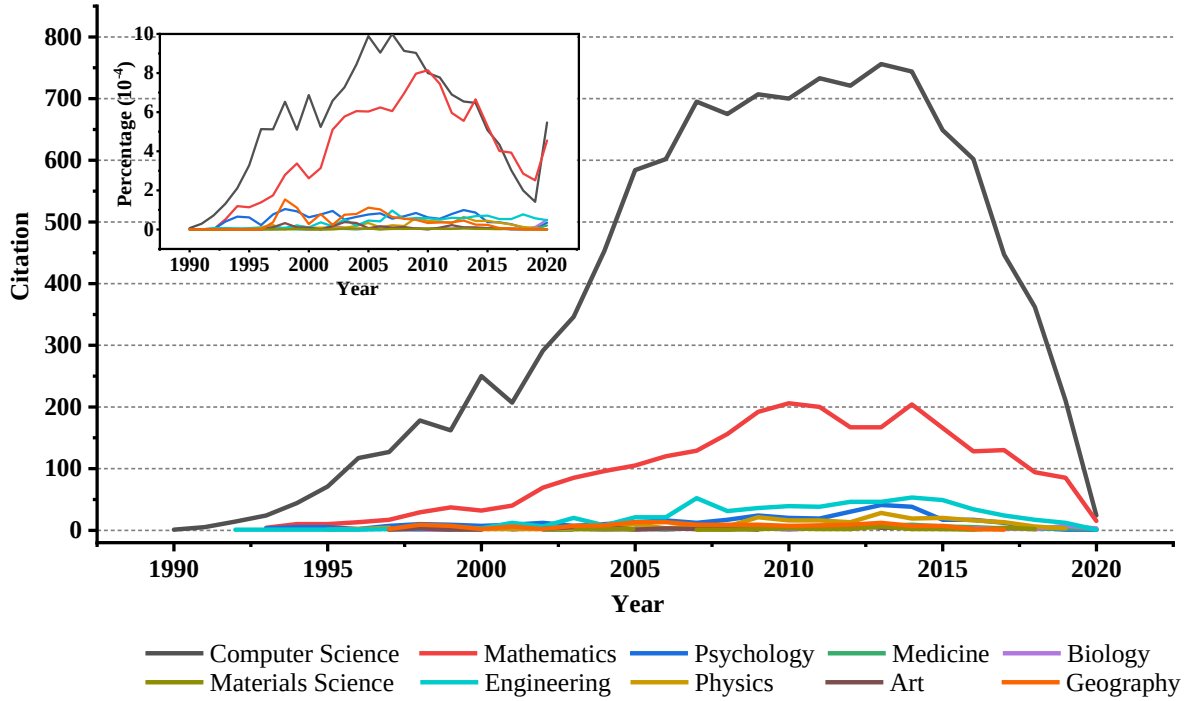
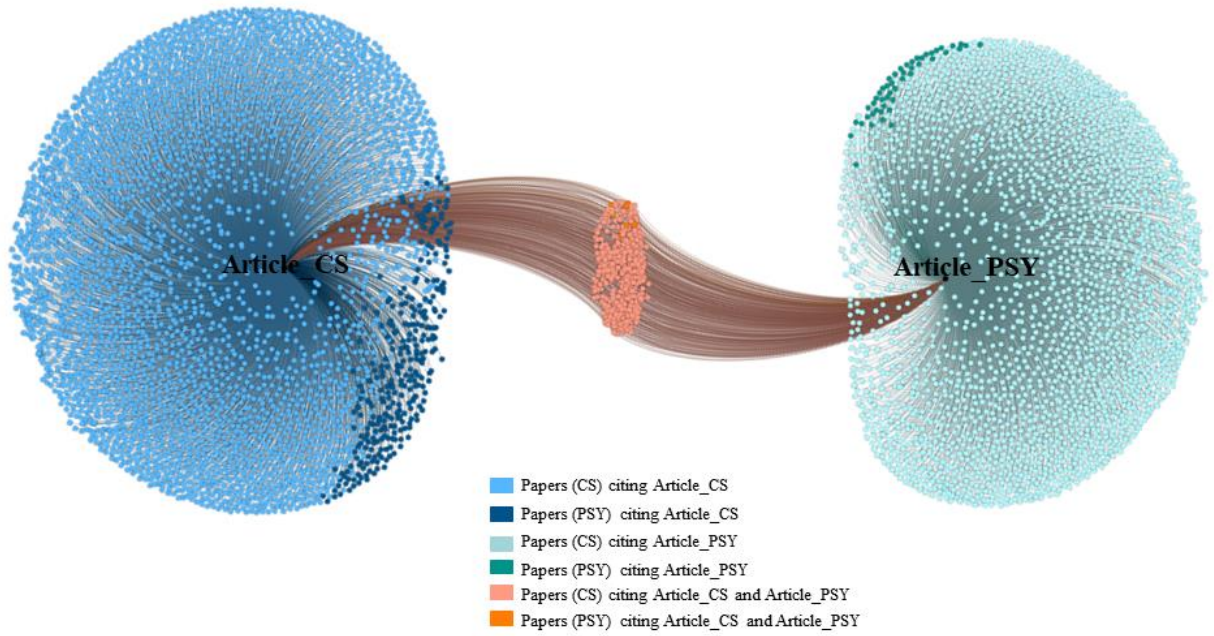


Figure 2 shows the citation relationship of the two papers in the fields of Computer Science and Psychology. Article_PSY has been cited more by Computer Science, but Article_CS has been cited less by Psychology, which indicates that the influence of disciplinary boundary of Computer Science on Psychology is different from that of that of Psychology on Computer Science.

Figure 2: The citation relationship of the two papers in the fields of Computer Science and Psychology



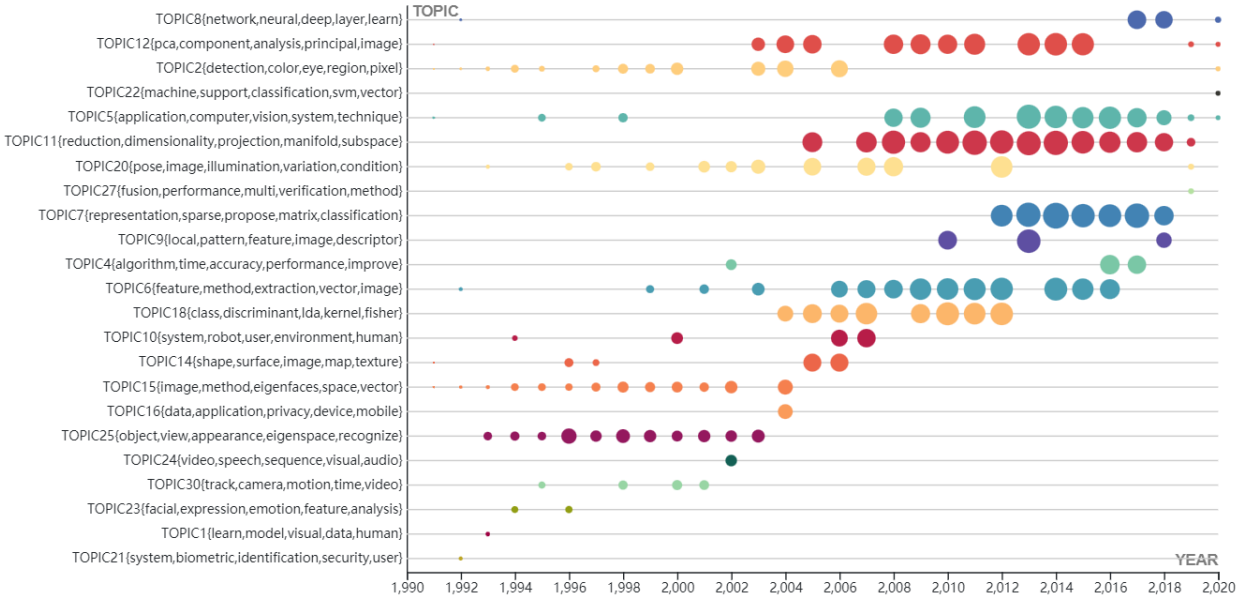
Analysis of Topic Evolution of Citations

By analyzing the yearly distribution of the top 5 popular topics generated from citations of Article_PSY and Article_CS, as shown in Figures 3 and 4, we examine various stages of the knowledge diffusion process across disciplines.

Figure 3: Evolution of the five most popular topics in each year of Article_PSY citations. (Note: The ordinate axis shows the top 5 keywords of each topic. The size of nodes represents the level of popularity of the topic in that year.)

As shown in Figure 3, the evolution of topics in the psychology diffusion channel has mainly experienced three stages. From 1990 to 1999, topics related to computer vision application, such as topic9 and topic22, have had strong growth. From 2003 to 2015, studies related to classical face recognition algorithms, such as topic2 and topic18, are becoming increasingly popular. After 2015, a facial signal processing method based on sparse representation (topic16) has become popular. At this stage, PCA, SVM, and other algorithms have gradually been replaced by manifold learning and dictionary learning.

Figure 4: Evolution of the five most popular topics in each year of Article_CS citations.



As shown in Figure 4, the evolution of topics in the computer science diffusion channel has also experienced three stages mainly. From 1990 to 2002, growing topics are primarily concerned about the artificial selection of human features, such as the color of face images (topic2). Meanwhile, increasing applications of face recognition in the way of human-robot interaction of topic10 and topic30 also emerge. From 2003 to 2012, topics related to PCA and manifold learning are growing in popularity. After 2012, the sparse representation-based classification algorithm of topic7 begin to emerge. Recently, deep learning technology that automates face recognition in topic8 has been further applied.

Conclusion

Taking a study published in different fields as a case, we compare the impact of a study disseminated in different disciplines to reveal the disciplinary boundaries during the knowledge diffusion process. Results show that depending on different diffusion disciplines, the academic impact of the same study can obviously differ at the topic level.

From the perspective of citation topics, it can be observed that the research topics triggered by the two articles have been considerably different over years. First, different topics have remained highly popular for a long time in the diffusion of the two articles, such as SVM (topic18) and LBP (topic24) in Article_PSY citations and sparse representation (topic7) and manifold learning (topic11) in Article_CS citations. Second, the point of popular time of topics in different disciplines is different. For example, the popularity time of topics about PCA, SVM, is 2002 in the diffusion of Article_PSY, earlier than that of Article_CS. Third, the focus of research topics affected by the same study diffused in different disciplines is different. For example, the topics of citations of the two articles have different emphases. Lots of topics in citations of Article_PSY are directly related to face recognition methods such as LDA, PCA, and manifold learning, while feature extraction of multimodal data is more abundant in citations of Article_CS such as eye region and image eigenfaces.

The in-depth differential analysis of the knowledge flow between disciplines is helpful for us to understand the communication of knowledge at a micro level and more importantly, to reveal the existing disciplinary boundary. In the future, we will further identify research topics in different disciplines and determine how the same knowledge affects disciplines differently.

References

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research, 3*(Jan), 993-1022.

Dirks, Arthur L. (1996). Organization of knowledge: The emergence of academic specialty in America. Published on-line by author. Retrieved June 25,2021 from: http://webhost.bridgew.edu/adirks/ald/papers/orgknow.htm

Smelser, N. J., & Baltes, P. B. (Eds.). (2001). *International encyclopedia of the social & behavioral sciences* (Vol. 11). Amsterdam: Elsevier.

Turk, M., & Pentland, A. (1991a). *Eigenfaces for recognition. Journal of cognitive neuroscience, 3*(1), 71-86.

Turk, M., & Pentland, A. (1991b, January). Face recognition using eigenfaces. In *Proceedings 1991 IEEE computer society conference on computer vision and pattern recognition* (pp. 586-587). IEEE Computer Society.

Zhai, Y., Ding, Y., & Wang, F. (2018). Measuring the diffusion of an innovation: A citation analysis. *Journal of the Association for Information Science and Technology, 69*(3), 368-379.