https://doi.org/10.51408/issi2025_172

A look behind metrics for knowledge integration: Some notable cases

Pei-Shan Chi[1], Wolfgang Glänzel[2]

[1]*peishan.chi@kuleuven.be*, [2]*wolfgang.glanzel@kuleuven.be*
Faculty of Economics and Business, ECOOM, KU Leuven, Naamsestraat 61, 3000, Leuven (Belgium)

Introduction

In earlier papers, we have used the analysis of cited references to study cognitive aspects of interdisciplinarity (IDR) in scientific research. We assumed IDR being an expression of knowledge integration that can be traced by analysing cited references, which in turn are considered a form of use of scientific information in the framework of documented scholarly communication. Yet, measure based on cited references tend to overestimate cognitive links in favour of methods and instruments used (e.g., Glänzel & Thijs, 2017). The same applies to IDR measures, if those are based on citation links. In particular, we found that while, at the nano level, the distinction between IDR and multi-disciplinarity is straightforward, the distinction between IDR and cross-disciplinarity (CDR) remains a challenge. We used variety and disparity measures to describe important characteristics of IDR, but found striking examples, notably of high disparity, in which the extent of knowledge integration is questionable. Papers in archaeology and religion, in which advance imaging technologies or instruments were used and referred to in the bibliography without true integration of the underlying knowledge into the research, may just serve as an example. The citers of these studies typically remained in the field of archaeology or religion. This forced us to assume that the role of users (citers) also play an important role in the understanding of knowledge integration.

Basic approach

We proceed from the assumption that true knowledge integration takes place if and only if some new research is established combining sources from different disciplines or is used in research of one of the disciplines referred to, but information is used by other fields and not only by this discipline. Figure 1 gives an illustration of three typical examples. Top-left corner: two disciplines are cited, but only one of those cite the results. Top-right corner: the same references, but both disciplines are using the results. Bottom: knowledge from two disciplines is combined in a different discipline and cited there. The first case may not reflect true integration of knowledge.
This leads us directly to the following important question.

- What typical aspects need to be considered to improve the meaningfulness of IDR metrics for cited and citing literature?

In the following, we propose some metrics-based method to answer this question and give examples of specific cases of knowledge use vs. integration.

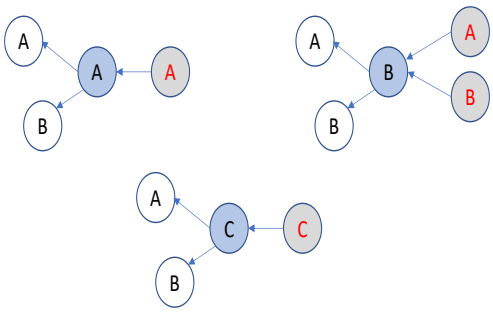

Figure 1. Three examples of information use from and in different disciplines (blue: source, white: references, grey: citations).

Methodology

As already proposed, e.g., by Stirling (2007) and Leinster & Cobbold (2012), we will apply two standard measures commonly used in bibliometric research on IDR: variety (V_S) and disparity (D_S). In particularly, we will use V_S and D_S for all source items in conjunction with *Characteristic Scores and Scales* (CSS) classes to obtain scale-independent measures (cf. Glänzel & Debackere, 2022). To identify potentially problematic cases, we will select papers with outstanding disparity but low variety because these use knowledge from few but very distant disciplines. In a second step, we retrieve all citations to these source items. Both papers and citations are taken from recent volumes of Clarivate Web of Science Core Collection. For each source item, we also calculate the V_C and D_C values based on citing items. Finally, we determine the profile similarity of reference items of each source paper and the citing set of papers (s_{rc}). While V_C and D_C are used to further separate cases, the last indicator will be used to help answer the research question.

Results

To trace the appearance of the cases such as in the model sketched in Fig. 1 based on a systematic approach, we have set a filter on the D_S of all papers indexed for the year 2019 in the Web of Science Core Collection (WoS). We focused on the highest CSS class of disparity (cf. Glänzel & Debackere, 2022), i.e., we selected those papers that used information from different but not related fields. In turn, we focussed on two cases of patterns of citations received by these papers, high and low V_C and reasonable D_C values, if V_C is large. Finally, we use the (dis-)similarity of citation and reference sets of each paper as kind of validation of our results but also to detect outliers. A low similarity between reference and citation profiles would be unrealistic but we can expect moderate to large similarity according as use of knowledge in citations differs from that used in references. Thus, instead of just presenting statistics, we intended to lock "behind" these cases to better understand the mechanism of knowledge use, diffusion and integration.

In the following, we give some examples together with their indicator values. Before we give a small example set, we point to two interesting archetypes of IDR-related knowledge diffusion. DOI: 10.1371/journal.pone.0239831 ("*The length of a scroll: Quantitative evaluation of material reconstructions*") in ancient religion cited literature from religion and physical chemistry (imaging technology) with high D_S, but results are apparently only relevant for religion (low V_C). The single-authored DOI: 10.1093/isd/ixz006 in entomology ("*A systematist's guide to estimating Bayesian phylogenies from morphological data*") tells against the myth that IDR requires co-operation of researchers with different professional background. Both cited and citing papers represent sets of broad and similar subject profiles including entomology, evolutionary biology, genetics heredity, ecology, palaeontology, zoology, anatomy, morphology, and mathematics-/computer science in the references.

Table 1. A sample representing ten IDR/CDR papers with different disparity/variety values of their references and citations with moderate to strong similarity of cited and citing literature.

DOI	D_s	D_c	V_s	V_c	s_{rc}
10.1016/j.nimb.2018.05.002	3.77	3.89	6.48	7.36	0.89
10.1089/ast.2017.1746	3.08	4.27	5.18	6.32	0.90
10.1016/j.envres.2018.09.039	3.51	4.37	6.53	9.45	0.82
10.1080/09296174.2017.1405719	3.62	2.52	6.89	4.00	0.59
10.1016/j.enpol.2018.07.040	3.18	2.42	4.57	4.17	0.48
10.1039/c8an01059e	4.49	2.48	12.25	5.56	0.79
10.1177/0022429418799362	3.11	1.06	6.44	3.00	0.70
10.1039/c8an01526k	3.25	3.41	8.05	10.75	0.91
10.1016/j.jvs.2018.04.029	2.96	3.61	5.72	10.89	0.75
10.1016/j.saa.2018.09.051	3.77	3.26	8.41	9.92	0.45

Table 1 gives a small part of records with interesting quadruples of indicator values in which we found remarkable cases. Some will be discussed here. DOI: 10.1080/09296174.2017.1405719 ("*The Stylometric Impacts of Ageing and Life Events on Identity*") in quantitative linguistics with moderate profile similarity of cited and citing literature and strong disparity uses literature from a large range of subjects in neuroscience and behavioral sciences, psychology, linguistics, literature, ecology, computer science and some related fields, with citation impact on telecommunications, computer science, linguistics, electrical and electronic engineering. Cited and citing literature show different foci. DOI:

10.1016/j.enpol.2018.07.040 ("*Costs and benefits of saving unprofitable generators: A simulation case study for US coal and nuclear power plants*") uses knowledge from environmental science, electrical and electronic engineering, economics, cardiology, energy and fuels, computer science, while it impacts on economics, energy and fuels, chemical engineering, environmental sciences and mathematics. Again, the different foci lower similarity of profiles. DOI: 10.1177/0022429418799362 ("*Music Performance Anxiety and Perceived Benefits of Musical Participation Among Older Adults in Community Bands*") combines knowledge from music, psychology, neurosciences, education, gerontology, medical sciences, health care and impacts the same disciplines however with a somewhat narrower scope.

Conclusions

We have briefly discussed five noticeable cases obtained from the application of interdisciplinarity metrics. These examples show that it is worthwhile looking "behind" the indicators to correctly interpret IDR-related phenomena. In the analysis of larger sets, we found documents that used relevant literature without true integration of knowledge. Others produced knowledge outside used literature. In future research, we will develop further filters to detect papers and distinguish types with typical and atypical patterns of knowledge integration and diffusion on a largescale. We expected to give further insight into the mechanisms of creating new knowledge relevant even beyond the disciplinary scope of literature used for the research.

References

Glänzel, W. & Debackere, K. (2022), Various aspects of interdisciplinarity in research and how to quantify and measure those. *Scientometrics*. 127(9), 5551–5569.

Glänzel, W. & Thijs, B. (2017). Using hybrid methods and 'core documents' for the representation of clusters and topics: the astronomy dataset. *Scientometrics*. 111(2), 1071–1087.

Leinster, T., & Cobbold, C.A. (2012). Measuring diversity: The importance of species similarity. *Ecology*, 93(3), 477–489.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15), 707–719.