How Activity of a Researcher is influenced by Conducting Interdisciplinary Research

Asahi Kiyokawa[1], Fuyuki Yoshikane[1], Shuntaro Kawamura[2] and Takafumi Suzuki[3]

[1] asahi@klis.tsukuba.ac.jp, fuyuki@slis.tsukuba.ac.jp
University of Tsukuba, 1-2, Kasuga, Tsukuba, Ibaraki 305-8550 (Japan)

[2] n-kawa@ka2.so-net.ne.jp
Graduate School of Education, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033 (Japan)

[3] t_suzuki@nii.ac.jp
Faculty of Sociology, Toyo University, 5-28-20, Hakusan, Bunkyo-ku, Tokyo 112-8606 (Japan)

Introduction

Interdisciplinary research is actively conducted in various fields. While interest in interdisciplinary research grows, studies to clarify its very nature have also been conducted. Rinia et al. (2002) indicated that the level of interdisciplinarity is not always positively correlated with the level of impact. Lariviere & Gingras (2010) also indicated that (1) correlation between interdisciplinarity and impact varies by research field and (2) journals with the extreme level of interdisciplinarity have a low impact. These studies use indices including the number of citations in an attempt of quantitatively clarifying the nature of interdisciplinary research. However, there are few studies that review the relationship between interdisciplinary research and productivity. This study focuses on the influence on researchers of interdisciplinary research, in particular the influence on productivity, and attempts to clarify it through a quantitative approach.

Data

We selected environmental sciences as a target field. Data of papers in this field was downloaded from Scopus, and authors were divided into groups of conducting and not conducting interdisciplinary research as follows:

Group conducting interdisciplinary research
1. Identify papers with the keywords (*author keywords* and *index keywords*), such as interdisciplinary research and interdisciplinary studies, using the keyword search.
2. Obtain a list of authors based on 1 and search papers by their author IDs.

Group not conducting interdisciplinary research
1. Identify papers published in 2003 without the keywords (*author keywords* and *index keywords*) related to interdisciplinary research and sample 100 papers from them.
2. Obtain a list of authors based on 1 and search papers by their author IDs.

With five years before conducting interdisciplinary research (as for the second group, before 2003) as the former period and subsequent five years as the latter period in regards to each author, papers within each period were identified. The numbers of authors of the first and second groups are 221 and 361. The numbers of papers published by them are 3,174 and 6,718, respectively.

Methodology

Classification and regression analysis were attempted with the machine learning method, Random Forests (RF)

(Breiman, 2001). RF is known to perform extremely well in classification tasks in other fields using large amounts of data, but few studies have used this method in bibliometrics. By utilizing RF, we classified authors conducting and not conducting interdisciplinary research, moreover, verified which variables classification or regression depends on. RF was adopted since it is possible to avoid the problem of multicollinearity among variables through its learning process. The following viewpoints were considered:

(1) What differences are observed between groups conducting and not conducting interdisciplinary research?

(2) What differences are observed between before and after conducting interdisciplinary research? Is there anything different from authors who did not conduct it?

(3) Are factors influencing productivity different between groups conducting and not conducting interdisciplinary research?

The number of published papers (complete count: PC, first-author count: PF), the number of journals containing those papers NJ, the number of their coauthors CA, and the number of years elapsed after publishing the first paper before conducting interdisciplinary research (as for the second group, before 2003) C are measured for each author. Variables except C are transformed into logarithm. Subscripts attached to the variables signify the former ("1") or latter ("2") period. For instance, PC_1 represents the number of papers (complete count) in the former period. For each task, 50% of the data was used for training.

Results

Classification between conducting and not conducting interdisciplinary research

The classification error rate for the group conducting interdisciplinary research is relatively high, i.e., 19.0%; however it is 13.2% as a whole, suggesting that the difference between the two groups appears in classification results. Shifting our eyes to the values of contribution in classifying, those of the variables in the latter period are positioned higher than those in the former period. In particular, CA_2 has a significant influence. NJ_2 and PC_2 are also positioned higher. Thus, it is considered that the tendencies differ by period.

Classification between former and latter periods

Data on authors was divided into the former and latter periods, and classification between them was attempted based on the variables, which makes it possible to assume the aspects influenced by conducting interdisciplinary research. The classification error rate for the group conducting interdisciplinary research (14.9%) is lower than that for the group not conducting it (20.4%). That is to say, as for the first group, there is a greater difference regarding the tendency between the two periods (before/after conducting interdisciplinary research). For the value of contribution in classifying, PC is ranked 2nd or higher for both the groups. The difference in contribution is relatively large among variables for the group not conducting interdisciplinary research, while variables except PF are balanced in the case of the group conducting it. It is indicated that researchers who have conducted interdisciplinary research change in various aspects not only in productivity.

Regression of productivity

Finally, we applied RF regression for each group using the number of papers *PC* as the response variable, in order to review the influence of conducting interdisciplinary research on productivity. Variables except those relating to productivity were entered as explanatory variables. The explanatory ratio was 0.47 for the group conducting interdisciplinary research, and 0.55 for the group not conducting it. The regression results of *PC* for each group are indicated in Figure 1. There are differences in contribution of each variable between the two groups. As to the group conducting interdisciplinary research, the contribution of CA_2 and NJ_2 (variables after conducting interdisciplinary research) are particularly significant; and as to not conducting interdisciplinary research, the values of contribution of variables are similar. Thus, the influence of interdisciplinary research is assumed to be strong, particularly on the number of their coauthors and the number of journals of their papers.
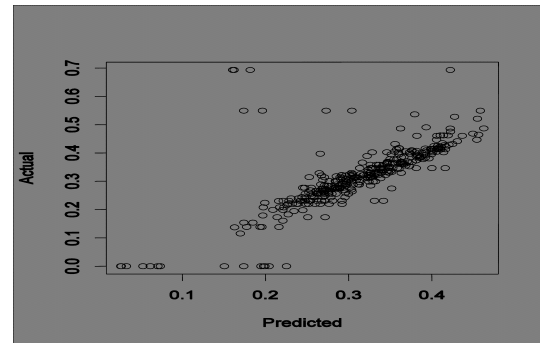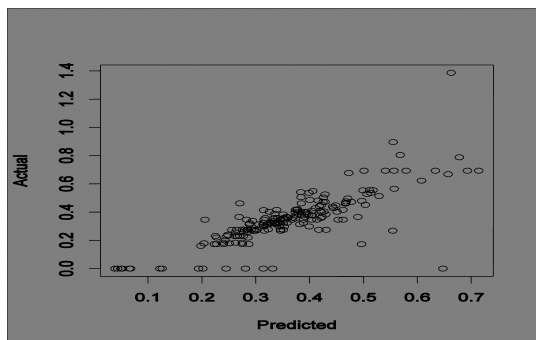

Figure 1. Regression of productivity: the first group (above) and the second group (below).

Conclusions

In this study, how conducting interdisciplinary research influences researchers' activity was analyzed. The results of our analysis imply that interdisciplinary research might influence not only the number of papers, which is generally emphasized in the context of research evaluation, but also aspects such as the increase in the number of coauthors as well as in the number of journals of their papers. In the classification of conducting and not conducting interdisciplinary research as well as regression of productivity, the variables in the former period (before conducting it) were less effective. This result would also become a clue to find the influence of interdisciplinary research.

References

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.

Lariviere, V. & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, 61, 126-131.

Rinia, E.J., van Leeuwen, T.N. & van Raan, A.F.J. (2002). Impact measures of interdisciplinary research in physics. *Scientometrics*, 53, 241-248.