

Nishizawa, Masaki

Associate Professor, Principles of Informatics Research Division

[TITLE]

Identifying the origins of excellent research, based on data on research papers

[MAIN BODY]

My research goal is to enable appropriate investments in R&D that offer great future potential. To that end, I'm examining what kind of support and funds have been provided to research activities that led to high-quality achievements to probe the origins of research excellence.

Tracking the history of successful research

Specifically, I'm examining the past history of scientific research using databases on publicly available research papers.

For example, I investigate the history of grants provided to research groups (units like departments in categories such as faculty, department, and major) involved in excellent research. This provides information on how research activities got underway and how they progressed to success.

Although considering research activities at the initial stages won't tell us what research activities will succeed, it provides the insight needed for research-supporting organizations to determine priority areas for research support.

Examining correlations of researches using language processing technologies

To identify research excellence and to examine their genealogy, I examine how often research papers are cited and the history of papers written by the same researchers.

When a research paper is cited as a reference in another paper, we have to determine whether the paper was used in a positive or negative way. We also have to determine the main authors of the referenced paper. Since there may be two or more researchers with the same name, we check institutional affiliations and paper titles for accurate researcher identification. Correlations among research papers are examined by comparing the similarities of key words in papers based on linguistic processing.

Data on the numbers of research papers and the numbers of key words obtained through the aforementioned processes shows a relatively clean statistical distribution. If we exclude or use words that fall outside this statistical distribution, various correlations become clear.

Visualizing contributions to Japan's international competitiveness from specific universities and the characteristics of these contributions

The meanings of key words change over time, and the same words may have different meanings in different fields. I've not yet taken these variations into consideration. After a given research topic is addressed and the necessary research done, it takes some time for a certain number of research reports to be made available. Furthermore, research papers aren't referenced in other papers until two or three years after publication. That means examination of the use of research papers as references has to occur within a long-range perspective.

Despite these difficulties, my research results enable us to roughly see the significance of contributions by specific universities to the enhancement of Japan's international competitiveness and the degree of contributions made by specific departments of such universities. The recent data indicates large overall contributions from the University of

Tokyo and Kyoto University. Close examination by research field reveals that some single-department colleges and regional universities have strengths in certain fields. If we further analyze this data, it might explain why some research activities achieved greater success than expected.

(Interviewed and summarized by Atsushi Saito)