학술 연구논문 데이터에 기반한 시각화

이현창[*] · 신성윤[**]

[*]원광대학교, [**]군산대학교

Data Visualization based on Academic Research Papers

HyunChang Lee[*], SeongYoon Shin[**]

[*]Wonkwang University

[**]Kunsan National University

E-mail : [*]hclglory@gmail.com, [**]s3397220@kunsan.ac.kr

ABSTRACT

Citation of academic research papers is a very important result for academic researchers, and their utilization is becoming an important evaluation factor. Most papers are composed of authors' keywords. However, there may be some papers with little relevance between the textual content and the presented keywords. Therefore, it is necessary to extract and present important keywords through objective methods for titles and abstracts of theses. In this paper, we present the development results of important keywords through data visualization for academic research papers.

Keywords

co-word analysis, python, wordcloud

I. Introduction

In order to collect data on the research papers, the data of the altmetrics site were used to test the related data. We also used Python for experimental development. Python is a computer engineering languageand and it is a widely used high-level programming language for general-purpose programming. An interpreted language, Python has a design philosophy that emphasizes code readability, and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages. The language provides constructs intended to enable writing clear programs on both a small and large scale.

II. Research method

Python features a dynamic type system and automatic memory management and supports multiple programming paradigms, including object-oriented, imperative, fuctional programming, and procedural styles. It has a large and comprehensive standard library.

Word cloud is a visual representation of text data, typically used to depict keyword meta data(tag) on web sites, or to visualize free form text. Python can use the function of wordcloud make a directive visualization that the most important information could be got. That is the reason for the wordcloud function's exist.

III. Analysis result

To install the wordcloud package firstly call to the terminal and install it. This is the most important step. If it was missed that the function of wordcloud would not be operated.

IV. Conclusion

This thesis inroduced the wordcloud and a programming tool named 'Jupyter' which could programming the Python code in a web environment. It means 'Jupyter' could edict the

Fig. 1. Development result.

codes and operate the codes in a web environment. And wordcloud is a convinient and popular form to make data analysis nowadays. Finally, the result of a wordcloud image showed with a broeser in a web enviroment has been achieved.

Reference

[1] Li X, Thelwall M, Giustini D. Validating online reference managers for scholarly impact measurement. Scientometrics. 2011;91(2):461-71.

[2] Haustein S, Larivière V. A multidimensional analysis of Aslib proceedings - using everything but the impact factor. Aslib Journal of Information Management. 2014;66(4):358-80.

[3] Costas R, Zahedi Z, Wouters P. Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. Journal of the Association for Information Science and Technology. 2015;66(10):2003-19.