Analysis of Koreans' Consciousness of 2016 by Using Wordcloud

Seungkook Roh [a*], Ik jeong [a], Hyunjin Kim[b], Kibo Sim [b], Dongwook Kim [c]
[a]*Policy Research Division, KAERI, 111, Daedeok Daero 989beon-gil, Yusung-gu, Daejeon, 34507*
[b]Korea Nuclear Energy Promotion Agency
[c]*KAIST (Korea Advanced Institute of Science and Technology)*

[*]*Corresponding author: skroh@kaeri.re.kr*

1. Introduction

In 2011, the Fukushima accident and its consequent radiation caused the image of nuclear power in Korea, the closest neighbor country to Japan, to deteriorate(Tanter, 2013).

In 2012, the case of Korea Hydro & Nuclear Power's supply chain corruption further lowered the public confidence in nuclear power(Kim, 2016).

Since then, the 2016 Gyeongju earthquake and its motif "Pandora" have contributed to the general public's image of the existing nuclear power plant as an evasive facility.

Therefore, it is necessary for the Korean government and the atomic energy industry to grasp the current image of the Korean people on nuclear energy and to consider policy measures for enhancing the acceptance of nuclear energy over other agenda. Therefore, this study utilizes text mining analysis method, which is a natural language analysis method, rather than existing formal survey method to accurately grasp how Koreans perceive unconsciously the image of nuclear power.

Based on the results of the analysis, we propose a policy plan for enhancing Korea's nuclear acceptance.

2. Methods and Results

2.1 Data

The data used in this study were extracted from the questionnaire results produced jointly with agency A, which is responsible for enhancing nuclear acceptance.

After completing this questionnaire, the Korea Research Center Co., Ltd., a survey company, conducted a survey based on the questionnaire.

The Korea Research Center conducted a face-to-face questionnaire survey of 1,009 adults aged 20 years or older residing in Korea for duration of about a month in November 2016 (95% ± 3.1% sampling error). The characteristics of the sample can be deduced that it is the opinion on nuclear power which is located in the consciousness of the Korean people as of November 2016 because it properly reflects the population distribution and tendency of the current population in Korea.

2.2 Methodology - Word cloud

A word cloud is a visual representation of the frequency of words used in a document. Because words with high frequency are given larger area, one can understand the core contents of the document at a glance(Roh, 2017).

A word cloud is also called a tag cloud. A tag is a tag attached to a clothing or article to describe the material or handling method. A keyword that is added to describe content in a web page or a social network service (SNS) is called a tag. A tag cloud displays the importance of a tag on a website using different text size or text color(van Atteveldt, 2008).

Depending on the nature of the content you are trying to express, it may be divided into a text cloud and a data cloud. While the document cloud is a visual representation of the words in the document, the data cloud is a representation of the numeric information

using different sizes and colors instead of text information. For example, the size and color of a company's name are determined by the size of a country or the color of a company. (Naver Encyclopedia of Knowledge - Data visualization (Big data, 25. 2. 2013., Communication Books)

2.3 Results

For this study, the respondents were asked to answer: 'Please describe the things you want to know about nuclear power.'

We have obtained 357 sentences for the above inquiry except for responders that did not respond or responded with invalid responses. All 357 sentences were then converted to unstructured data. Several preprocessing steps were performed prior to the actual text mining analysis.

First, we performed extraction of nouns from the clauses and phrases described. In this process, the R package 'extractNoun' was utilized(Sonmale & Ambole). The similarities and errors of the second extracted nouns were adjusted. For example, "information" and "disclosure" extracted from "local residents" and "disclose information" correspond to "information disclosure" for items described as "residents" and "local residents". In addition, "radioactive waste", "radioactive waste" and other precise terms of nuclear energy, "radioactive waste" to repeat the process was modified. In addition, words that describe similar meanings such as "transparent management", "transparency", "transparency" and "human adverse effect" and "human influence" were respectively considered as "transparency" and "human influence" Such word matching was repeatedly performed.

As a result, the words mentioned more than once were 161 nouns in total, and the total sum of the frequency of those words was 1,180 times. The most frequently used words were 'safety', 'information disclosure', 'publicity', 'danger', 'radioactive waste' and so on. Table 2 below shows the top 13 most frequently spoken words, including 'Nuclear', 'Nuclear power generation' and 'Nuclear power plant'.

Fig. 1. Koreans think about nuclear power expressed in word cloud

Table. 1. Ranking of exposure frequency

Ranking	Exposure frequency(times)	words	Ratio
1	200	safety	16.949%
2	174	nuclear power generation	14.746%
3	97	Disclosure	8.220%
4	64	nuclear power plant	5.424%
5	49	promotion	4.153%
6	41	Risks	3.475%
7	36	Radioactive waste	3.051%
8	33	radiation	2.797%
9	31	objectivity	2.627%
10	30	nuclear power	2.542%
11	28	Basic education	2.373%
12	25	Local resident	2.119%
13	24	Waste treatment process	2.034%

3. Conclusions

So far, we have examined using text mining analysis techniques on the things public wants to know about

nuclear power. In order to improve image of the nuclear energy and the contents analyzed in the previous section, we provide the following suggestions on the policy formulation of the government and the nuclear energy industry.

The first is the expansion of related technology research to expand the safety of non-power generation and radioactive waste management in nuclear power generation. The beginning of public confidence in nuclear safety requires public's awareness on the progress of and innovation in R & D.

The second is building a new PR strategy for the nuclear power generation process and operation. This means that the nuclear industry needs to focus more on detailed information from the basic theory of nuclear power generation to the whole development process. It is necessary to establish a marketing strategy so that the site can be easily retrieved from internet portals such as Naver, Daum, and Google, portals most widely used by Internet users.

Finally, disaster response to nuclear power plant accident should be taught to the public by the government and the industry. After the screening of the Pandora movie, the Korean people became more interested in how to cope with serious accidents at nuclear power plants. Therefore, the government and the nuclear industry need to be more active in spreading to the public on the existing public evacuation tips.

REFERENCES

[1] Kim, S. B. (2016). Dangling the Carrot, Sharpening the Stick: How an Amnesty Program and Qui Tam Actions Could Strengthen Korea's Anti-Corruption Efforts. *Nw. J. Int'l L. & Bus., 36*, 235.

[2] Roh, S. (2017). Big Data Analysis of Public Acceptance of Nuclear Power in Korea. *Nuclear Engineering and Technology*.

[3] Sonmale, M. M. S., & Ambole, R. H. Survey on Named Entity Recognition System over Twitter Data.

[4] Tanter, R. (2013). After Fukushima: a survey of corruption in the global nuclear power industry. *Asian Perspective, 37*(4), 475-500.

[5] van Atteveldt, W. H. (2008). Semantic network analysis: Techniques for extracting, representing, and querying media content.