



## Original Article

# Identification of public concerns about radiation through a big data analysis of questions posted on a portal site in Korea



So Yun Jeong, Jae Wook Kim, Han Young Joo, Young Seo Kim, Joo Hyun Moon\*

Department of Nuclear Engineering, Dankook Univ., 119, Dandae-ro, Dongnam-gu, Cheonan, Chungnam, 31116, Republic of Korea

## ARTICLE INFO

## Article history:

Received 10 November 2020

Received in revised form

20 November 2020

Accepted 30 November 2020

Available online 5 December 2020

## Keywords:

Concerns

Radiation

Korean public

Big data analysis

TF-IDF analysis

## ABSTRACT

This paper analyzed the primary concerns about radiation among the Korean public with a big data analysis of questions posted at the section of “Knowledge iN” on the portal site NAVER in Korea from January 2010 to August 2020. First, we extracted questions about radiation and categorized them into the three categories with TF-IDF analysis: “Medical,” “Career Counseling,” and “General Interest”. The “Medical” category includes questions about radiation diagnosis or treatment. The “Career Counseling” category includes questions about entering college and the prospect of finding jobs in radiation-related fields. The “General Interest” category includes questions about terminology and the basic knowledge of radiation or radioisotopes. Second, we extracted common questions for each category. Finally, we analyzed the temporal change in the numbers of questions for each category to confirm whether there is any correlation between radiation-related events and the number of questions. The analysis results demonstrate that major radiation-related events have little relevance to the number of questions except during March 2011.

© 2020 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Radiation is everywhere: in the soil, in the air, and even in the human body [1]. As of the end of 2018, the total number of radiation workers in Korea was 44,122, and approximately 0.09% of the entire Korean population have radiation-related jobs [2]. Excluding those people, most of the public are likely to experience radiation in hospitals. People who travel abroad frequently are exposed to cosmic radiation. Although everyone experiences radiation in their everyday lives, few persons are aware that they live with natural radiation. Furthermore, many persons have a vague unease concerning radiation due to previous nuclear power plant (NPP) accidents.

Many researchers have conducted surveys of the public perception of nuclear power and radiation. Kim et al. investigated the impact of the Fukushima NPP accident on citizens' acceptance of nuclear energy in 43 countries [3]. Goodfellow et al. argued that the public perception of nuclear power was an essential factor in the construction of NPPs, based on the research on the risk perception of nuclear power for 40 years [4].

Korea and other countries, including China [5], the Netherlands [6], the United Kingdom [7], and Hong Kong [8], investigated public opinions on nuclear power generation. Furthermore, Yim et al. examined the influence of the public's academic background on its perception and attitudes toward nuclear power [9].

Several independent surveys of perception of radiation were conducted that were not as robust as those of perceptions of nuclear power. Kenny et al. assessed the Ireland public's current level of understanding of the risks from ionizing radiation [10]. Miura et al. conducted a survey about the perception of radiation health care risk as a predictor for mental health conditions after a nuclear disaster [11]. Surveys of the perception of radiation protection were also conducted. Florig analyzed the stances of public organizations on radiation protection [12]. Tanha et al. evaluated the status and perception of radiation protection for the public and radiation workers, respectively, in Afghanistan [13].

Recent studies have expanded the survey objects and analyzed big data extracted from mass media, social networks, and portal sites. Several papers have analyzed big data posted on NAVER, which is a leading portal site in Korea. Nam et al. analyzed the characteristics of knowledge generation and participant behavior in the “Knowledge iN” section on NAVER [14]. Park et al. used NAVER Café to investigate the temporal trend of search words related to

\* Corresponding author.

E-mail address: [jhmoon86@dankook.ac.kr](mailto:jhmoon86@dankook.ac.kr) (J.H. Moon).

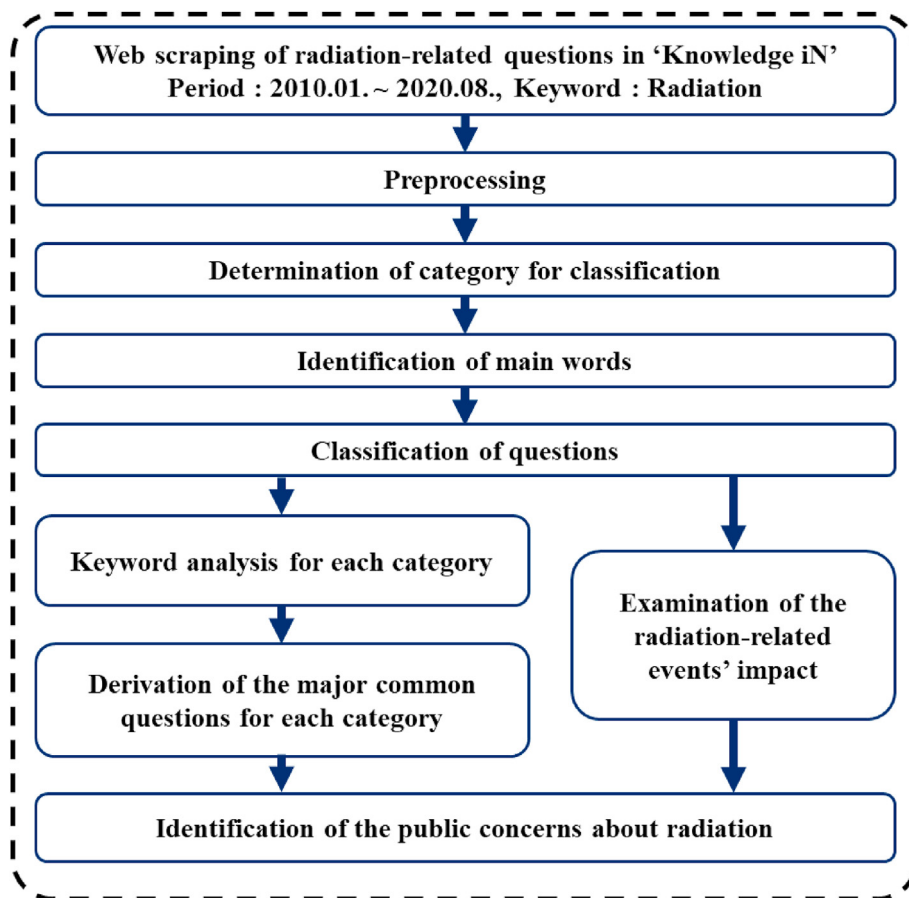


Fig. 1. Overview of the analysis procedure.

the oral health of infants and pregnant women [15]. Park et al. analyzed regional characteristics of sharing knowledge by investigating the geographic locations of knowledge users and the motivation of knowledge providers in “HERE,” the location-based knowledge service site on NAVER [16].

Currently, NAVER operates a knowledge information sharing site called “Knowledge iN.” The “Knowledge iN” service started in October 2002, where one user asks a question and other users can answer the question. On “Knowledge iN,” the questioner could gain knowledge or information he or she wants to know from others. All Q&As posted on “Knowledge iN” have been stored in the database and could be accessed at any time by keyword searching [17]. Among questions registered on “Knowledge iN” are those about radiation.

Inquiries by questioners are just small, individual curiosities. In contrast, an extensive collection of inquiries could reveal public perceptions about an issue in society. “Knowledge iN,” a knowledge-sharing service of NAVER, is an open platform where any NAVER user can register and answer a question. Accordingly, we thought that, by analyzing the frank questions on “Knowledge iN,” we could discover what ordinary people really think about radiation. Therefore, in this paper, we collected and analyzed all questions about radiation posted on “Knowledge iN” from January 2010 to August 2020 to identify the primary concerns about radiation among the Korean public. We used R, a big data analysis software program. Fig. 1 illustrates the overview of the analysis procedure of this paper.

## 2. Materials and method

### 2.1. Web scraping

First, with the R program, we extracted questions that included the keyword “radiation” in a single part of the question from questions registered on “Knowledge iN” from January 2010 to August 2020.

### 2.2. Preprocessing

Before starting the primary analysis, questions first extracted should be preprocessed to obtain reliable analysis results. First, we screened out questions that could not be analyzed. For example, many questions included the keyword “radiation” only in their answer, which was irrelevant to the question’s intention. Hence, we extracted questions that included the keyword “radiation” in both their title and content.

### 2.3. Classification of questions by category

#### 2.3.1. Determination of category for classification

Because we could not read all questions extracted one-by-one for classification, we identified the categories for each question by considering the most frequently mentioned words. We used the “NLP4kec” package to extract words (nouns). After excluding the unnecessary words among the extracted words, we listed the top 100 words most frequently mentioned in questions. Consequently,

we identified three categories: “Medical,” “Career Counseling,” and “General Interest.”

### 2.3.2. Identification of main words

For the three categories, we performed term frequency–inverse document frequency (TF-IDF) analysis. TF-IDF is one of the most widely used terminology-weighting algorithms in the field of information retrieval. TF-IDF indicates the importance of the appearance of a specific word in a document [18]. TF illustrates how often the same word occurs in a single document, and IDF represents the scarcity and importance of the word in the entire document [18]. TF-IDF is a mechanism that determines the ranking of the words by their importance. The TF-IDF weight increases if the frequency of a specific word in a single document is high while the number of documents in which the word appears is small [18]. We chose TF-IDF for the selections of the main words to find words appropriately representing each classification category. TF-IDF can provide words with high relevance to the main words for each category, and we can classify questions by confirming whether the corresponding words are included in the questions.

### 2.3.3. Questions classification

Questions that included the words for the categories of “Medical” and “Career Counseling” were classified into the two categories automatically by the algorithm. In contrast, questions in the “General Interest” category were classified manually. The “General Interest” category questions include a range of questions about radiation or radioisotopes, so it is challenging to identify the main words for this category because few words repeatedly emerged across questions. Neither keyword analysis nor the TF-IDF algorithm could find words representing the category. Hence, we inevitably classified them manually. We matched a single question to the multiple categories if the question included the words for the multiple categories. Finally, we excluded questions that do not belong to the three categories from the analysis.

## 2.4. Identification of the common questions for each category

### 2.4.1. Keyword analysis for each category

We performed a keyword analysis to extract the common questions for each category. “Knowledge iN” does not have a standard form for questions, and any user can register questions on the site. Hence, it was challenging to count the number of questions by category by just looking at the sentences of the question because we were not sure that the sentences of the question precisely matched a category even if the subject of questions matches the category. Therefore, we decided to extract the common questions using the keywords identified through keywords analysis for each category. And we performed keyword analysis of the questions registered in March 2011 when the questions in the “General Interest” category were the most registered. For the keyword analysis, we extracted the 50 nouns using the “NLP4kec” package in order of frequency.

### 2.4.2. Derivation of the major common questions for each category

We extracted common questions based on the keywords identified by the keyword analysis by category. Among the registered questions, those including three or more keywords identified by keyword analysis were selected as common questions. For the three categories, we searched questions most registered on “Knowledge iN.” For the “General Interest” category, we added the common questions most registered on March 2011.

## 2.5. Examination of radiation-related events' impact

Finally, we investigated whether a correlation between the number of questions of the same subjects and major domestic and foreign radiation-related events exists. After plotting the number of questions by category over time on a graph, we marked when major incidents occurred on the same graph for ease of understanding. Table 1 presents the major domestic and foreign radiation-related incidents.

The Fukushima NPP accident was nuclear-related rather than radiation-related. However, it was included in Table 1 because several questions were registered on “Knowledge iN” after the accident, which also significantly influenced public perception.

## 3. Result

### 3.1. Web scraping and preprocessing

The total number of questions that included the keyword “radiation” scraped from “Knowledge iN” from January 2010 to August 2020 was 159,085, with an average of 1243 questions per month. Only questions that included the keyword “radiation” in both their title and content were retained for analysis. The total number of questions first sorted was 31,882, with an average of 249 questions per month. Fig. 2 illustrates the number of questions first sorted per month. The dark dotted line between the two gray dotted lines indicates 249, which is the average number of questions, and the top and bottom lines illustrate the standard deviations  $\pm 1\sigma$  ( $\pm 47.57$ ).

Fig. 2 illustrates that the month when questions were the most registered was March 2011, when the Fukushima NPP accident occurred. Since 2016, the number of questions per month trended upward, although the monthly numbers of questions fluctuated every month.

### 3.2. Classification of questions by category

#### 3.2.1. Determination of category for classification

We performed a keyword analysis to classify questions. Table 2 presents the 100 most frequently mentioned words over the entire analysis period. Most of the 100 words were relevant to radiation treatment, cancer treatment, and insurance, based on which we created the “Medical” category. There were many words about entering college, increasing significantly and cyclically during the entrance exam season. Accordingly, we created the “Career Counseling” category. There were also other words relevant to radiation terminologies such as “radioactive,” “radioactivity,” “alpha ray,” “beta ray,” and “gamma ray,” and less frequently mentioned than those in the two categories. For those words, we created the third category, “General Interest.” Consequently, we established the three categories, “Medical,” “Career Counseling,” and “General Interest” for further analysis.

#### 3.2.2. Identification of main words

We identified the main words best representing each category based on TF-IDF analysis. As depicted in Figs. 3 and 4, the main words of the “Medical” category were “treatment,” and “radiation,” and the main words of the “Career Counseling” category were “radiation,” “department,” and “university.” The words of relevance are connected by lines. Cor\_value (correlation value) indicates the degree of relevance between the words, expressed by the thickness of each line. The thicker the line is, the more relevant the words connected with the line.

Table 3 presents the words with high relevance extracted through TF-IDF analysis. The keywords relevant to the main words

**Table 1**  
Major incidents or issues that occurred related to radiation.

Time (Month-Year)	Incidents or events
March 2011	· Fukushima nuclear power plant accident.
November 2011	· Detection of abnormal radiation level on a road in Wolgye-dong, Nowon-gu, Seoul.
March 2012	· The first anniversary of the Fukushima accident.
July 2012	· Enforcing of act on protective action guidelines against radiation in the natural environment.
July 2015	· First disposal of low- and intermediate-level radioactive waste in Korea.
April 2017	· Radiography worker exposure accident in Yeosu-si.
May 2018	· Detection of Radon in daily supplies.
July 2019	· Unauthorized disposal of exempt radioactive waste generating from decommissioning of research reactors at KAERI.
August 2019	· Enforcing of the amendment to the act on protective action guidelines against radiation in the natural environment.
January 2020	· Radiation worker exposure accident at Seoul Semiconductor Co., Ltd.
	· Radioactive material release accident at KAERI.

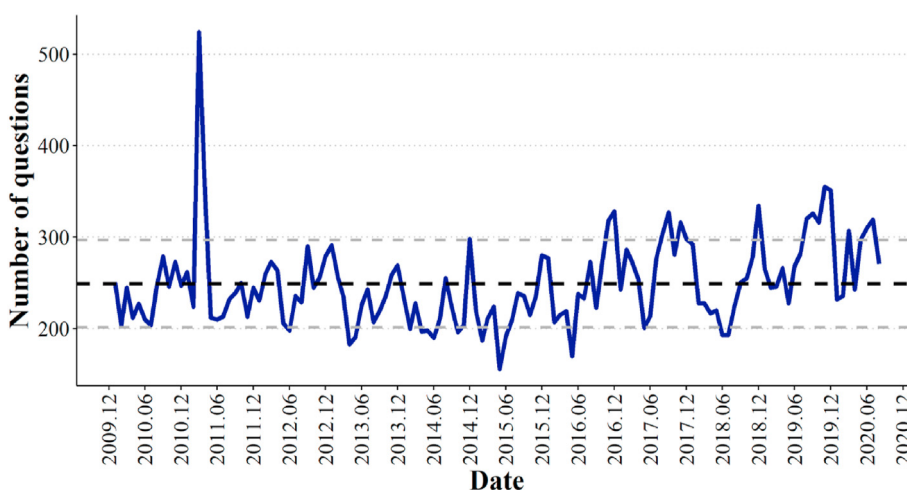


Fig. 2. Number of radiation-related questions by month.

for the “Medical” category were “CT,” “X-ray,” “exposure,” “photographing,” “treatment,” “side effects,” “anti-cancer,” “surgery,” “hospital,” “tumor,” “recurrence,” and “pregnancy.” The keywords relevant to the main words for the “Career Counseling” category were “department,” “employment,” “college,” “university,” “school,” “health,” “physical,” “grade,” and “specialty.”

3.2.3. Questions classification

For the “Medical” and “Career Counseling” categories, we classified questions by the keywords extracted through TF-IDF analysis. For the “General Interest” category, there was a wide range of topics, many duplicated words, and several questions not related to the category topic. Hence, we classified questions in this category manually. The criteria applied for classification were terminology, basic knowledge about radiation, and radiation protection/shielding.

We matched a single question to the multiple categories if the question included the words for multiple categories. The examples included in the multiple categories were “Is there a university hospital specializing in radiation treatment?” or “How much exposure do I receive due to x-rays when I graduate from the department of radiology and become a radiographer?” As those questions include the keywords in both “Medical” and “Career Counseling” categories, those questions belong to two categories simultaneously. Duplicate classification was enabled because questions were classified by words alone, rather than by context.

Fig. 5 illustrates the number of questions in each category by month. The red line represents the numbers of questions in the “Medical” category, the green represents that of questions in the

“Career Counseling” category, and the blue represents that of questions in the “General Interest” category.

For the “Medical” category, the total number of questions was 22,680, with a monthly average of 177. The number of questions in the “Medical” category consistently ranked first, with most questions primarily relevant to cancer treatment, radiation treatment, and anti-cancer.

For the “Career Counseling” category, the total number of questions was 11,794, with a monthly average of 92. The number of questions in the “Career Counseling” category trended upward, although the monthly numbers of questions fluctuated. The number of questions in this category increased significantly and cyclically in the entrance exam season every year. In Korea, the timing to apply for rolling admission is September, and the timing to apply for regular admission is predominantly between December and January next year. As the entrance exam season arrives, the students, including high school seniors, ask questions about the curriculum and employment prospects in college radiation departments before submitting their applications.

For the “General Interest” category, the total number of questions was 1,773, with a monthly average of 12. For the “General Interest” category, several questions per month were registered, except during March 2011.

3.3. Identification of major common questions for each category

3.3.1. Keyword analysis for each category

We performed a keyword analysis to extract the common questions for categories in which NAVER users expressed the

**Table 2**  
**Top 100 most frequently mentioned keywords in questions registered on “Knowledge iN”.**

Number	Word	Frequency	Number	Word	Frequency
1	Radiation	45,172	51	Tumor	4977
2	Treatment	44,603	52	Metastasis	4851
3	Surgery	34,485	53	Concern	4817
4	Diagnosis	32,220	54	Payment	4758
5	Insurance	29,792	55	Relation	4723
6	Surgery fee	27,033	56	Health	4651
7	Renewal	24,401	57	Peace	4477
8	Disease	24,166	58	Burn	4475
9	Hospital	19,041	59	Time	4422
10	Hospitalization	18,676	60	Content	4395
11	Collateral	18,500	61	Radiology	4328
12	Abnormality	17,070	62	Way	4313
13	Check	16,257	63	Actual expense	4312
14	Anti-cancer	13,723	64	Problem	4152
15	Degree	13,359	65	Exclusion	4093
16	Buying insurance	13,197	66	Breast cancer	4053
17	Guarantee	11,791	67	Month	4020
18	Medical expenses	10,454	68	Obstacle	4002
19	Person	10,336	69	Impediment	4001
20	Death	9922	70	Selection	3943
21	Injury	9679	71	Outpatient	3890
22	Expiry	9405	72	Say	3879
23	Insurance premium	9336	73	Need	3672
24	Sum	9233	74	Thanks	3660
25	Special contract	8777	75	Family	3620
26	Case	8738	76	Hospital bill	3602
27	Sickness	8643	77	Comprehensive	3563
28	Thought	8595	78	Infarction	3500
29	Transplant	7894	79	Add	3487
30	Request	7713	80	Pain	3473
31	Outpatient	7401	81	Myocardium	3442
32	Medical expense	7332	82	Patient	3416
33	Possibility	6960	83	Teacher	3391
34	Condition	6833	84	Charge	3388
35	Daily payment	6790	85	Symptom	3364
36	General practice	6762	86	Father	3356
37	Fracture	6655	87	Plan	3355
38	Particular	6461	88	Department	3290
39	CT	6154	89	Brain tumor	3288
40	Period	5896	90	Prescription	3270
41	Accident	5756	91	Mother	3238
42	University	5633	92	Basic	3146
43	Life	5284	93	After	3139
44	Result	5216	94	University	3118
45	Doctor	5178	95	School	3114
46	Contract	5167	96	Occurrence	3087
47	Medical care	5118	97	Teeth	3046
48	Drug	5103	98	X-ray	3033
49	Scanning	5092	99	Responsibility	2937
50	Part	5006	100	Acute	2900

greatest interest and curiosity. Before extracting questions, we identified the words in each category that were frequently mentioned through keyword analysis. Tables 4 and 5 presents the top 50 frequently mentioned words by category and ‘General Interest on March 2011’. For the “Medical” category, “radiation treatment,” “surgery,” “insurance,” and “hospital” were the most mentioned. For the “Career Counseling” category, “college,” “radiology,” and “department” were the most mentioned. Finally, for the “General Interest” category, “radiation,” “radioactivity,” “radioactive,” and “energy” were the most mentioned, and on March 2011, “radiation,” “Japan,” “exposure,” and “earthquake” were the most mentioned.

### 3.3.2. Derivation of the major common questions for each category

Common questions were extracted by combining words with high frequency in keyword analysis. These questions including three or more words were drawn by considering keywords

identified through keyword analysis. In the “General Interest” category, we also extracted common questions registered at the time of the Fukushima accident. Tables 6–8 present the lists of the common questions by category, and Table 9 is a list of common questions registered at that time of Fukushima accident.

Most of the “Medical” category questions were relevant to radiation treatment and insurance buying/processing due to radiation treatment. Subsequently, there were various medical care questions, such as the exposure dose due to radiography and metastasis to another organs due to radiation treatment.

For the “Career Counseling” category, most questions were relevant to entering college, such as applying for the department of radiology, admission, transfer, and graduation. Furthermore, there were many questions about jobs in radiation-related fields.

For the “General Interest” category, the most common question was “difference in meanings of radiation, radioactive, and radioactivity”. Questioners wanted to obtain information on terminology

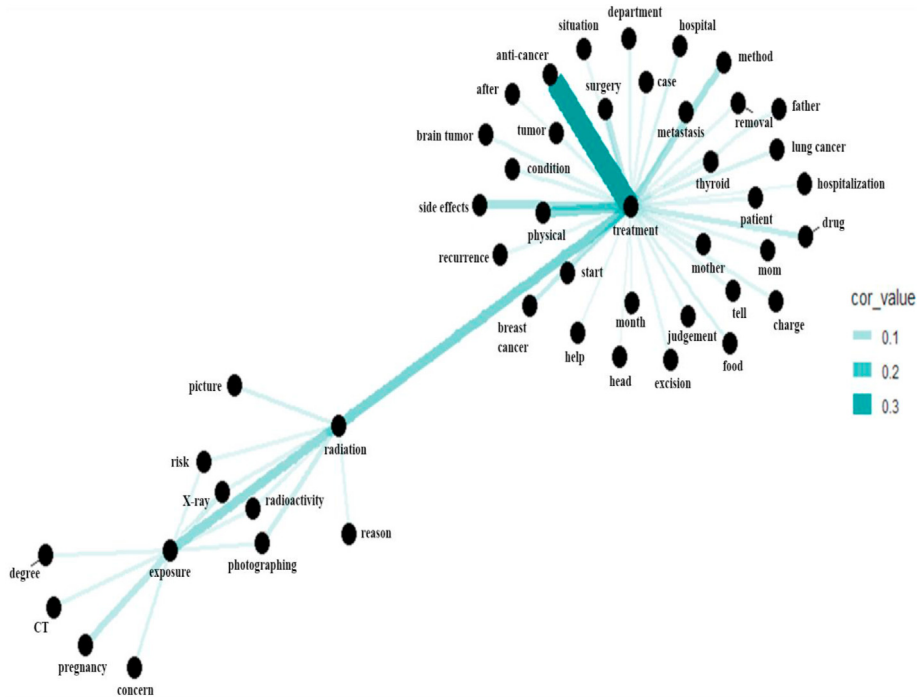


Fig. 3. TF-IDF analysis results for the "Medical" category.

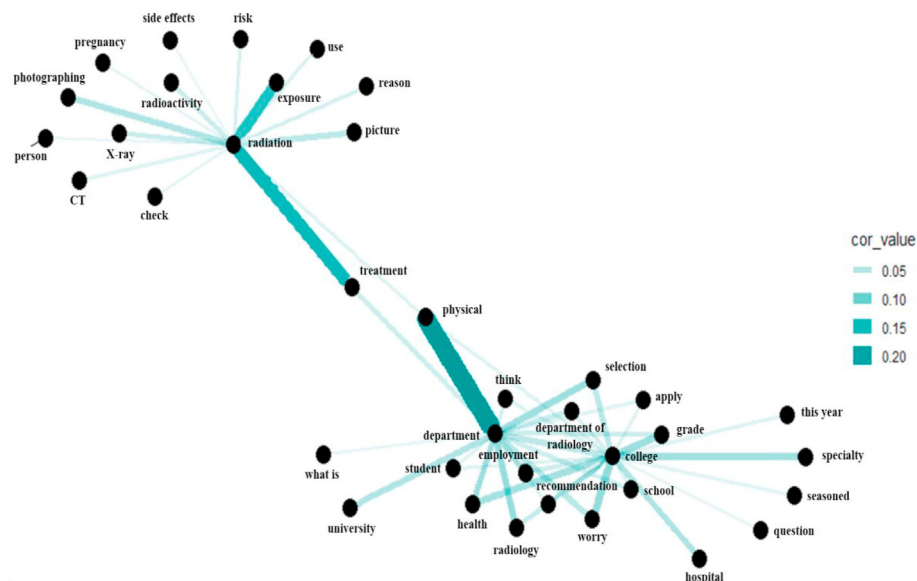


Fig. 4. TF-IDF analysis results for the "Career Counseling" category.

**Table 3**  
Keywords to the main words in the "Medical" and "Career Counseling" categories.

Category	Keywords
Medical	· CT, X-ray, exposure, photographing, treatment, side effects, anti-cancer, surgery, hospital, tumor, recurrence, pregnancy
Career Counseling	· Department, employment, college, university, school, health, physical, grade, specialty

and unit related to radiation. They also asked questions about explanations or differences in the accident at the Chernobyl, Three Mile, and Fukushima NPPs. In March 2011, when the Fukushima NPP accident occurred due to the great east Japan earthquake, the

number of common questions in the "General Interest" category increased significantly. Questions over this period were primarily concerns about radiation safety.

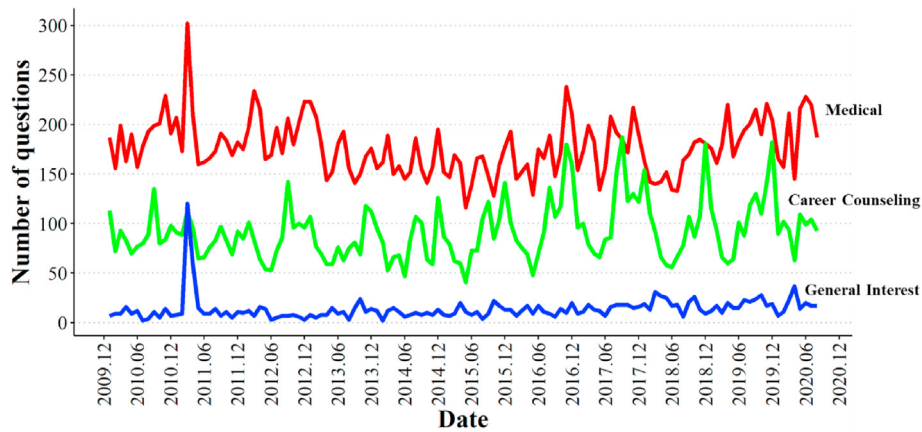


Fig. 5. Number of questions in each category by month.

Table 4  
Top 50 frequently mentioned keywords for 'Medical' and 'Career Counseling' categories.

Number	Medical		Career counseling	
	Word	Frequency	Word	Frequency
1	Treatment	44,603	Radiation	14,270
2	Surgery	34,485	College	5633
3	Radiation	33,969	Degree	5618
4	Diagnose	32,137	Over	5388
5	Insurance	29,671	Think	4932
6	Surgery fee	27,033	Case	4472
7	Renewal	24,360	Radiology	4308
8	Disease	24,107	Department	3290
9	Hospital	19,041	Possibility	3250
10	Hospitalization	18,659	University	3118
11	Collateral	18,498	School	3114
12	Abnormal	16,308	Request	2766
13	Check	14,558	Hygiene	2732
14	Anti-cancer	13,723	Condition	2648
15	Join	13,171	Department of radiology	2613
16	Guarantee	11,746	Time	2502
17	Degree	11,676	Medical care	2381
18	Medical expenses	10,452	Relation	2250
19	Death	9851	General	2178
20	Injury	9662	Doctor	2173
21	Expiry	9395	Employment	2160
22	Insurance premium	9312	Result	2132
23	Sum	9215	Peace	2132
24	Special contract	8755	Physical	2093
25	Person	8640	Radiographer	2072
26	Sickness	8578	Thanks	2030
27	Pay	8403	Tell	2,015
28	Case	8172	Study	1982
29	Transplant	7872	Part	1968
30	Outpatient	7398	Period	1940
31	Treatment expense	7332	Health	1936
32	Daily pay	6788	Death	1878
33	Fracture	6643	Problem	1864
34	Think	6499	Content	1831
35	Particular	6402	Life	1762
36	Request	6372	Apply	1709
37	Condition	6308	Grade	1697
38	General	6164	Specialty	1671
39	CT	6153	Way	1638
40	Period	5621	Worry	1637
41	Accident	5560	Student	1553
42	Possibility	5325	License	1547
43	Drug	5093	Science	1503
44	Scan	5092	Clinic	1477
45	Life	5007	Concern	1466
46	Tumor	4977	Need	1456
47	Doctor	4945	Selection	1455
48	Metastasis	4819	Graduate	1384
49	Medical care	4584	After	1352
50	Health	4423	Exam	1342

**Table 5**  
**Top 50 frequently mentioned keywords for ‘General Interest’ category and ‘General Interest on March 2011’.**

Number	General interest		General interest on march 2011	
	Word	Frequency	Word	Frequency
1	Radiation	3541	Radiation	258
2	Radioactivity	944	Radioactivity	120
3	Element	615	Japan	94
4	Material	547	Material	58
5	Radioactive	412	Country	48
6	Energy	338	Exposure	48
7	Nuclear energy	309	Earthquake	36
8	Exposure	297	Damage	35
9	Decay	293	Influence	32
10	Measurement	287	Person	27
11	Same position	279	Uranium	21
12	Person	276	Nuclear power plant	21
13	Japan	238	Count	20
14	Degree	233	Wind	19
15	Explain	232	Iodine	18
16	Reason	218	Explosion	18
17	Atom	206	Condition	16
18	Use	200	Nuclear energy	16
19	Request	189	Spill	16
20	Emit	184	Travel	15
21	Influence	184	Concern	14
22	Electron	184	Radioactive	14
23	Problem	182	Nature	14
24	Possibility	176	Possibility	13
25	Gamma ray	176	Leak	13
26	Particle	174	Emit	13
27	Half-life	172	Risk	13
28	Risk	166	If	12
29	Earth	162	Recently	11
30	Uranium	156	Difference	11
31	Utilize	154	Big earthquake	10
32	Age	152	Safe	10
33	Neutron	151	Abnormal	10
34	Power plant	146	Korea	10
35	Time	146	Problem	9
36	Experiment	142	Power plant	9
37	Human body	140	Earth	9
38	Fact	139	Air	8
39	Nuclear power plant	137	U.S.A	8
40	Unit	136	Anxiety	8
41	Damage	136	Cesium	8
42	Stability	135	Notify	8
43	Country	134	Westerlies	8
44	Difference	134	Case	7
45	Discovery	133	Tokyo	7
46	Way	133	My country	7
47	Contamination	133	Human body	7
48	Carbon	130	Some	7
49	Think	129	Nuclear fission	7
50	Nucleus	123	Fukushima	7

3.4. Examination of radiation-related events' impact

Major events are marked in Fig. 6 to confirm whether there was a correlation between the frequency of questions of the same subjects and major domestic and foreign radiation-related events. Major events are presented in Table 1 of Section 2.5.

The examination illustrates that major radiation-related events had little relevance to the frequency of questions except during

March 2011. Even during May 2018, when the radon mattress scandal outbreaked in Korea, fewer questions that included the keyword “radiation” were registered than in the previous month because the public focused more on other keywords such as “radon,” “bed,” and “anion,” rather than “radiation”. Furthermore, the number of questions in the “Career Counseling” category has a similar pattern every year, increasing significantly around the entrance exam season. The number of questions in the “General

**Table 6**  
**Common questions in the “Medical” category.**

Question	
Q1	I am receiving radiation treatment for anti-cancer. Is there any cautions I should be aware of?
Q2	I have been diagnosed with cancer and am getting radiation treatment. Can I buy cancer insurance and cover treatment expense by the insurance?
Q3	I had a CT scan as a health examination. How much is the radiation dose?
Q4	If I am on radiation treatment, can it metastasis to another organs?



**Table 7**  
Common questions in the “Career Counseling” category.

Question	
Q1	I am a high school student who wants to enter the department of radiology. What should I prepare for?
Q2	Please tell me the university or college where department of radiology is located.
Q3	Can I become a radiographer even if I graduate from a health related department?
Q4	Can I get a job easily after graduating from the department of radiology?

**Table 8**  
List of common questions in the “General Interest”.

Question	
Q1	Please explain the difference between radiation, radioactive and radioactivity.
Q2	Please explain the major accidents occurred in other countries nuclear power plants such as Chernobyl and Fukushima accidents.
Q3	Please explain how to age-dating of the Earth using radioactive isotopes.
Q4	Please explain about the units of radiation.
Q5	Is radiation emitted from human body?

Interest” category was just 12 per month, which is too small to confirm whether there is any influence due to radiation-related events.

**4. Discussion**

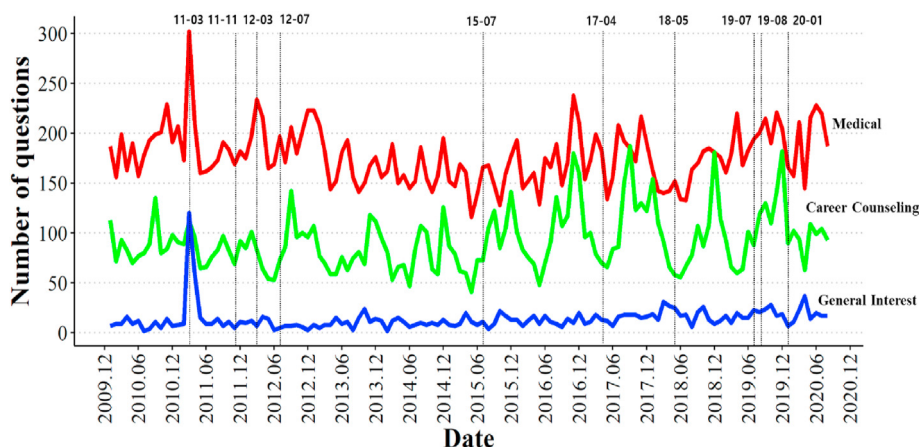
This paper analyzed the public concerns about radiation by collecting and analyzing all questions about radiation posted on the NAVER “Knowledge iN” site in Korea. This paper extracted questions that included the keyword “radiation” in their title and content among questions registered on “Knowledge iN.” The analysis period was from January 2010 to August 2020. After classifying the pre-processed questions into radiation use categories, common questions for each category were extracted. Finally, we investigated whether a correlation between the number of questions of the same subjects and major domestic and foreign radiation-related events existed.

**Table 9**  
Common questions in the “General Interest” category at the time of March 2011.

Question	
Q1	Is Korea influenced by radiation released from Fukushima nuclear power plants?
Q2	Is it safe to travel abroad such as Japan?
Q3	When does radioactive material arrive in Korea by wind?
Q4	Is it safe from radiation exposure if I take a thyroid protectant (iodine)?

Based on the most frequent words in questions, we established three categories for further analysis: “Medical,” “Career Counseling,” and “General Interest.” Moreover, we found that the main words in the “Medical” category were “radiation,” and “treatment.” The main words in the “Career Counseling” category were “radiation,” “department,” and “university.”

We then extracted the common questions for each category by analyzing keywords to determine the most frequently used words. Based on these words, we extracted common questions as subjects that people frequently asked. In the “Medical” category, most questions were relevant to radiation treatment. In the “Career Counseling” category, questions were primarily about entering the college radiology department. Questions in the “General Interest” category were primarily relevant to the terminology and basic knowledge of radiation and background radiation around questioners. Furthermore, as a result of analyzing questions registered in March 2011, when the number of questions was the highest,



**Fig. 6.** Number of questions per month and timing of radiation-related events.

most were about the effects of radioactive material released from Fukushima NPP on the Korean people and the biological effects of radiation. There were also questions about travel abroad and radiation dose.

Finally, we confirmed whether there was a correlation between the frequency of questions of the same subjects and major domestic and foreign radiation-related events. The analysis reveals that major radiation-related events have little relevance to the frequency of questions except during March 2011.

As people become more sensitive to radiation, we expect that more questions about radiation will be registered on portal sites. People are more interested in radiation treatment, which is directly related to public health. For the “Career Counseling” category, the same pattern would be maintained in the future, where the number of questions increases significantly repeatedly during the entrance exam season every year. For obtaining greater knowledge in a radiation-related field, it is necessary to work out the proper strategies considering examinees’ interests and when the number of questions increases. Furthermore, it is necessary to have standard answers available to common questions that many people are curious about as a useful public relation strategy. As a future work, we will expand the analysis objects to the questions posted in the other social media such as “Blog”, “Cafe”, “Twitter” and “Facebook”, to understand the public concerns better.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korean government (Ministry of Science and ICT). (No. 2020M2D2A2062436).

#### References

- [1] C. Grupen, M. Rodgers, *Radioactivity and Radiation: what They Are, what They Do, and How to Harness Them*, Springer, UK, 2019.
- [2] Nuclear Safety and Security Commission, 2018 Report on Occupational Radiation Exposures in Korea, 2020. KoFONS/RR-006/2020.
- [3] Y. Kim, M. Kim, W. Kim, Effect of the Fukushima nuclear disaster on global public acceptance of nuclear energy, *Energy Pol.* 61 (2013) 822–828, <https://doi.org/10.1016/j.enpol.2013.06.107>.
- [4] M.J. Goodfellow, H.R. Williams, A. Azapagic, Nuclear renaissance, public perception and design criteria: an exploratory review, *Energy Pol.* 39 (2011) 6199–6210, <https://doi.org/10.1016/j.enpol.2011.06.068>.
- [5] G. He, A.P.J. Mol, L. Zhang, Y. Lu, Public participation and trust in nuclear power development in China, *Renew. Sustain. Energy Rev.* 23 (2013) 1–11, <https://doi.org/10.1016/j.rser.2013.02.028>.
- [6] K. Mulder, The dynamics of public opinion on nuclear power. Interpreting an experiment in The Netherlands, *Technol. Forecast. Soc. Change* 79 (2012) 1513–1524, <https://doi.org/10.1016/j.techfore.2012.04.018>.
- [7] M.J. Goodfellow, P. Dewick, J. Wortley, A. Azapagic, Public perceptions of design options for new nuclear plants in the UK, *Process Saf. Environ. Protect.* 94 (2015) 72–88, <https://doi.org/10.1016/j.psep.2014.12.008>.
- [8] D.N. Yin Mah, P. Hills, J. Tao, Risk perception, trust and public engagement in nuclear decision-making in Hong Kong, *Energy Pol.* 73 (2014) 368–390, <https://doi.org/10.1016/j.enpol.2014.05.019>.
- [9] M.S. Yim, P.A. Vaganov, Effects of education on nuclear risk perception and attitude: Theory, *Prog. Nucl. Energy* 42 (2003) 221–235, [https://doi.org/10.1016/S0149-1970\(03\)80010-0](https://doi.org/10.1016/S0149-1970(03)80010-0).
- [10] E. Kenny, B. Byrne, M. Lewis, D.M. King, Perception of medical radiation risk in Ireland: results of a public survey, *Phys. Med.* 68 (2019) 96–103, <https://doi.org/10.1016/j.ejmp.2019.10.033>.
- [11] I. Miura, M. Nagai, M. Maeda, M. Harigane, S. Fujii, M. Oe, H. Yabe, Y. Suzuki, H. Takahashi, T. Ohira, S. Yasumura, M. Abe, Perception of radiation risk as a predictor of mid-term mental health after a nuclear disaster: the Fukushima health management survey, *Int. J. Environ. Res. Publ. Health* 14 (2017) 1–13, <https://doi.org/10.3390/ijerph14091067>.
- [12] H.K. Florig, An analysis of public-interest group positions on radiation protection, *Health Phys.* 91 (2006) 508–513, <https://doi.org/10.1097/01.HP.0000232650.11844.74>.
- [13] M.R. Tanha, F.R. Khalid, C. Hoeschen, Assessment of radiation protection and awareness level among radiation workers and members of the public in Afghanistan - a pilot study, *J. Radiol. Prot.* 39 (2019) N1–N7, <https://doi.org/10.1088/1361-6498/ab2204>.
- [14] K.K. Nam, A. Arbor, M.S. Ackerman, Questions in , knowledge in ? A study of naver ’s question answering community, university of Michigan, CHI ’09 proc. Twenty-sixth, Annu. SIGCHI Conf. Hum. Factors Comput. Syst. (2009) 779–788, <https://doi.org/10.1145/1518701.1518821>.
- [15] J.E. Park, J.W. Cho, J.H. Jang, Keyword trends for mother–child oral health in Korea based on social media big data from naver, *Healthc. Inform. Res.* 26 (2020) 212–219, <https://doi.org/10.4258/hir.2020.26.3.212>.
- [16] S. Park, Y. Kim, U. Lee, M.S. Ackerman, Understanding localness of knowledge sharing: a study of naver KiN “Here,” *MobileHCI 2014* -, Proc. 16th ACM Int. Conf. Human-Computer Interact. with Mob. Devices Serv (2014) 13–22, <https://doi.org/10.1145/2628363.2628407>.
- [17] S. Jung, *Communication for ethics and corrective action in cyberspace*, 10, Cybercommunication Academic Society, 2002, pp. 75–104.
- [18] A.A. Jalal, B.H. Ali, Text documents clustering using data mining techniques, *Int. J. Electr. Comput. Eng.* 11 (2021) 664–670, <https://doi.org/10.11591/ijece.v11i1.pp664-670>.