J Korean Soc Dent Hyg 2022;22(6):495-502
https://doi.org/10.13065/jksdh.20220056

JKSDH

pISSN : 2287-1705
eISSN : 2288-2294

Original Article

Research trends in dental hygiene based on topic modeling and semantic network analysis

Yun-Jeong Kim[1] · Jae-Hee Roh[2]

[1]Department of Dental Hygiene, Kwangju Women's University

[2]Department of Pet Health, Kwangju Women's University

Corresponding Author: Jae-Hee Roh, Department of Pet Health, Kwangju Women's University, 40 Gwangjuyeodai-gil, Gwangsan-gu, Gwangju-si, 62396, Korea. Tel: +82-62-950-3935, Fax: +82-62-950-3841, E-mail: snowida@kwu.ac.kr

ABSTRACT

Objectives: The purpose of this study was to analyze research trends in dental hygiene using topic modeling and semantic network analysis. **Methods:** A total of 261 published studies were collected 686 key words from the Research Information Sharing Service (RISS) by 2019-2021. Topic modeling and semantic network analysis were performed using Textom. **Results:** The most frequently and frequency-inverse document frequently key words were 'dental hygienist', 'oral health', 'elderly', 'periodontal disease', 'dental hygiene'. N-gram of key words show that 'dental hygienist-emotional labor', 'dental hygienist-elderly', 'dental hygienist-job performance', 'oral health-quality of life', 'oral health-periodontal disease' etc. were frequently. Key words with high degree centrality were 'dental hygienist (0.317)', 'oral health (0.239)', 'elderly (0.127)', 'job satisfaction (0.057)', 'dental care (0.049)'. Extracted topics were 5 by topic modeling. **Conclusions:** Results from the current study could be available to know research trends in dental hygiene and it is necessary to improve more detailed and qualitative analysis in follow-up study.

Key Words: Dental hygiene, Research trends, Semantic network analysis, Topic modeling

Introduction

The *Journal of the Korean Society of Dental Hygiene* is one of 334 registered (candidate) journals in the field of medicine and pharmacology [1] and, as a representative academic exchange of dental hygiene, publishes the largest number of articles in the field of dentistry annually [2]. The editorial board is committed to managing article quality by organizing article reviews and editorial member workshops, selecting journal publication support projects for the *Journal of Korean Federation of Science and Technology Societies*, listing in the Directory of Open Access Journal (DOAJ), joining the Korean Association of Medical Journal Editors (KAMJE), and maintaining journal listings in the National Research Foundation of Korea and KoreaMed [3]. Along with these efforts at the academic society level, meaningful work to suggest future directions by promoting better research progress through the analysis of accumulated research achievements is essential for academic development [4].

With this need, research trend analysis provides valuable information on future research directions by identifying trends in relevant disciplines [5]. Research trend analysis using conventional methods has been performed as a content analysis method in which researchers directly read, code, and analyze the documents, which requires subjective interpretations and has limited external validity [6]. Recently, however, an increasing number of studies have applied text analysis, a big data analysis method that minimizes the problem of subjective interpretation by researchers by understanding the special relationship and meaning of collected data, ensuring research objectivity, and effectively analyzing large amounts of information in a short time to provide new perspectives

Received November 09, 2022 **Revised** November 22, 2022 **Accepted** November 22, 2022

and predictions [7-9].

Semantic network analysis, widely used in social science research or pedagogy, is a method of analyzing meaning by extracting words from data composed of text or language and identifying connectivity relationships based on the simultaneous appearance relationships between words [10]. As semantic network analysis was developed in social network service, it was initially called social network analysis; since then, this method has been used in combination with semantic, language, keyword, and simultaneous appearance network analyses [11]. The present study applied semantic network analysis as it was appropriate for identifying research trends by understanding the semantic relationship of major terms. Semantic network analysis represents the degree of connection between keywords as the degree of centrality, in which the higher the number of links connected to nodes, the higher the result value [12].

Topic modeling helps identify major research topics or trends by extracting latent topics based on words used in large volumes of documents or text in data [11]. This method designates the number of topics and arranges the words in order according to the probability of belonging to each topic; the keyword representing the topic reveals related keywords according to the distribution probability of each word [13]. The types of analysis methods include latent semantic analysis (LSA), probabilistic latent semantic analysis (PLSA), and latent Dirichlet allocation (LDA), among which LDA is the most used [14]. LDA is a form of text mining based on natural language processing that identifies the relationship between words and concepts [15]. Topic modeling combined with semantic network analysis allows the extraction of the leading research topics and the easy determination of research trends by identifying the importance and influencing relationships of the topic [11].

In 2010, the Korea Citation Index (KCI) listed only three studies that applied topic modeling and semantic network analysis in the field of pedagogy in social science; however, >70 studies have been published every year since 2020. In nursing, discussions have proposed academic development with the publication of >20 studies [16]. However, studies [17-21] on the research trends and concepts of dental hygiene for articles published in the *Journal of the Korean Society of Dental Hygiene* have mainly applied the content analysis method, which is a conventional research trend analysis.

Therefore, the present study investigated the connectivity between keywords by applying a new approach, topic modeling and semantic network analysis, to major keywords, to explore the meaning of the extracted topics and identify the research trends in dental hygiene academic articles published in the *Journal of Korean Society of Dental Hygiene*.

Methods

1. Research subjects and method

This study selected academic articles published in the *Journal of the Korean Society of Dental Hygiene* between 2019 and 2021 to identify research trends in dental hygiene. Ninety-four articles in 2019, 89 articles in 2020, and 78 articles in 2021 were collected from the Research Information Sharing Service (RISS) and 686 English keywords presented in the articles were analyzed.

2. Data analysis

For the collected keywords, spacing and singular and plural numbers were unified using the Notepad++ program. All keywords spaced within the keywords were refined by concatenating them as a single keyword. The top 50 keywords were extracted using Textom version 6.0 (The IMC Inc., Korea). The results were calculated through keyword appearance frequency analysis (TF; term frequency), weighted analysis of main keywords (TF-IDF; term frequency-inverse document frequency), co-occurrence frequency (the degree of closeness between keywords, N-gram), and latent Dirichlet allocation (LDA)-based topic modeling. The number of topics was presented in five interpretable categories by repeatedly performing topic modeling [22]. In addition, to identify the semantic network structure, the binary matrix file of one mode was converted and the centrality of the main keywords was analyzed.

Results

1. Top appearance keyword analysis

<Table 1> presents 21 keywords appearing five times or more and the results of the weighting analysis of major keywords in articles published in the *Journal of the Korean Society of Dental Hygiene*. The highest keyword frequency and weight were observed for 'dental hygienist', followed by 'oral health', 'elderly', 'periodontal disease', and 'dental hygiene' in descending order.

Table 1. Term frequency and term frequency-inverse document frequency of main key words

Rank	Keywords	TF	TF-IDF
1	Dental hygienist	71	92.43
2	Oral health	47	80.58
3	Elderly	23	55.87
4	Periodontal disease	12	36.96
5	Dental hygiene	11	34.83
6	Dental care	8	27.88
6	Dental caries	8	27.88
6	Periodontitis	8	27.88
6	Job satisfaction	8	27.88
6	Turnover intention	8	27.88
11	Self efficacy	7	25.33
11	Depression	7	25.33
11	Adolescents	7	25.33
11	Dental hygiene student	7	25.33
15	Toothbrushing	6	22.64
15	Satisfaction	6	22.64
15	Oral health behavior	6	22.64
18	Infection control	5	19.78
18	Emotional labor	5	19.78
18	Mental health	5	19.78
18	Qualitative research	5	19.78

TF: term frequency, TF-IDF: term frequency-inverse document frequency

2. Simultaneous appearance keyword analysis

<Table 2> shows the results of arranging the simultaneous appearance frequency of keywords observed three or more times. When 'dental hygienist', which had the highest keyword frequency, appeared, 'emotional labor', 'elderly', and 'job performance' also appeared simultaneously (three times each). When 'oral health', which ranked second in the keyword frequency analysis, appeared, 'quality of life' and 'periodontal disease' appeared simultaneously (three times each). 'Elderly', which ranked third in the keyword frequency analysis, had the highest simultaneous appearance frequency with 'oral health' (four times).

3. Centrality analysis

<Table 3> shows the centrality index of the main keywords, which was used to identify the semantic network structure of the articles published in the *Journal of the Korean Society of Dental Hygiene*. The order of the centrality of the major keywords was as follows: 'dental hygienist (0.317)', 'oral health (0.239)', 'elderly (0.127)', 'job satisfaction (0.057)', and 'dental care (0.049)'.

Table 2. N-gram of main keywords

Rank	Keywords 1	Keywords 2	N
1	Elderly	Oral health	4
2	Dental hygienist	Emotional labor	3
3	Mastication	Oral health	3
4	Dental hygienist	Elderly	3
5	Dental hygienist	Job performance	3
6	Dental hygiene	Dental hygienist	3
7	Oral health	Quality of life	3
8	Oral health	Periodontal disease	3

Table 3. Centrality index of main keywords

Rank	Keywords	Degree centrality	Rank	Keywords	Degree centrality
1	Dental hygienist	0.317	11	Depression	0.036
2	Oral health	0.239	12	Dental caries	0.034
3	Elderly	0.127	13	Dental hygiene student	0.034
4	Job satisfaction	0.057	14	Oral health behavior	0.034
5	Dental care	0.049	15	Emotional labor	0.034
6	Tooth brushing	0.044	16	Adolescent	0.034
7	Periodontal disease	0.042	17	Professionalism	0.034
8	Periodontitis	0.042	18	Dental hygiene	0.031
9	Turnover intention	0.042	19	Knowledge	0.031
10	Self efficacy	0.042	20	Adolescents	0.029

4. Topics and keywords extracted from articles published in the *Journal of the Korean Society of Dental Hygiene*

The topics extracted through topic analysis and the top three keywords included in the topics are presented in <Table 4>. The five derived topics were as follows: Topic 1, 'oral health'; Topic 2, 'oral health in the elderly'; Topic 3, 'periodontal tissue health'; Topic 4, 'mental health and oral health'; and Topic 5, 'dental caries management'. In addition, the research weights were identified in the following order: 'oral health in the elderly', 'mental health and oral health', 'oral health', 'periodontal tissue health', and 'dental caries management'.

Table 4. Extracted topics and keywords through the topic modeling

Number	Topic name	Weight	Keywords
1	Oral health	0.090	Dental hygienist, Oral health, Dental hygiene
2	Elderly oral health	0.101	Dental hygienist, Oral health, Elderly
3	Periodontal health	0.086	Oral health, Periodontal disease, Dental hygienist
4	Mental health and oral health	0.097	Dental hygienist, Oral health, Depression
5	Management of dental caries	0.039	Dental hygienist, Dental caries, Oral health

5. Intertopic distance map (IDM) of the identified topics

The five topics that appeared in the articles published in the *Journal of the Korean Society of Dental Hygiene* were visualized with IDM <Fig. 1>. The five topics were distributed in the first, second, and fourth quadrants. Topic 5 was distributed in the first quadrant, Topics 1–3 in the second quadrant, and Topic 4 in the fourth quadrant. Topics 1 and 3 overlapped to a large extent in the first quadrant. The IDM according to each topic is shown in <Fig. 2>.

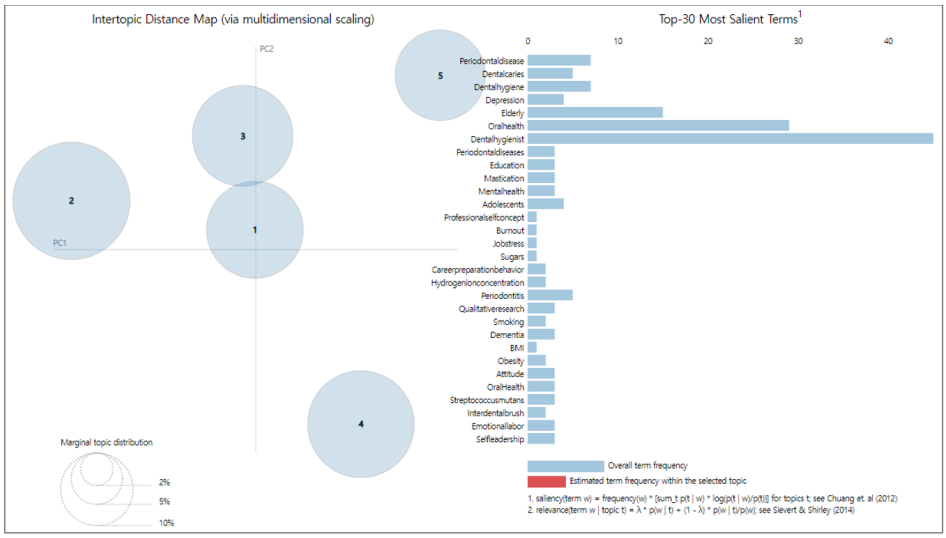

Fig. 1. Intertopic distance map of 5 topics

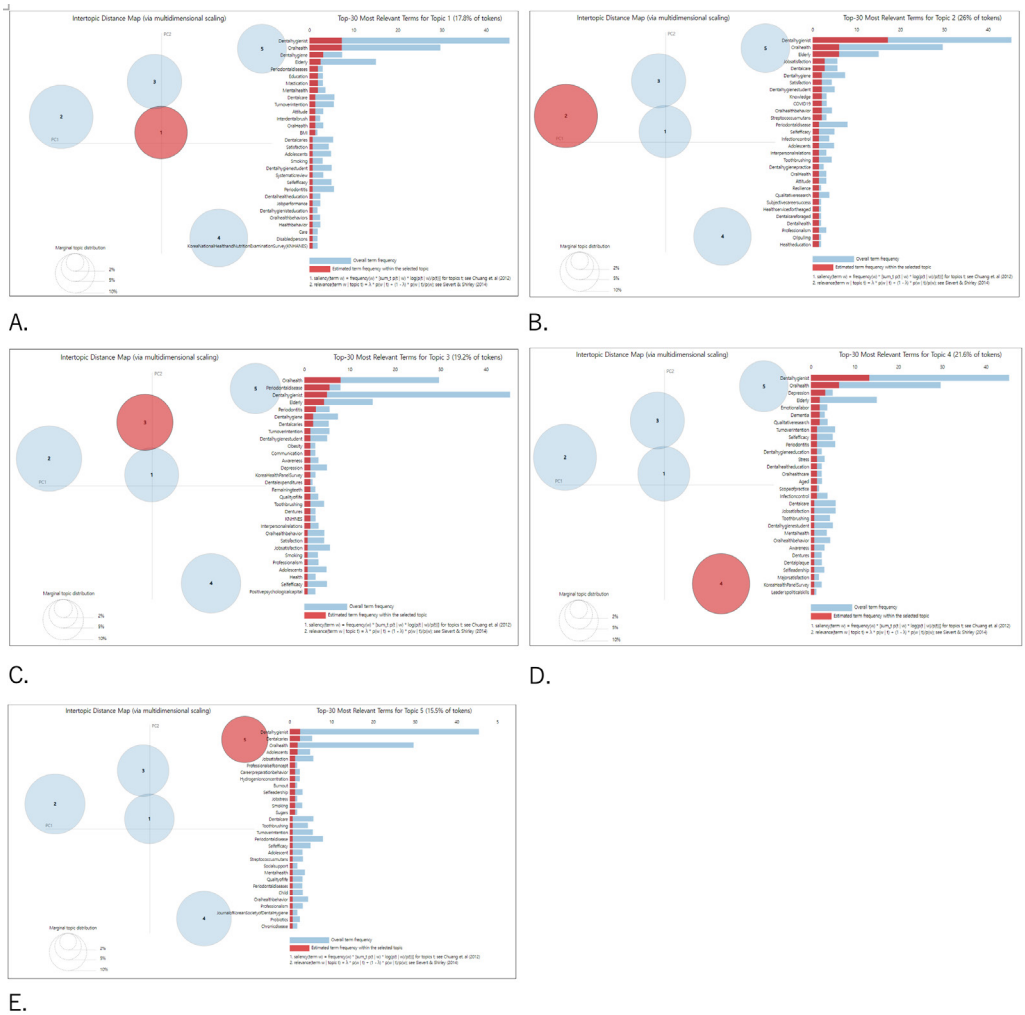

Fig. 2. Intertopic distance map of topic 1-5
A. Topic 1, B. Topic 2, C. Topic 3, D. Topic 4, E. Topic 5

Discussion

Since researchers select keywords to adequately represent and express the purpose and topic of their studies, it is easy to understand the overall knowledge structure and analyze the research [23]. Therefore, the present study empirically summarized the research trends of articles published in the *Journal of the Korean Society of Dental Hygiene* in the last 3 years based on topic modeling and semantic network analysis, in which the analysis subjects comprised the keywords in each article. Since keywords with high frequencies of appearance are likely to be important within the data containing those keywords [12] and the keywords of the documents can be extracted using the weighting analysis result of TF-IDF [9], keywords were selected through TF and TF-IDF analysis. Both TF and TF-IDF showed the highest frequencies for 'dental hygienist', followed by 'oral health', 'elderly', 'periodontal disease', and 'dental hygiene', in this order, indicating the overall relatedness of the studies.

Comparison of the keywords with the top 10 frequencies of appearance in previous studies showed that a previous study [20] analyzing articles published in the *Journal of the Korean Society of Dental Hygiene* from 2016 to 2018 included 'adolescents' and 'knowledge', while 'dental care' and 'turnover intention' were included in the results of this study, demonstrating the change in research trends over time.

The results of the N-gram analysis of the relationship between commonly used and related words showed that 'dental hygienist' and 'emotional labor', 'elderly', and ' job performance' appeared simultaneously and were highly related. Moreover, 'oral health' was widely used with 'quality of life' and 'periodontal disease'.

Since centrality analysis is an index that can identify keywords that play significant roles in a network and their impact on the semantic network [9], this method is widely used in research trend analysis [24]. Centrality analysis includes degree, closeness, and betweenness centrality, among which degree centrality is the most critical value, indicating the extent to which other words are connected in the network [25]. In addition, keywords with a high degree of centrality generally show a high frequency of appearance, which can be understood as a central concept of the text and, thus, a topic of interest in academic articles [26].

The results of the centrality analysis of the main keywords in this study revealed that the top three keywords ('dental hygienist', 'oral health', 'elderly') were the same as those in the frequency of keyword appearance, indicating that the relevant keywords had a high frequency of appearance and degree centrality. Thus, 'dental hygienist', 'oral health', 'elderly', 'job satisfaction', and 'dental care', which are high-ranked keywords, play major roles in forming the semantic structure and context of articles published in the *Journal of the Korean Society of Dental Hygiene* and are keywords with significant influence.

The results of the topic modeling analysis suggested five derived topics. As researchers can designate in advance the number of topics according to the interpretability and validity of the calculated topics [22], the present study determined the number of topics by performing several simulation processes to view the overall research trends for the past 3 years. The weights of the five derived topics were in the following order: Topic 2, 'oral health in the elderly'; Topic 4, 'mental health and oral health'; Topic 1, 'oral health'; Topic 3, 'periodontal tissue health'; and Topic 5, 'dental caries management'. The top ranking of Topic 2 demonstrates the high interest in 'oral health in the elderly' due to the increasing population and lifespan of the elderly as Korea rapidly transitions from an aging to an aged society, suggesting that considerable research has been performed.

The IDM, which represents the weights and distances between topics, can identify the degree of relevance of each topic with other topics and the similarity between topics [27]. The results of IDM analysis in the present study showed that Topic 1 ('oral health') overlapped with Topic 3 ('periodontal tissue health'). This research area was likely also affected because dental hygienists, who are clinicians, mainly perform dental plaque and calculus removal, expert brushing, and patient education for patients with periodontal diseases [28].

This study is meaningful in that it presents the research trends of articles published in the *Journal of the Korean Society of Dental Hygiene* over the past three years through TF, TF-IDF, N-gram, centrality analysis, and LDA-based topic modeling utilizing big data analysis methods, which differ from the research methods previously conducted on research trends in dental hygiene. However,

the introduction of text analysis tools such as R, Krtext, Python, Ucinet, Netminer, and Excel and training on their use at the academic society level can contribute to the academic development of dental hygiene. However, one limitation of this study was the inclusion only of articles published in the *Journal of the Korean Society of Dental Hygiene* in the past 3 years. Therefore, follow-up studies are needed to expand the selection and period of various domestic and foreign articles and apply additional text analysis methods to identify research trends and their meaningful implications in dental hygiene.

Conclusions

The results of the topic modeling and semantic network analysis of 261 academic articles published in the *Journal of the Korean Society of Dental Hygiene* from 2019 included the following:

1. The keyword appearance frequency and weight in descending order were 'dental hygienist', 'oral health', 'elderly', 'periodontal disease', and 'dental hygiene'.

2. 'Dental hygienist' appeared simultaneously with 'emotional labor', 'elderly', and 'job performance', while 'oral health' appeared simultaneously with 'quality of life' and 'periodontal disease'.

3. The ordered list of the centrality of the major keywords was as follows: 'dental hygienist (0.317)', 'oral health (0.239)', 'elderly (0.127)', 'job satisfaction (0.057)', and 'dental care (0.049)'.

4. A total of five topics were derived in the following order: Topic 2, 'oral health in the elderly'; Topic 4, 'mental health and oral health'; Topic 1, 'oral health'; Topic 3, 'periodontal tissue health'; and Topic 5, 'dental caries management.'

These results provide meaningful implications for identifying research trends in dental hygiene. Subsequent follow-up studies may provide comprehensive in-depth and qualitative analyses.

Conflicts of Interest

The authors declared no conflicts of interest.

Acknowledgements

This research was supported by Research Funds of Kwangju Women's University in KWUI22-061.

Authorship

Conceptualization: YJ Kim, JH Roh; Data collection: YJ Kim; Formal analysis: YJ Kim, JH Roh; Writing-original draft: YJ Kim, JH Roh; Writing-review&editing: YJ Kim, JH Roh

References

1. www.kci.go.kr. http://www.kci.go.kr/kciportal/main.kci [Internet]. Korea Citation Index[cited 2022 Nov 06]. Available from: https://www.kci.go.kr/kciportal/po/statistics/poStatisticsMain.kci.
2. www.kci.go.kr. http://www.kci.go.kr/kciportal/po/search/poSereSear.kci [Internet]. Korea Citation Index[cited 2022 Nov 06]. Available from: https://www.kci.go.kr/kciportal/po/search/poSereSearList.kci.
3. Kim YJ. Trend analysis of articles published in the Journal of Korean Society of Dental Hygiene, from 2016 to 2018. J Korean Soc Dent Hyg 2020;20(5):733-41. https://doi.org/10.13065/jksdh.20200067
4. Yook DI. Text mining-based analysis for research trends in vocational studies. J Korea Academia-Industrial Soc 2017;18(3):586-99. https://doi.org/10.5762/KAIS.2017.18.3.586

5. Choi JE. Keyword network analysis of trends in research on young children's play. KALCI 2019;19(14):605-26. https://doi.org/10.22251/jlcci.2019.19.14.605
6. Park HW, Leydesdorff L. Understanding the KrKwic: a computer program for the analysis of Korean text. JKDAS 2004;6(5):1377-87.
7. Kang SJ, Jung HY, Lee YS. A semantic network analysis on parents' perception of children's play space: focusing on playground and kids cafe. Educ Res 2018;38(2):281-304. https://doi.org/10.18023/kjece.2018.38.2.011
8. Reardon S. Text-mining offers clues to success. Nature 2014;509(7501):410. https://doi.org/10.1038/509410a
9. Lee GS. Semantic network analysis on preschooler safety in the COVID-19 using big data. Korean J Child Educ 2021;30(4):197-213. https://doi.org/10.17643/KJCE.2021.30.4.11
10. Lee SS. A content analysis of journal articles using the language network analysis methods. JKOSIM 2014;31(4):49-68. https://doi.org/10.3743/KOSIM.2014.31.4.049
11. Hwang SI, Park YW. An analysis of arts management-related studies' trend in Korea using topic modeling and semantic network analysis. JAMP 2019;50:5-31. https://doi.org/10.52564/JAMP.2019.50.5
12. Choi JH, Park JK, Kim MY. Analysis of research trends related to diagnosis of ASD through keyword network analysis: focusing on domestic academic journals published from 2011-2020. J Behavior Analysis Support 2021;8(1):115-35. https://doi.org/10.22874/kaba.2021.8.1.115
13. Hwang SI, Park JB, Kim MK. An analysis of humanities contents_related studies' trends in Korea focused on topic modeling and semantic network analysis. Humanities Contents 2020;56:123-41. https://doi.org/10.18658/humancon.2020.03.123
14. Ahn DJ. Analysis of the adolescent obesity research trend using the text mining in Korea and China[Doctoral dissertation]. Cheongju: Korea National of Education University, 2020.
15. YI IS, Na EY. A study on the journal analysis of cognitive field using text mining(2000-2017). J Humanities Social Sci 2016;9(3):415-26. https://doi.org/10.22143/HSS21.9.3.32
16. www.kci.go.kr. http://www.kci.go.kr/kciportal/po/search/poArtiSear.kci [Internet]. Korea Citation Index[cited 2022 Nov 06]. Available from: https://www.kci.go.kr/kciportal/po/search/poArtiSearList.kci.
17. Jang JH, Won BY, Jang GW, Kim SK, Oh SH, Kim YJ, et al. Trend analysis of research in the Journal of Korean Society Dental Hygiene from 2001 to 2015. J Korean Soc Dent Hyg 2017;17(4):693-704. https://doi.org/10.13065/jksdh.2017.17.04.693
18. Kang BW, Ahn SY, Kim SK, Yoo YS, Yoo EM, Lee SM. The research trends of papers in the journal of Korean society of dental hygiene. J Korean Soc Dent Hyg 2010;10(6):991-1000.
19. Kim YJ. Comparison of author key words and Medical Subject Heading terms in the Journal of Korean Society of Dental Hygiene from 2001 to 2015. J Korean Soc Dent Hyg 2018;18(6):1047-55. https://doi.org/10.13065/jksdh.20180090
20. Kim YJ. Analysis of authors' key words published in the Journal of Korean Society of Dental Hygiene across 3 years (2016 to 2018). J Korean Soc Dent Hyg 2019;19(6):1059-66. https://doi.org/10.13065/ jksdh.20190091
21. Kim YJ. Trend analysis of articles published in the Journal of Korean Society of Dental Hygiene, from 2016 to 2018. J Korean Soc Dent Hyg 2020;20(5):733-41. https://doi.org/10.13065/jksdh.20200067
22. Nahm CH. An illustrative application of topic modeling method to a farmer's diary. Cross Cultural Studies 2016;22(1):89-135.
23. Kim JH, Choi WS, Chung MR. Research trends and knowledge structure of Korean Society for Early Childhood Education through an analysis of keyword network. Educ Res 2017;37(3):269-88. https://doi.org/10.18023/kjece.2017.37.3.013
24. Bang SW. Analysis of research trends on online Korean language education-using topic modeling and semantic network analysis-. J Int Network Korean Language Culture 2021;18(1):1-30. https://doi.org/10.15652/ink.2021.18.1.001
25. Hwang SI, Hwang DR. A study on the research trends in arts management in Korea using topic modeling and semantic network analysis. J Arts Manage Policy 2018;48:5-29. https://doi.org/10.52564/JAMP.2018.47.5
26. Lee HR. Use of text network analysis in early childhood education research. JKOAECE 2019;24(1):293-314. https://doi.org/10.20437/KOAECE24-1-12
27. Jin MR, Ko HK. Analysis of trends in mathematics education research using text mining. J Korea Soc Math Ed Ser E 2019;33(3):275-94. https://doi.org/10.7468/jksmee.2019.33.3.275
28. Moon SE, Hong SH, Kim YJ, Kim SY, Cho HE, Kang HJ, et al. A comparative study of the perceptions of dental hygienists and dentists of nonsurgical periodontal therapy: application of a co-orientation model. J Korean Soc Dent Hyg 2020;20(1):107-16. https://doi.org/10.13065/jksdh.20200011