

# Research on the Discourse of Libraries During COVID-19 in YouTube Videos Using Topic Modeling and Social Network Analysis

**Euikyung Oh** 

Department of Library and Information Science, Sangmyung University,  
Seoul, Korea  
E-mail: oheui@smu.ac.kr

**Ok Nam Park\*** 

Department of Library and Information Science, Sangmyung University,  
Seoul, Korea  
E-mail: ponda@smu.ac.kr


## ABSTRACT

This study explored issues related to the library in the COVID-19 era in YouTube videos in Korea. This study performed social network analysis and topic modeling analysis by collecting 479 YouTube videos, 20,545 words, and 8,379 channels related to COVID-19 and the library from 2019 to 2020. The study results confirmed that YouTube, a social media platform, was used as an important medium to connect users and physical libraries and provide/promote online library services. In the study, major topics and keywords such as quarantine, vlog, and library identity during the COVID-19 pandemic, library services and functions, and introductions and user guides of libraries were derived. Additionally, it was identified that videos about COVID-19 and the library are being produced by various actors (news and media channels, libraries, government agencies, librarians, and individual users). However, the study also identified that the actor network is fragmented through the channel network, showing a low density or weak linkage, and that the centrality of the library in the actor network is weak.

**Keywords:** COVID-19, library, YouTube, social media, network analysis, topic modeling

**Received:** February 17, 2023  
**Accepted:** April 9, 2023

**Revised:** March 27, 2023  
**Published:** September 30, 2023

**\*Corresponding Author:** Ok Nam Park  
 <https://orcid.org/0000-0001-7046-2081>  
**E-mail:** ponda@smu.ac.kr



All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

## 1. INTRODUCTION

Since the first confirmed case of COVID-19 in December 2019, the world has been struggling with the COVID-19 pandemic. In South Korea, the first confirmed case was reported in January 2020, and since then the COVID-19 situation has continued, with several rapidly spreading and subsiding periods for more than three years.

The spread of COVID-19 has introduced a contactless environment in all sectors of society. As the situation continues longer than expected, “contactless” is being accepted as a core keyword in the new normal era. Additionally, due to social distancing measures introduced to prevent the spread of COVID-19 infection, participation in events and gatherings of large numbers of people was restricted, and accordingly, the use of cultural institutions such as libraries, art galleries, and museums was greatly reduced.

Concerning libraries, it was inevitable that some services had to be suspended, with opening hours adjusted and closed for a certain period according to the social distancing stage. Namely, libraries had to go through a repeated cycle of closures and re-openings. However, this did not mean libraries ceased their services during closures. They actively introduced new alternatives, such as using the library’s digital content, strengthening existing online services, implementing contactless loan services such as book drive-throughs, and continuing existing library programs using a real-time video platform. Also, in the reopening stage, libraries have maintained their services and have made efforts to change the so-called 4S (Space, Service, Safety, Sanitization) and to adapt to the changes by reconstructing their space to minimize face-to-face contact, and setting safety measures and sanitary rules to protect users from infection and allow them to use the libraries safely (Dobrevă & Anghelescu, 2022). Currently, the possibility of COVID-19 becoming endemic is being raised. It has come to a point where we must focus on the challenges and changes experienced so far and settle into a new form of libraries rather than going back to the pre-COVID-19 outbreak situation.

Therefore, it is necessary to investigate the discourses surrounding the library in the context of a COVID-19 disaster, namely, the library’s disaster response, the services provided by the library, the library usage behavior, and the perception of and demand for the library, and use the results as data that enable the establishment of a direction for the new library. Although these discourses are also dealt with in various media and in studies, they can occur

more freely and in a more diverse way on social media. Social network analysis can identify the influence of library services performed through social media in contactless situations and can provide guidelines on how libraries can use social media in the future.

Based on this background, this study explores the “COVID-19 and the library” issue over the past two years through social network analysis and topic modeling analysis. The analysis was limited to the current situation in South Korea, and data from YouTube among multiple social media were collected and analyzed.

## 2. LITERATURE REVIEW

### 2.1. Types and Characteristics of Social Media

Social media refers to a service platform that connects people who have signed up for a social networking service (SNS). SNS is a service that connects one person to another. There are various types of social media, and although the characteristics of each platform are different, they have in common features such as profile management, content production and sharing, presentation of opinions, and social network management. However, there are differences in user interaction methods and content types. Among them, YouTube is the world’s largest video sharing site, launched in 2005, where users can upload videos they have created, watch them with others, leave comments, and share them. These videos can be shared through other social media channels such as Twitter and Facebook. Recently, YouTube has been positioned as a means of active communication between content creators and users with the activation of information recommendations by algorithms and the activation of comments. Additionally, as the use of YouTube as an information search tool has rapidly increased, the number of monthly users has exceeded 1.5 billion, making it the second-largest search engine in the world after Google (Kong & Ahn, 2020).

### 2.2. Studies on Libraries’ Response to COVID-19

Since the COVID-19 pandemic, most research on libraries has been conducted on their responses to the COVID-19 situation and the information needs of users. Libraries provided information on personal hygiene and electronic information sources (Oméluzor et al., 2022), repaired existing services, or developed and supported new services (e.g., tele-reference service; Avila et al., 2022). Additionally, since users wanted their information usage pattern before COVID-19 to be maintained, they requested information on the latest topics and a diversity of tools

and services that can be accessed through contactless interaction (Harlow, 2022; Wahler et al., 2022). In addition, during this period, social media played an active role as an information communicator in various fields of society, and it was reported that when the library provided the latest information using social media, it was highly useful in resolving users' information needs (Kerns & Robertson, 2022). Additionally, it was confirmed that significant changes were made in digital applications, user support, and librarian education over the COVID-19 period (Barreto et al., 2022).

### 2.3. Social Network Analysis Study of Library Data in COVID-19

Social network analysis studies related to COVID-19 have also been conducted from various perspectives. At the beginning of the epidemic, social network analysis was performed to understand the spread of COVID-19-related information, trust in information, and social perception and frame (Hung et al., 2020; Kim et al., 2022; Yum, 2020). Over time, social network analysis was used to study changes and responses to sustain daily life in the COVID-19 situation, and social media data analysis studies on the library in relation to COVID-19 have appeared in the library field.

Park and Oh (2020) analyzed news reporting patterns and major issue changes using text-mining technology to recognize the library field's activities and changes in the environment surrounding the library in response to the spread of COVID-19. Based on 1,852 news reports and 227,983 library-related tweets, four issues were derived, namely, prolonged contactless situations, increases in e-book loans, improving expectations for online services and librarians, and reexamining library space needs. A follow-up study comprehensively summarized the responses of libraries in South Korea to the spread of COVID-19 and investigated user responses to library-related issues by analyzing 496,741 tweets related to libraries in 2019 and 2020. The analysis results revealed that there were four issues, namely, COVID-19 and lack of face-to-face service, e-books and electronic services, library operation and hosted events, and use of space and materials. It was also confirmed that the aspects mentioned in the tweets varied according to the closure period and partial opening period while libraries went through a period of four temporary closures and three partial openings during 2020 (Park & Oh, 2021).

Alajmi and Albudaiwi (2021) investigated the use of Twitter in public libraries during the first few months af-

ter the outbreak of the COVID-19 pandemic. The study analyzed 9,450 tweets posted by 38 public libraries in New York from December 2019 to April 2020; 85.5% of the tweets posted by the New York Public Library system included information about routine library services (information on remote library services available during the lockdown, social support information, etc.), and 14.5% were COVID-19 information. During the pandemic, most public libraries in New York City continued to operate as usual and supported the community in maintaining a sense of calm during the tense period.

Osakwe and Cortés (2021) analyzed information shared on Twitter in Spanish about the COVID-19 pandemic using a text mining approach. About 10,000 tweets were collected by searching for "Coronavirus," "COVID-19," "Corona," "#COVID19," and "#Coronavirus" from June 3 to June 10, 2020, and were categorized by topic. As a result, six themes were identified: (1) prevention measures, (2) epidemiology/surveillance, (3) economic impact, (4) optimizing the nursing workforce, (5) access to reliable information, and (6) a call for a response from the local government. The top trending hashtags were #COVID19 (n=7,098), #Coronavirus (n=6,394), and #SN-TESALUD (n=2,598).

Therefore, social network analysis studies of the COVID-19 discourse have mainly been conducted using Twitter. Since YouTube is characterized by active reactions and sharing between creators and users, users and users, and other social media platforms, the analysis of library discourse on YouTube is expected to capture meaningful changes in a broader and more multifaceted context than previous studies.

## 3. METHODS

### 3.1. Data Collection and Preprocessing

To perform social network and topic modeling analysis on libraries and COVID-19 on YouTube, data were collected for 24 months from January 2019 to December 2021, when COVID-19 broke out in Korea. Keyword network analysis program NetMiner4.4 (Cyrax Inc., Sungnam, Korea) was used to collect and analyze YouTube data. The keywords "COVID-19," "coronavirus," "library," and "libraries" were used to search for data.

The keywords used in the collected data were organized by word spacing, part of speech, and similar words. Preprocessing was performed where words were difficult to understand. Preprocessing methods were based on studies by Feinerer and Hornik (2014) and Oh and Park

(2018). Punctuation, numbers, symbols, stopwords, and words that were less than three letters long (e.g., !, \*, and, or) were removed. The terms “YouTube,” “video,” “Coronavirus,” and “COVID-19,” which were common in all videos, were excluded, and similar terms were gathered. After applying this process, 479 videos, 20,545 words, and 8,379 channels were collected. The video trend by year is seen in Fig. 1. Related videos continue to increase and decrease, but it can be seen that more YouTube videos were created in 2020 (287), when COVID-19 emerged and library access and face-to-face services were suspended, compared with 2021 (192), when library access and face-to-face services were resumed.

### 3.2. Analysis Method

In this study, first, keyword frequency extraction, centrality, co-occurrence keyword frequency analysis, and network analysis were conducted to analyze the keywords of YouTube videos related to libraries during the time of COVID-19. Second, topic modeling analysis was conducted to understand the topic of YouTube videos related to the library at the time of COVID-19. Third, frequency extraction, centrality, and network analysis were performed

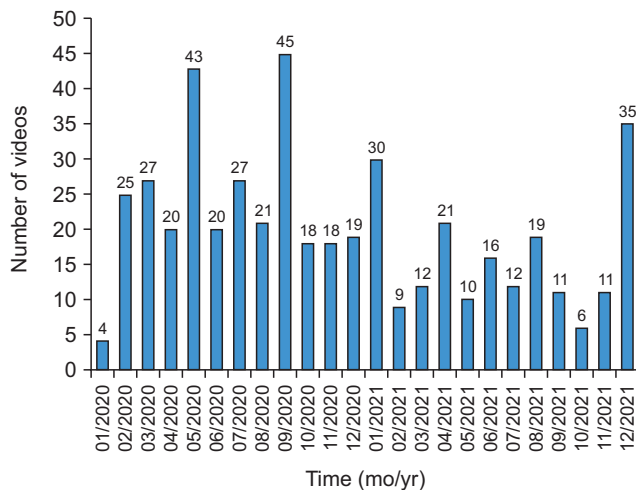


Fig. 1. Video trend by year.

to analyze the user community. The following Table 1 shows data items collected to analyze YouTube videos, terms, and channels.

#### 3.2.1. Social Network Analysis

Social network analysis is a method to quantitatively analyze the topological structure and diffusion process by modeling the relationship between individuals and groups as nodes and links. By analyzing the network consisting of nodes and links to analyze the density, distance, cohesion, connection degree, centrality, etc., it is possible to understand the connection between entities and the knowledge structure for a specific issue (Borgatti et al., 2013).

For social network analysis on YouTube video, keywords, channels, and replier/commenters were extracted, and relational properties between videos were extracted as a link. Additionally, keyword analysis, co-occurrence network analysis, and centrality analysis were performed to investigate the keywords and network between keywords appearing in the video and channels and the channel networks of the video and replier/commenters.

#### 3.2.2. Topic Modeling

Topic modeling analysis is an algorithm that automatically extracts a topic or topic group representing the texts based on the simultaneous use pattern of keywords from massive unstructured data (Blei, 2012). Topic modeling is a statistical model that derives the topic of document groups, and it consists of a probabilistic set of topics. In this study, Latent Dirichlet Allocation (LDA), introduced in 2003 in the seminal paper of Blei et al. (2003), was applied among topic modeling models. The LDA method estimates the distribution of terms and documents through the Bayesian technique, which assumes that there is a prior distribution of terms and documents, and infers main topics constituting the entire text data and keywords constituting the topics.

Topic modeling and social network analysis are being used in various fields to understand research trends and knowledge structures on issues. Zhang et al. (2012) performed social network analysis to analyze research trends

Table 1. Data collection method and content

Video	Video ID, video title, date/time, channel ID, channel name, views, likes, dislikes, comments, description
Words	- Word ID, word, part of speech, frequency - Related channel ID, terms used by the user in videos, comments/replies, total number of times used
Channel	- Channel ID, channel name - Video ID, video title, writing date/time, writer ID, writer name, comment/reply ID, number of comments/replies, whether there are comments/replies, full text

in the field of patient adherence. Their study carried out co-occurrence network analysis and social network analysis on 2,308 articles from 2000 to 2011 in the Web of Science. The study found that the research topic in the early stage reflected the general research content of the study, but in the later period, many new terms appeared and the research field was greatly expanded. Jussila et al. (2017) used social network analysis and topic modeling to analyze social big-data-related researchers and related topics. The study searched 58 articles related to social big data and compared the co-authorship network and citation network and major topics. Recently, topic modeling and social network analysis have also been used to analyze issues related to COVID-19 on social media. Zhang et al. (2021) collected tweets about the three anti-epidemic measures of COVID-19 (mask, vaccine, lockdown) on Twitter from February to October 2020, focusing on four cities in Canada and the US. The collected tweets were analyzed, focusing on human emotional responses. As a result, it was found that public sentiment about COVID-19 differed by time and place, and in general, people have positive feelings about COVID-19 and masks but have negative feelings on topics about vaccines and lockdowns. In other words, topic modeling and social network analysis are being used to identify discourses on various issues in various fields and are used as methods to analyze social media data.

In this study, keyword analysis and topic modeling analysis were conducted together. While keyword analysis is a research method that quantitatively identifies frequently occurring keywords and co-occurring keywords, topic modeling is a method that reversely classifies the topics that appear in a set of documents based on the

probability distribution, so they are complementary.

## 4. FINDINGS

### 4.1. Keyword Analysis

Social network analysis was performed for keyword analysis. Because of the analysis, it was found that 20,545 keywords were used in the 479 videos. The keywords most frequently used in COVID-19 and the library videos were “video” (2,348 times, 153 videos), “use” (461 times, 112 videos), “prevention” (943 times, 94 videos), “online” (333 times, 92 videos), “subscription” (442 times, 86 videos), “books” (209 times, 81 videos), “homepage” (174 times, 76 videos), “channel” (482 times, 75 videos), “progress” (272 times, 75 videos), “society” (596 times, 74 videos), “school” (966 times, 71 videos), “culture” (435 times, 71 videos), and “class/lecture” (424 times, 72 videos).

To analyze the co-occurrence network among keywords used on YouTube, 189 keywords appearing in more than 20 videos were extracted. Through this, 14,575 pairs of co-occurrence keywords were found, and 66 pairs of keywords that co-occurred in more than 10 videos were extracted to analyze the main keywords that appeared together. The keywords that appeared the most mainly consisted of content related to COVID-19 preventive measures such as “distance [-ing]” (41 times) and “quarantine, rules” (35 times), and content related to the library services, such as “loans, books” (20 times), “book, topic” (19 times), “operated, program” (14 times), and “lecture, content” (13 times). The top co-occurrence keywords are as follows in Table 2.

Centrality was measured to understand the location and network structure of keywords in the co-occurrence

**Table 2.** Top co-occurring keywords

Co-occurring keywords	No. of videos	Co-occurring keywords	No. of videos	Co-occurring keywords	No. of videos
Star, Gram	43	Video, Production	17	At home, Europe	13
Distance [-ing]	41	Loan, Service	17	Drive, Through	13
Subscription, Channel	35	Secondary, School	17	Era, Travel	13
Quarantine, Rules	27	Temporary, Closed	17	Children, Youth	13
Facilities, Use	24	Society, Region	14	Multiple, Use	13
Culture, Foundation	20	Operation, Program	14	Video, Program	12
Loan, Book	20	Wearing, Mask	14	Elementary, School	12
Online, Report	19	Lecture, Content	13	Method, Use	12
Book, Topic	19	Infection, Group	13	Lecture, Lecturer	12
District, Seodaemun	17	Travel, Europe	13	Culture, Art	11



network. Network centrality is an indicator that can determine how important a specific node is in the network structure or how central it is to the entire network structure. That is, the higher the centrality, the greater the influence within the network. Among the methods to determine centrality, this study used PageRank centrality (Cambridge Intelligence, Cambridge, UK). PageRank centrality is a method of assigning weight according to the relative importance of documents. Lim (2019)'s study performed a regression analysis on the number of views on YouTube videos focusing on wide network centrality indicators, and it was found that among the network centrality indicators, PageRank centrality appeared as the index that had a stable and positive effect on the number of image views.

Because of the PageRank centrality analysis ( $\alpha =$

0.85), keywords with high centrality were in the order of channel, culture, use, report, quarantine, children, school, positive test results, lecture, street, loan, book, society, travel, video, program, and homepage. These keywords were frequently used along with other keywords, indicating that these keywords were important terms in the network when paired with related terms. That is, it can be seen that YouTube videos related to these keywords were actively produced. It can be seen that the network of these keywords forms a cluster of four groups as shown in the following Fig. 2. Fig. 2 is a keyword network map visualized with Spring Map based on pathfinder networks (PFnet) performance data. A PFnet is a type of hierarchical clustering technique that forms a cluster by connecting nodes with nearest neighbors as a PFnet. Group 1 consists of keywords related to YouTube channels and individuals

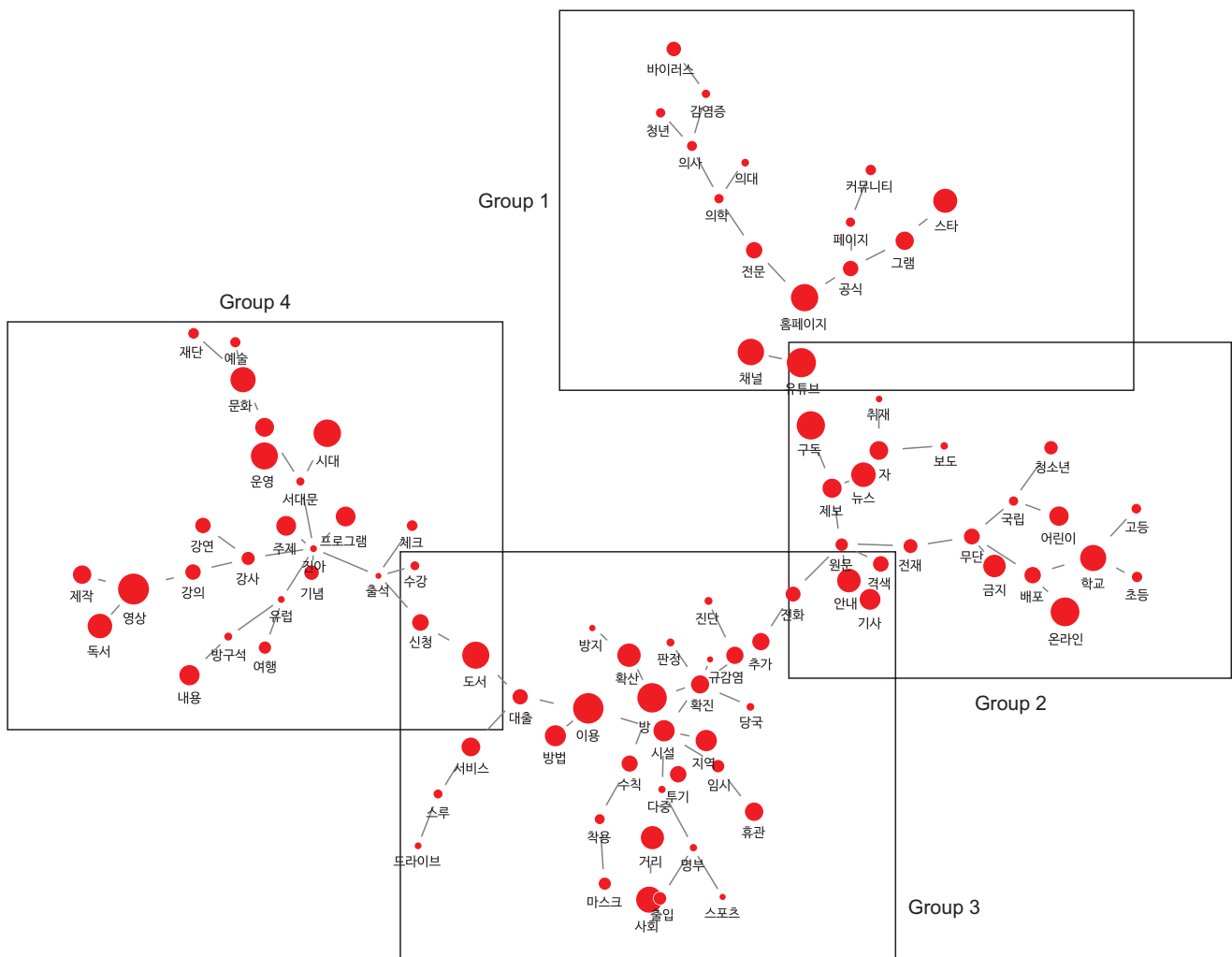


Fig. 2. Keyword network visualization.

(e.g., Instagram, administrators, homepages, channels, medical librarians), and Group 2 consists of keywords related to library users (e.g., school, youth, and children). Group 3 consists of keywords related to library quarantine and spread, and news about COVID-19 (e.g., news, COVID-19 positive, quarantine, report, and reporter), and Group 4 consists of keywords related to library work (e.g., loan, program, lecture, culture, reading). PageRank centrality, frequency, and the number of videos appearing from the most frequent keywords are shown in Table 3, and are arranged in descending order according to PageRank centrality.

#### 4.2. Video Analysis

YouTube videos related to library and COVID-19 issues are composed of various subjects and topics. In this study, topics were extracted based on the probability of keywords appearing in YouTube videos using the LDA technique, and the topic was derived by analyzing the documents related to the topic. Various K values (different numbers of topics) were used to derive meaningful results from the collected videos. The topic modeling parameters used in this study were as follows:  $\alpha$ : 0.1,  $\beta$ : 0.001, iteration: 5,000, and the final 20 topics were selected as K values. The probability of the appearance of 20 topics extracted from topic modeling among the total was analyzed. The topics that appeared more than 5% in frequency were Topic 1, Topic 3, Topic 4, Topic 9, Topic 10, Topic 12, Topic 15, Topic 17, Topic 19, and Topic 20, and 10 or more topics accounted for 65% of the total. Topic 9,

Topic 15, and Topic 3 appeared most frequently, and Topic 18 appeared least frequently. The label of each topic was selected by reviewing the probability of each topic, the top five terms for each topic, documents with a high probability (probability > 0.6) in each topic (the most relevant documents for each topic), and topics with similar topics, which were clustered to form the subject (Table 4). As a result of topic modeling, it was identified that YouTube videos are being produced under five major subjects: quarantine, vlogs, library identity during COVID-19, library services and functions, and library information and use guidelines.

The first major subject consists of topics related to the prevention of COVID-19 and includes Topics 2, 7, 6, 12, and 18. Topic 2 mainly deals with countermeasures of libraries such as library usage etiquette in the era of COVID-19. Topic 6 is about organizations applying the COVID-19 quarantine pass and related content, and Topic 7 mentions the closure of facilities due to COVID-19. Topic 17 focused on library disinfection, such as QR code verification for library access. Topic 18 consists of briefing videos on COVID-19 cases. Since the topics related to the prevention of COVID-19 contain videos containing news related to the library and general news about COVID-19, five topics were included, but all topics except Topic 12 accounted for less than 5% probability. The main keywords in Topic 12 are “prevention,” “facility,” “stage,” “spreading,” and “metropolitan area.”

The second major subject is vlog-related topics and includes Topics 3, 5, 13, and 15. The term “vlog” is a com-

**Table 3.** Analysis of high-frequency keywords

Term	PageRank centrality	Frequency	No. of appearances in YouTube	Term	PageRank centrality	Frequency	No. of appearances in YouTube
Channel	0.000535	482	75	Distance	0.000252	377	57
Culture	0.000373	435	71	Loan	0.000252	152	38
Use	0.000373	461	112	Book	0.000252	209	81
Report	0.000322	300	46	Society	0.000252	506	74
Reporter	0.000284	460	46	Travel	0.000252	223	25
Quarantine	0.000284	943	94	Video	0.000252	2,348	153
Children	0.000284	165	45	Europe	0.000252	176	14
School	0.000284	966	71	Program	0.000252	156	48
Confirmed	0.000284	822	43	Home Page	0.000214	174	76
Lecture	0.000252	95	38	News	0.000208	600	66
Class	0.000252	329	38	Formula	0.000207	167	34

**Table 4.** Topic modeling of YouTube videos

No.	Topic name	Probable terms	Probability	Subject
1	Education change and response in the time of COVID-19	Education, space, career, center, support	5.35	Library functions
2	Guide to using the library during social distancing	Social distancing, patient, society, occurrence, response	2.88	Quarantine
3	Librarian vlog	Video, Instagram, Facebook, YouTube, Life	5.35	Vlog
4	Class and lecture	Class, video, comments, youth, humanities	7.41	Library functions
5	Study in the library	Study, time, math, problem, concept	1.85	Vlog
6	Application of quarantine pass	Prevention, quarantine pass, vaccination, meeting, youth	3.09	Quarantine
7	Spread of COVID-19 infection and case update	Report, reporter, confirmation, infection, Seoul	4.73	Quarantine
8	Library tour during the time of COVID-19	Method, preparation, video, content, knowledge	3.70	Library introduction and use guide
9	Book circulation during the time of COVID-19	Service, loan, book, use, drive-through	9.05	Library functions
10	Introduction to schools and libraries in the time of COVID-19	School, student, university, class, grade	5.76	Vlog
11	Library lecture	Byeolmadang (Starfield), heart, society, Gunpo city, children, parents	3.91	Library functions
12	Quarantine due to the spread of COVID-19	Quarantine, facility, stage, metropolitan area, spread	5.14	Quarantine
13	Health information	Doctor, medical school, professor, digital, youth	3.09	Vlog
14	Library culture program	Culture, Seoul, art, neighborhood, foundation	4.94	Library functions
15	Users' library use	English, vlog, video, daily life, introduction	8.23	Vlog
16	Book curation	Literature, world, classic, earth, author	4.12	Library functions
17	Diverse trials of libraries during COVID-19	Reading, librarian, lifelong learning, citizenship, camping	6.99	Library's new identity
18	COVID-19 briefing	Confirmation, test, Daegu, quarantine, Ulsan	1.65	Quarantine
19	How to use the library during COVID-19	Children, video, performance, parliament, online	6.58	Library introduction and use guide
20	Library incentives during COVID-19	News, use, Daejeon, bookstore, application	6.17	Library's new identity

bination of “video” and “blog” and refers to video content that captures daily life as a video. Vlogs have become popular as a means of communicating with others by sharing individuals’ own daily lives online. Vlogs were categorized into those run by librarians or libraries and those run by users. Topic 5 and Topic 15 are vlogs created by users. Topic 5 contains the daily life of studying in the library, such as studying for exams and studying English during the COVID-19 outbreak, and Topic 15 contains the vlogs of users who use the library during COVID-19, such as dating in libraries and visiting the library in the time of

no face-to-face services. Topic 10 includes vlogs run by a librarian or library, containing the librarian’s book recommendations and the changed daily life of the librarian. Topic 13 contains a wide range of health information from medical librarians. Among these, the topic with the highest probability was Topic 15, and user vlogs about using libraries in various ways accounted for 8.23%, with librarians’ vlogs accounting for 5.35%. The main keywords of Topic 15 are “vlog,” “video,” “daily life,” “English,” and “introduction.”

The third major subject includes topics related to li-



brary orientation and use guides, and consists of Topic 8, Topic 10, and Topic 19. These topics are composed of videos that introduce the library space in contactless way during COVID-19 and give guidelines for using the library. Topic 8 focused on library tours during the COVID era, and Topic 10 is about the introduction to university, college, and school libraries during the pandemic, and is mainly directed at first-year student orientations and campus tours. Topic 19 is about library usage guidelines for children and consists of videos on how to use the library. Topic 19 had the highest probability (6.6%) and is composed of keywords such as “video,” “performance,” “online,” and “children.”

The fourth major subject consists of Topic 1, Topic 4, Topic 9, Topic 11, Topic 14, and Topic 16 and is about the main functions of libraries during COVID-19 (loan/returns, cultural programs, education, and lectures). Topic 1 mainly consists of library education videos about education and careers that have changed due to the COVID-19 pandemic, and the main keywords are “education,” “career path,” “center,” and “support.” Topic 4 includes library videos about humanities lectures, and the main keywords are “class,” “video,” “humanities,” “comment,” and “youth.” Topic 9 included introductions to loan/return services during COVID-19, and the main keywords were “service,” “loan,” “book,” “use,” and “drive-through.” Topic 11 was about classes/lectures delivered by the library, such as those about the message of child-rearing and happiness, and Topic 14 was about various library programs and consisted of keywords such as “culture,” “neighborhood,” and “art.” In Topic 16, it can be seen that keywords such as culture, world, and classic have become issues through book curation topics. Through this, it can be seen that the library was closed or operated limitedly in the early days of COVID-19, but as the non-face-to-face situation was prolonged, efforts were made to replace different functions of the library online. Topics 1, 4, 9, and 14 all exceeded 5%, and Topic 9 was the highest at 9% among the 20 topics, and Topic 4 was the second-highest at 7.4%.

The final subject emphasizes the changes of the library during the COVID-19 pandemic by reestablishing the identity of the library in that period. Topic 17 and Topic 20 fall into this category. Topic 17 shows the various efforts that libraries made to survive during the pandemic. For example, it includes videos about alternative uses of library space, such as camping in the library, the several efforts of librarians to provide digital library services, and online services of the library as a lifelong learning institution. The main keywords were “reading,” “librarian,” “life-

long learning,” “citizen,” “camping,” etc. Topic 20 includes videos dealing with contents such as bookstore rental services, late fee waivers, and book loans in advance as various efforts to revitalize library use which had been reduced due to COVID-19, and the main keywords are “use,” “bookstore,” and “application,” etc. Topic 17 accounted for 7%, and Topic 20 accounted for 6.2%, both accounting for more than 5% probability.

Table 4 shows the probability of the top five terms by topics. Terms in each topic are arranged in descending order based on the probability of occurrence of terms in the topic.

The following is the result of visualizing the topic modeling network and presenting it as a map (refer to Fig. 3). Topic 9, Topic 4, Topic 15, Topic 3, and Topic 19 showed the largest ego (node) size. Topic 9 (circulation/loan), with a high probability of appearance, was linked with Topic 14, Topic 17, Topic 20, and Topic 4. It can be seen that Topic 9 is connected to topics such as library services, diverse new attempts by libraries, and library incentives during COVID-19. Additionally, it can be seen that Topic 4 (class/lecture) is connected to various topics such as Topic 3 and Topic 15 (vlog), Topic 17 (different changes of the library), Topic 9 (library service), and Topic 19 (library use).

#### 4.3. Channel Analysis

The total number of channels that posted videos related to COVID-19 and the library issues and participated in comments and replies was 8,379. Channel analysis was carried out to identify influencers among the relevant users and to analyze the type of relationship created and the network flow. To this end, channels that posted the most videos, the number of video views and likes, the creator-commenter relationship, and the network between users (commenter-replier) were analyzed.

Of the 304 channels that uploaded videos, 55 were operated by libraries and two were operated by librarians. Looking at the distribution of the number of videos posted by channel, one channel posted 50 or more videos, 3 channels posted 11-20 videos, 5 channels posted 6-10 videos, 2 channels posted 5 videos, 10 channels posted 4 videos, 8 channels posted 3 videos, 39 channels posted 2 related videos, and 236 channels posted only one related video, showing that various channels posted videos related to COVID-19 and the library. Of these, 129 videos were posted by library channels, and three videos were posted by librarians. It was identified that 27.6% (132 videos) of the 479 videos were posted by libraries and librarians.

Because of analysis of the top 30 channels that up-

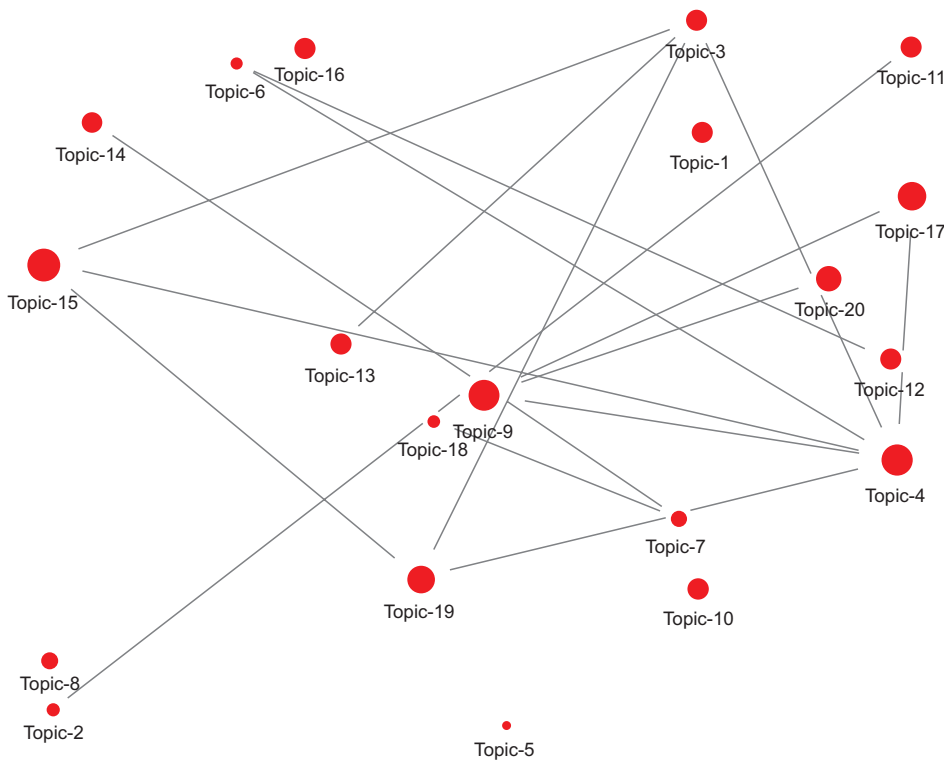


Fig. 3. Topic modeling network.

Table 5. Channel analysis of YouTube videos

	No. of node	No. of links	Density	Average degree	Clustering coefficient	Node connectivity	No. of isolated nodes
Upload-comment	7,152	6,578	0.00009	0.781	0.004	0	1,769
Comment-reply	4,704	4,088	0.00005	0.452	0.053	0	4,298

loaded three or more videos, 4 channels were operated by individuals, 13 channels were operated by libraries, seven channels were operated by broadcasting and news media centers, five channels were operated by government agencies, and one channel was operated by a company. In other words, it was identified that libraries were posting the most videos about COVID-19 and the library compared to other channels.

Network analysis of the uploader-commenter and commenter-replier links for all channels was conducted, and the research results are as follows (Table 5).

First, the results of the network analysis between the video uploader and commenter are as follows. Of the 7,152 nodes, 1,769 isolated nodes, which is 24.7% of the nodes, were not connected to other nodes, and the number of links was 6,589, which was less than the number of nodes, showing that many nodes did not form links with each other. The density was 0.00009 and the clustering

coefficient was 0.003, which is very low, and the average connectivity was close to 0, indicating that the network's cohesiveness is very low. Among the top 30 channels that posted the most videos, most of them had few or no comments except for news channels or personal vlog channels, so the network density is, inevitably, low. Because of network analysis between video creators and commenters, among the top 30 channels, only three channels were operated by libraries: Seocho Cultural Foundation (5th), Gyeonggi-do Cyber Library (13th), and Seodaemun-gu Library (30th). The rest were run by broadcast and news media centers (5 channels), schools (1 channel), government agencies (3 channels), and individuals (18 channels), with channels run by individuals making up the largest proportion. In particular, the average number of comments on posted videos was insignificant in the library channels, and it was found that libraries did not write comments on other videos related to "COVID-19 and the

library.” The COVID Seocho Cultural Foundation wrote 38 comments on 2 of 479 videos, and Gyeonggi-do Cyber Library and Seodaemun-gu Library wrote only one comment on one video each.

Second, the result of the network analysis of video commenters and repliers is as follows. Of the 4,704 nodes, 4,298 were isolated nodes, most of the nodes (91%) were not connected to other nodes, and the number of links was 4,088, which is less than the number of nodes. The density was 0.00005, the clustering coefficient was 0.053, and the average connectivity was close to 0, indicating that the network’s cohesiveness was very low as in the creator-commenter network. Of the top 30 channels with a large ego-network size and degree, only two were operated by libraries: the Seocho Cultural Foundation and Seodaemun-gu Library. The remaining 28 channels were operated by individuals. In other words, compared to individuals who are active in writing replies to comments on the video library, channels have a limited number of comments and are not active in communication activities such as replying to comments.

The following Fig. 4 shows the visualization of the network by selecting Fruchterman and Reingold’s layout (cooling coefficient 35; national length coefficient 1.0; maximum iterations 500) among the node layout algorithms. It shows that both the uploader-commenter and commenter-reply channel networks are distributed in multiple clusters, and many isolated networks exist at the bottom. In addition, it can be seen that the connectivity between different clusters or between nodes within a cluster is insufficient. Among the personal vlogs, vlogs about dating or studying in the library in daily life had the ego-

centered network with the largest ego size and the highest degree in the network cluster, so they are in the center surrounded by other users.

## 5. DISCUSSION

The results and implications of this study are as follows. First, co-occurrence and network analysis results showed that the keywords most used in “COVID-19 and the library” videos were “video,” “use,” “prevention,” “online,” “subscription,” “book,” “homepage,” “channel,” “progress,” “society,” “school,” “culture,” and “class/lecture.” The keywords that appeared the most from the co-occurrence keyword analysis mainly consisted of contents related to library services, such as “distancing,” “quarantine, rules,” “loan, book,” “book, topic,” “operated, program,” and “lecture, content.” The analysis of centrality showed that the keywords with high centrality were “channel,” “culture,” “use,” “report,” “prevention,” “children,” “school,” “confirmation,” “lecture,” “distance,” “loan,” and “book.” Through this, it was identified that the words most used in videos related to COVID-19 and the library consisted of terms related to COVID-19 news and library services or functions, and mainly videos related to these keywords were produced.

Second, 20 topics were extracted through topic modeling, and five subjects (quarantine, vlog, the library’s new identity in the time of COVID-19, library service and functions, library information and usage guide) were extracted. The subject related to COVID-19 quarantine included general news about COVID-19 and library-related news, such as library usage etiquette during COVID-19, disease prevention rules, library quarantine, closure and

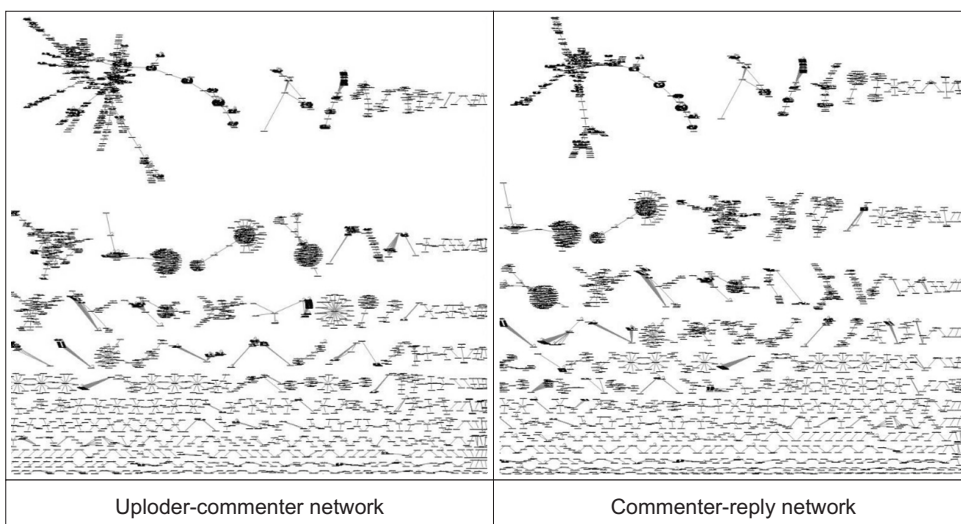


Fig. 4. Actor network.

partial reopening of facilities, COVID-19 prevention passes, and COVID-19 case updates. The subject related to vlogs is divided into vlogs run by the library and by the users. They contain topics about the daily life of librarians or library use in the daily life of users. The subject related to library information and usage guides included topics about the orientation of the library and how to use it. The subject of the library's main functions is composed of topics such as library classes/lectures, educational programs, cultural programs, and loans/returns. The last library-identity-related subject includes topics related to various changes and concerns of the library in responding to COVID-19, including topics such as providing incentives for library use during that time, changing library space, and emphasizing online services. In other words, it can be seen that differentiated videos related to libraries have been produced to respond to COVID-19, such as in library work, library identity, and library use.

Third, through the topic and keyword analysis, it was found that the library is attempting various countermeasures to overcome the changes during the COVID-19 situation and simultaneously find a sustainable operation plan afterward. Various methods were proposed, for instance, reorganizing the physical library space, encouraging the use of materials by allowing users to purchase desired books directly through bookstores, and reducing late fees. Additionally, it can be seen that libraries are thinking about ways to establish a library identity despite COVID-19, such as emphasizing online library services in the era of lifelong learning.

Fourth, through YouTube topics related to library tours and orientation, library circulation, and library functions and services, it was identified that in the early days of COVID-19, the library was closed or operated limitedly, but as the non-face-to-face situation was prolonged, efforts have been made to replace various library functions online. Varied library contactless services, such as online lectures and services of cultural programs, were provided through the YouTube platform. Simultaneously, libraries made continuous efforts to induce constant user interest in the physical library space and materials and reduce the impact on the library through videos on library tours and orientations, library loans/returns, and library usage.

Fifth, the discourse on library services during COVID-19 is being created by various participants such as libraries, news and media channels, government agencies, and individual users. Additionally, the role of the library as a producer of YouTube videos was confirmed through channel analysis. Of the 304 channels with 479 videos,

55 channels (18%) were operated by libraries, and library channels uploaded 26.9% (129) of all videos. Among the 30 channels with the highest number of videos posted, 13 (43.3%) were operated by libraries, accounting for the largest proportion. In other words, libraries were found to play a role in developing diverse services for users during the COVID-19 situation.

Sixth, it can be seen that a powerful influence is emerging on issues related to COVID-19 and the library. Additionally, it was found that the user density was low, and the connection between various actors like uploaders, commenters, and repliers was weak. Both the network between users who posted a comment-reply, and the network between the uploader and commenter showed low density and clustering coefficients, and many isolation networks showed a fragmented network that was loosely connected and divided into several clusters. Also, the connections between and within clusters were weak. This means that networks are created only in close friendships, and there is no external attempt to get out of the friendships. People who leave comments do so only on the videos they are interested in rather than commenting on many videos, meaning that the channel where the video is uploaded and the people who leave comments are different.

Seventh, findings showed that the centrality and density of the channels the library operated in the channel network were low. The library channel did not appear to comment on any other videos. The average number of comments/replies on the uploaded videos was also low. It can be seen that the videos created in the library channels are getting very few comments and very few likes, which means that they are not getting much attention reaching a wider audience. From this, it could be inferred that library channels are only used to upload videos related to the "COVID-19 and the library" topic and not as interactive channels to keep in touch with other users and other video publishers interested in this issue. This shows that the library plays a role as a content producer on YouTube, but not as an influencer. This implies that videos related to COVID-19 and the library may not reach more users as expected. This could be for a number of reasons. The library's YouTube channel may have had a low profile. It is also possible that the COVID-19 videos generated by libraries did not reflect users' interests. For this, it is vital to understand users' needs related to COVID-19.

Cooperation with various channel operators is vital, especially for channels with high influencing power or delivery rate. Efforts as active participants are required, such



as increasing the number of subscribers to channels operated by libraries or librarians, subscribing to channels that post videos related to other COVID-19 and the library topics, or writing comments and replies. Additionally, it is necessary to examine whether the contents in library and COVID-19 topics reflect users' library interests.

## 6. CONCLUSION

Due to COVID-19, the library has faced many challenges. This study explored the issue of "COVID-19 and the library" that appeared in YouTube videos. This study performed social network analysis and topic modeling analysis by collecting 479 YouTube videos, 20,545 terms, and 8,379 channels related to COVID-19 and the library from 2019 to 2020. Keyword network analysis, topic modeling analysis, and actor network analysis were performed.

This study can summarize that YouTube, a social media medium, was identified as an important platform for connecting users and the physical library, and for providing/promoting online library services. The library has provided user services in many ways in the COVID-19 situation and has been trying to connect online and offline libraries. The influence and importance of social media is growing day by day. Content production and dissemination are important to more actively inform the library's diverse changes and efforts to respond to COVID-19 through YouTube. For this, the library should cooperate with various channel operators related to the COVID-19 and the library issue, continue varied activities on the network as an active participant, and reflect on video user interests and needs.

This study is limited in that the data used in the analysis were only from 2019-2020 during the COVID-19 period. In the future, it will be necessary to check how the library discourse has changed through YouTube using time series analysis before and after the COVID-19 outbreak period. Additionally, this study was conducted on YouTube videos in South Korea. Conducting a comparative analysis using overseas YouTube will enable the comparison of differences in discourses. Future research should conduct time-series studies of how library discourse has changed on YouTube across different countries. This study is also limited in that it focuses only on YouTube and does not consider other social media platforms such as Facebook, where discussions about libraries and COVID-19 may occur. Therefore, the findings of this study may not be representative of the discourse about libraries and COVID-19 on social media.

Finally, this study did not compare the differences between the channel groups. A study that compares and analyzes the differences between major groups (libraries, librarians, individuals, government agencies, schools, news, and newspapers) that uploaded videos is necessary to reveal the characteristics of actors. However, the findings here have implications in that they identify that social media platforms such as YouTube are playing a more important role in disseminating information about library responses to the pandemic.

## CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

## REFERENCES

- Alajmi, B. M., & Albudaiwi, D. (2021). Response to COVID-19 pandemic: Where do public libraries stand? *Public Library Quarterly*, 40(6), 540-556. <https://doi.org/10.1080/01616846.2020.1827618>
- Avila, S., Norris, S. A., Murphey, R. M., Tong, M., & Harrison II, R. H. (2022). The tele-reference model: Adopting virtual tools to enhance reference services during COVID-19 and beyond. In M. Lamba (Ed.), *Technological advancements in library service innovation* (pp. 1-22). IGI Global.
- Basurto, L. F., Martínez-Camacho, H., & Calderón-Swain, A. (2022). Technological scenarios for the new normality in Latin American academic libraries. *IFLA Journal*, 48(4), 538-547. <https://doi.org/10.1177/03400352211035412>
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84. <https://doi.org/10.1145/2133806.2133826>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022. <https://dl.acm.org/doi/10.5555/944919.944937>
- Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. Sage.
- Dobreva, M., & Angheliescu, H. (2022). Libraries and COVID-19: Opportunities for innovation. *IFLA Journal*, 48(1), 3-8. <https://doi.org/10.1177/03400352221077748>
- Feinerer, I., & Hornik, K. (2023). *tm: Text mining package: A framework for text mining applications within R*. <https://cran.r-project.org/web/packages/tm/index.html>
- Harlow, S. (2022). Assessing demographics and needs of online students before and during the COVID-19 pandemic: Lessons from academic libraries. *Journal of Library & Information Services in Distance Learning*, 16(1), 20-37. <https://doi.org/10.1080/15393099.2022.2088888>



- org/10.1080/1533290X.2022.2056279
- Hung, M., Lauren, E., Hon, E. S., Birmingham, W. C., Xu, J., Su, S., Hon, S. D., Park, J., Dang, P., & Lipsky, M. S. (2020). Social network analysis of COVID-19 sentiments: Application of artificial intelligence. *Journal of Medical Internet Research*, 22(8), e22590. <https://doi.org/10.2196/22590>
- Jussila, J., Menon, K., Gupta, J., & Kärkkäinen, H. (2017, July 3-4). Who is who in big social data? A bibliographic network analysis study. In A. Skaržauskienė & N. Gudeliene (Eds.), *Proceedings of the 4th European Conference on Social Media* (pp. 161-169). Mykolas Romeris University.
- Kerns, H., & Robertson, S. (2022). Academic libraries versus the doom scroll: Engaging with at-home users on social media during COVID-19. *Journal of Electronic Resources Librarianship*, 34(1), 21-29. <https://doi.org/10.1080/1941126X.2022.2028426>
- Kim, S., Lim, H. W., & Chung, S. Y. (2022). How South Korean internet users experienced the impacts of the COVID-19 pandemic: Discourse on Instagram. *Humanities and Social Sciences Communications*, 9, 75. <https://doi.org/10.1057/s41599-022-01087-7>
- Kong, M. Y., & Ahn, H. (2020). Information adoption on YouTube: Examining the influence of YouTube genre. *Journal of the Korea Society of Computer and Information*, 25(6), 131-141. <https://doi.org/10.9708/jksci.2020.25.06.131>
- Lim, Y. S. (2019). Considering YouTube video network centrality indices. *Journal of the Korean Data Analysis Society*, 21(6), 3169-3178. <https://doi.org/10.37727/jkdas.2019.21.6.3169>
- Oh, J. S., & Park, O. N. (2018). Topics and trends in metadata research. *Journal of Information Science Theory and Practice*, 6(4), 39-53. <https://doi.org/10.1633/JISTaP.2018.6.4.4>
- Omeluzor, S. U., Nwaomah, A. E., Molokwu, U. E., & Sambo, A. S. (2022). Dissemination of information in the COVID-19 era in university libraries in Nigeria. *IFLA Journal*, 48(1), 126-137. <https://doi.org/10.1177/03400352211037700>
- Osakwe, Z. T., & Cortés, Y. I. (2021). Impact of COVID-19: A text mining analysis of Twitter data in Spanish language. *Hispanic Health Care International*, 19(4), 239-245. <https://doi.org/10.1177/15404153211020453>
- Park, T. Y., & Oh, H. J. (2020). A study on library service in the post-COVID era through issues on media. *Journal of Korean Library and Information Science Society*, 51(3), 251-279. <https://doi.org/10.16981/kliss.51.3.202009.251>
- Park, T. Y., & Oh, H. J. (2021). Analysis of library user needs in the new normal era: Focusing on social media. *Journal of Korean Library and Information Science Society*, 52(2), 303-330. <https://doi.org/10.16981/kliss.52.2.202106.303>
- Wahler, E. A., Spuller, R., Ressler, J., Bolan, K., & Burnard, N. (2022). Changing public library staff and patron needs due to the COVID-19 pandemic. *Journal of Library Administration*, 62(1), 47-66. <https://doi.org/10.1080/01930826.2021.2006985>
- Yum, S. (2020). Social network analysis for coronavirus (COVID-19) in the United States. *Social Science Quarterly*, 101(4), 1642-1647. <https://doi.org/10.1111/ssqu.12808>
- Zhang, J., Xie, J., Hou, W., Tu, X., Xu, J., Song, F., Wang, Z., & Lu, Z. (2012). Mapping the knowledge structure of research on patient adherence: Knowledge domain visualization based co-word analysis and social network analysis. *PLoS One*, 7(4), e34497. <https://doi.org/10.1371/journal.pone.0034497>
- Zhang, Q., Yi, G. Y., Chen, L. P., & He, W. (2021). Text mining and sentiment analysis of COVID-19 tweets. *arXiv*. <https://doi.org/10.48550/arXiv.2106.15354>