# The Multidimensional Evolution of Literacy and Its Social Role: An Empirical Study Through Network Analysis

Younghee Noh*, Jeong-won Roh**, Ho-won Kim***

ARTICLE INFO

*Article history:*
Onine first 31 January 2025

*Keywords:*
Literacy,
reading comprehension,
social big data, digital literacy,
media literacy,
data literacy

ABSTRACT

*Purpose*: The purpose of this study is to empirically investigate the multidimensional transformation of literacy as a core competency adapting to the digital environment in contemporary society and to propose directions for improving literacy policies and educational programs.
*Methodology*: Using social big data collected from Naver and Daum between 2021 and 2024, the study employed text mining, network analysis, CONCOR analysis, QAP regression, and topic modeling. Key keywords were identified through word frequency and TF-IDF analysis, while N-gram analysis and network visualization were used to examine the relationships between major topics.
*Findings*: The analysis revealed that digital literacy, media literacy, and data literacy emerged as core themes, and their networks demonstrated statistically significant structural patterns. QAP regression analysis confirmed that the literacy-related network exhibited clear patterns compared to a random network, and CONCOR analysis indicated that major topics were closely interconnected.
*Implication*: Policy recommendations include collaborating with community stakeholders to develop localized literacy programs, enhancing digital infrastructure to reduce the information gap, and offering lifelong learning opportunities. Subsidized digital literacy workshops can help address disparities in underserved areas. Integrated approaches in policy and education are essential, with tailored programs and policy support to reduce information inequality.

## 1. Introduction

The 21st century society is experiencing rapid digital technological advancements and an explosive increase in information. Consequently, individuals and society are required to possess competencies for solving various problems in complex informational environments. In this context, Literacy now encompasses not only reading and writing but also information search, interpretation, and critical thinking within digital contexts. Particularly in a modern society where digital media and social

* Professor, Department of Library & Information Science, Konkuk University, Korea (First Author)(iru4u@kku.ac.kr)
** Ph.D. candidate in Library and Information Science at Konkuk University(Corresponding Author)(njw3227@naver.com)
***Ph.D. candidate in Library and Information Science at Konkuk University(Co-Author)(hwonidot@gmail.com)

networks have become routine, the ability to critically evaluate and create new value from information is essential. Therefore, interest in subcategories such as digital literacy, media literacy, and data literacy has increased.

One of the critical issues in contemporary society is the overabundance of information and the spread of misinformation. The digital environment enables anyone to create and distribute information, leading to the proliferation of fake news, misinformation, and malicious rumors. This phenomenon necessitates individuals' critical acceptance and accurate judgment of information, underscoring literacy as a social competence. Moreover, addressing the digital divide and information inequality is a significant challenge, emphasizing the importance of expanding information accessibility and bridging competency gaps through literacy education.

In the field of education, literacy plays a vital role. As digital learning environments expand, the ability of students to effectively utilize diverse information and gain a deeper understanding of learning content becomes increasingly important. Consequently, literacy education has transformed from traditional text-based education to an integrated approach that includes digital media usage and critical thinking. This shift positions literacy not merely as a learning tool but as an essential educational objective aimed at nurturing responsible citizens capable of correctly interpreting and utilizing information.

However, existing studies have primarily focused on individual analyses of specific literacy types. Comprehensive research on the multidimensional interactions and structural characteristics of literacy in response to socio-technical changes is relatively scarce. Given that literacy evolves as a multidimensional concept alongside social and technological transformations, a multifaceted analysis of key topic interactions is necessary.

This study aims to elucidate the interconnectivity and structural characteristics of various topics related to literacy using social big data. Data collected from major digital platforms like Naver and Daum were analyzed through text mining, network analysis, and topic modeling to identify key discussions and trends in literacy in digital environments. Furthermore, statistical significance was verified through QAP correlation and regression analyses to empirically demonstrate literacy's role as a structural competence in contemporary society.

By exploring the relationships among literacy's diverse subdomains and its significance in social and educational contexts, this study is expected to provide systematic and comprehensive approaches to literacy policy and educational program development. Additionally, by proposing concrete strategies to enhance information utilization and critical thinking capabilities required in the digital age, the study aims to contribute to positioning literacy education as a practical tool for empowering individuals and society.

## 2. Literature Review

Literacy has evolved from traditional reading and writing skills into a more complex and multidimensional concept in modern society. This evolution is confirmed through various studies.

First, the concept of literacy has expanded to include information comprehension and critical thinking abilities, enabling individuals to function effectively in complex informational environments.

For instance, Pegado et al. (2014) analyzed the impact of reading skills on visual processing and reported that learning to read enhances visual precision and invariance.

Second, digital and media literacy have become increasingly important in contemporary society. Arshad et al. (2022) developed a three-factor model for measuring media literacy, highlighting its significance. Similarly, Papadopoulou et al. (2023) explored the integration of formal and informal literacy learning through digital media.

Third, the effectiveness of literacy education has been demonstrated in several studies. The literacy environment at home plays a crucial role in early childhood development. Deitcher et al. (2024) reported that parental literacy beliefs and home literacy activities positively influence children's early literacy skills. Additionally, Ihmeideh & Al-Maadadi (2020) emphasized the positive impact of family literacy programs on early literacy development in kindergarten settings.

Fourth, the impact of social inequality on literacy competence is noteworthy. Nag et al. (2024) conducted a meta-analysis examining the relationship between home learning environments and children's language and literacy skills in low- and middle-income countries, finding significant correlations between educational characteristics of the home and children's outcomes.

Finally, mental health literacy has also emerged as an important research topic. Yang et al. (2024) analyzed the impact of mental health literacy on seeking professional psychological help, reporting that higher mental health literacy increases social support and reduces the stigma associated with seeking psychological assistance.

Building on Ko Yoon-jung's (2023) insights on big data literacy, this study extends the analysis to include how literacy adapts to the digital environment, particularly post-COVID-19. Additionally, studies such as Papadopoulou et al. (2023) on digital media integration and Ihmeideh & Al-Maadadi (2020) on family literacy highlight the need for inclusive approaches, which this research incorporates through network and sentiment analyses. Kil Ho-yeon & Lee Ji-seon (2022) analyzed the usage patterns of literacy concepts by the general public, revealing that 'literacy' is often used to denote competence in various domains, while 'reading comprehension' focuses on ability. Additionally, Kim Seul-ki & Kim Tae-young (2022) applied a liberal arts course on big data analysis using textual programming languages at teacher training institutions, finding significant results in all areas and interpreting it as an improvement in competence.

The next sections will detail the methodology, data collection, and analysis results. The study concludes with practical suggestions for improving literacy education and policy based on empirical findings from social big data analysis.

## 3. Research Design and Methodology

### 3.1 Research Procedure and Content

This study was conducted through a series of steps to analyze the multidimensional changes in literacy and derive policy and educational implications. First, a review of existing literature on literacy was conducted to establish a theoretical foundation. Second, data was collected from major digital platforms such as Naver and Daum, covering the period from 2021 to 2024. Third, text

mining techniques were employed, including word frequency and TF-IDF analyses, to extract core keywords related to literacy. Fourth, a 2-mode matrix analysis was performed to examine literacy trends among educational stakeholders. Fifth, sentiment analysis and CONCOR analysis were conducted to investigate key discourses and societal perceptions related to literacy. Sixth, LDA (Latent Dirichlet Allocation) topic modeling was utilized to extract major topics and analyze core themes in literacy. Finally, the findings were discussed, and conclusions as well as policy and educational recommendations were presented.

| Step | Contents | Ressearch Details |
|---|---|---|
| Step 1 Preliminary Research | Review of Related Literature | • Conducted a review of literature related to literacy and big data |
| Step 2 Data Collection | Data Collection | • Collected social big data on literacy using TEXTOM |
| Step 3 Data Analysis | Analysis of Literacy Keywords through Text Mining | • Extracted core literacy keywords through word frequency analysis<br>• Identified significant literacy keywords using TF−IDF analysis<br>• Derived co−occurrence word networks through N−gram analysis |
| | Analysis of Literacy Trends by Educational Stakeholders | • Examined stakeholder−specific literacy trends through 1−mode matrix analysis |
| | Sentiment and Discourse Analysis | • Identified emotional perceptions of literacy through sentiment analysis<br>• Grouped literacy discourses using CONCOR analysis |
| | Topic Analysis | • Derived major topics through LDA topic modeling<br>• Identified literacy topics through Word−level Semantic Clustering |
| Step 4 Results Analysis and Conclusion | Discussion | • Discussed significant findings based on data analysis |
| | Conclusion and Recommendations | • Proposed improvement strategies and implications for future research |

<Fig. 1> Research Procedure and Content

2) Data Collection Method

Data was collected from the two major platforms, Naver and Daum, using the keyword "literacy" (both in Korean and English). The data collection period spanned three years, from January 1, 2021, to December 31, 2024, to capture comprehensive trends in literacy-related discussions. The collected data underwent a rigorous preprocessing stage, including text cleaning, removal of duplicates, and elimination of stop words. Keywords were selected based on their relevance, determined through TF-IDF analysis, which prioritized terms with high document-specific frequency and low overall frequency across the corpus.

Table 1. Analysis Data Information

| Category | Details |
|---|---|
| Tool | TEXTOM |
| Platforms | Naver (blogs, news, Q&A, web documents), Daum (blogs, news) |
| Period | January 1, 2021 - December 31, 2024 |
| Keywords | "Literacy" |

This study collected social big data to explore literacy trends and conducted text mining analysis, matrix analysis, and topic analysis according to the research objectives. First, to identify the main keywords of literacy, word frequency analysis and TF-IDF analysis were conducted to extract the top-ranking keywords. Word frequency analysis was performed to identify frequently occurring keywords within the overall dataset, while TF-IDF analysis was used to calculate the weighted importance of words within documents, allowing for the identification of key terms that are not easily captured by simple frequency. N-gram analysis was then conducted to extract sequential word patterns for each keyword, and named entity recognition analysis was used to derive topic-specific keywords. Subsequently, a 2-mode matrix was employed to analyze literacy trends by educational actors, constructing rows and columns to generate a literacy activity network. Sentiment analysis was used to examine emotional words within the documents, and CONCOR analysis was employed to group and categorize discourses. Finally, LDA topic modeling and Word-level Semantic Clustering techniques were applied to extract and analyze the main topics within literacy and the key topics across various fields.

Table 2. Analysis Framework

| Category | Analysis Method | Description |
|---|---|---|
| Text Mining Analysis | Word Frequency Analysis | Extracted core keywords from literacy data |
| | TF-IDF Analysis | Identified important keywords within literacy data |
| | N-Gram Analysis | Derived co-occurrence word networks within educational data |
| | Named Entity Recognition Analysis | Identified names of disciplines, organizations, regions, and keywords |
| Matrix Analysis | 2-Mode Matrix Analysis | Analyzed literacy networks by collaborative stakeholders |
| Sentiment and Discourse Analysis | Sentiment Analysis | Identified emotional perceptions of literacy |
| | CONCOR Analysis | Grouped literacy discourses |
| Topic Analysis | LDA Topic Modeling Analysis | Extracted major topics related to literacy activities |
| | Word-Level Semantic Clustering | Derived research-specific topics within literacy activities |

3) Research Questions

In this study, the following research questions were raised and are discussed based on the analysis conducted:

RQ 1. What are the key trends in literacy, and how is it expanding in the digital and media environment? Keywords were derived using word frequency, TF-IDF, and N-Gram analyses, with a focus on digital and media literacy.

RQ 2. What are the relationships and structural characteristics among key themes in the literacy network? 2-Mode matrix analysis and network property analyses were conducted to explore these interactions.

RQ 3. What are the patterns of emotional vocabulary associated with literacy, and what is the distribution of positive and negative emotions? Sentiment analysis was used to measure these patterns and their underlying factors.

RQ 4. Does the literacy-related topic network exhibit statistically significant structures? QAP correlation and regression analyses were performed to validate statistical significance.

RQ 5. What are the implications for literacy education and policy formulation? Recommendations were drawn based on network analysis, CONCOR, and clustering results.

The overarching aim of these questions is to propose a multidimensional approach and integrated policy directions to position literacy as a key competency in the digital society.

## 4. Research Results

### 4.1 Results of Literacy Keyword Analysis through Text Mining

1) Extracting Core Literacy Keywords through Word Frequency Analysis

To identify trends in literacy, a word frequency analysis was conducted to extract the top core keywords. Excluding dependent nouns without specific meanings, such as "and," "of," and "related," the frequencies and percentages of the top 50 keywords were analyzed, as shown in <Table 3>.

The analysis revealed that the most frequently mentioned keyword was "literacy," which demonstrates that it is the central theme of this study and the subject of extensive discussion. Following this, "digital" and "media" recorded high frequencies, indicating that literacy is closely linked to digital and media environments in modern society. Other keywords, such as "ability," "education," and "data," were also prominent, highlighting that literacy extends beyond basic reading and writing skills to encompass various forms of information utilization and education.

Additionally, frequent mentions of "reading comprehension" and "understanding" suggest that literacy is deeply associated with these cognitive skills. Interestingly, common terms like "coffee" and education-related words such as "platform" and "class" also appeared, indicating that literacy is discussed in diverse contexts, including everyday life and education.

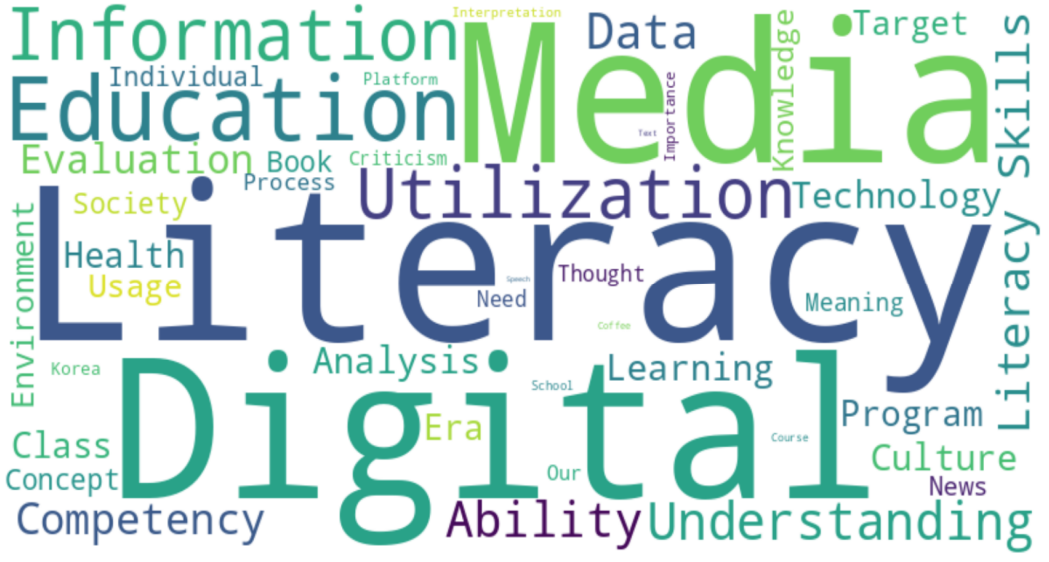
In conclusion, the word frequency analysis suggests that literacy has established itself as a concept encompassing digital media, information utilization, education, and social contexts. This underscores its growing importance in contemporary society.

Table 3. Top 50 Keywords in Literacy through Word Frequency Analysis and Word Cloud

| no. | Keyword | N | % | no. | Keyword | N | % |
|---|---|---|---|---|---|---|---|
| 1 | Literacy | 8,385 | 13.476 | 26 | Us/Our | 218 | 0.35 |
| 2 | Digital | 2,062 | 3.314 | 27 | Book | 216 | 0.347 |
| 3 | Media | 1,553 | 2.496 | 28 | Coffee | 194 | 0.312 |
| 4 | Ability | 1,413 | 2.271 | 29 | News | 191 | 0.307 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5 | Education | 1,101 | 1.769 | 30 | Knowledge | 190 | 0.305 |
| 6 | Data | 1,066 | 1.713 | 31 | Concept | 187 | 0.301 |
| 7 | Information | 828 | 1.331 | 32 | Character/Text | 179 | 0.288 |
| 8 | Reading Comprehension | 726 | 1.167 | 33 | Criticism | 172 | 0.276 |
| 9 | Understanding | 707 | 1.136 | 34 | Program | 172 | 0.276 |
| 10 | Utilization | 531 | 0.853 | 35 | Individual | 167 | 0.268 |
| 11 | Speech | 450 | 0.723 | 36 | Culture | 167 | 0.268 |
| 12 | Capability | 383 | 0.616 | 37 | English | 163 | 0.262 |
| 13 | Meaning | 343 | 0.551 | 38 | Reinforcement/Strengthening | 161 | 0.259 |
| 14 | Era | 309 | 0.497 | 39 | Media | 156 | 0.251 |
| 15 | Technology | 307 | 0.493 | 40 | Finance | 150 | 0.241 |
| 16 | Text | 301 | 0.484 | 41 | Thought | 150 | 0.241 |
| 17 | Society | 261 | 0.419 | 42 | Interpretation | 147 | 0.236 |
| 18 | Necessity | 261 | 0.419 | 43 | Diversity | 138 | 0.222 |
| 19 | Evaluation | 257 | 0.413 | 44 | Target/Subject | 130 | 0.209 |
| 20 | Health | 243 | 0.391 | 45 | School | 129 | 0.207 |
| 21 | Usage | 242 | 0.389 | 46 | Korea | 128 | 0.206 |
| 22 | Importance | 232 | 0.373 | 47 | Learning | 127 | 0.204 |
| 23 | Process | 224 | 0.36 | 48 | Platform | 126 | 0.203 |
| 24 | Analysis | 220 | 0.354 | 49 | Environment | 126 | 0.203 |
| 25 | Meaning | 219 | 0.352 | 50 | Class | 118 | 0.19 |

[Word Cloud]




2) Extracting Key Literacy Keywords Through TF-IDF Analysis

In this study, TF-IDF (Term Frequency-Inverse Document Frequency) analysis was performed to address the limitations of relying solely on word frequency analysis, such as the overemphasis of commonly used terms in all documents (No Yeon-Woo et al., 2015). Words that frequently appear in specific documents but are rare across all documents have higher TF-IDF values. As a result, keywords that were relatively significant in literacy-related documents were identified.

Through this analysis, the study effectively pinpointed terms that are prominent in specific documents but may not be evident through simple frequency analysis.

Keywords such as "coffee," "health," and "finance" indicate that literacy is discussed in various contexts of daily life. This suggests that literacy is expanding beyond basic literacy skills to include

practical competencies relevant to everyday life. In particular, financial literacy emerges as an essential component in modern society for making sound economic decisions and managing personal finances.

Additionally, keywords like "data," "digital," and "media" highlight the close relationship between literacy and the ability to utilize and interpret information in digital environments. This reflects the growing emphasis on literacy as an essential skill for information use in the digital age.

Keywords such as "education," "reading comprehension," "understanding," and "capability" represent fundamental aspects of literacy, demonstrating its critical role in educational approaches and learning processes across documents. Meanwhile, terms like "society," "evaluation," and "process" suggest that literacy is significantly addressed in social contexts where learning, evaluation, and practice are interconnected.

Keywords like "news" and "book" emphasize literacy's relevance to understanding and critically engaging with specific content media. This underscores literacy's importance as a critical thinking skill necessary in navigating the overwhelming information in contemporary society. Furthermore, "English" highlights the role of foreign language literacy as a key learning area, emphasizing its importance in global communication.

Overall, the analysis demonstrates that literacy extends beyond basic reading and writing abilities to play an essential role in various domains, particularly in digital environments and social contexts, reflecting its increasing importance in modern society.

Table 4. Top 30 Keywords Through TF-IDF Analysis

| no. | Keyword | TF-IDF | no. | Keyword | TF-IDF |
|---|---|---|---|---|---|
| 1 | Data | 1.843 | 16 | Era | 1.958 |
| 2 | Digital | 0.948 | 17 | Text | 1.95 |
| 3 | Media | 1.228 | 18 | Evaluation | 2.098 |
| 4 | Education | 1.046 | 19 | Society | 2.065 |
| 5 | Ability | 0.713 | 20 | Book | 2.482 |
| 6 | Information | 1.169 | 21 | Necessity | 2.01 |
| 7 | Reading Comprehension | 1.284 | 22 | Usage | 2.142 |
| 8 | Understanding | 1.145 | 23 | News | 2.703 |
| 9 | Utilization | 1.432 | 24 | Process | 2.295 |
| 10 | Health | 3.049 | 25 | Meaning | 2.319 |
| 11 | Coffee | 3.693 | 26 | Analysis | 2.26 |
| 12 | Speech | 1.567 | 27 | Importance | 2.137 |
| 13 | Capability | 1.757 | 28 | Finance | 3.288 |
| 14 | Technology | 2.006 | 29 | Us/Our | 2.178 |
| 15 | Meaning | 1.778 | 30 | English | 2.898 |


[Word Cloud: Top 50 Keywords]

3) Analysis of Keyword Influence Relationships Through N-Gram Analysis

N-gram analysis is primarily used to understand the relationships between consecutive words in sentence structures. It is particularly useful for identifying which nouns are frequently used together within a given text and what kind of influence relationships they exhibit. In this study, N-gram analysis was conducted on the entire social big data related to literacy to uncover the associations between keywords. The results of the N-gram analysis showed that specific types of literacy, such

as digital literacy, media literacy, and data literacy, frequently appeared together, emphasizing their interrelationships and the potential for educational applications.

First, the Bigram analysis revealed that the most frequently occurring combinations were "digital literacy," "media literacy," and "data literacy," in that order. These results indicate a high level of societal interest in the ability to understand and utilize information in digital environments, highlighting the need for literacy education related to digital, media, and data contexts.

In the Trigram analysis, repetitive combinations such as "digital literacy literacy," "media literacy literacy," and "data literacy literacy" were frequently observed. This repetition suggests that the concept of literacy is heavily emphasized and repeatedly mentioned across various contexts, reflecting its expanded meanings. Additionally, the frequent appearance of combinations like "literacy literacy education" indicates that literacy is closely linked to education and is strongly tied to efforts to enhance individual capabilities.

The Tetragram analysis produced combinations such as "literacy literacy digital literacy," "media literacy literacy education," and "digital literacy literacy education." These combinations suggest that digital, media, and data literacy play a key role not only in acquiring information but also in practical education and capacity-building processes.

Particularly noteworthy keywords derived from the N-gram analysis include "digital," "media," "data," "education," and "capability." These results show that literacy has expanded into a concept closely tied to critical thinking and information utilization in digital environments, as well as to enhancing individual competencies. For example, combinations like "capability building" and "digital technology" highlight that literacy is recognized as an important factor in improving individuals' social and technical skills.

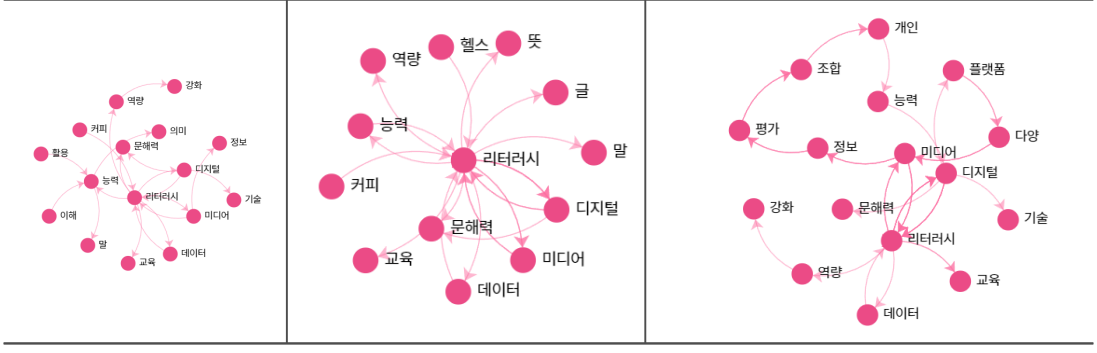
The results of this analysis emphasize that various types of literacy have become essential societal competencies in the digital age, underscoring the need to develop integrated literacy education programs to support these competencies. Given the central roles of digital and media literacy, it is important to consider their complementary relationships in the design of policy and educational strategies. Additionally, the high frequency of data literacy-related terms suggests that the ability to critically understand and utilize information in a data-driven society is indispensable.

This study carries significance in that it proposes policy and educational directions to support various types of literacy, based on an expanded concept of literacy. By doing so, it aims to provide practical approaches for enhancing both individual and societal capabilities.

Table 5. Top 20 Consecutive Words Derived from N-gram Analysis

| Bigram(2) | | | Trigram(3) | | | | Tetragram(4) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Word 1 | Word 2 | Frequency | Word 1 | Word 2 | Word 3 | Frequency | Word 1 | Word 2 | Word 3 | Word 4 | Frequency |
| Literacy | Literacy | 4250 | Digital | Literacy | Literacy | 1005 | Digital | Literacy | Literacy | Digital | 241 |
| Digital | Literacy | 1011 | Media | Literacy | Literacy | 782 | Media | Literacy | Literacy | Education | 144 |
| Media | Literacy | 791 | Data | Literacy | Literacy | 425 | Media | Literacy | Literacy | Media | 141 |
| Data | Literacy | 428 | Literacy | Literacy | Education | 340 | Data | Literacy | Literacy | Data | 131 |
| Literacy | Education | 340 | Literacy | Literacy | Digital | 308 | Digital | Literacy | Literacy | Education | 108 |
| Literacy | Digital | 309 | Literacy | Literacy | Media | 173 | Literacy | Literacy | Digital | Literacy | 88 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Capability | Speech | 182 | Literacy | Literacy | Literacy | 160 | Literacy | Digital | Literacy | Literacy | 88 |
| Understanding | Capability | 177 | Literacy | Literacy | Comprehension | 147 | Literacy | Literacy | Media | Literacy | 85 |
| Literacy | Media | 173 | Literacy | Literacy | Data | 141 | Literacy | Media | Literacy | Literacy | 84 |
| Utilization | Capability | 173 | Coffee | Literacy | Literacy | 106 | Literacy | Literacy | Literacy | Literacy | 82 |
| Literacy | Comprehension | 147 | Literacy | Literacy | Capability | 101 | Literacy | Literacy | Digital | Comprehension | 61 |
| Literacy | Data | 141 | Literacy | Literacy | Competence | 97 | Literacy | Literacy | Competence | Enhancement | 51 |
| Digital | Comprehension | 111 | Literacy | Digital | Literacy | 88 | Media | Information | Evaluation | Combination | 46 |
| Coffee | Literacy | 108 | Health | Literacy | Literacy | 85 | Information | Evaluation | Combination | Individual | 45 |
| Capability | Enhancement | 103 | Literacy | Media | Literacy | 85 | Evaluation | Combination | Individual | Capability | 45 |
| Capability | Meaning | 102 | Literacy | Literacy | Meaning | 70 | Capability | Digital | Literacy | Literacy | 44 |
| Literacy | Capability | 101 | Literacy | Literacy | Text | 69 | Digital | Platform | Various | Media | 43 |
| Media | Information | 98 | Capability | Literacy | Literacy | 62 | Platform | Various | Media | Information | 40 |
| Literacy | Capability | 97 | Literacy | Literacy | Speech | 61 | Various | Media | Information | Evaluation | 40 |
| Digital | Technology | 96 | Literacy | Digital | Comprehension | 61 | Literacy | Literacy | Digital | Technology | 39 |




## 4.2 Analysis of Relationships in Literacy Activities

To analyze the relationships between keywords related to literacy, a 1-mode matrix analysis was conducted, generating a co-occurrence matrix where both rows and columns consisted of identical words (Matrix Word analysis). Based on the cosine similarity analysis, the keywords most closely associated with literacy were "understanding," "utilization," "speech," and "information." This indicates that literacy is closely connected to the understanding and utilization of information, as well as communication skills. Additionally, "capability" and "education" showed high similarity scores, suggesting that literacy plays a significant role in enhancing individual capabilities and supporting educational practices. The similarity with "digital" and "media" was also relatively high, highlighting that literacy is an essential skill in digital and media environments in modern society.

Overall, literacy plays a central role in the understanding and utilization of information, digital and media technologies, and educational practices. These findings underscore the need for strategies to strengthen literacy-based competencies and their application in society.

Table 6. Matrix Analysis Results

| | Literacy | Digital | Media | Capability | Education | Data | Information | Comprehension | Understanding | Utilization | Speech | Competence | Meaning | Era | Technology | Text | Society | Necessity | Evaluation | Health |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Literacy | 0 | 0.0292 | 0.0245 | 0.0623 | 0.048 | 0.0141 | 0.0512 | 0.0491 | 0.0676 | 0.0584 | 0.0547 | 0.0489 | 0.0554 | 0.0461 | 0.0424 | 0.0478 | 0.0461 | 0.0523 | 0.0451 | 0.0101 |
| Digital | 0.0292 | 0 | 0.0268 | 0.0671 | 0.0534 | 0.0095 | 0.0636 | 0.0569 | 0.0594 | 0.0726 | 0.0516 | 0.0637 | 0.0486 | 0.0622 | 0.0705 | 0.0326 | 0.0491 | 0.0516 | 0.0772 | 0.0057 |

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Media | 0.0245 | 0.0268 | 0 | 0.0654 | 0.0502 | 0.003 | 0.0653 | 0.0385 | 0.0708 | 0.0441 | 0.0474 | 0.03 | 0.0519 | 0.0399 | 0.036 | 0.0313 | 0.0449 | 0.042 | 0.0749 | 0.0049 |
| Capability | 0.0623 | 0.0671 | 0.0654 | 0 | 0.0927 | 0.0299 | 0.1501 | 0.1222 | 0.2181 | 0.1688 | 0.1725 | 0.0907 | 0.1743 | 0.1214 | 0.1066 | 0.1613 | 0.1058 | 0.1271 | 0.1563 | 0.023 |
| Education | 0.048 | 0.0534 | 0.0502 | 0.0927 | 0 | 0.0155 | 0.0832 | 0.0802 | 0.0939 | 0.1029 | 0.0778 | 0.1007 | 0.0815 | 0.0662 | 0.081 | 0.0495 | 0.0834 | 0.0762 | 0.0722 | 0.0098 |
| Data | 0.0141 | 0.0095 | 0.003 | 0.0299 | 0.0155 | 0 | 0.0135 | 0.0118 | 0.0277 | 0.0421 | 0.0242 | 0.0435 | 0.0304 | 0.021 | 0.0248 | 0.02 | 0.0115 | 0.0218 | 0.0116 | 0.0003 |
| Information | 0.0512 | 0.0636 | 0.0653 | 0.1501 | 0.0832 | 0.0135 | 0 | 0.1024 | 0.1855 | 0.1389 | 0.1334 | 0.0903 | 0.1088 | 0.0987 | 0.0895 | 0.0672 | 0.0994 | 0.1296 | 0.1467 | 0.034 |
| Comprehension | 0.0491 | 0.0569 | 0.0385 | 0.1222 | 0.0802 | 0.0118 | 0.1024 | 0 | 0.136 | 0.0778 | 0.1275 | 0.063 | 0.1185 | 0.0894 | 0.0691 | 0.1497 | 0.0804 | 0.1167 | 0.081 | 0.0127 |
| Understanding | 0.0676 | 0.0594 | 0.0708 | 0.2181 | 0.0939 | 0.0277 | 0.1855 | 0.136 | 0 | 0.1846 | 0.1801 | 0.1161 | 0.1806 | 0.1147 | 0.115 | 0.1672 | 0.1201 | 0.152 | 0.1149 | 0.0401 |
| Utilization | 0.0584 | 0.0726 | 0.0441 | 0.1688 | 0.1029 | 0.0421 | 0.1389 | 0.0778 | 0.1846 | 0 | 0.1299 | 0.1485 | 0.1501 | 0.0885 | 0.1489 | 0.0818 | 0.0787 | 0.1377 | 0.0917 | 0.0205 |
| Speech | 0.0547 | 0.0516 | 0.0474 | 0.1725 | 0.0778 | 0.0242 | 0.1334 | 0.1275 | 0.1801 | 0.1299 | 0 | 0.0595 | 0.153 | 0.1119 | 0.0825 | 0.1657 | 0.0941 | 0.1342 | 0.0816 | 0.0199 |
| Competence | 0.0489 | 0.0637 | 0.03 | 0.0907 | 0.1007 | 0.0435 | 0.0903 | 0.063 | 0.1161 | 0.1485 | 0.0595 | 0 | 0.0927 | 0.121 | 0.0857 | 0.0503 | 0.1017 | 0.1039 | 0.0882 | 0.0143 |
| Meaning | 0.0554 | 0.0486 | 0.0519 | 0.1743 | 0.0815 | 0.0304 | 0.1088 | 0.1185 | 0.1806 | 0.1501 | 0.153 | 0.0927 | 0 | 0.1087 | 0.1073 | 0.1814 | 0.0824 | 0.1103 | 0.0807 | 0.0193 |
| Era | 0.0461 | 0.0622 | 0.0399 | 0.1214 | 0.0662 | 0.021 | 0.0987 | 0.0894 | 0.1147 | 0.0885 | 0.1119 | 0.121 | 0.1087 | 0 | 0.0805 | 0.0751 | 0.103 | 0.1594 | 0.0543 | 0.0068 |
| Technology | 0.0424 | 0.0705 | 0.036 | 0.1066 | 0.081 | 0.0248 | 0.0895 | 0.0691 | 0.115 | 0.1489 | 0.0825 | 0.0857 | 0.1073 | 0.0805 | 0 | 0.0498 | 0.068 | 0.0989 | 0.0868 | 0.0093 |
| Text | 0.0478 | 0.0326 | 0.0313 | 0.1613 | 0.0495 | 0.02 | 0.0672 | 0.1497 | 0.1672 | 0.0818 | 0.1657 | 0.0503 | 0.1814 | 0.0751 | 0.0498 | 0 | 0.0725 | 0.0874 | 0.039 | 0.0071 |
| Society | 0.0461 | 0.0491 | 0.0449 | 0.1058 | 0.0834 | 0.0115 | 0.0994 | 0.0804 | 0.1201 | 0.0787 | 0.0941 | 0.1017 | 0.0824 | 0.103 | 0.068 | 0.0725 | 0 | 0.087 | 0.0691 | 0.0201 |
| Necessity | 0.0523 | 0.0516 | 0.042 | 0.1271 | 0.0762 | 0.0218 | 0.1296 | 0.1167 | 0.152 | 0.1377 | 0.1342 | 0.1039 | 0.1103 | 0.1594 | 0.0989 | 0.0874 | 0.087 | 0 | 0.0735 | 0.0263 |
| Evaluation | 0.0451 | 0.0772 | 0.0749 | 0.1563 | 0.0722 | 0.0116 | 0.1467 | 0.081 | 0.1149 | 0.0917 | 0.0816 | 0.0882 | 0.0807 | 0.0543 | 0.0868 | 0.039 | 0.0691 | 0.0735 | 0 | 0.0031 |
| Health | 0.0101 | 0.0057 | 0.0049 | 0.023 | 0.0098 | 0.0003 | 0.034 | 0.0127 | 0.0401 | 0.0205 | 0.0199 | 0.0143 | 0.0193 | 0.0068 | 0.0093 | 0.0071 | 0.0201 | 0.0263 | 0.0031 | 0 |

### 4.3 Network Analysis Results

#### 1) Network Properties

The structural descriptive statistics of the entire network were analyzed, and relevant indices were measured based on the interconnectivity among network nodes, focusing on the 1-mode selected keyword, "literacy." The network analysis revealed a total of 20 nodes and 190 edges, with one connected component, indicating that all nodes are part of a single network. In terms of network centralization metrics, the values for degree centrality, closeness centrality, and betweenness centrality were all 0. This suggests that no specific node occupies a significantly central position in the network. Additionally, the network diameter was 1, and the reciprocity was 0, meaning that all nodes are equally interconnected within the network. In summary, the keywords related to literacy are uniformly connected, forming a balanced network structure where no single topic dominates or exclusively leads the discourse.

Table 7. Network Analysis Attributes and Values

| Network Measures | Value |
|---|---|
| Nodes | 20 |
| Total Edges | 190 |
| Diameter | 1 |
| Degree Centralization | 0 |
| Closeness Centralization | 0 |
| Betweenness Centralization | 0 |
| Connected Components | 1 |
| Overall Reciprocity | 0 |

#### 2) Centrality Analysis

Centrality analysis quantifies the relative importance of specific words (nodes) within a network.

By analyzing centrality, it is possible to identify more influential and meaningful words from a structural perspective. The network analysis was calculated based on the [Matrix] 1-mode selected keyword.

As a result of the centrality analysis, centrality values measuring the influence of specific words within the network were derived. Through this analysis, it was possible to identify words that occupy relatively significant positions within the literacy network. Key words such as literacy, capability, education, information, and understanding showed high centrality, confirming their pivotal roles within the network.

The word "literacy" acts as an important hub, leading the flow of information through its connections with various other words in the network. "Capability" and "education" were identified as key concepts closely related to the practical effects of literacy. Additionally, "information" and "understanding" represent the fundamental functions of literacy, with high centrality values indicating their role as essential elements that shape the core themes of literacy.

These findings highlight the need for strategic approaches centered on specific keywords when conducting literacy-related research and policy formulation. Furthermore, they can serve as valuable foundational data for designing literacy education and capability development programs.

Table 8. Centrality Analysis Results

| Keyword | Degree Centrality | Closeness Centrality | Betweenness Centrality | Eigenvector Centrality | PageRank | Clustering Coefficient |
|---|---|---|---|---|---|---|
| Literacy | 1 | 1 | 0 | 0.224 | 0.264 | 1 |
| Digital | 1 | 1 | 0 | 0.224 | 0.096 | 1 |
| Media | 1 | 1 | 0 | 0.224 | 0.075 | 1 |
| Capability | 1 | 1 | 0 | 0.224 | 0.08 | 1 |
| Education | 1 | 1 | 0 | 0.224 | 0.051 | 1 |
| Data | 1 | 1 | 0 | 0.224 | 0.046 | 1 |
| Information | 1 | 1 | 0 | 0.224 | 0.05 | 1 |
| Comprehension | 1 | 1 | 0 | 0.224 | 0.038 | 1 |
| Understanding | 1 | 1 | 0 | 0.224 | 0.045 | 1 |
| Utilization | 1 | 1 | 0 | 0.224 | 0.035 | 1 |
| Speech | 1 | 1 | 0 | 0.224 | 0.03 | 1 |
| Competence | 1 | 1 | 0 | 0.224 | 0.024 | 1 |
| Meaning | 1 | 1 | 0 | 0.224 | 0.026 | 1 |
| Era | 1 | 1 | 0 | 0.224 | 0.022 | 1 |
| Technology | 1 | 1 | 0 | 0.224 | 0.023 | 1 |
| Text | 1 | 1 | 0 | 0.224 | 0.021 | 1 |
| Society | 1 | 1 | 0 | 0.224 | 0.019 | 1 |
| Necessity | 1 | 1 | 0 | 0.224 | 0.02 | 1 |
| Evaluation | 1 | 1 | 0 | 0.224 | 0.02 | 1 |
| Health | 1 | 1 | 0 | 0.224 | 0.016 | 1 |

3) Ego Network Analysis

An ego network analysis extracts a network focused on a specific ego (central) node, analyzing its connections with other nodes (alters) linked to it. In this study, "literacy" was selected as the ego node. The results of the ego network analysis revealed an ego network density of 0.55555, indicating that the nodes in the network are connected at a relatively high density. The number of brokers was measured at 152, and the ego between centrality value of literacy was 0, indicating that it does not play an intermediary role between connected nodes as the ego node.

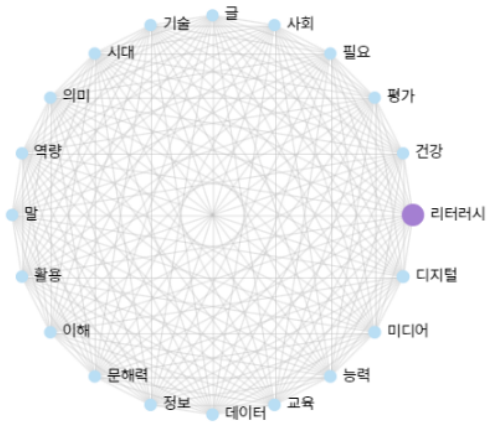
Examining the connection characteristics of the alter nodes included in the ego network, all nodes exhibited the same degree centrality (1), closeness centrality (1), and betweenness centrality (0). This reflects a simple structure where literacy is directly connected to all alter nodes. Regarding eigenvector centrality, the keywords "digital" (0.096), "media" (0.075), "capability" (0.08), "education" (0.051), and "information" (0.05) showed the highest values, highlighting the significant roles of digital and media-related keywords within the literacy network.

The PageRank values also ranked "digital" (0.096), "media" (0.075), and "capability" (0.08) as the highest, indicating that these keywords function as major concepts closely connected to literacy. The clustering coefficient for all nodes was 1, demonstrating a highly dense structure where each alter node is fully connected.

These results suggest that literacy functions as a central hub closely linked to digital, media, capability, education, and information, with strong associations among keywords due to the high density of the network. Based on this analysis, an integrated approach emphasizing the importance of digital and media literacy is recommended for literacy-related research and the development of educational programs.

Table 9. Ego Network Analysis Results

| Ego | Size | Ties | Pairs | Ego-Density | Broker | Ego Between |
|---|---|---|---|---|---|---|
| Literacy | 19 | 190 | 342 | 0.55555 | 152 | 0 |
| | | | Ego | | | |
| Alter | Degree Centrality | Closeness Centrality | Betweenness Centrality | Eigenvector Centrality | PageRank | Clustering Coefficient |
| Literacy | 1 | 1 | 0 | 0.224 | 0.264 | 1 |
| Digital | 1 | 1 | 0 | 0.224 | 0.096 | 1 |
| Media | 1 | 1 | 0 | 0.224 | 0.075 | 1 |
| Capability | 1 | 1 | 0 | 0.224 | 0.08 | 1 |
| Education | 1 | 1 | 0 | 0.224 | 0.051 | 1 |
| Data | 1 | 1 | 0 | 0.224 | 0.046 | 1 |
| Information | 1 | 1 | 0 | 0.224 | 0.05 | 1 |
| Comprehension | 1 | 1 | 0 | 0.224 | 0.038 | 1 |
| Understanding | 1 | 1 | 0 | 0.224 | 0.045 | 1 |
| Utilization | 1 | 1 | 0 | 0.224 | 0.035 | 1 |
| Speech | 1 | 1 | 0 | 0.224 | 0.03 | 1 |
| Competence | 1 | 1 | 0 | 0.224 | 0.024 | 1 |
| Meaning | 1 | 1 | 0 | 0.224 | 0.026 | 1 |
| Era | 1 | 1 | 0 | 0.224 | 0.022 | 1 |
| Technology | 1 | 1 | 0 | 0.224 | 0.023 | 1 |
| Text | 1 | 1 | 0 | 0.224 | 0.021 | 1 |
| Society | 1 | 1 | 0 | 0.224 | 0.019 | 1 |
| Necessity | 1 | 1 | 0 | 0.224 | 0.02 | 1 |
| Evaluation | 1 | 1 | 0 | 0.224 | 0.02 | 1 |
| Health | 1 | 1 | 0 | 0.224 | 0.016 | 1 |




4) Shortest Path Analysis

Shortest path analysis examines the shortest route between a starting word (node) and a target word within the network. Here, a path refers to the sequence of connecting nodes required to reach another node.

The shortest path analysis between literacy and comprehension revealed that the two keywords are directly connected, indicating a strong interrelationship. This suggests that literacy is an expanded concept of comprehension, encompassing key elements such as the ability to understand and utilize information. This relationship implies that strengthening comprehension is an essential component of literacy education.

The shortest path analysis between literacy and capability also showed a direct connection, demonstrating a close association. This highlights the role of literacy in not only enhancing information comprehension but also contributing significantly to the development of various capabilities required by individuals and society. The direct connection between literacy and capability indicates that the development of core skills—such as problem-solving, critical thinking, and communication—based on information utilization can be a major goal of literacy education. Consequently, it is necessary to develop educational programs that account for the relationship between literacy and capability.

The shortest path analysis between literacy and speech revealed a direct connection, indicating that literacy is deeply associated with verbal expression and communication skills. Literacy goes beyond the ability to read and understand information; it also plays a crucial role in verbal abilities, such as effectively conveying and expressing information through speaking.

Therefore, literacy education needs to be expanded to include the enhancement of verbal communication skills, such as speaking. This can support effective communication in various contexts, helping individuals to express themselves competently in diverse situations.

Table 10. Shortest Path Analysis Results

| Start Word | End Word | Shortest Path | Interpretation |
|---|---|---|---|

| | | | |
|---|---|---|---|
| Literacy | Comprehension | Directly Connected | Indicates a strong interrelationship, showing literacy as an expanded concept of comprehension. |
| Literacy | Capability | Directly Connected | Highlights literacy's role in developing various capabilities required by individuals and society. |
| Literacy | Speech | Directly Connected | Reflects literacy's deep association with verbal expression and communication skills. |

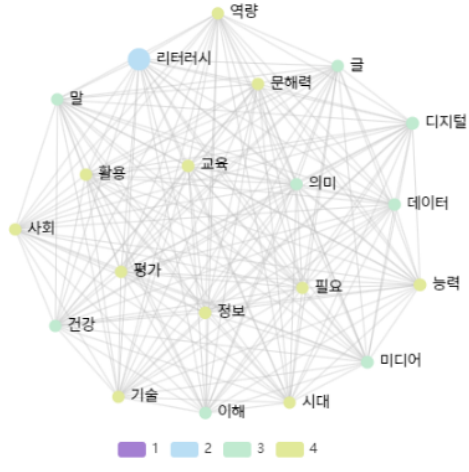
5) CONCOR Analysis

CONCOR analysis is a block modeling technique used to identify block structures within a network. By repeatedly performing correlation analyses among words, CONCOR groups blocks with high correlations. For the analysis, the number of iterations required for convergence was set to 100 (iteration = 100), and the number of clusters was set to 4.

The results of the CONCOR analysis showed that each cluster was formed based on specific themes or meanings. Literacy was positioned alone in Cluster 2, demonstrating its role as a core concept with a central function in the network. Cluster 3 included "digital," "media," "data," "understanding," "speech," "meaning," "text," and "health," reflecting a close association with information comprehension and expression in the digital environment, as well as health-related topics. Cluster 4 included "capability," "education," "information," "comprehension," "utilization," "competence," "era," "technology," "society," "necessity," and "evaluation," encompassing themes related to capability building, technological changes, and societal demands.

These results indicate that literacy is closely connected to various fields, including the digital environment, information utilization, societal needs, and capability development. This highlights the need for education and policy support based on these connections.

Table 11. CONCOR Analysis Results

| Cluster | Words | Visualization Results |
|---|---|---|
| 2 | Literacy |  |
| 3 | Digital | |
| 3 | Media | |
| 3 | Data | |
| 3 | Understanding | |
| 3 | Speech | |
| 3 | Meaning | |
| 3 | Text | |
| 3 | Health | |
| 4 | Capability | |
| 4 | Education | |
| 4 | Information | |
| 4 | Comprehension | |
| 4 | Utilization | |
| 4 | Competence | |
| 4 | Era | |
| 4 | Technology | |
| 4 | Society | |
| 4 | Necessity | |
| 4 | Evaluation | |



6) Clustering Analysis Results

Clustering analysis, conducted using a hierarchical clustering approach, identifies small-scale groups within a network. This method calculates patterns of relationships that each node forms with others,

classifying nodes with similar relational patterns into the same cluster. In this study, Louvain, Leiden, Girvan-Newman, and Clauset-Newman-Moore (CNM) clustering analyses were performed.

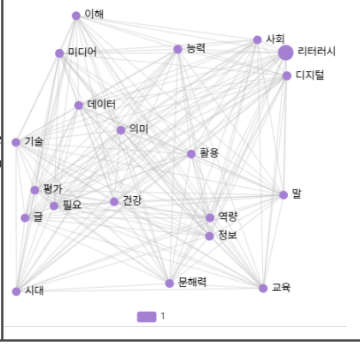
The clustering analysis results revealed that literacy was identified as a core node within the main cluster across all methods, including Louvain, Leiden, Girvan-Newman, and CNM analyses. Literacy was grouped with key concepts such as digital, utilization, evaluation, society, understanding, competence, and education.

Although there were slight differences in cluster classifications depending on the method used, common themes like digital, media, data, education, competence, and societal needs consistently emerged as key concepts closely linked to literacy. In particular, Louvain and Leiden analyses grouped literacy with digital, evaluation, and utilization, emphasizing the importance of information utilization and evaluation skills as critical components of literacy. Similarly, Girvan-Newman and Clauset-Newman-Moore analyses also included literacy in the central cluster, underscoring its essential role in the digital era.

These findings highlight that literacy is strongly connected to various societal, educational, and technological topics and plays a central role within the network. They suggest that an integrated approach encompassing digital literacy, information utilization skills, and evaluation capabilities is necessary for literacy-related policy-making and the development of educational programs.

Table 12. Clustering Analysis Results

| Louvain | | Leiden | | Girvan-Newman | | Clauset-Newman-Moore | | Visualization Results |
|---|---|---|---|---|---|---|---|---|
| Cluster | World | Cluster | World | Cluster | World | Cluster | World | |
| 1 | Literacy | 1 | Utilization | 1 | Literacy | 1 | Utilization | Louvain |
| 1 | Digital | 1 | Evaluation | 2 | Utilization | 1 | Evaluation | |
| 1 | Media | 1 | Speech | 2 | Evaluation | 1 | Society | |
| 1 | Capability | 1 | Health | 2 | Speech | 1 | Understanding | |
| 1 | Education | 1 | Media | 2 | Health | 1 | Competence | |
| 1 | Data | 1 | Capability | 2 | Media | 1 | Speech | |
| 1 | Information | 1 | Meaning | 2 | Capability | 1 | Text | Leiden |
| 1 | Comprehension | 1 | Necessity | 2 | Meaning | 1 | Necessity | |
| 1 | Understanding | 1 | Society | 2 | Necessity | 1 | Technology | |
| 1 | Utilization | 1 | Understanding | 2 | Society | 1 | Comprehension | |
| 1 | Speech | 1 | Competence | 2 | Understanding | 1 | Health | |
| 1 | Competence | 1 | Text | 2 | Competence | 1 | Education | Girvan-Newman |
| 1 | Meaning | 1 | Technology | 2 | Text | 1 | Literacy | |
| 1 | Era | 1 | Comprehension | 2 | Technology | 1 | Era | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Technology | 1 | Education | 2 | Comprehension | 1 | Media | Clauset-Newman-Moore |  |
| 1 | Text | 1 | Literacy | 2 | Education | 1 | Capability | | |
| 1 | Society | 1 | Era | 2 | Era | 1 | Digital | | |
| 1 | Necessity | 1 | Digital | 2 | Digital | 1 | Meaning | | |
| 1 | Evaluation | 1 | Information | 2 | Information | 1 | Information | | |
| 1 | Health | 1 | Data | 2 | Data | 1 | Data | | |



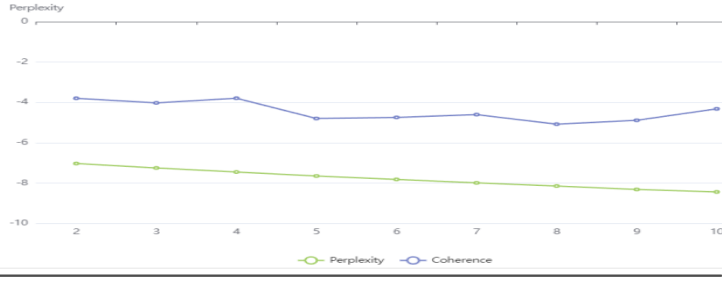
## 4.4 Results of Literacy Key Topic Analysis

### 1) Optimal Topic Performance Analysis

The optimal number of topics was determined based on Perplexity and Coherence scores. Lower Perplexity scores and higher Coherence scores indicate better model performance. As a result of analyzing the scores, the optimal number of topics was found to be 4. At this number, the Perplexity score was -7.4532, which was relatively low, and the Coherence score was -1.8981, indicating the highest cohesion. This suggests that having 4 topics is the most appropriate choice in terms of model performance.

Although Perplexity scores gradually decrease as the number of topics increases, Coherence scores peak at 4 topics and then show a significant decline. Therefore, this analysis provides insight that 4 topics are the most suitable criterion for optimizing the quality of topic modeling.

Table 13. Optimal Topic Performance Analysis

| N. of Topics | Perplexity Score | Coherence Score | Visualization Results |
|---|---|---|---|
| 2 | −7.0329 | −1.9002 |  |
| 3 | −7.2527 | −2.0135 | |
| 4 | −7.4532 | −1.8981 | |
| 5 | −7.6523 | −2.4003 | |
| 6 | −7.8215 | −2.3738 | |
| 7 | −7.9923 | −2.3034 | |
| 8 | −8.1495 | −2.5412 | |
| 9 | −8.3185 | −2.4452 | |
| 10 | −8.4416 | −2.1615 | |



### 2) LDA Topic Modeling Analysis Results

LDA (Latent Dirichlet Allocation) topic modeling is an algorithm that groups words with similar meanings to identify themes within a large corpus of documents. In this study, LDA topic modeling was used to identify emerging topics within literacy-related data and analyze trends. In topic modeling,

the size of a circle represents the frequency of a given topic, which can be interpreted as its significance.

The analysis revealed that literacy plays an important role in diverse areas such as the utilization of digital technology, data analysis, health, and finance, with clear trends emerging around these topics. Particularly, digital literacy and data literacy are emphasized as essential competencies in modern society, highlighting the growing need for information utilization and critical thinking skills based on these literacies. Additionally, there is a need to expand literacy education to emphasize specialized areas such as health literacy and financial literacy, focusing on deepened education and problem-solving skills relevant to daily life.

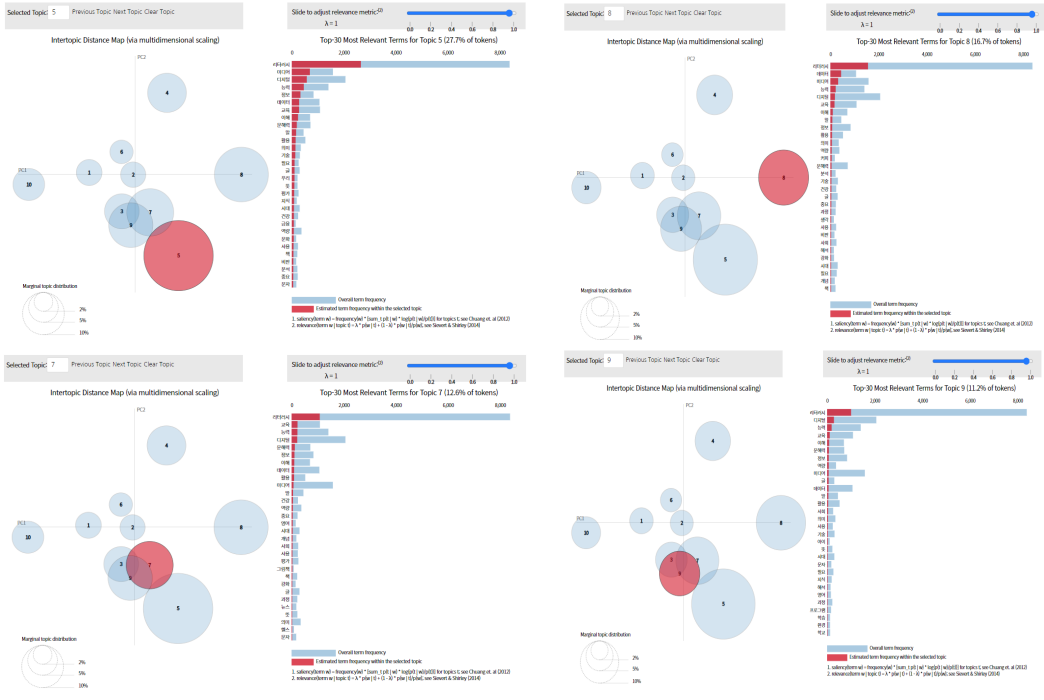
These findings suggest that literacy education should not be limited to a specific age group but should span the entire life cycle. Tailored programs are needed to enhance reading comprehension and critical thinking for children and adolescents, while focusing on practical problem-solving skills for adults and the elderly. It is especially important to develop literacy programs that align with local characteristics and foster sustainable literacy education environments through the cooperation of various stakeholders in the community.

Moreover, national-level policy support and interdisciplinary collaboration are essential to maximize the effectiveness of literacy education. Governments, educational institutions, corporations, and local communities must work together to disseminate the societal value of literacy and establish a comprehensive and systematic education system that equips citizens with the diverse competencies required in a digital society. Through these efforts, literacy can become a key competency that contributes to improving individuals' quality of life and solving societal challenges.

Table 14. LDA Topic Modeling Analysis Results

| Topic | Words Included (Topic Inclusion Probability) | Key Features |
|---|---|---|
| 1 | Literacy (0.096), Digital (0.038), Capability (0.025), Education (0.012), Comprehension (0.012), Media (0.011), Understanding (0.01), Utilization (0.01), Information (0.009), News (0.008), Technology (0.006), Data (0.006), Book (0.006), Process (0.005), Evaluation (0.004), Speech (0.004), Use (0.004), Text (0.004), Korea (0.004), Competence (0.004), Child (0.004), Finance (0.003), Operation (0.003), Era (0.003), Analysis (0.003), Writing (0.003), Progress (0.003), Knowledge (0.003), Newbie (0.003), Culture (0.003) | Literacy, digital skills, and information utilization are central themes. |
| 2 | Literacy (0.108), Digital (0.023), Capability (0.018), Media (0.015), Education (0.014), Understanding (0.013), Comprehension (0.012), Data (0.011), Information (0.008), Era (0.007), Competence (0.006), Use (0.006), Technology (0.006), Utilization (0.006), Meaning (0.005), Analysis (0.005), Society (0.005), Process (0.004), Media (0.004), Speech (0.004), Problem (0.004), Concept (0.004), Knowledge (0.004), Writing (0.004), Access (0.003), Strengthening (0.003), Our (0.003), Science (0.003), Evaluation (0.003), Customer (0.003) | Digital literacy and data analysis are emphasized as key capabilities for the digital era. |
| 3 | Literacy (0.131), Digital (0.054), Media (0.024), Capability (0.018), Information (0.016), Utilization (0.012), Comprehension (0.01), Data (0.009), Understanding (0.009), Era (0.008), Education (0.008), News (0.006), Competence (0.005), Evaluation (0.005), Speech (0.005), Book (0.005), Content (0.005), Society (0.005), Our (0.004), Smartphone (0.004), Program (0.004), Meaning (0.004), Analysis (0.004), Technology (0.004), Necessity (0.004), Use (0.003), Criticism (0.003), Concept (0.003), Platform (0.003), Individual (0.003) | Focuses on digital content and news literacy. |
| 4 | Literacy (0.114), Media (0.035), Digital (0.023), Education (0.021), Comprehension (0.017), Information (0.015), Coffee (0.012), Capability (0.009), Data (0.008), Understanding (0.008), Utilization (0.007), Competence (0.006), Program (0.006), Café (0.005), Class (0.005), Process (0.005), Meaning (0.004), Student (0.004), Era (0.004), Daegu (0.004), Environment (0.004), Necessity (0.004), Media (0.004), Thought (0.004), Importance (0.004), Speech (0.003), Society (0.003), Evaluation (0.003), Book (0.003), English (0.003) | Features local and lifestyle-focused literacy, including coffee, Daegu, and the environment. |
| 5 | Literacy (0.155), Media (0.041), Digital (0.034), Capability (0.027), Information (0.019), Data (0.017), Education (0.016), Understanding (0.014), Comprehension (0.011), Speech (0.01), Utilization (0.009), Meaning (0.008), Technology (0.008), Necessity (0.006), Writing (0.006), Our (0.006), Meaning (0.006), Evaluation (0.006), Knowledge (0.005), Era (0.005), Health (0.005), Finance (0.005), Competence (0.005), Culture (0.005), Use (0.004), Book (0.004), Criticism (0.004), Analysis (0.004), Importance (0.003), Text (0.003) | Covers specialized literacy, such as financial literacy, health, and cultural dimensions. |
| 6 | Literacy (0.077), Digital (0.023), Education (0.023), Media (0.019), Capability (0.015), Data (0.014), Comprehension (0.012), Utilization (0.01), Information (0.009), Understanding (0.008), Competence (0.007), Era (0.007), Society (0.007), Speech (0.006), Strengthening (0.006), Health (0.005), Analysis (0.004), Importance (0.004), Process (0.004), News (0.004), Necessity (0.004), Korea (0.004), Support (0.004), Program (0.004), Finance (0.004), Our (0.004), Meaning (0.004), Research (0.003), Individual (0.003) | Focuses on capability enhancement and policy support, highlighting strengthening and supporting literacy. |
| 7 | Literacy (0.138), Education (0.029), Capability (0.029), Digital (0.028), Comprehension (0.016), Information (0.014), Understanding (0.013), Data (0.013), Utilization (0.012), Media (0.01), Speech (0.007), Health (0.007), Competence (0.006), Importance (0.006), English (0.006), Era (0.005), Concept (0.005), Society (0.005), Use (0.004), Evaluation (0.004), Picture Book (0.004), Book (0.004), Strengthening (0.004), Writing (0.004), Process (0.003), News (0.003), Meaning (0.003), Health (0.003), Text (0.003) | Emphasizes literacy education for children and adolescents, featuring concepts like picture books and health. |
| 8 | Literacy (0.15), Data (0.043), Media (0.031), Capability (0.021), Digital (0.018), Education (0.017), Understanding | Centers on data utilization and |

| | | |
|---|---|---|
| | (0.01), Speech (0.008), Information (0.007), Utilization (0.007), Meaning (0.006), Competence (0.005), Coffee (0.005), Comprehension (0.005), Analysis (0.005), Technology (0.004), Health (0.004), Writing (0.004), Importance (0.004), Process (0.004), Thought (0.003), Use (0.003), Criticism (0.003), Society (0.003), Interpretation (0.003), Strengthening (0.003), Era (0.003), Necessity (0.003) | real-life literacy, such as health and practical applications. |
| 9 | Literacy (0.144), Digital (0.041), Capability (0.027), Education (0.016), Understanding (0.012), Comprehension (0.011), Information (0.01), Competence (0.01), Media (0.009), Writing (0.009), Data (0.009), Speech (0.008), Utilization (0.007), Society (0.006), Meaning (0.006), Use (0.005), Technology (0.004), Child (0.004), Meaning (0.004), Era (0.004), Text (0.003), Necessity (0.003), Knowledge (0.003), Interpretation (0.003), English (0.003), Process (0.003), Program (0.003), Learning (0.003) | Highlights literacy in educational contexts, focusing on schools and learning. |
| 10 | Literacy (0.077), Digital (0.071), Comprehension (0.018), Capability (0.016), Education (0.012), Information (0.011), Media (0.009), Evaluation (0.009), Era (0.008), Data (0.008), Competence (0.007), Understanding (0.007), Speech (0.006), Necessity (0.006), Technology (0.006), Process (0.005), Writing (0.005), Society (0.005), Utilization (0.005), Meaning (0.004), Our (0.004), Use (0.004), Finance (0.004), Platform (0.004), Knowledge (0.003), Program (0.003), News (0.003), Problem (0.003), Lecture (0.003) | Focuses on platform utilization and problem-solving skills, featuring topics like financial literacy. |



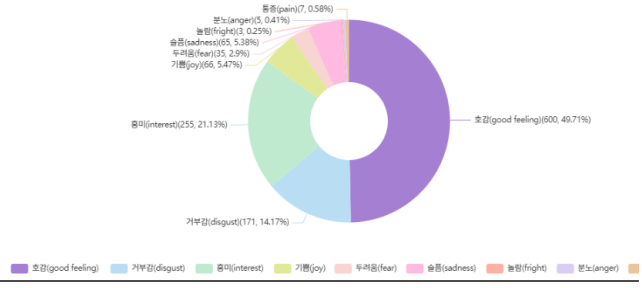

## 4.5 Sentiment Analysis Results

Based on the sentiment dictionary provided by TEXTOM, the frequency of sentiment-related words was analyzed. TEXTOM's sentiment dictionary is designed for Korean text analysis and consists of seven emotions (joy, interest, affection, pain, sadness, anger, fear, surprise, and aversion). It is primarily based on Robert Plutchik's emotional framework but also considers the characteristics of the Korean language and domestic research trends.

The analysis results indicate a high level of positive social interest in topics related to literacy. Particularly, the high proportions of affection and interest suggest that literacy-related activities are perceived positively. However, the presence of negative emotions such as aversion, sadness, and fear highlights areas requiring improvement in literacy policies and education. These negative emotions may stem from issues such as inequality in literacy education, lack of access to information, and social disparities.

Therefore, policy efforts are required to further promote positive perceptions while reducing negative emotions. Specifically, it is necessary to enhance accessibility to alleviate aversion and develop psychological support programs to address fear and sadness. Through these efforts, literacy policies and programs can gain greater social trust and support.

Table 15. Sentiment Analysis Results

| Sentiment Category | Emotion Classification | Frequency (Count) | Proportion (%) | Visualization Results |
|---|---|---|---|---|
| Total | | 1207 | 100 |  |
| good feeling | Positive | 600 | 49.71 | |
| disgust | Negative | 171 | 14.17 | |
| interest | Positive | 255 | 21.13 | |
| joy | Positive | 66 | 5.47 | |
| fear | Negative | 35 | 2.9 | |
| sadness | Negative | 65 | 5.39 | |
| fright | Negative | 3 | 0.25 | |
| 분노(anger) | Negative | 5 | 0.41 | |
| 통증(pain) | Negative | 7 | 0.58 | |



## 4.6 QAP Correlation and Regression Analysis Results

The Quadratic Assignment Procedure (QAP) correlation analysis is a method used to analyze the similarity between two matrices within network data, offering a more precise understanding of the relationships between literacy-related topics. Unlike traditional correlation coefficient analysis, QAP correlation analysis tests the statistical significance of relationships based on the structural connections between nodes. Similarly, QAP regression analysis differs from standard regression analysis by using randomization to assess significance. After conducting an ordinary least squares (OLS) regression, the matrices are randomized multiple times, and repeated regression analyses are performed to evaluate significance.

The QAP correlation analysis between the literacy-related network and the randomized network produced a correlation coefficient of 0.12 and a p-value of 0.015, indicating a statistically significant correlation. The positive correlation coefficient, combined with its statistical significance, implies that literacy-related topics form structurally meaningful relationships compared to the randomized network. This suggests that key literacy-related topics—such as digital literacy, education, and social communication—are closely interconnected within the actual network. These findings highlight the importance of considering interactions between specific topics when designing literacy policies and programs.

For QAP regression analysis, the literacy network was set as the dependent variable, and the randomized network was used as the independent variable. The analysis yielded an R-squared value of 0.0146 and a p-value of 0.016. While the low R-squared value suggests that the randomized network does not fully explain the structure of the literacy network, the statistically significant p-value indicates a meaningful relationship between the two. The low R-squared value, coupled with a significant p-value, underscores the independent and unique structural characteristics of the literacy network. This result suggests the existence of specific patterns among literacy-related topics,

rather than random relationships, warranting further in-depth research into these structural features.

The results of both QAP correlation and regression analyses confirm that the relationships between literacy-related topics exhibit clearer patterns compared to a randomized network and that these relationships are statistically significant. This finding highlights the necessity of an integrated approach in literacy policy and program design, taking into account the close connections between topics such as digital literacy, education, and data. Additionally, the literacy network's pivotal role across diverse fields emphasizes the need for multifaceted literacy education that strengthens practical problem-solving skills alongside policy support. Future research should focus on the interactions between specific topics and develop specialized programs tailored to local communities or institutions.

By incorporating these structural insights into literacy education and policy-making, it becomes possible to create more effective and targeted strategies that address societal needs while enhancing individual and collective capacities in a digital era.

Table 16. QAP Correlation and Regression Analysis Results

| QAP Correlation Results | | QAP Regression Results | | QAP Regression With Permutation Results | |
|---|---|---|---|---|---|
| Metric | Value | Metric | Value | Metric | Value |
| Correlation | 0.12088912263583615 | Coefficient | 0.12088912263583 | Original R-squared | 0.0146141799716622 |
| p-value | 0.0155592149262673 | Intercept | 0.063955755883681 | p-value | 0.016 |
| | | R-squared | 0.0146141799716622 | | |

## 5. Discussion

This study confirmed that literacy plays a significant role in modern society, extending beyond basic comprehension to encompass digital skills, information utilization, social communication, and personal capacity building. Particularly, the results from network analysis and QAP regression analysis demonstrated that literacy is closely interconnected with various fields, necessitating policy and educational approaches based on this interconnectedness. Literacy education, therefore, should be designed to strengthen information interpretation and utilization skills in digital environments, along with the development of tailored programs across the life cycle. Furthermore, efforts to spread positive perceptions and alleviate negative emotions through psychological and social support measures are required. Such an integrated approach will enable literacy to enhance individual and societal capacities and contribute to sustainable development. This section discusses the findings in relation to the research questions posed earlier.

First, What are the main trends in literacy, and what is its core role in digital environments?

Text mining and N-gram analyses revealed that the main trends in literacy are associated with topics such as digital technology, media, and data utilization. High-ranking terms like "digital," "media," and "data" in the word frequency analysis indicate that literacy in modern society is an essential competence for navigating digital environments and utilizing information. The combinations derived from the N-gram analysis, such as "digital literacy," "media literacy," and "data literacy," demonstrate how literacy has expanded beyond traditional comprehension to emphasize critical thinking, information interpretation, and utilization skills. These findings highlight the need for literacy education programs aligned with the digital era. Specifically, as the importance of digital literacy

grows, an integrated educational model focused on fostering information utilization skills and critical thinking is required. Additionally, addressing information overload in digital environments necessitates emphasizing critical reception skills and the ability to identify reliable information sources.

Second, What are the interconnections and structural characteristics of literacy across various fields?

Matrix and network attribute analyses identified the relationships and structural characteristics of key literacy-related keywords. The network analysis revealed that keywords with the highest similarity to literacy were "understanding," "utilization," "information," and "capability," indicating that literacy plays a crucial role beyond comprehension in fostering information understanding, utilization, and personal capacity building. Centrality analysis showed that all keywords exhibited equal degree centrality, confirming a balanced network structure without reliance on specific keywords. QAP correlation and regression analyses further supported the existence of meaningful structural patterns in the literacy network compared to random networks. The significant correlation coefficient of 0.12 and p-value of 0.015 indicate strong interactions among key literacy-related topics. These results suggest that literacy policies should adopt approaches that account for the interconnectedness of topics such as information understanding, digital utilization, and social communication.

Third, How is literacy perceived socially, and what are the implications of positive and negative emotional elements?

Sentiment analysis revealed that the most prominent emotions in literacy-related documents were "affection" (49.71%) and "interest" (21.13%), indicating generally positive social perceptions of literacy. This suggests that literacy activities are perceived as engaging and beneficial. However, negative emotions such as "aversion" (14.17%), "sadness" (5.39%), and "fear" (2.9%) were also present. These negative emotions may stem from issues such as inequality in literacy education, limited access to information, and digital divides. Therefore, policy and educational support are required to amplify positive perceptions while mitigating negative emotions. Specifically, programs to enhance information accessibility, reduce digital divides, and provide psychological support are necessary. These efforts will help literacy-related policies gain social trust and support while fostering a more inclusive literacy environment.

Fourth, Do literacy-related topic networks exhibit statistically significant structures?

The results from QAP correlation and regression analyses confirmed that literacy-related topic networks form statistically significant structures. The correlation coefficient and p-value within the significance threshold indicate strong interactions among key topics such as digital literacy, education, information utilization, and social communication. Compared to random networks, the literacy network demonstrates unique patterns and structural characteristics, underscoring literacy's multifaceted role in modern society. These findings emphasize the importance of literacy in strengthening problem-solving, critical thinking, and communication skills beyond mere information delivery.

Fifth, What are the implications for literacy education and policy development?**

Based on the results of network analysis, CONCOR analysis, and clustering analysis, literacy policies and educational programs require an integrated approach that considers connections with various fields, with a focus on digital and media literacy. Literacy extends beyond basic comprehension to encompass practical competencies in digital information utilization, social communication, and

fields like finance and health. This necessitates the development of tailored programs reflecting these realities. Digital literacy, media literacy, and data literacy emerged as central elements in the network, underscoring the importance of literacy education suited for digital environments. Furthermore, collaborations with local communities to develop region-specific literacy programs and policy efforts to enhance information accessibility are crucial for addressing social inequalities. Through such measures, literacy education can serve as a vital tool for improving quality of life and strengthening problem-solving capacities at both individual and societal levels.

## 6. Conclusion and Recommendations

This study analyzed trends in literacy-related topics and clarified the relationships among key topics through network structures. The analysis results confirmed that literacy plays an essential role in various areas such as information utilization in digital environments, education, and social communication. Additionally, it was found that closely connected interactions are formed around literacy within the network. QAP correlation and regression analyses verified that the literacy network has statistically significant structural patterns rather than random structures, showing strong connections among key topics. The results of CONCOR and clustering analyses revealed that digital literacy, media literacy, and data literacy occupy central positions in the network. These findings suggest that literacy policies and educational programs should be designed based on these characteristics.

First, when establishing literacy policies, digital, media, and data literacy should be treated as core elements, and an integrated educational approach that considers connections with various topics is necessary. In particular, programs that strengthen information utilization skills and critical thinking suitable for the digital era need to be developed.

Second, literacy education should not be limited to specific age groups but should be provided across the entire life cycle. For children and adolescents, literacy education should focus on enhancing basic literacy and digital utilization skills, while for adults and seniors, tailored education programs that strengthen real-life problem-solving skills should be implemented.

Third, it is necessary to develop and operate literacy programs that reflect regional characteristics through collaboration with local communities. By fostering cooperation among various stakeholders in the region, sustainable literacy education environments can be created, reducing information disparities and enhancing social inclusivity.

Fourth, national-level policy support and interdisciplinary collaboration are essential to maximize the effectiveness of literacy education. Governments, educational institutions, businesses, and civil society should work together to systematically establish literacy education frameworks and promote policies that improve information accessibility.

This study reaffirmed the multifaceted roles and social value of literacy and emphasized that literacy serves as a crucial means of improving individuals' quality of life and contributing to solving social issues. Based on these findings, policy and educational recommendations were provided to strengthen literacy’s role in enhancing personal and societal capacities. Future research should focus on in-depth analyses of individual topics related to literacy and case studies on literacy applications in various environments. Such efforts are expected to enable the development of more effective

literacy policies and programs.

## References

Han, S., & Kim, T. (2021). News big data analysis of 'media literacy' using topic modeling analysis. Journal of the Korea Contents Association, 21(4), 26-37. https://doi.org/10.5392/JKCA.2021.21.04.026

Jung, D. H., & Kim, S. (2021). A study for activation of non-face-to-face education based on CONCOR big data analytics: Data literacy perspectives. Journal of Internet Electronic Commerce Research, 21(6), 119-132. https://doi.org/10.37272/JIECR.2021.12.21.6.119

Kil, H., & Lee, J. (2022). A study on the use of literacy concept in big data articles. Journal of Reading Research, 65, 151-199.

Kim, H. Y. (2022). Imbalance of intergenerational representation in big data and politics (participation): Focused on the information gap and digital literacy. Journal of Legal Studies, (76), 1-27. https://doi.org/10.17248/knulaw..76.202201.1

Kim, S., & Kim, T. (2022, January). A study on the effect of big data analysis lecture on data literacy of teacher training college students using text programming language. Proceedings of the 2022 Winter Conference of the Korean Association of Information Education, Jeju National University Teacher Education Center/Online Webinar.

Kim, Y. H., & Kim, W. (2021). Social network analysis on digital literacy using Youtube big data. Journal of Learner-Centered Curriculum and Instruction, 21(18), 351-361.

Ko, Y. (2023). A study on improvement of big data literacy: Focusing on 'Big data analysis' courses at H University. Humanities and Social Sciences Research, 66(2), 25-36. https://doi.org/10.17939/hushss.2023.66.2.002

Moon, M.-H. (2021). Exploring data literacy education of early childhood teachers using big data. Journal of Learner-Centered Curriculum and Instruction, 21(22), 357-372.

Noh, Y., Kim, D., Han, J., Yook, M., Lim, J., Bok, K., & Yoo, J. (2015). Hot topic prediction scheme considering user influences in social networks. Journal of the Korea Contents Association, 15(8), 24-36. https://doi.org/10.5392/JKCA.2015.15.08.024

[ About the author ]

**Younghee Noh** has an MA and PhD In Library and Information Science from Yonsei University, Seoul. She has published more than 50 books, including 3 books awarded as Outstanding Academic Books by Ministry of Culture, Sports and Tourism (Government) and more than 120 papers, including

one selected as a Featured Article by the Informed Librarian Online in February 2012. She was listed in the Marquis Who's Who in the World in 2012-2016 and Who's Who in Science and Engineering in 2016-2017. She received research excellence awards from both Konkuk University (2009) and Konkuk University Alumni (2013) as well as recognition by "the award for Teaching Excellence" from Konkuk University in 2014. She received research excellence awards form 'Korean Y. Noh and Y. Shin International Journal of Knowledge Content Development & Technology Vol.9, No.3, 75-101 (September 2019) 101 Library and Information Science Society' in 2014. One of the books she published in 2014, was selected as 'Outstanding Academic Books' by Ministry of Culture, Sports and Tourism in 2015. She received the Awards for Professional Excellence as Asia Library Leaders from Satija Research Foundation in Library and Information Science (India) in 2014. She has been a Chief Editor of World Research Journal of Library and Information Science in Mar 2013 ~ Feb 2016. Since 2004, she has been a Professor in the Department of Library and Information Science at Konkuk University, where she teaches courses in Metadata, Digital Libraries, Processing of InterSnet Information Resources, and Digital Contents.

**Jeong-won Roh** is currently pursuing a Ph.D. in Library and Information Science at Konkuk University while balancing research and practical experience in the IT field. Her primary research interests include information management systems and data analysis. She focuses on interdisciplinary studies combining IT and Library and Information Science.

**Ho-won Kim** is a Ph.D. candidate in Library and Information Science at Konkuk University, conducting research on information system development and data processing using IT technologies. His interests lie in database management and information classification, and he is actively engaged in applying cutting-edge IT technologies to research in Library and Information Science.