A Comparative Analysis of Research Trends in the Information and Communication Technology Field of South and North Korea Using Data Mining

Jiwan Kim
News Technology Team, SBS A&T, Seoul, Korea
E-mail: kimsabino@sbs.co.kr

Hyunkyoo Choi
Korea Institute of Science and Technology Information, Seoul, Korea
E-mail: hkchoi@kisti.re.kr

Jeonghoon Mo*
Department of Industrial Engineering, Yonsei University, Seoul, Korea
E-mail: J.mo@yonsei.ac.kr

ABSTRACT

The purpose of this study is to compare research trends in the information and communication technology (ICT) field between North and South Korea and analyze the differences by using data mining. Frequency analysis, clustering, and network analysis were performed using keywords from seven South Korean and two North Korean ICT academic journals published for five years (2015-2019). In the case of South Korea (S. Korea), the frequency of research on image processing and wireless communication was high at 16.7% and 16.3%, respectively. North Korea (N. Korea) had a high frequency of research, in the order of 18.2% for image processing, 16.9% for computer/Internet applications/security, and 16.4% for industrial technology. N. Korea's natural language processing (NLP) sector was 11.9%, far higher than S. Korea's 0.7 percent. Student education is a unique subject that is not clustered in S. Korea. In order to promote exchanges between the two Koreas in the ICT field, the following specific policies are proposed. Joint research will be easily possible in the image processing sector, with the highest research rate in both Koreas. Technical cooperation of medical images is required. If S. Korea's high-quality image source is provided free of charge to N. Korea, research materials can be enriched. In the field of NLP, it calls for proposing exchanges such as holding a Korean language information conference, developing a Korean computer operating system. The field of student education encourages support for remote education contents and management know-how, as well as joint research on student remote evaluation.

Keywords: data mining, South Korea, North Korea, information and communication technology, research trends
1. INTRODUCTION

A new concept of the Fourth Industrial Revolution was presented at the annual meeting of the World Economic Forum in 2016 to solve global economic problems. The Fourth Industrial Revolution is a next-generation industrial revolution in which advanced information and communication technologies (ICT) are integrated into the economy and society as a whole, resulting in innovative changes. The core of the Fourth Industrial Revolution is defined as ICT (Telecommunications Technology Association, 2016). ICT increases the productivity of the industry and accelerates changes in the industrial structure. Many countries around the world regard ICT as an important factor in national economic development. Since the 2000s, the Republic of Korea (S. Korea) has designated ICT as a growth engine for national economic development, and has emerged as one of the top 10 global trade powers by expanding ICT infrastructure and investing heavily. The Democratic People’s Republic of Korea (N. Korea) is emphasizing the importance of science and technology in building a socialist economic power, with the goal of achieving economic development beyond international sanctions. Referring to the case of S. Korea, N. Korea is pursuing a quantum jump strategy to promote economic development by focusing on ICT. Due to the common point that the two Koreas focus on the ICT field for economic development, ICT can be evaluated as a field with great potential for academic exchange, technology exchange, and economic cooperation. ICT academic exchanges between the two Koreas were relatively active compared to other fields. In 2002, scholars from S. Korea, N. Korea, and China met to jointly publish the International standard information technology dictionary (Kim et al., 2015). The joint publication of the dictionary was a good milestone for inter-Korean exchanges. But the exchanges are not sustainable due to political problems. Inter-Korean summits were held three times in 2018, and inter-Korean exchanges were expected to expand further. However, in 2020 N. Korea destroyed the inter-Korean joint liaison office located in the Kaesong Industrial Complex, and peaceful exchanges remain further away. However, in order to prepare for academic exchanges and technological cooperation with N. Korea that can resume at any time, an objective understanding of N. Korea’s ICT trends and academic research topics is continuously required. The purpose of this study is to compare and analyze the trends of ICT research between the two Koreas using academic journals, and to provide basic data and cooperation methods to identify areas that can be continuously developed in joint research and exchanges. Comparative analysis of academic research trends between countries informs us of commonalities, differences, and theoretical maturity. It allows us to analyze whether the same topic presents different issues for different conditions, politically and culturally (Choi & Kim, 2011).

2. BACKGROUND AND LITERATURE REVIEW

Research on ICT in N. Korea which is conducted in S. Korea generally analyzes policy and technology trends, by using articles on propaganda from newspapers (Rodong Sinmun), new year’s addresses and lessons from N. Korean leaders, policies emphasized at the Congress of the Workers’ Party of N. Korea, and academic journals as basic data. Lee (2002) analyzed the trend of N. Korea’s research by industry using data from the academic journal Technology Innovation, published in the 1990s. Since Kim Il-sung emphasized “field support research,” N. Korea has focused on research on solving technical problems at production sites. Scientific and technological journals in N. Korea have also become closely related to solving problems at production sites. The number of “technology innovation” research paper posts decreased in 1993. It was evaluated that this was due to the collapse of the socialist countries and the stagnation of foreign trade, which reduced research and development activities at the production site. Despite the decrease in the number of research paper posts, the tendency to post research papers in light industry, agriculture, and fisheries rather increases. This is interpreted as the result of declaring “light industry firstism” and “agricultural firstism” through severe food shortages and consumer disasters during the arduous march. The amount of research papers submitted by companies decreased to less than half, and instead the amount of research papers submitted by research institutes and universities increased significantly. This is evaluated as a result of decreased technological innovation activities of factories and increased field support from universities and research institutes. It was revealed that the increase in the number of research paper posts by universities and research institutes did not lead to an improvement in the quality level of research, and only produced theoretical papers due to a lack of research conditions. Since 1999, when the arduous march ended, it has been confirmed that technological innovation activities have been gradually increasing, and efforts have been made to disseminate academic journals and scientific information.
Choi et al. (2021) extracted science and technology terms from newspaper (Rodong Sinmun) articles in the last 10 years (2011-2020) and identified major science and technology interests in N. Korea through term frequency and network analysis. When the subject of the article was divided into 16 fields, the number of articles in the construction field was the largest, and the terms in the mechanical and informatics fields were relatively numerous. N. Korea admitted that it did not meet the goal of the Five-Year Plan for National Economic Development at the 8th Congress of the Workers’ Party of N. Korea, held in January 2021. Although it was evaluated that science and technology did not contribute to economic development, the party clarified its attitude to continue its path of emphasis on science and technology. The implementation tasks of the Five-Year Plan strategy are support for scientific research, such as strengthening of science and technology education in elementary and secondary schools, fostering base universities, establishing an academic unification system, expanding graduate programs, and establishing a science and technology distribution network. A plan was also established to prevent waste of research funds as much as possible by criticizing scientists’ passivity and establishing a strict deliberation system for R&D. The frequency of keywords related to ICT in Kim Jong-un’s reign (2012-2020) was high in the order of “program,” “automation,” “new century industrial revolution,” and “integrated production system.” N. Korea is paying attention to the digital economy, and there is a latest trend that emphasizes Artificial Intelligence (AI) and big data, which are the basis of the Fourth Industrial Revolution.

Lim (2019) studied North Korea’s response to the Fourth Industrial Revolution and its achievements. The study was conducted through the perception of scholars, media, and a research journal Economic Research for seven years (2012-2018). In Lim’s paper, N. Korea is attempting a single leap from the third industry to the fourth industry stage based on ICT, and N. Korea is promoting the Fourth Industrial Revolution in a unique way with the precision of computer numerical control (CNC) machines. In addition, N. Korea is promoting differentiation from the previous regime by developing technologies such as AI software, virtual reality (VR), and industrial information. Research institutes under the National Academy of Sciences were newly reorganized to focus on high-tech research and profit-making projects, and a strategy for commercialization. By implementing remote education and improving the education system, scientific and technological talents were fostered. Based on these efforts, some factories have been unmanned and automated. In addition, the number of research paper posts in SCI journals in the field of physics and ICT is gradually increasing. However, it is believed that there will be many restrictions on economic development due to international sanctions caused by nuclear and intercontinental ballistic missile (ICBM) factors, restrictions on technology exchanges in closed societies, and difficulties in importing advanced parts and equipment.

Baxter et al. (2021) conducted a network analysis using a N. Korean scientific journal article to study the structure of the nuclear research community from the beginning of the nuclear program to 2018. The network has shown an increase in the internal density of research cooperation, which shows that N. Korea has improved knowledge transfer over time and solidified its nuclear research community. The study confirmed that other areas of nuclear research increased interconnectivity over time.

Research data related to N. Korea are difficult to access due to the closeness of the N. Korean regime. In particular, in the face of military confrontation between the two Koreas, N. Korea is extremely reluctant to disclose information in the ICT field. It is not easy to obtain specific and objective statistics and up-to-date information. To overcome these limitations, research using N. Korea’s open academic journals is a very important means of identifying the field, objectively and transparently. Research using N. Korea’s academic journals is being conducted in various fields, and it overcomes the limitations of subjective interpretation and qualitative review by researchers.

3. METHODOLOGY

Scholars who study science and technology in quantitative ways have developed tools for bibliographic methods using text (Leydesdorff, 1987, 1989). Quantitative research has mainly focused on various relations such as co-authorships and co-occurring keywords (Héllsten et al., 2020). In this study, co-occurring keywords are used among quantitative research methods. Keywords were extracted from the title and body of ICT academic journals, and research trends between the two Koreas were compared and analyzed using keyword network analysis. Keyword network analysis is a network-based content analysis method that extracts meaningful words from large text sets, and selects keywords to grasp the context based on semantic connection relationships (Lee, 2018). The ICT fields covered in this study follow the method defined and classified by the OECD. The OECD defines the ICT fields.
as “A combination of manufacturing and services industries that capture, transmit and display data and information electronically” (OECD, 2002). This definition is based on International Standard Industrial Classification (ISIC) Rev. 3, and has the following principles:

1. (For manufacturing industries) Must be intended to fulfil the function of information processing and communication including transmission and display;
2. (For manufacturing industries) Must use electronic processing to detect, measure, and/or record physical phenomena or control a physical process;
3. (For services industries) Must be intended to enable the function of information processing and communication by electronic means. (OECD, 2002)

The analysis targets and scope of this study were determined based on the above criteria, and are as shown in Table 1. In the case of S. Korea, among Korea Citation Index (KCI) registered academic journals published in Korean, all ICT academic journals with sections such as “engineering” and divisions such as “electronic/information and communication engineering” were selected as subjects of the study. In the case of N. Korea, Information Science, published by Science and Encyclopedia Publishing House (SEPH) and Information Science and Technology, published by Central Information Agency for Science and Technology (CIAST) were selected as subjects of the study. As of 2022, there are only two ICT academic journals published in N. Korea. N. Korea does not have any ICT academic journals published in English. Most of the English academic papers were published via co-authors while staying abroad. In the field of computer science, 10 papers were published in 2015 and four papers in 2016 (Choi & Noh, 2017). Compared to S. Korea, academic papers in English are extremely insufficient in quantity. Therefore, the comparison of English academic papers is meaningless.

3.1. Data Collection and Refinement

To build N. Korean paper data, the NK Tech webpage (www.nktech.net), operated by the Korea Institute of Science and Technology Information (KISTI), was crawled with the Python program to build a list of papers to be analyzed. Next, by visiting the Information Center on North Korea under the Ministry of Unification, located in the National Library of Korea, and checking the papers in each journal, keyword data was constructed by extracting

| Table 1. Periodic academic journals of ICT between South and North Korea (2015-2019) |
|-------------------------------------------------|----------|-------------------------------------------------|
| Academic journals                              | Papers   | Journal publisher                               |
| S. Korea                                        |          |                                                  |
| Journal of Korea Multimedia Society            | 757      | Korea Multimedia Society (KMMS)                 |
| Journal of Broadcast Engineering                | 396      | The Korean Institute of Broadcast and Media Engineers (KIBME) |
| Journal of the Acoustical Society of Korea     | 312      | The Acoustical Society of Korea (ASK)           |
| Journal of The Korea Institute of Electronic Communication Sciences | 798 | The Korea Institute of Electronic Communication Sciences (KIECS) |
| Journal of Korea Institute of Information, Electronics, and Communication Technology | 395 | The Korea Institute of Information, Electronics, and Communication Technology (KIECT) |
| Journal of the Korea Institute of Information and Communication Engineering | 1,360 | The Korea Institute of Information and Communication Engineering (KIICE) |
| Journal of Korean Institute of Communications and Information Sciences | 1,337 | Korean Institute of Communications and Information Sciences (KICS) |
| N. Korea                                        |          |                                                  |
| Information Science                            | 1,158    | Science and Encyclopedia Publishing House (SEPH) |
| Information Science and Technology              | 864      | Central Information Agency for Science and Technology (CIAST) |

ICT, information and communication technology.
technical nouns from the title and body of the paper. Since the keyword network analysis method aims to extract keywords and analyze relationships between keywords, objective criteria setting and consistency for keyword extraction are very important (Yang, 2019). In the case of N. Korean published papers, there are often no abstracts and selected keywords due to the short content of one to two pages per piece. Therefore, from the introduction, body, and conclusion of the paper, subject-related terminologies that appear more than three times were extracted. In the case of S. Korea, the journals to be analyzed were searched on the webpage (www.kci.go.kr) operated by KCI, and the paper list was downloaded in the form of an Excel file. Terminology was extracted from the title, abstract, and keywords of the papers. The refinement of data is divided into correcting, controlling, and removing (Lee, 2014). First, in the correction stage, all spaces in the compound nouns are attached, and special symbols were unified. Second, different spellings and similar words were unified into one. Among several words with the same meaning, they were unified into one word that appeared statistically more often. Third, among the words with a high frequency of appearance, words representing general concepts such as “research,” “method,” and “design” were removed. Fourth, the terminologies used in S. Korea and N. Korea are different in terms of notation and form, so a S. Korean-N. Korean-English comparison table was created, and a data list was constructed by matching words with English words.

3.2. Data Analysis

The refined keywords were coded in Excel, and the number of papers and the frequency of keywords by year were analyzed using the Knowledge Matrix Plus ver. 0.80 program (Korea Institute of Science and Technology Information, 2016). It is difficult to analyze the network because the data are too vast. Therefore, the formula 1 was used which was devised by Donohue (1973) and modified by Sun (1992). This formula removes insignificant keywords and selects high-frequency keywords that are meaningful for analysis.

\[
    n = \frac{-1 + \sqrt{1 + 4D}}{2}
\]

\(D\) is the total number of keywords, and the value \(n\) is the number of high frequency keywords. By checking the number of appearances of the top \(n\)-th keyword, a keyword with more than that number was selected as the final keyword to be analyzed. For keyword network analysis, the two-mode network database of “papers and keywords” was converted into a one-mode network database of “keywords and keywords,” and this data was analyzed by Gephi 0.9.2 (Bastian et al., 2009) and visualized by R program.

Degree centrality is an indicator of how much one node is connected to another. The combination between keywords becomes a research topic. Keywords with high degree centrality are often used simultaneously with other keywords. This means that there are many research topics directly related to those keywords (Kho, 2014). This paper used Eigenvector centrality, which modified degree centrality. Eigenvector centrality is a concept that extends degree centrality by considering not only the number of connected nodes but also how important the connected nodes are (Bonacich, 2007). The modularity algorithm of the Gephi 0.9.2 program was used (Blondel et al., 2008). Using this algorithm, keywords with high relevance were clustered to classify research topics and to understand the meaning of keywords. Modularity refers to the degree to which a particular node is close to one community and far from another, and is used to distinguish communities from each other in a network. Modularity resolution is derived by adjusting the value from 0 to 1. The number of communities increases as the value converges to 0, and the number of communities decreases as the value increases to 1. The researcher adjusts the resolution value to determine whether it has become an appropriate community. In general, a value of 0.3 or more is considered good clustering (Oh, 2017). The classification of clustering topics was based on S. Korea’s ICT R&D technology classification system. The classification system is defined in Article 14 of the Information Communication and Broadcasting R&D Management Regulations of Ministry of Science and ICT Notice No. 2021-23. The ICT R&D technology classification system is divided into four levels (Sections-Divisions-Groups-Classes). In this study, the main topics of clustered communities were directly identified in the papers and the topics were set based on Groups. However, when it was difficult to see the overall picture of the ICT research trend because some subject was divided in detail based on Groups, the subject was defined as Divisions. In order to compare the research trends and interests of the two Koreas, the relationship and characteristics between keywords were expressed as structured images. The network visualization method used the Kamada–Kawai algorithm. The Kamada–Kawai algorithm is characterized by placing a node with high centrality in the center of the image. The size of the node represents degree centrality, and the thickness of the link represents the frequency of...
co-occurring between the two words.

4. RESULTS

4.1. Publishing Papers and Clustering

As shown in Table 2, over the past five years, S. Korea academic papers in the field of ICT have been published 2.6 times more in quantity than in N. Korea. The amount of N. Korea academic papers published annually from 2015 to 2019 remains constant at around 400, but the amount of S. Korea academic papers tends to gradually decrease over time. This phenomenon reflects the tendency of S. Korean researchers to publish papers written in English rather than Korean. In the case of S. Korea, 11 communities were detected as a result of clustering with the modularity resolution value of 0.8. Based on the detected community results, clusters were divided into 11 major research fields of S. Korea’s ICT: image processing, wireless communication, sound signal processing, computer/Internet application/security, Internet of Things (IoT), convergence media, autonomous driving/location based service (LBS)/intelligent transport system (ITS), industrial technology, networks, neural network/big data, and energy. In the case of N. Korea, eight communities were detected as a result of clustering with the modularity resolution value of 0.6. Major research fields in N. Korea’s ICT were divided into a total of eight fields: image processing, computer/Internet application/security, industrial technology, natural language processing (NLP), neural network/decision-making, wireless communication, student education, and sound signal processing.

As shown in Fig. 1 and Fig. 2, image processing is the subject that occupies the highest percentage of the ICT academic papers of the two Koreas. There are 893 papers (16.7%) in S. Korea and 369 papers (18.2%) in N. Korea. S. Korean academic papers have been studied in the following order: wireless communication, 874 papers (16.3%), sound signal processing, 483 papers (9.0%), computer/Internet application/security, 474 papers (8.9%), IoT, 370 papers (6.9%), convergence media, 344 papers (6.4%), autonomous driving/LBS/ITS, 241 papers (4.5%), industrial technology, 238 papers (4.4%), networking, 214 papers (4.0%), neural network/big data, 192 papers (3.6%), energy, 116 papers (2.2%), and others, 916 papers (17.1%). N. Korean academic papers have been studied in the following order: computer/Internet application/security, 342 papers (16.9%) industrial technology, 331 papers (16.4%), NLP, 240 papers (11.9%), neural network/decision-making, 176 papers (8.7%), wireless communication, 130 papers (6.4%), student education, 113 papers (5.6%), sound signal processing, 64 papers (3.2%) and others, 257 papers (12.7%).

4.2. Characteristics of ICT Research in South Korea

4.2.1. Image Processing

Of the total 5,355 papers in S. Korea, 893 papers (16.7%) were in the field of image processing, accounting for the largest frequency. According to the top keyword of S. Korea’s research for five years, the “image” keyword topped the list with 281 times (Table 3). In 2017, the “deep

| Year | S. Korea | | | | | | N. Korea | | | | | | Total | | | | | | Total |
|      | Papers  | 1,189 | 1,107 | 1,083 | 1,033 | 943 | 5,355 | 412 | 399 | 406 | 408 | 397 | 2,022 |
|      | Total number of keywords (D) | 4,968 | 4,913 | 4,679 | 4,668 | 4,248 | 17,442 | 3,040 | 2,869 | 2,750 | 2,833 | 3,130 | 11,594 |
|      | Number of high frequency keywords (n) | 69.9 | 69.5 | 67.9 | 67.8 | 64.6 | 131.5 | 54.6 | 53.0 | 51.9 | 52.7 | 55.4 | 107.1 |
|      | Number of appearances of the “n”th keyword | 8 | 7 | 7 | 6 | 6 | 22 | 4 | 4 | 4 | 4 | 4 | 12 |
|      | The final keywords number to be analyzed | 86 | 87 | 74 | 91 | 86 | 134 | 87 | 86 | 70 | 80 | 91 | 113 |

a) Deduplication.
“learning” keyword ranked high with 18 times, and the frequency increases at a rate close to double every year. It topped the list with 59 times in 2019 (Fig. 3). The “convolutional neural network” (CNN) keyword has also ranked high since 2017, with 47 times in 2018 and 42 times in 2019. Since 2018, the “deep neural network” (DNN) keyword has doubled in frequency in a year. Four keywords, “image,” “deep learning,” “CNN,” and “DNN,” have been clustered under the subject of image processing, and the frequency of these keywords has increased rapidly since 2017, indicating that deep learning image processing research is becoming very active. “VR” and “high efficiency video coding” (HEVC) were also clustered under the subject of image processing. Five keywords out of the top 20 keywords were identified as keywords related to image processing (Table 3). “HEVC” is a keyword used intensively in the image processing field, as Eigenvector centrality is the lowest at 0.910 (Table 3).

4.2.2. Wireless Communication
S. Korea’s wireless communication industry was selected as a main industry and strategically fostered to overcome the economic crisis in the 1990s and upgrade the industrial structure. For this policy reason, the world’s best communication infrastructure has been established so far. Due to the revitalization of the communications industry, the research frequency was 874 (16.3%), accounting for the second highest proportion of research (Fig. 1). Although “multiple-input multiple-output” (MINO) and “orthogonal frequency-division multiplexing” (OFDM) have ups and downs in ranking by year, they have been ranked high in S. Korea’s top keywords for five years (Table 3). The “MIMO” and “OFDM” keywords have relatively low Eigenvector centrality values compared to other keywords. This is because “MIMO” and “OFDM” are not widely used apart from in wireless communication due to technological characteristics. “WLAN,” “LTE,” and “5G” have ups and downs in keyword rankings by year (Fig. 3). This is the result of reflecting the rapid changes and developments in S. Korea’s wireless communication.

4.2.3. Sound Signal Processing
The papers clustered on the subject of sound signal processing consist of underwater sound and sound navigation and ranging (SONAR), 142 papers (30%), voice recognition, 56 papers (12%), media sound, 56 papers (12%), architectural sound, 20 papers (4%), architectural sound, 20 papers (4%), and others, 189 papers (39%). “SONAR” has been seen for three consecutive years with 14 to 15 frequency since 2017, and “underwater acoustic communication” is 13 times in 2019 (Fig. 3). However, the top 20 keywords for five years do not include the keyword for sound signal processing (Table 3). This result seems to have been
derived because there is no keyword bias. The topics of research on sound signal processing were not concentrated to one side, and various topics were widely studied.

4.2.4. Computer/Internet Application/Security

474 papers (8.9%) were clustered on computer/Internet application/security subjects. This subject makes for the fourth highest number of research papers. It consists of computer and Internet applications, 230 papers (48%), security and encryption, 219 papers (46%), and Blockchain, 25 papers (5%). Computer and Internet application papers consist of various topics such as computer hardware and software, data communication, and Internet services. Therefore, security and encryption comprise the largest number of papers on a single topic in computer/Internet application/security cluster. Secret key distribution using quantum computing and Fintech simple certification using Blockchain are being covered as the latest topics. "Authentication" is a keyword related to security and encryption, and is ranked in the top 20 with 66 times in Table 3. “Cloud service,” “software,” and “interface” are also keywords clustered under this subject.

4.2.5. IoT

IoT, an element technology of the Fourth Industrial Revolution, is being used in various ways regardless of

Table 3. Top 20 keywords for ICT research in South and North Korea for 5 years (2015-2019)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Keywords</th>
<th>S. Korea Freq</th>
<th>Eigenvector centrality</th>
<th>Rank</th>
<th>Keywords</th>
<th>N. Korea Freq</th>
<th>Eigenvector centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image</td>
<td>281</td>
<td>1.000</td>
<td>1</td>
<td>Image</td>
<td>153</td>
<td>0.998</td>
</tr>
<tr>
<td>2</td>
<td>IoT</td>
<td>238</td>
<td>1.000</td>
<td>2</td>
<td>Database</td>
<td>71</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>Deep Learning</td>
<td>122</td>
<td>1.000</td>
<td>3</td>
<td>Pixel</td>
<td>63</td>
<td>0.954</td>
</tr>
<tr>
<td>4</td>
<td>CNN</td>
<td>117</td>
<td>1.000</td>
<td>4</td>
<td>Evaluation Index</td>
<td>63</td>
<td>0.883</td>
</tr>
<tr>
<td>5</td>
<td>Big Data</td>
<td>95</td>
<td>0.994</td>
<td>5</td>
<td>AHP</td>
<td>58</td>
<td>0.953</td>
</tr>
<tr>
<td>6</td>
<td>Noise</td>
<td>91</td>
<td>1.000</td>
<td>6</td>
<td>Student</td>
<td>53</td>
<td>0.988</td>
</tr>
<tr>
<td>7</td>
<td>Smart Phone</td>
<td>90</td>
<td>1.000</td>
<td>7</td>
<td>Camera</td>
<td>51</td>
<td>0.916</td>
</tr>
<tr>
<td>8</td>
<td>Sensor</td>
<td>86</td>
<td>1.000</td>
<td>8</td>
<td>Neural Network</td>
<td>48</td>
<td>0.991</td>
</tr>
<tr>
<td>9</td>
<td>MIMO</td>
<td>86</td>
<td>0.964</td>
<td>9</td>
<td>Genetic Algorithm</td>
<td>48</td>
<td>0.949</td>
</tr>
<tr>
<td>10</td>
<td>UAV</td>
<td>84</td>
<td>1.000</td>
<td>10</td>
<td>Corpus</td>
<td>42</td>
<td>0.945</td>
</tr>
<tr>
<td>11</td>
<td>Monitoring System</td>
<td>82</td>
<td>1.000</td>
<td>11</td>
<td>Decision Making</td>
<td>42</td>
<td>0.693</td>
</tr>
<tr>
<td>12</td>
<td>Vehicle</td>
<td>81</td>
<td>1.000</td>
<td>12</td>
<td>Literature</td>
<td>41</td>
<td>0.192</td>
</tr>
<tr>
<td>13</td>
<td>WSN</td>
<td>79</td>
<td>0.993</td>
<td>13</td>
<td>Korean</td>
<td>40</td>
<td>0.946</td>
</tr>
<tr>
<td>14</td>
<td>Machine Learning</td>
<td>77</td>
<td>1.000</td>
<td>14</td>
<td>Delphi</td>
<td>38</td>
<td>0.877</td>
</tr>
<tr>
<td>15</td>
<td>VR</td>
<td>71</td>
<td>1.000</td>
<td>15</td>
<td>Voice</td>
<td>35</td>
<td>0.955</td>
</tr>
<tr>
<td>16</td>
<td>Energy</td>
<td>70</td>
<td>1.000</td>
<td>16</td>
<td>Evaluation</td>
<td>34</td>
<td>0.968</td>
</tr>
<tr>
<td>17</td>
<td>User</td>
<td>69</td>
<td>1.000</td>
<td>17</td>
<td>DEA</td>
<td>34</td>
<td>0.837</td>
</tr>
<tr>
<td>18</td>
<td>OFDM</td>
<td>68</td>
<td>0.980</td>
<td>18</td>
<td>Histogram</td>
<td>34</td>
<td>0.794</td>
</tr>
<tr>
<td>19</td>
<td>HEVC</td>
<td>68</td>
<td>0.910</td>
<td>19</td>
<td>Process</td>
<td>33</td>
<td>0.889</td>
</tr>
<tr>
<td>20</td>
<td>Authentication</td>
<td>66</td>
<td>0.994</td>
<td>20</td>
<td>Similarity</td>
<td>31</td>
<td>0.960</td>
</tr>
</tbody>
</table>

ICT, information and communication technology; IoT, Internet of Things; CNN, convolutional neural network; MINO, multiple-input multiple-output; UAV, unmanned aerial vehicle; WSN, wireless sensor network; VR, virtual reality; OFDM, orthogonal frequency-division multiplexing; HEVC, high efficiency video coding; AHP, analytic hierarchy process; DEA, data envelopment analysis.
specific industrial groups. In Table 3, the frequency of the “IoT” keyword is 238 times, which is located at the top level, but the number of clustered papers is relatively small at 370 (6.9%). On the other hand, the Eigenvector centrality of the “IoT” keyword is 1.000. It can be seen that the keyword “IoT” is not limited to a single clustered topic, but is very widely used in other fields such as wireless communication and industrial technology (Fig. 4). The “IoT” keyword was clustered together with the “security” and “network” keywords, and it is confirmed that security of the IoT is a major topic in S. Korea.

4.2.6. Others
Convergence media consists of topics such as development of game and educational content using VR, mobile environments using smart phones, increasing user convenience, and ultra high definition (UHD) next-generation broadcasting systems. Convergence media research is characterized by combining technology and cultural content. Keywords such as “user” and “contents” were clustered into convergence media. The “VR” keyword clustered under the subject of image processing can also be seen in convergence media papers.

In the subject of autonomous driving/LBS/ITS, it was confirmed that there were many papers on V2X, monitoring systems using drones, and location recognition and estimation. In Fig. 3, the “unmanned aerial vehicle” (UAV) keyword has been on the upward curve since 2016, and the “vehicle” keyword has also been steadily ranked except for 2019. “UAV” is ranked as the 10th keyword and “vehicle” is the 12th keyword with the highest frequency (Table 3). Industrial technology has clustered topics such as monitoring systems of smart factories using IoT sensors, smart agriculture using machine learning, integrated railway control, railway wireless networks, and weather forecasting systems. Industrial technology is being applied in many ways regardless of the subject. Wired networks are a field where technical perfection is very high and the activity of research is gradually reduced due to the rapid development of wireless communication technology. The network field was composed of topics such as software defined networking (SDN) and network traffic control and analysis. Neural network/big data is clustered separately from the field of computer/Internet application/security, and is mainly focused on algorithm techniques, performance analysis, and improvement. It has been confirmed
that keywords of the neural network, such as “DNN,” “CNN,” “deep learning,” and “machine learning,” are used in all fields of ICT such as image processing, industrial technology, and convergence media, rather than a single field. Energy is a recent emerging field due to carbon neutrality. Research topics such as energy storage system (ESS), solar power generation, smart grids, and wireless power have been clustered.

The ICT research trend in S. Korea has the following characteristics. First, there are many convergence studies in ICT research. The Eigenvector centrality of 14 keywords among the top 20 keywords in S. Korea is 1.000, which has very high connectivity and importance (Table 3). As shown in Fig. 4, the frequency of co-appearance of the top 20 keywords is also high, and it can be confirmed that they are used in various fields. In particular, it can be seen that the frequency of co-appearance in image processing, IoT, and wireless communication fields is high. This shows that ICT research in S. Korea is active in interdisciplinary convergence research. Second, it can be seen that ICT research in S. Korea is very sensitive to trends. As shown in Fig. 3, it can be confirmed that keywords follow the latest trend every year. New keywords are shown every year. “IoT” and “image” keywords remain at the top, but the trend line of many keywords is upward.

4.3. Characteristics of ICT Research in North Korea

4.3.1. Image Processing

Of the total 2,022 papers in N. Korea, 369 papers (18.2%) were in the field of image processing, accounting for the largest frequency. Among the keywords clustered by image processing, “image” was the most common with 153 times, followed by “pixel,” 63 times, “camera,” 51 times, and “histogram,” 34 times (Table 3). Four out of the top 20 were composed of keywords related to image processing. Among the image processing papers, the number of research papers related to object detection is 169 papers (46%). There have been a number of studies related to object detection using CNN and object surface detection. It can be seen that some research efforts are following the latest trends. There were 54 papers related to text or document image processing written by humans. There are 21 papers (6%) related to image encryption, such as secret images and watermarking. There are 33 papers (9%) in the subject of medical imaging, such as computed tomography (CT) or X-ray image processing, research on separating blood cells from images, vein recognition, and lung tissue separation. Ryugyong Dental Hospital in Munsu Street, Pyongyang, developed a bone density measurement and implant design program using X-ray images of patients. Ryugyeong Dental Hospital was selected as the first medical institution to be an exemplary unit for infor-
matization at the National Exhibition of IT Successes 2016 (Rodong Sinmun, 2018).

### 4.3.2. Computer/Internet Application/Security

There were 342 N. Korean papers (17%) clustered on computer/Internet application/security. This cluster comprises the second highest number of research papers. It consists of security and encryption, 144 papers (42%), web services, 84 papers (25%), software, 59 papers (17%) and databases, 37 papers (11%). In N. Korea, which has military-first policies, the largest number of papers being related to security and encryption shows that there is a high research interest in hacking and system attacks. Among the web service studies, it was confirmed that the ad hoc on-demand distance vector (AODV) protocol, which was actively studied in S. Korea in the early 2000s, was studied in N. Korea in 2015. This proves that N. Korea lags behind S. Korea by more than a decade in web service technology gap. It can be understood that the web infrastructure that operates the web service is still backward. Since the 2000s, N. Korea has been intensively fostering the software industry rather than investing in IT infrastructure or hardware, using it as a means to overcome economic difficulties by overcoming poor manufacturing and acquiring foreign currency. Due to the high technology level of development personnel and relatively low labor costs, software development organizations are dispatched to China, and they are generating profits by developing software modules through overseas contracts (Kim & Lee, 2015). The "database" keyword was ranked second with 71 times, and the Eigenvector centrality showed the largest value at 1.000. Among the computer/Internet application/security keywords, only "database" was ranked in the top 20 (Table 3).

### 4.3.3. Industrial Technology

The number of papers on industrial technology in N. Korea is the third largest with 331 papers (16.4%). That is not much different from the number of papers in the first and second fields. The reason for the large number of papers related to industrial technology is N. Korea's policy and factory university system. In April 2018, N. Korea announced the "All-out Concentration Policy for Economic Building" at the plenary meeting of the Central Committee of the Workers' Party of N. Korea. With the aim of the socialist economy and the knowledge economy, the strategy planned for unmanned and automated factory production processes and modernized management activities of the...
enterprise (Park, 2018). A factory university, a higher education institution established in factories and enterprises, is a regular educational institution that allow workers in the relevant enterprise or nearby areas to work and study at the same time (Cho, 2004). In addition to the purpose of training technical personnel and expanding higher education, factory universities also serve to solve technical problems arising from factories and improve productivity with creative ideas as a window for industry-academic cooperative research (Yoo, 2021). Among the keywords clustered in industrial technology are “evaluation index,” “analytic hierarchy process” (AHP), “decision-making,” “Delphi,” “data envelopment analysis” (DEA), and “process” (Table 3). The proportion of research was in the order of technology management, 80 papers (24%), improvements in production processes at factories, 64 papers (19%), technology evaluations, 51 papers (15%), and process management systems, 29 papers (9%). As shown in Fig. 5, the keywords AHP and decision-making have ups and downs in ranking according to frequency by year. However, the “evaluation index” and “Delphi” keywords have steadily risen from 2016 to 2019. The keyword “process” has emerged rapidly since 2018, which is believed to be related to the policy announced at the plenary meeting of the Central Committee of the Workers’ Party of N. Korea.

4.3.4. Natural Language Processing

The frequency of research on NLP in N. Korea was significantly higher than that of S. Korea (Fig. 6). The number of NLP papers in N. Korea is 240 (11.9%), the fourth most frequently studied topic. On the other hand, S. Korea has 35 papers (0.6%) on NLP. In both the absolute number of papers and the ratio of NLP papers to all ICT papers, N. Korea was found to be much higher. Research on machine translation was the most common with 63 papers (26%). Among machine translation, there are 27 Chinese studies and 20 English studies. In addition, QnA, 34 papers (14%) and information retrieval, 28 papers (12%) were followed. The keywords clustered by NLP are “corpus,” “literature,” and “Korean,” which contain three of the top 20 keywords (Table 3). The AI Technology Research Institute at Kim Il-sung University in N. Korea has developed a machine translation system that can translate English, Chinese, Japanese, Russian, German, French, and Spanish (2021). It is said that the translation accuracy is almost at the level of an expert, and the translation speed is three times faster than before (Kang, 2021).

4.3.5. Student Education

Student education clustered into N. Korea’s ICT research field is not classified in S. Korea (Fig. 6). There are
113 papers on student education. Papers on the subject of student evaluation were the most common with 49 papers (43%), followed by the subject of remote education with 26 papers (23%). The main research contents were learning evaluation of the remote education system and the evaluation of teaching ability. The Kim Jong-un regime set the goal of “Human Resources Development of National Science and Technology,” revised the Education Act in 2015 and revised the constitution in 2019, and promoted modernization and informatization of education. As national computer networks are equipped and computers are distributed to schools, various attempts are being made, such as developing educational programs using remote education systems and establishing remote schools. In 2016, the university entrance examination was conducted remotely by computer (Kim & Kim, 2020). N. Korea is using ICT infrastructure to maximize educational effects and make national investments to improve the educational gap and the environment.

4.3.6. Others

The number of papers clustered with neural network/decision-making is ranked fifth with 176 papers (8.7%). The “neural network” keyword is ranked 7th with 48 frequency, and Eigenvector centrality is 0.991 (Table 3). The relatively high Eigenvector centrality of the keyword “neural network” shows that it is used as an application case in various fields rather than in research on the neural network alone. In particular, the “neural network” keyword was most prominent in the field of image processing. “Neural network” is also widely used in research on the industrial technology and NLP. Wireless communication is the second most studied topic in S. Korea with 874 papers (16.3%), but in N. Korea, it is ranked sixth with 130 (6.4%), so its research performance is relatively low. The field of wireless communication consists of wireless sensor network (WSN), 31 papers (24%), OFDM, 28 papers (22%), and code-division multiple access (CDMA), 10 papers (8%). Areas that require large-scale infrastructure investment such as wireless communication are very low in terms of research frequency and ratio compared to S. Korea. Network infrastructure requires huge capital inputs and overseas imports of equipment. However, N. Korea is under sanctions for its nuclear testing and ICBM launches. Therefore, large-scale infrastructure investment is difficult and the number of related research efforts is inevitably small. The number of papers clustered by sound signal processing is the smallest at 64 papers (3.2%). S. Korea’s research on sound signal processing has a variety of subjects such as underwater sound and SONAR, voice recognition, and media sound. On the other hand, in N. Korea the number of voice recognition research is 47 papers (73%), accounting for a very high percentage of one subject. As shown in Table 3, the “voice” keyword clustered by sound signal processing was ranked as the only top 20 entry with 35 frequency. In addition to the clustered field, it was confirmed whether there were keywords related to the Fourth Industrial Revolution. At the 7th Congress of the Workers’ Party of N. Korea held in May 2016, N. Korea instructed to focus on high-tech technologies, mainly based on information technology, nanotechnology, and biotechnology, and aimed to build a knowledge-based economic powerhouse. At this time, N. Korea used the word “New Century Industrial Revolution,” a concept similar to the Fourth Industrial Revolution (Lim, 2019). N. Korea media are promoting that smart homes using IoT and VR were introduced at the “National Exhibition of IT Successes 2018” (Kang, 2018). In the analysis of ICT research papers, two VR papers and one IoT paper could be found. However, keywords related to the Fourth Industrial Revolution such as big data, smart home, smart city, and UAV were not found.

The ICT research trend in N. Korea has the following characteristics. First, there is a higher tendency for independent research than interdisciplinary convergence research. It can be seen that the Eigenvector centrality is relatively low compared to the top 20 keywords in S. Korea. Of the top 20 keywords in S. Korea, 14 have Eigenvector centrality of 1.000, but only one in N. Korea has Eigenvector centrality of 1.000 (Table 3). As such, the connectivity and importance of keywords are low. Second, there was a higher tendency to be influenced by Workers’ Party of N. Korea policies than by research topics in the global state-of-the-art technology trends. Researchers are conducting policy-related studies in the Workers’ Party of N. Korea, such as industrial technology advancement, remote education, and NLP. There were many studies that directly helped the economy of people’s livelihoods.

5. SUGGESTIONS FOR INTER-KOREAN COOPERATION

The N. Korea leader’s new year’s address has mentioned modernization of the light industry and agricultural and fishery industries using ICT technology since 2014. In addition, ICT business profits are paid to developers to encourage commercialization of ICT technology. Through these contents, it is confirmed that N. Korea has a high...
government interest and demand for economic development using ICT (Seo, 2018). The research situation and the stage of technology development are different for each detailed ICT field. Therefore, it is necessary to check the demand in each ICT field in N. Korea. In addition, specific exchange strategies should be established for each field, rather than unconditionally transplanting the developed S. Korea's ICT technology into the poor environment of N. Korea.

Image processing has a very high research rate in both Koreas and research is actively being conducted. It is expected that joint research will be possible in the future. In particular, if medical imaging technology is cooperated in consideration of the poor medical environment in N. Korea, it is expected that the N. Korea authorities will try to cooperate first and favorably. In addition, it is proposed to provide UHD sports videos that S. Korea has the right to broadcast to N. Korea for research. If this happens, research materials in the field of N. Korea's image processing, which focuses on text and document images, can be enriched. Since a sports video has no political content and is not affected by sanctions against N. Korea, research exchanges are likely to proceed very stably.

Computer/Internet application/security research is also expected to be very active as a joint research subject between the two Koreas. This proposes to create a special economic zone for ICT and to build a cloud center that can gather a common database from areas where there is no significant difference in opinion such as language, ancient history, medicine, forests, and fisheries. The database aims at the unique culture of the Korean peninsula. Cloud centers in special economic zones cooperate by investing in S. Korea hardware and capital and dispatching N. Korea management personnel. With the help of international organizations such as UNESCO, there is a high probability of cooperation. In the field of security and encryption, N. Korea has a high national interest and technology. Since security and encryption technologies can be easily used in the military, exchanges and cooperation are very difficult due to sanctions against N. Korea. It may be possible to hold cyber security contests in third countries to evaluate security and encryption skills and to start exchanges by organizing a network of human resources cooperation.

The frequency of industrial technology research in N. Korea and the demand for manufacturing development are high. If capital is supported and production facilities are equipped with new models, it is expected to be of great help in improving manufacturing efficiency and revitalizing the industry. If S. Korea’s IoT technology is combined with N. Korea’s production process, it will be able to secure remarkable growth engines. However, manufacturing support that may be used as munitions should be excluded. Humanitarian assistance in manufacturing should be selected with experts. Industrial technology is a field that requires voluntary investment and support in the private industry. It is more efficient to conduct industrial technology exchanges through technology transfer of research institutes operated by private companies. If there is a high risk from investment in N. Korea, business operators will not invest in or support it. Prior to private exchanges, the two Korean governments should use the case of the Kaesong Industrial Complex as a teacher and eliminate investment risks in advance.

N. Korea’s linguistic theory of Juche (self-reliance) defines language as a product of society and considers it important to protect ethnicity through language (Jeon & Kim, 2019). Language research is active in N. Korea. Because language policy focuses on social meaning and function and considers language an important tool for social development, it is presumed that many studies on NLP using Korean in the field of ICT were influenced by language policies. N. Korea focuses on its own computer operating system and language processing technology without relying on foreign software (Ko et al., 2007). The development of open-source-based software using its own language information is based on the activation of NLP research. Taking advantage of N. Korean language research, we would like to propose exchanges such as promoting Korean language information conferences, developing PC and mobile operating systems based on Korean, and importing Chinese machine translation technology. NLP research is also a field that can be relatively less affected by political situations or sanctions against N. Korea. If we continue to exchange information academic conferences between the two Koreas, which were held at the Korean Language Information Society in the past, it will be a good example for exchange and cooperation in other fields.

N. Korea is very interested in national investment to improve the educational environment and reduce the educational gap by using ICT infrastructure to build a strong country. In April 2020, N. Korea enacted the Remote Education Act and strengthened the legal foundation for Human Resources Development of National Science and Technology (Yang, 2020). Any N. Korean citizen was granted the right to become a student. It was stipulated as the duty of all institutions to ensure remote education by preparing educational facilities and learning places. In or-
order to upgrade educational conditions across N. Korea, a network system was established and remote education was started, and research on education evaluation methods is also steadily being conducted. If the N. Korean education system is opened up, it is predicted that it will be opened to the extent that there is no threat to the socialist system. In addition, since the obligation to remote education is imposed by law, it is judged that there will be sufficient demand for the establishment of an advanced education system. The cooperation plan related to education is as follows: free support of educational contents in S. Korea, teaching know-how to operate Korea National Open University (KNOU) in S. Korea, a joint study on student remote evaluation, a joint development of educational contents using AR and VR, the establishment of a unification education broadcasting station, and so on.

It is expected that private exchanges between the two Koreas in the field of wireless communication will not be easy. There is no possibility that N. Korea will try to cooperate first because the opening of radio wave resources and disclosure of the current status are burdensome for maintaining the N. Korea regime. However, N. Korea’s research in wireless communication is poor and its technological completeness is not high, so N. Korea recognizes it as a field that requires investment first. Therefore, it is necessary to first sign an agreement between the governments and promote cooperative projects through incentives such as technology transfer and free support. Specifically, these include installation of 4G infrastructure equipment, maintenance training of wireless network systems, and personnel exchanges.

6. CONCLUSION

This study compared and analyzed research trends in the ICT fields of the two Koreas using academic papers published for five years from 2015 to 2019. In the case of S. Korea, the frequency of research on image processing and wireless communication was high, while N. Korea was in the order of image processing, computer/Internet application/security, and industrial technology. In N. Korea, NLP was found to have a much higher frequency of research than in S. Korea. Student education in N. Korea is a unique field that has not been clustered in S. Korea. Based on these research results, specific exchange strategies for each ICT field should be established. In the fields of image processing and computer/Internet application/security, the research activities of the two Koreas are high, so joint inter-Korean research is expected to be possible quickly and actively. Research in the NLP field makes the most of the advantages of N. Korea’s language policy, so that S. Korea can receive and utilize know-how. It is necessary to conduct private business exchanges by confirming the demand for industrial technology. It has been confirmed that there is sufficient demand for the establishment of a remote education system, and educational content exchange is necessary. The wireless communication field needs free infrastructure support after an intergovernmental agreement. The public-private joint plan will bring about the development of inter-Korean ICT exchange and cooperation.

The limitations of this study and the characteristics of N. Korea’s ICT research publications are as follows.

First, in N. Korea’s ICT research, topics and keywords that can be used for military weapons were not well expressed. This means that the N. Korean authorities are tightening control over ICT. Due to the disclosure of the latest ICT information that can be used as a military weapon, it is difficult to accurately estimate the level of ICT in N. Korea only with publication research keywords. Second, N. Korean papers are composed of one to three pages, and there is a lack of evidence for research methodology and experiments. Therefore, it is difficult to judge the quality level of the studies. Third, it is believed that the policy of the N. Korean authorities and the intention of the N. Korean Kim Jong-un may have strongly affected research content and trends. It is necessary to check the impact of N. Korea’s ICT policy on research activities through follow-up studies. Fourth, specific cooperation methods were proposed for each field as described above, but the United Nations currently comprehensively prohibits science and technology cooperation with N. Korea. Therefore, the process needs to find ways to bypass sanctions against N. Korea. A strategy should be established from the exemption field (Choi et al., 2020). This study is meaningful in that it compared the trends of ICT research between the two Koreas through quantitative analysis without relying on the subjective interpretation of the researcher. It is expected to be used as basic data for the establishment of the government’s exchange policy and private cooperation.

CONFLICTS OF INTEREST

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


Choi, H. K., Byun, H. M., Noh, K. R., Jung, E. S., & Lee, S. H. (2020). Research on inter-Korean cooperation measures such as unique technologies and strong technologies that lead the recent change in North Korea and human resources related thereto. Korea Information Society Development Institute.


Leydesdorff, L. (1989). Words and co-words as indicators of...
Yoo, Y. K. (2021). North Korea, modernizing production with "example factory." https://www.pressian.com/pages/articles/ 2021051309585314611#0DKU