

A Big Data Analysis of Public Interest in Defense Reform 2.0 and Suggestions for Policy Completion

Kim, Tae Kyoung, Army Officer, Rep. of Korea

Kang, Wonseok, Professor, Management, Korea Military Academy, Rep. of Korea

Abstract

This study conducted a big data analysis study through text mining and semantic network analysis to explore the perception of defense reform 2.0. The collected data were analyzed with the top 70 keywords as the appropriate range for network visualization. Through word frequency analysis, connection centrality analysis, and an N-gram analysis, we identified issues that received much attention such as troop reduction, shortening of military service period, dismantling of the border area unit, and returning wartime operational control. In particular, the results of clustering words through CONCOR analysis showed that there was a great interest in pursuing the technical group, concerns about military capacity reduction, and reorganization of manpower structure. The results of the analysis through text mining techniques are as follows. First, it was found that there was a lack of awareness about measures to reinforce the reduced troops while receiving much attention to the reduction of troops in Defense Reform 2.0. Second, it was found that it is necessary to actively communicate with the local community due to the deconstruction and movement of the border area units, such as the decrease of the population of the region and the collapse of the local commercial area. Third, it was judged that it is necessary to show substantial results through the promotion of barracks culture and the defense industry, which showed that there was less interest than military structure and defense operation from the people and the introduction of active policies. Through this study, we analyzed the public's interest in defense reform 2.0, which is a representative defense policy, and suggested a plan to draw support for national policy.

Key words: Defense Reform 2.0, Big Data Analysis, Text Mining, Defense Policy, Defense Issues

* Received: Aug. 3rd, 2022 Revised: Sep. 21th, 2022 Accepted: Oct. 18th, 2022

This work was supported by Korea Military Academy (ROK Army Soldiers' Value and Culture Research Center) Research Grant in 2021.

1. Introduction

The Republic of Korea Armed Forces is promoting reforms in various fields within the military through Defense Reform 2.0, which was released on July 27, 2018. The reason why such a reform plan was promoted is due to the lack of military service resources due to population cliffs, military-political involvement, and corruption of defense, which led to deepening restrictions on defense conditions due to the deterioration of public trust. In addition, the uncertainty caused by North Korea's threat, the expansion of the influence of neighboring countries, and the spread of transnational and non-military threats have diversified security threats and increased uncertainty (Ministry of National Defense of South Korea, 2019). The existing defense reform was changed and replaced by the Lee Myung-bak administration with the National Defense Reform Basic Plan 2009-2020, the National Defense Reform Basic Plan 2011-2030(307) and 2012-2030, and the National Defense Reform Basic Plan of Park Geun-Hye Government has been led to 2014-2030, which was followed by the change of the security environment, starting with the National Defense Reform 2020 promoted by the Roh Moo-hyun administration. Although it is inevitable to revise and supplement the basic plan of defense reform according to changes in the security environment, the method of implementation changed and replaced by each government is not fully equipped with the plan, the driving force is weakened, and the fatigue of reform is at risk of being terminated at the level of improvement, not reform (Choi, 2019). Therefore, Defense Reform 2.0 is essentially required to have an executive power that can overcome the inertia of defense reform that has been followed by each government (Kim & Jeong, 2014). Such policy implementation power is essential for the people's support, and if the public's awareness of defense policies such as Defense Reform 2.0 is poor, the government will also lose the implementation power of policy promotion (Kim, 2017). In addition, if public opinion on the president's policy promotion is poor, even the ruling party may not be able to actively support the policy, which may eventually make it impossible to implement the policy (Lee, 2018). Therefore, it is very important to grasp the public's interest in policy, and the public's interest in the current policy can be used as basic data for public relations activities to elicit policy support from the people. In addition, since defense policy has an influence that can determine the fate of the country, it is necessary to share policy values by communicating with the people through promotional activities not only for policies that are currently exposed but also relatively unnoticed (Cho, 2009).

Therefore, this study aims to identify and analyze the public's interest in defense issues for about 28 months from the announcement of the policy. Through this, the government intends to present activities necessary for the successful completion of defense reform 2.0 and barracks innovation and provide basic data for establishing public relations strategies in the future.

2. Theoretical Background

2.1 Defense Reform 2.0

The Moon Jae-in government's defense reform plan, first announced on July 27, 2018, aims to move toward a military that can take a leading response to all-round security threats with the early implementation of a strong military force that supports the Republic of Korea with peace and prosperity and is operating in line with an elite and advanced country based on advanced science and technology. In addition, the promotion stance of Defense Reform 2.0 is to strengthen the constitution and foundation for leading defense capabilities, actively utilize science and technology in the era of the 4th industrial revolution, and pursue reforms that meet the needs of the nation and society. The military is working as a decisive attitude to accept the defense reform as an order of the people who can no longer delay the national defense reform. The field of promotion is divided into the military structure, defense operation, barracks culture, and defense business. A total of 270.7 trillion won will be invested from 2019 to 2023, including 94.1 trillion won for defense improvement and 176.6 trillion won for combat power operations (Ministry of National Defense of South Korea, 2019). In order to prepare for all-around security threats with state-of-the-art science and technology-based future-oriented military structure, the military's command, unit, combat power, and force structure will be reorganized. In the field of barracks culture, policies such as reforming the military justice system, abolishing the guardhouse system, raising hospital salaries, improving human rights in the barracks, enhancing expertise and flexibility, strengthening defense R&D capabilities and defense competitiveness, and fostering defense brokerage (Editorial Department, 2019). As a result of research on national defense reform 2.0, Kang & Kim (2019) argued that future-oriented policies should be pursued through the elite adjustment of reserve forces, the practical modernization of armed forces, and the reform of the training system, and the increase

of related defense expenses is needed for the 'reduction of reserve power', which is part of the national defense reform 2.0 promotion trend. Yoon & Lee (2019) proposed the conversion of awareness of reserve forces, the enactment of partial mobilization laws, the strengthening of mobilization support capabilities for the fourth industry, and the realization of allowances for the elite power reserve. Jeong et al. (2018) said that through the military judicial reform, which is being promoted dramatically through the National Defense Reform 2.0, the democratization of the military should be achieved and the stigma of human rights can be removed.

2.2 Big Data Research

2.2.1 Concept of Big Data

Big data is data generated in a digital environment, which is massive in size, has a short generation cycle, and has large-scale data including text and video data as well as numerical data. The big data environment has increased the amount of data compared to the past, and the types of data have become diverse so that people can analyze and predict their thoughts and opinions through location information and SNS as well as their actions. Big data refers to atypical data that is difficult to collect or manage when defined in a narrow range. In McKinsey (2011), Big Data is defined as data of a scale that exceeds the scope of storage, management, and analysis of existing methods. On the other hand, Korea IDC defined Big Data as a next-generation technology and architecture designed to extract value from various kinds of data at low cost and to support data superfast collection, discovery, and analysis. Gartner Inc. defined big data as crude oil in the 21st century and defined it as a phenomenon in which various types of data are generated so quickly that companies cannot handle it. In McKinsey's case, big data is interpreted in a narrow range focusing on the size of data, and Gartner and IDC in Korea interpreted big data in a wider range focusing on how to utilize such large-scale data, that is, how to perform (Han & Kang, 2016). The predictability of big data is analyzed in large quantities such as information and opinions that have already been freely shared in the web-based environment, unlike the existing research methods that have been verified after establishing hypotheses, and it is used in various fields as a more practical and realistic research method (Kim, 2020).

2.2.2 Component of Big Data

Big data is generally based on 3V (Volume, Velocity, Variety) and is described by adding 1V (Value) or 1C (Complexity). Gartner's Doug Laney, who first paid attention to big data, first proposed it 12 years ago, and since then, several people have attempted to expand the components of big data by adding various Vs such as validity, integrity, value, and visibility. IBM defined Big Data as an opportunity to obtain insights that could not be answered in the past with new types of data with 3V, and added veracity to explain Big Data. SAS presented 4V as a component of big data by adding new value to 3V and Gartner added complexity to 3V and presented another component of big data. Scale refers to the amount of data generated. The size of the data is showing an unprecedented rapid increase. Although 'large size' cannot be clearly defined, more data is usually defined as big data based on petabytes or zetabytes. And these criteria are expected to surpass petabytes and zeta bytes later as data is continuously generated. Diversity refers to various types of data. Diversity is associated with managing complex and diverse forms of data that include all structured, semi-structured, and unstructured data (IBM, 2012). Existing data were dominated by simple structured numbers or text-based structured data, which refers to data stored in fixed fields. For example, when shopping, addresses, phone numbers, and payment information are entered into a pre-generated frame, which can be referred to as structured data and managed and analyzed in a traditional way. Semi-structured data is not stored as a fixed field, but refers to data including metadata or schema such as XML or HTML. That is, it is not data input in a certain frame, but data stored in various computer languages. However, in big data analysis, the majority of unstructured data analysis is difficult to put into the framework. Data generated on social network services (SNS) such as Twitter and Facebook are representative types of unstructured data, and real-time location information transmitted to smartphones, various log records, and multimedia (music, photos, videos, etc.) are also unstructured data. These unstructured data account for more than 90% of big data analysis and include various types of data, not existing structured data such as photos and videos (Kim, 2013). Speed refers to the movement of data. IBM defined speed as 'streaming data analysis' in its 2011 report that enables decision-making in a matter of a second (IBM, 2011). A complex meaning is added to the concept of data movement. That is, it refers to the speed at which data is generated and processed, and the meaning of this speed is divided into three categories.. It refers to the speed at which data is stored in a server within a company, the speed at which useless or meaningless parts of the generated data are processed and available, and the

speed at which refined data is analyzed and meaning is extracted to achieve the final purpose. Veracity refers to the certainty of data. IBM explained this veracity in addition to 3V, which means the level of trust that can be given to certain types of data (IBM, 2011). This veracity acts as an important factor in big data analysis. The unstructured data produced in modern times are less reliable, and for example, there was an incident in which a murder case posted on Facebook spread at a tremendous rate, which eventually turned out to be false. In other words, it is important to dig up accurate and reliable information from a huge amount of big data and process it again to make it into high reliability data. Value is added to 3V because through Big data analysis, we can derive valuable results that could not be derived previously.(Dijcks, 2013). There is meaningful information hidden amongst a large and non-traditional data, and we can identify it by Big data analysis. By adding these two Vs to the existing 3V, Douglas said good big data should meet all of 5V(Douglas, 2012).

2.2.3 Case of Using Big Data

Analyzing big data is changing the real world. Things that were impossible in the past are now being realized by the development of big data analysis technology. Big data is a large, unstructured pile of data that was previously difficult to systematically collect, manage and analyze. Big data analysis technology is to analyze the global data exponentially due to the increase of users of various IT, mobile devices, and SNS, and to create new added value. McKinsey, a global consulting firm, summarizes the background of big data rise in 'Big Data: The Next Frontier for Innovation, Competition, and Productivity as follows.

First, it is the increase in customer data tracking and collection behavior of companies. Customer data is tracked through various media such as the Internet and smartphones, and information on user information and consumer behavior can be collected not only online but also offline. British retailer Tesco collects more than 1.5 billion customer data every month. Second, the reduction in storage media, camera modules, and display prices has led to the spread of multimedia content use and an increase in information related to it. High-definition videos already account for more than 50% of the total Internet traffic and are expected to increase to more than 70% in the near future. Third, along with the rapid spread of SNS such as Twitter and Facebook, it is an explosion of unstructured data such as text. On average, more than 90 content is uploaded per user per month on Facebook alone, and 24 hours of video are uploaded every minute on YouTube. Information distributed on SNS is unstructured data, and

additional data processing is required to process it, thereby increasing the complexity of the data. Fourth, it is an increase in data generated in inter-object communication networks due to the development of communication technologies such as Machine to Machine (M2M) and the Internet of Things (IoT). Through activation of M2M and IoT, users did not generate data, and the infrastructure itself generated a large amount of data (Kim, 2013). In this background, IT service companies such as HP, Samsung SDS, LG CNS, IBM, etc. have become more and more desperate to create new added value, and as a result, they have become interested in big data business as a new source of revenue for the cloud era. In addition, general companies are now judged to be ahead of the market competition, which is becoming increasingly intense by strengthening business analysis/prediction capabilities through data secured by the company, not simply information.

Big data is expanding/evolving through remarkable results in marketing areas such as providing customized services, preventing tax evasion, intelligent transportation guidance systems, and production management areas such as production process optimization, demand prediction, and credit card electronic payment fraud.¹⁾

3. Research Method

3.1 Research Subjects and Data Collection Methods

This study conducted a big data analysis using text data to explore awareness of national defense issues. Data collection used Naver, Daum, Google, and SNS services Facebook and Twitter, the largest portal sites at home and abroad, while Naver selected the entire web, blog, news, cafe, knowledge, and academic information, and Google collected web documents and news channels. The analysis keyword is Defense Reform 2.0, and the data analysis period is 38 months from July 27, 2018, when Defense Reform 2.0 was announced, to September 27, 2021, for data collection, and the collected data was conducted focusing on the top 70 keywords (Sung, 2020). Specific information on the analysis data is shown in Table 1.

1) Kim Yeon-jin (2013), "aCorporate Innovation through the Introduction of an Advanced Analysis System Based on Big Data," PRACTICE & Small and Medium Business Forum, p.44.

<Table 1> Information on analysis data

Division	Content
Scope of collection	Naver (web documents, blogs, news, cafes, intellectuals, entire academic information), Daum (web documents, blogs, news, cafes), Google (web documents, news, Facebook), Twitter
Collection period	From July 27, 2018 to September 27, 2021 (38 months)
Collection tools	TEXTOM
Search word	Defense Reform 2.0
Analysis tool	Ucinet 6

3.2 Research Tools and Methods

In this study, big data analysis was conducted through text-mining and semantic network analysis. Text mining refers to a technology (Sung, 2020) that analyzes information indicating interest in unstructured text data in the context beyond the level of keywords and finds valuable information by finding patterns or relationships, and can support decision-making (Kim & Kim, 2019).

Semantic network analysis is a powerful methodology for finding the form and semantic structure of the relationship between variables by grasping the actual meaning, not within the researcher's framework (Lee, 2016). In this study, data that was not directly related to Defense Reform 2.0 were deleted from the collected data, and the process of combining words that had the same meaning but were written differently was conducted. In addition, words with incorrect spacing and incomplete words were refined, and words with single syllables whose meaning could not be grasped were deleted. Examples of data purification are shown in Table 2.

<Table 2> Examples of data purification

Words with the same meaning (integration)	Words that are unnecessary for analysis (delete)
Conversion of wartime operational control (conversion of operational control, the transition of wartime operational control, and conversion of wartime operations)	- Words of unclear meaning
Period of service (duration of military service, service period)	
Moon Jae In Government (Moon Jae-Government)	
Implementation of military agreements (Military agreements)	- News, journalists, today, this day, this time, part, and the middle. Words unrelated to Defense Reform 2.0.
Use mobile phone (cell phone)	
Shin, Won Sik (Shin, Won-sik, Wonsik Shin, Shin Won Sik, Sin, Wonsik)	

Looking at the Table 2, where refined words are organized, words written in other words of the same meaning, such as the transition of wartime operational control and service period, were completely corrected, such as the implementation of government or military agreements, and words recognized as other words were integrated. In addition, words whose meaning was unclear as a single syllable were deleted, and words that were not directly related to Defense Reform 2.0, such as news, reporters, today, this day, and part of the analysis, were also deleted.

TEXTOM, which provides big data analysis functions, was used for data collection and analysis in various fields, and its reliability and functionality were confirmed. Through the data collection function provided by Textom, text data were collected from channels such as Naver, Daum, Google, and Twitter, and frequency analysis, connection centrality analysis, and N-gram network graph analysis were performed, and the words were clustered and analyzed through Ucinet6.

4. Research Results

4.1 Data Collection

Since Defense Reform 2.0 was announced for 38 months from July 27, 2018, to September 27, 2021, a total of 7,069 data extracted from Naver, Daum, Google, and Twitter have been collected, focusing on keywords for Defense Reform 2.0, 3,457 from Naver, 3,442 from Daum, and 1,169 from Google and 1 from Twitter.

<Table 3> Data Collection Results

Channel	Section	Collection amount(number)
Naver	Web document	1,000
	Blog	718
	News	109
	Cafe	530
	Intellectual	10
	Academic information	90
Daum	Web document	829
	Blog	910
	News	727
	Cafe	976
Google	Web document	1,000
	News	169
	Facebook	0
Twitter	Twitter	1

As a result of removing duplicates based on the collected data, 4,140 data were derived. In the purification module, Espresso K was used to reflect proper and complex nouns in the result values. The specific data collection results are shown in Table 3.

4.2 Word Frequency and Connection Centrality Analysis Results

Table 4 shows the results of word frequency analysis and connection centrality analysis for Defense Reform 2.0. The word frequency analysis is listed in the collected data based on the number of appearances of specific words. Looking at the results of the word frequency analysis in Table 4, the top 10 words were defense reform 2.0 (1,914), defense ministry (1,308), propulsion (1,152), military (590), government (505), reduction (459), army (447), conversion (407), Moon Jae-in (360), presentation (358), president (350), defense (319), shortening (264), wartime operational control (264), and troop (245), Minister (234), Service Period (230), People (229), Minister of National Defense (220), and Preparation (210) were surveyed in 10-20th place.

The centrality is an indicator that expresses the degree of the center of an actor in the entire network, which is a measure of the relative importance of a vertex or node in a graph or social network. Among them, the degree of connection centrality is higher as the number of nodes directly connected to one node and the number of nodes is higher (Lee Sun-hee, Seok-ryu, 2020). As shown in Table 4, Connection centrality was found to be the top 10 words: Defense Reform 2.0 (0.06), Defense Ministry (0.08), Promotion (0.057), Military (0.042), Army (0.035), Government (0.030), Defense (0.024), Announcement (0.024), Reduction (0.022), and President (0.018). Next, the connection centrality index was high in the order of planning (0.018), hosting (0.016), military personnel (0.016), minister (0.016), conversion (0.015), Moon Jae In (0.015), unit (0.015), military (0.015), border region (0.015), and task (0.015).

<Table 4> Frequency and connection centrality of extracted words

Word	Frequency	Connection centrality	Word	Frequency	Connection centrality
Defense Reform 2.0	1,914	0.106	Jeong Kyung Doo	182	0.010
Ministry of National Defense	1,308	0.080	2022	181	0.010
Propulsion	1,152	0.057	Smart Defense Innovation	178	0.007
Military	590	0.042	Threat	176	0.008
Government	505	0.030	General	174	0.006
Reduction	459	0.022	Military unit	172	0.008
Army	447	0.035	Industrial revolution	169	0.010
Conversion	407	0.015	Expansion	167	0.012
Moon Jae-in	360	0.015	2020	165	0.012
Presentation	358	0.024	Seo Wook	162	0.006
President	350	0.018	Change	162	0.012
Defense	319	0.024	Republic of Korea	160	0.011
Shortening	264	0.001	North Korea	160	0.014
wartime operational control	264	0.008	Air force	160	0.013
Force	245	0.013	National security	159	0.010
Minister	234	0.016	Response	156	0.010
Service period	230	0.010	Policy	155	0.009
People	229	0.013	Construction	154	0.006
Defense minister	220	0.011	Soldiers	150	0.013
Preparedness	210	0.012	High-tech	143	0.008
Host	210	0.016	Success	141	0.006
Civilian worker in the military	210	0.016	Fulfillment	141	0.005
Reorganization	207	0.013	National forces	135	0.008
Dismantle	203	0.011	Era	130	0.007
Plan	200	0.018	Strong army	129	0.003
Foundation	197	0.012	Female soldier	129	0.007
Task	196	0.015	2018	128	0.010
Future	195	0.011	Combat power	128	0.009
Peace	195	0.011	Realization	127	0.007
Unit	193	0.015	Smart	126	0.008
Base	192	0.007	Song Yeongmu	121	0.007
Armed forces	192	0.015	Korea-U.S. alliance	121	0.006
Core	189	0.011	Blue House	121	0.009
Soldier	188	0.012	Korean Peninsula	120	0.008
Border area	187	0.015	Candidate	120	0.008

4.3 N-gram Network Analysis and Visualization Results

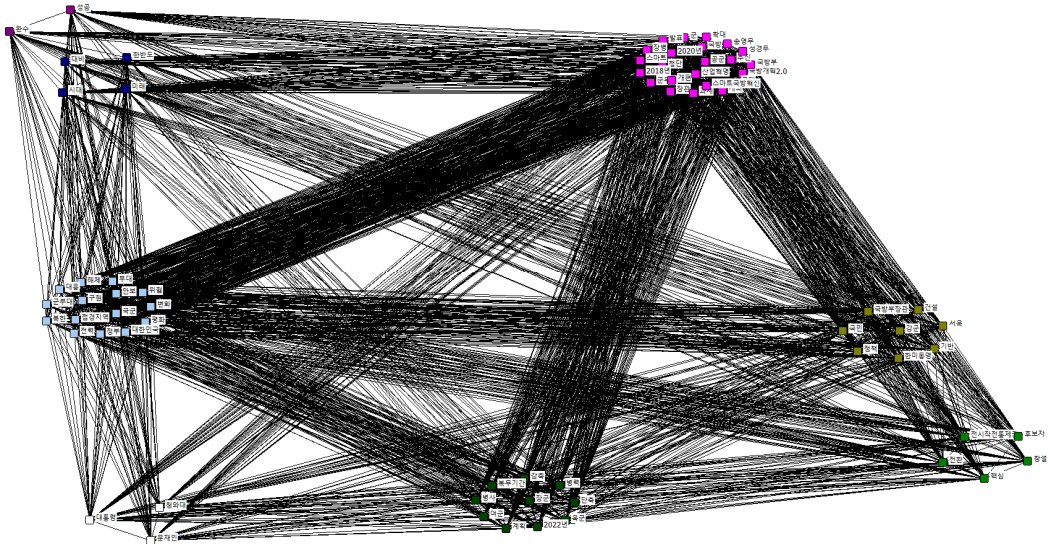
In order to visually confirm the connection relationship within the network, visualization was performed through N-gram network analysis and ego network graph for Defense Reform 2.0. The N-gram network analysis indicates the amount of two words appearing in a sentence together, which means that the higher the frequency, the more the number of times the corresponding two words appear at the same time, and through this, the relationship between words can be identified. The Ego network means that the larger the size and smaller the word, the lower the frequency of appearance, and the closer it is to the central keyword, the higher the frequency of simultaneous appearance (TEXTOM, 2016).

The results of N-gram network analysis in this study are as follows: Wartime Operational Control/Conversion (265), Moon Jae-In/President (181), Moon Jae-in/Government (147), National Defense Reform 2.0/Promotion (145), Ministry of National Defense/Defense Reform 2.0 (139), Service Period/ Shortening (129), Forces/Reduction (93), Strong Army/Construction (82), Korea-US Alliance/Base (78), and Success/Fulfillment (72). Next, Government/Defense Reform 2.0(69), Smart Defense Innovation/Promotional Inspection Meeting(69), Minister of National Defense and Defense(67), Disruption/Promotion(65), Wartime Operation/Conversion(62), Public/Trust(59), Propulsion/People(55), General/Reduction(52), Active/Promotion(51) ranked 10~20th.

4.4 CONCOR Analysis Results

CONCOR (Convergence of iterated correlations) is a typical structural isotope measurement method using correlation, and it is a method using the correlation between actors to derive the relationship pattern between actors (TEXTOM, 2019). Through CONCOR analysis, words with high correlation can be rearranged and grouped, and analyzed. Figure 3 is a visualization of the semantic network of the CONCOR analysis of Defense Reform 2.0, and the details of grouping are shown in Table 5.

Group 1 included 23 words such as expansion, smart defense innovation, air force, military personnel, presentation, reorganization, minister, hosting, soldiers, military, defense, the industrial revolution, Song Young-moo, 2018, smart, advanced, 2020, Jeong Gyeong-du, promotion, task, and defense reform 2.0.



<Figure 1> Visualizing the semantic network of CONCOR analysis of Defense Reform 2.0.

<Table 5> Words included in each group of CONCOR Analysis

Group	Words that are included
1st Group (23 words)	Expansion, Smart Defense Innovation, Air Force, Military Personnel, Presentation, Reorganization, Minister, Hosting, Soldiers, Military, Defense, Industrial Revolution, Song Young-Moo, 2018, Smart, Advanced, 2020, Jeong Gyeong-du, Promotion, Task, Defense Reform 2.0.
2nd Group (15 words)	Dismantling, Security, Response, Realization, Government, Peace, ROK Military, Combat Power, Unit, Threat, Change, North Korea, Military Unit, Republic of Korea, Border Region
3rd Group (10 words)	Troops, Army, Service Period, Shortening, Soldiers, Generals, Female Soldiers, Plans, Reductions, 2022
4th Group (8 words)	Seo Wook, Minister of National Defense, People, Korea-U.S. Alliance, Foundation, Construction, Combat Power, Policy
5th Group (5 words)	Wartime Operational Control, Transition, Core, Creation, Candidate
6th Group (4 words)	Times, Preparedness, Future, Korean Peninsula
7th Group (3 words)	Moon Jae-In, President, Blue House
8th Group (2 words)	Success, Completion

Group 2 included 15 words such as dismantling, security, response, realization, government, peace, ROK military, power, unit, threat, change, North Korea, military unit, Republic of Korea, and the border region, and named "concerns about power reduction of defense reform 2.0."

Group 3 included 10 words for troops, army, service period, shortening, soldiers, generals, female soldiers, plans, reductions, 2022, and designated as 'National Defense Reform 2.0's manpower restructuring'.

Group 4 included eight words: Seo Wook, the Minister of National Defense, the people, the Korea-U.S. alliance, the foundation, construction, power, and policy, and named it "the importance of the Korea-U.S. alliance under Defense Reform 2.0".

Group 5 included five words: wartime operational control, transition, core, creation, and candidate, and were named "the wartime operational control of Defense Reform 2.0".

Group 6 included four words: the times, preparedness, future, and the Korean Peninsula, and named it "the purpose of defense reform 2.0".

Group 7 included three words: Moon Jae In, the president, and the Blue House, and named it "the subject of defense reform 2.0".

Group 8 included two words of success and completion, and was named "will achieve defense reform 2.0".

5. Discussion and Conclusion

5.1 Discussion

This study aims to understand which areas of policy the public is paying attention to and which policies are exposed to the public by using the text mining technique, one of the big data analysis techniques. The results of this study are as follows.

First, in the word frequency analysis conducted with the keyword of Defense Reform 2.0, the Ministry of National Defense (1,308), the government (505), Moon Jae-in (360), the president (350), and the minister (234), which are the parties to the defense reform, are appearing in many cases and are at the top of the list. In addition, the reduction (459) was ranked 6th and the troop (245) was ranked 15th, which showed that much interest was being paid to the reduction of standing forces and the reduction of general capacity in progress as part of the National Defense

Reform 2.0.

In the case of standing troops, the government is gradually reducing the number from 599,000 in 2018 to 500,000 in 2022, and the mobilization reserve army is also trying to reduce the number from 1.3 million to 950,000 in the government term. This reduction of military force seems to be a move to maintain and expand defense capacity by actively coping with the shortage of military service resources due to the demographic cliff. The reduction of troops reduced the number of troops and generals, with the number of army operations from three to two, the Corps from eight to six, the division from 39 to 34, and the number of generals from 436 in 2017 to 360 in 2022 (Korea Ministry of National Defense, 2019).

Also, the shortening (264) and the service period (230) were ranked in the top 20, and the defense reform 2.0 showed that the service period was frequently exposed to the website with the attention of many active soldiers or prospective enlistees who are performing military service duties as the Army was shortened from 21 months to 18 months, the Navy was shortened from 23 months to 20 months, and the Air Force was shortened from 24 months to 22 months (Korea Ministry of National Defense, 2019).

In the connection centrality analysis, the border area (0.015), which was relatively low in the frequency analysis, ranked 19th in the 20th place, and North Korea (0.014), which was 48th, ranked 21st. In defense reform 2.0, which is trying to cope with the diversification of security threats, many factors are still considered for the border area that is in contact with North Korea, which is the main enemy and the biggest threat. In the border area, the connection center is high due to the economic activity of local residents and the problem of regional population decrease as the units that existed in the region are dismantled as part of the reorganization of the unit.

Next, the N-gram network analysis showed that wartime operational control/conversion (264) was the number one, and it was found that there was a great interest in the transition of wartime operational control based on the expansion of the leading defense capability by defense reform 2.0. The transition of wartime operational control is directly related to the ROK-US alliance, which can be regarded as the biggest axis of defense, so it is necessary to be done at the same time, it is confirmed that the defense reform 2.0 has been shown on the website about what and how much contribution it can make to the transition of wartime operational control.

Finally, as a result of clustering words through CONCOR analysis, 8 groups were formed: **the 1st** group was 'pursuing the technical army of defense reform 2.0', 2nd group was 'concerning the combat power reduction of defense reform 2.0', 3rd group was 'reorganizing the manpower

structure of defense reform 2.0', 4th group was 'importance of ROK-US alliance under defense reform 2.0', 5th group was 'conversion of wartime operational right of defense reform 2.0', 6th group was 'purpose of defense reform 2.0', 7th group was named as 'subject of defense reform 2.0', and 8th group was named as 'will to achieve defense reform 2.0'.

In terms of discussion by the group, the first group, 'Searching for the Technical Group of Defense Reform 2.0', includes words such as Smart Defense Innovation, Industrial Revolution, Smart, and Advanced Technology. Through this, it can be seen that the direction of Defense Reform 2.0 is to build a high-tech army using core technologies of the era of the Fourth Industrial Revolution. It is considered that it was used in various policy fields within the National Defense Reform 2.0 because it is the most convincing method to maintain or increase military power in the reduction of troops.

The 2nd group, 'Concerns about the reduction of combat power in Defense Reform 2.0', includes words such as troops, dismantling, combat power, response, change, North Korea, border area, and security, and it is thought that there is concern about whether the reduction of the military scale conducted through Defense Reform 2.0 will lead to the weakening of power. Even in the changing security environment, North Korea is still the biggest threat that exists, and the dismantling of troops in the border area can lead to the weakening of the military's immediate response capability, which can cause anxiety about security, and there are people who think that the defense reform 2.0 can lead to the weakening of combat power through policies such as manpower reduction.

The 3rd group, the Reorganization of the Human Resources Structure of Defense Reform 2.0, includes words such as service period, shortening, general, reduction, female army, military force, soldier, and plan, which can be judged to be of great interest in the military's human resource structure. In the case of shortening the service period, it has attracted great interest from prospective enlistees, and it seems that the gradual expansion of the female army personnel has attracted attention because it is linked to the issue of gender equality, which has recently become an issue in various fields. In addition, it is believed that the bold reduction of the general personnel has caused a lot of interest in the Internet.

The 4th group, the importance of the ROK-US alliance under the Defense Reform 2.0, includes the ROK-US alliance, the foundation, the strong forces, construction, policy, Seo-wook, the Minister of National Defense, and the people. It is judged that the Defense Minister Seo-wook's remark in the dialogue with the United States that the construction of the

strong forces that the Defense Reform 2.0 ultimately aims to achieve should be based on the ROK-US alliance, which plays a pivotal role in defense.

In the 5th group, 'The Transition of Wartime Operational Control of Defense Reform 2.0', it can be seen that it is one of the ultimate goals to achieve the stable transition of wartime operational control through Defense Reform 2.0, and in the 6th group 'The Purpose of Defense Reform 2.0', it can be seen that Defense Reform 2.0 has the purpose of preparing for the future within the Korean Peninsula. The 7th group was formed into the same group as the current regime that is promoting defense reform 2.0 as the subject of defense reform 2.0, and there were many comments that promised or prayed for the success, completion of defense reform 2.0 through the 8th group 'will to achieve defense reform 2.0'.

5.2 Conclusion

This study was conducted with the purpose of grasping the public's interest in the policy after the announcement of the National Defense Reform 2.0 and grasping the factors to be considered for the successful completion of the policy in the future.

Through word frequency analysis, connection centrality analysis, and an N-gram analysis, we identified issues that received much attention such as troop reduction, shortening of service period, dismantling of the border area unit, and returning wartime operational control. As a result of clustering words through CONCOR analysis, we were able to confirm that there was a great interest in pursuing the technical group, concerns about combat power reduction, and restructuring of manpower structure. The results of this study suggest that the implementation of the National Defense Reform 2.0 is as follows.

First, there is a lack of awareness of alternative measures to meet the reduced troops, while receiving much attention to the reduction of troops in Defense Reform 2.0. It is believed that the reduction of troops may have been an inevitable choice due to the cliff of the population, but at the same time, the people recognize that the reduction of troops promoted by the National Defense Reform 2.0 is voluntary and active. However, if the plan for improving defense power is not fully known, there is a concern that this choice will lead to weakening defense capabilities. And if this does not lead to the support of the people, it will be hampered by the promotion of defense reform 2.0, which has to put a huge budget. Therefore, it is necessary to draw support for policy by emphasizing the areas linked to strengthening defense

power among various fields of defense reform 2.0. Second, it is necessary to actively talk about the problem of the decrease of the population of the region and the collapse of the local commercial area due to the dismantling and movement of the border area units. The dismantling and relocation of the units determined by the Defense Reform 2.0 can be a direct burden for the people who work there and the natural commercial area where the military unit is long-established. Therefore, it is necessary to minimize the decrease in public trust due to the implementation of policies through sufficient dialogue with the local community. Third, it is necessary to show substantial results through the military structure field, which is attracting much attention, the barracks culture field, which is shown to receive less attention than the defense operation field, and the publicity and active policy introduction of the defense industry field. In particular, in the case of barracks culture, if the right barracks culture is established through the improvement of the past human rights violations and poor service conditions that greatly reduced the trust of the people, the pride of the soldiers themselves, and the reliability of the military organization from the outside are improved inside, it will serve as a solid support base for various policies of the military as well as the successful completion of defense reform 2.0. As such, we could understand what areas are most interested in defense reform 2.0, which is a representative national defense policy of the current regime, by using the text mining technique, one of the big data analysis techniques, and suggested a plan to draw public support for the policy. And this study suggested a method of grasping public perception with much less cost and time than the existing methods of grasping public opinion on policies such as face-to-face, telephone interview, and ARS survey. Also, information that can be extracted through a large number of questions in the case of existing methods could be derived using various analysis methods based on collected data. Furthermore, this study can be a starting point for comparing changes in perceptions of major defense policies of each government through big data analysis. However, there were some limitations that were not solved in the analysis process of data. First, in the process of data purification, it was made a lot of effort for the researcher manually integrate and delete words that are not related to similar words, but there is still a possibility that objectivity was violated as the subjectivity of the researcher intervened in this process. In addition, although the data was collected through various channels to secure objectivity, the basic statistical data on the authors did not exist due to the nature of the limitations of most big data analysis studies, so this opinion could not be determined as 'perception' representing public opinion but be determined as 'interest' from

people. Therefore, it is expected that more reliable research will be possible if this study is utilized in connection with the results of actual surveys for the public based on basic statistical data in the future. Also, based on the results of the survey on the experience of barracks culture innovation, the results of the survey by actual soldiers, the matrix of barracks innovation exposed to media, the analysis of emotions and time series, etc., will be compared to feel the difference between the atmosphere felt in the field and the atmosphere formed in the media. This study lacks due to the limitation of the amount of data available to the researcher and the limitation of the author's analytical ability, but it is expected that more meaningful research will be carried out if the research is conducted to comprehensively understand the interest of the people who are more actively exposing their feelings to the media by expanding the amount of analysis data and extending the search period.

Reference

- [1] Cho Young Gap (2009) "Promotion Strategy for Securing the Base for Army Policy Support" Seoul: Ministry of National Defense.
- [2] Choi Byungwook (2019). Promotion of Defense Reform, What should we do?: Cases and lessons of reform of the US Army in the post-Cold War era. *Defense Policy Research*, 35(2), 121-143.
- [3] Defense Ministry of National Defense (2019), *Defense Reform 2.0*.
- [4] Dijcks J. P. (2013) *Oracle: Big Data for the Enterprise*, California, CA: Oracle Corprtation. 3-6
- [5] Douglas, L. (2012). The Importance of "Big Data": A Definition. Gartner. Online: <http://www.gartner.com/resId>.
- [6] Editorial Department (2018). Defense Department, basic direction for Defense Reform 2.0. *Defense and technology*, 474, 8-11.
- [7] Han Chan-ok, Kang Won-seok (2016). Research on the field of big data utilization in the county, *Business Management Research*, 9(1).
- [8] Jeong Jeong-gyun, Park Chul, and Park Sang-jae. (2018). A Review on the Reform of the Military Judicial System: Focusing on Defense Reform 2.0. *Journal of Convergence Security*, 18(2), 113-121.
- [9] Kang Yong-gu, Kim Tae-sung (2019). The evaluation of the promotion of the Korean military reserves and future-oriented policy direction: Focusing on the reserve power policy of Defense Reform 2.0. *Journal of North-East Asian Studies*, 24(3), 31-56.
- [10] Kim So-ra (2020). A fusion study on the recognition of anorak fashion using big data analysis. *The Korean Society of Science and Arts*, 38(4), 43-55.
- [11] Kim Sung-yeon and Jeong Hyo-myeong (2014). The effect of policy preference on party support. *Korean Political Science Review*, 48(1), 119-140.
- [12] Kim Sung-yeon (2017). The Effect of Policy Voting in the 20th General Election: Korean Party Studies Review, 16(3), 47-78.
- [13] Kim Yeon-jin (2013). Enterprise Innovation through the introduction of a big data-based advanced analysis system, *PRACTICE & Small Business Forum*, p. 44.
- [14] Kim Yong-sang (2017). The direction of barracks culture innovation in the era of the

- Fourth Industrial Revolution. *Journal of Spiritual & Mental Force Enhancement*, 48, p.307-371.
- [15] Kim Ji-sook (2013). "Consideration of Big Data Utilization and Analysis Techniques", Korea University Graduate School, p. 6.
- [16] Kim Taehyun and Kim Seokwoo (2019). Analysis of the characteristic elements of the use of big data in social perception of the introduction of the high school credit system. *Education Innovation Research*, 29(2), 225-242.
- [17] Lee Jung Jin. (2018). Effect of the president's support for the president, ruling party, and opposition parties on the president's policy promotion. *21st Century Political Science Review*, 28(2) and 121-143.
- [18] McKinsey Global Institute(2011), Big data: The next frontier for innovation, competition, and productivity.
- [19] Ministry of National Defense (2019), Defense Reform 2.0.
- [20] Noh Hoon. (2013). Diagnosis of the "Basic Plan for Defense Reform 2012-2030" and future defense reform strategies." *Strategic Research* 57, p.111.
- [21] Sung Kwang Sook. (2020). Big data analysis of social media on Generation Z fashion. *Journal of the Korea Fashion & Costume Design*, 22(3), 49-61.
- [22] Sunhee Lee, Pomegranate. (2020). Public opinion analysis on yoga through big data text mining analysis. *Martial Arts Research*, 14 (1), 223-238.
- [23] TEXTOM. (2019). CONCOR Analysis using UCINET.
<http://textom.co.kr/home/sub/view.php?id=report&no=23>
- [24] TEXTOM. (2016). Temstorm Update 1 - Improve and add Textorm visualization capabilities (page menu English, Word Tree, Ego Network). Use big data paper TEXTOM easily and quickly. Accessed December 3, 2020, <https://m.blog.naver.com/PostView.nhn?blogId=textom&logNo=220879473633&proxyReferer=https:%2F%2Fwww.google.com%2F>
- [25] Won Hong-gyu (2016). Recognition of military service and barracks life of new generation soldiers. *Military Forum*, 86, p.101-129.
- [26] Yoon Jin Young, Lee Se Young (2019). A review of innovative improvement measures for elite reserve power in connection with Defense Reform 2.0. *Korean Military Research Discussion*, 8(2), 77-105.
- [27] Yoon Jong-sung (2015). A Study on the Formation of Defense Culture in Defense Reform. *Strategic Research*, 22(2), p.105-138.

About the Authors

Kim, Tae Kyoung(1st author), Tae Kyoung Kim is working as a platoon leader in the 2nd Ammunition Depot Security. His research interests include big data analysis, military organizational analysis, and development plans.

Kang, Wonseok(corresponding author), Wonseok Kang is working as a professor at the Department of Psychology & Management of Korea Military Academy. He received his Ph.D. from the Samuel Ginn College of Engineering at Auburn University. His current research interests include risk management, big data analysis, military service system, R&D investment strategy and etc.