

## 빅데이터 기반 함정 시운전 종목명 분석

이형신\*, 서형필, 백용관, 이상일  
국방기술품질원

### Analysis of Sea Trial's Title for Naval Ships Based on Big Data

Hyeong-Sin Lee\*, Hyeong-Pil Seo, Yong-Kawn Beak, Sang-Il Lee  
Defence Agency for Technology and Quality(DTaQ)

**요약** 본 연구에서는 효율적인 함정 시운전을 위하여 빅데이터 기법인 워드 클라우드를 활용하여 한미 해군의 시운전 목적과 주안점을 다각적으로 파악하여 다음과 같은 결과를 도출하였다. 첫째, 한미 해군 시운전 종목을 키워드 클렌징을 통해 추출된 단어를 비교한 결과 한국 해군은 시운전을 단일 장비에 대한 시험의 개념으로 수행하며, 미 해군은 시스템에 초점을 둔 통합 시운전을 진행한다는 것을 알 수 있었다. 둘째, 한미 해군 시운전 연관도 분석 결과 약 66.6%가 유사한 항목으로 분석되었으며, 그중 2종목 이상 중복된 종목이 112종목이었다. 한국 해군 시운전 종목 252종목 대비 44%가 중복된 종목으로 미 해군 시운전 종목으로 통합시 89종목(전체 35%)이 축소 가능하다고 분석되었다. 함정은 여러 장비가 동시 다발적으로 작동하는 복합 시스템이다. 현재 한국 해군 시운전과 같이 개별 장비의 기능, 성능 확인에 중점을 두고 수행하는 것은 시운전 대상이 지나치게 많아져 시운전 기간이 증가하게 된다. 또한, 그로 인한 일정 및 평가비용 증가로 필요한 예산이 필연적으로 증가한다. 향후 미 해군의 시운전과 같이 통합 시스템적인 평가를 통한 효율적이면서 정확한 시운전을 위하여 추가적인 연구가 필요하다고 판단된다.

**Abstract** The purpose and main points of the ROK-US Navy were analyzed from various angles using the big data technology Word Cloud for efficient sea trials. First, a comparison of words extracted through keyword cleansing in the ROK-US Navy sea trial showed that the ROK Navy conducted a single equipment test, and the US Navy conducted an integrated test run focusing on the system. Second, an analysis of the ROK-US Navy sea trials showed that approximately 66.6% were analyzed as similar items, of which more than two items were 112 items. Approximately 44% of the 252 items of the ROK Navy sea trials overlapped, and that 89 items (35% of the total) could be reduced when integrated into the US Navy sea trials. A ship is a complex system in which multiple equipment operates simultaneously. The focus on checking the functions and performance of individual equipment, such as the ROK Navy's sea trials, will increase the sea trial period because of the excessive number of sea trial targets. In addition, the budget required will inevitably increase due to an increase in schedule and evaluation costs. In the future, further research will be needed to achieve more efficient and accurate sea trials through integrated system evaluations, such as the U.S. Navy sea trials.

**Keywords** : Sea Trial, Naval Ship, Word Cloud, Big Data, Inspection

---

\*Corresponding Author : Hyeong-Sin Lee(Defence Agency for Technology and Quality)

email: hslee2@dtaq.re.kr

Received July 14, 2020

Accepted November 6, 2020

Revised September 14, 2020

Published November 30, 2020

## 1. 서론

한국은 1977년 국산 호위함 설계, 건조를 시작으로 자국 조선소에서 군함을 건조하여 획득을 진행해오고 있으며, 한국전쟁 후 미국의 공여로 미 해군의 퇴역함정을 다수 인수하여 1999년도까지 기어링급 구축함을 운용하는 등 획득 절차에 대한 뿌리도 미 해군에 있다고 할 수 있다. 국내에서 획득되는 모든 해군함정은 미 해군과 마찬가지로 시운전(Sea Trial)을 거쳐 정부품질관리활동을 통해 해군에 인도된다.

한국 해군함정은 최근 건조된 울산급 호위함, 차기상륙함 등에서 조선소에서 해군으로의 인도 지연 사례의 상당 부분이 시운전 기간 중 발생함에 따라, 획득의 최종 단계이며 전체 함 건조 기간의 30%를 차지하는 시운전에 대한 분석이 필요함을 군, 관, 연, 산 모두 공통적으로 인식하였다. 특히 한-미 해군 함정 시운전 목록 비교를 통해 한국 해군 함정 시운전 제도를 진단할 수 있는 척도가 마련될 수 있을 것이다.

최근 연구의 동향을 살펴보면 빅데이터 기법을 활용하여 국민들의 여론과 동향을 파악하여 정책 개발과 예측에 활용하는 연구들이 늘고 있다. 본 연구에서는 한미 해군 함정의 시운전 종목명을 빅데이터 기법인 워드 클라우드를 활용하여 시운전의 목적과 주안점을 다각적으로 파악하고자 하였다.

## 2. 본론

### 2.1 이론적 배경

최근의 정보기술 진화로 데이터의 기하급수적인 증가를 경험하고 있으며, 그 방대한 데이터를 통해 현대사회가 가지고 있는 다양한 질문들에 대한 나은 답을 찾을 것이라 기대되고 있어, 빅데이터 인기 또한 증가하고 있다. 그로 인한 다양한 종류의 빅데이터를 분석하여 활용하려는 노력이 많은 분야에서 이루어지고 있다.

빅데이터는 정형과 비정형으로 분류될 수 있는데, 데이터는 형태에 따라 고정된 구조 형태로 구성된 데이터를 정형데이터(structured data)로, 정해진 구조가 없을 때는 비정형 데이터(unstructured data)로 구분한다. 정형 데이터를 이용한 마이닝을 데이터 마이닝(data mining), 비정형 데이터를 이용한 마이닝을 텍스트 마이닝(text mining)이라고 한다. 우리가 정보전달의 수단으로 가장 많이 쓰는 텍스트는 비정형 데이터에 속한다.

마이닝이란 데이터로부터 통계적인 의미가 있는 개념이나 특성을 추출하고 데이터 간의 패턴이나 추세 등의 고품질의 정보를 끌어내는 과정이다.

텍스트 분석은 Fig. 1과 같이 일반적으로 문서 수집, 파싱(Parsing) 및 필터링(Filtering), 구조화, 빈도 분석 및 유사도 분석의 순서로 수행된다. 구조화된 텍스트 문서는 분석 목적에 따라 다양한 형태로 활용될 수 있으며,

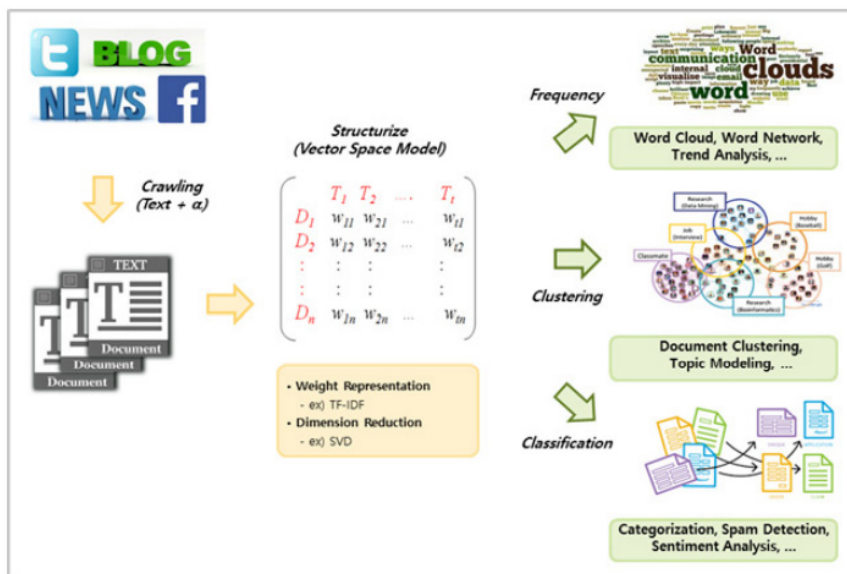


Fig. 1. Overview of text analytics related techniques and applications

대부분 전통적 데이터 마이닝의 대표적 기법인 빈도 분석(Frequency Analysis), 군집화(Clustering), 그리고 분류(Classification)의 주요 개념을 활용한다.

구체적으로 빈도 분석을 활용한 텍스트 마이닝의 대표적인 예로는 특정 문서 또는 문서 집합의 용어를 빈도에 따라 상이한 크기로 도식화하여 나타내는 워드 클라우드 분석(Word Cloud) 분석, 용어 간 동시출현 관계를 도식화하여 나타내는 용어 망(Word Network) 분석, 그리고 빈출 어휘의 기간별 추이를 나타내는 추세(Trend) 분석 등이 있다. [1]

## 2.2 분석 대상

분석 대상은 한미 해군 합정 시운전 목록이다. 미국 해군 시운전 목록 표준은 INSURVINST 4730.1E[2]의 Appendix A "Trial or Inspection Schedule"에 명시되어 있다. 시운전은 한국과 마찬가지로 건조자시운전과 인수시운전으로 구분되며 미국은 건조자시운전과 인수시운전 목록이 같다. [2]는 2005년도에 개정된 표준이나, F 버전(2008년)에는 시운전 목록이 없고, 최근 개정판인 G 버전(2013년)은 비공개 자료이므로 [2]를 사용하였다. 한국은 표준안 부재로 OO급 호위함 사례를 바탕으로 분석하였으며 건조자시운전과 인수시운전 목록이 약간 상이하여 인수시운전 기본으로 비교하였다.

## 2.3 자료수집 방법

데이터는 한미 해군 합정 시운전 목록을 분석 대상으로 하고, 한국 해군 합정 시운전 목록은 영어로 번역하여 목록화를 병행하였다. 일반적으로 인터넷 기반으로 수행하는 빅데이터 수집보다는 직접 확인할 수 있는 정제된 데이터 확보가 가능함에 따라 조사 목적에 부합하지 않는 키워드에 대한 삭제과정 등 전처리가 필요하지 않았다. 수집된 자료는 Table 1과 같이 전체 1,407개 단어 중 84% 유효하여 1,190 단어를 분석에 사용하였으며, 한국 해군 합정 시운전 목록은 전체 1,525개 단어 중 85%가 유효하여 1,299개 단어를 분석에 사용하였다.

Table 1. Data for Word cloud Analysis

	Total word	Deleted word	Effective word	Effective ratio
ROKN AT List	1,525	226	1,299	0.85
USN AT List	1,407	217	1,190	0.84

## 2.3 분석 절차 및 방법

번역 등을 거쳐 최종적으로 수집된 데이터를 대상으로 텍스트 마이닝 기법의 하나인 워드 클라우드 분석을 활용하였다. 워드 클라우드는 텍스트 데이터를 간결히 하여 시각적으로 나타내는 시각화 기법으로 다양한 문맥에서 가장 빈번하게 나타난 단어들을 추출한다. 출현 빈도가 높으면 크게, 낮으면 작게 나타나는 특징이 있다. 이는 텍스트 속에서 키워드를 통해 중요한 내용을 직관적으로 알 수 있게 만들어주어 신속하지만 비교적 정확한 정보를 제공해주는 장점을 가지고 있다.

워드 클라우드의 생성은 R 프로그램(Version 4.0.2)을 사용하여 분석하였다. R 프로그램에서 텍스트마이닝의 "tm", 텍스트 스템밍(stemming)의 "SnowballC", 워드 클라우드를 만들어주는 "wordcloud", 그리고 워드 클라우드의 색상을 정리하는 "RColor Brewer" 패키지를 사용하여 최종 워드 클라우드를 생성하였다.

본 연구에서는 분석의 유의미성을 고려하여 화제어 분석 키워드 빈도수를 수집하였으며 20순위까지의 분석 결과는 Table 2, Table 3과 같았다.

한미 해군 시운전 목록 빈도수 분석 결과 미 해군 시운전 목록에서는 System, Test, Inspect 순으로 분석되었으며, 한국 해군 시운전 목록에서는 Test, Operation, Inspection 순으로 분석되었다. 분석된 키워드 간의 관계를 조합하며 한미 해군 시운전 주안점을 분석하였다.

첫 번째로 가장 많이 반복된 1, 2위 키워드를 조합해 보면 미 해군은 시스템 검사(System Test), 한국 해군은 작동 검사(Operation Test)로 미 해군의 경우 시운전을 합정의 시험장비가 운용될 때 그와 연관된 장비가 정상적으로 작동하는지 연계하여 평가하는 반면 한국 해군은 시험장비만 정상적으로 작동하는지에 대하여 시운전을 수행한다는 것을 알 수 있다.

두 번째로 한국 해군 시운전 목록에는 Table 4와 같이 단일 장비 명칭이 많이 분포해있는 것이 파악되었는데 이것은 첫 번째 분석결과와 연계하여 분석해보면 레이더 성능 시험(Radar Operation Test) 15번, 발전기 성능 시험(Generator Operation Test) 7번 조합이 가능하며 중간 키워드를 수신(Receiver), 최대(Max)등과 같이 변경하여 시운전 수행을 가정한다면 직·간접 연관된 평가가 여러번 반복하여 수행 가능한 것으로 분석할 수 있다.

Table 2. Core keyword and Frequency analysis for USN

Rank	Keyword	N	Rank	Keyword	N
1	System	51	11	GCCS-M	11
2	Test	47	12	Space	10
3	Inspect	39	13	Alarm	10
4	Inspection	38	14	Fire	9
5	Demonstration	36	15	Deck	8
6	Check	24	16	Demo	8
7	Equipment	24	17	INSURV	8
8	Ship	16	18	Power	8
9	Station	13	19	Valve	8
10	Operate	12	20	Air	7

Table 3. Core keyword and Frequency analysis for ROKN

Rank	Keyword	N	Rank	Keyword	N
1	Test	159	11	Radar	15
2	Operation	110	12	Sensitivity	11
3	Inspection	47	13	Receiver	10
4	Function	34	14	GFE	10
5	Equipment	39	15	Room	10
6	System	38	16	Transmitter	8
7	Measure	22	17	Ammo	8
8	Ship	18	18	Fire	8
9	Operate	17	19	Generator	7
10	Component	15	20	Pump	6

Table 4. Ratio of equipment to core keyword on the ROKN-USN AT List

Item	Total	Similar item	Other
Number	252	168	84
Rate(%)	100	66.6	33.3

2.4 워드 클라우드 분석 결과

Fig. 2는 INSURVINST 4730.1E[2]의 Appendix A “Trial or Inspection Schedule”에 명시된 미국 해군의 시운전 종목명으로 워드 클라우드 분석을 한 것이다. 위의 워드 클라우드를 분석할 결과는 다음과 같다. 시운전 수행 개념에서 본다면 미해군 시운전은 (a)에서 보이는 바와 같이 검사와 시연, 시험과 확인의 개념으로 시운전을 수행한다는 점을 볼 수 있다. 한국 해군의 경우 (b)에서 보이는 바와 같이 시운전 제목부터 OO계통내 작동시험, 성능시험 등의 제목이 대다수이며, 검사(Inspection)와 시연(Demonstration), 확인(Check)의 개념보다는 시험(Test)의 개념을 강조한다고 볼 수 있다. 또한 (a)를 보면 System과 Inspection 이란 문자가 크게 보이는 것을 볼 수 있다. 이는 단일 장비에 대한 시험의 개념으로 수행하는 한국 시운전 보다는 시스템에 초점을 두고 검사의 개념으로 시운전에 접근한다는 것을 알 수 있다.

시운전의 내용적인 측면에서 접근해볼 때 미해군 시운전은 갑판, 비행갑판, 도어 밸브, 격납고, 항공기 축전지 등 주 선체와 전력 공급 계통, 파이프로 이송되는 유체를 제어하는 밸브, 항공자산에 대한 지원설비 등을 중요시하는 것으로 볼 수 있다. 반면 한국 해군 시운전은 미 해군과 같은 선체에 대한 시운전 보다는 개별 장비의 기능,

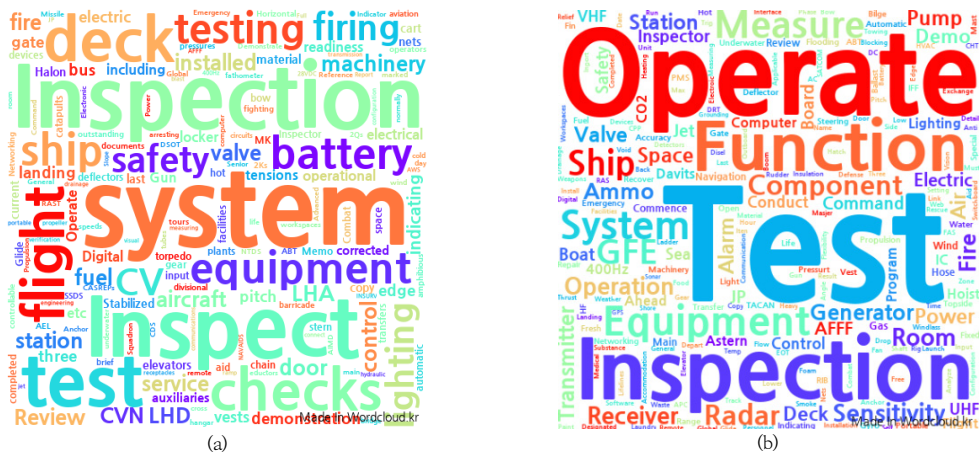


Fig. 2. Word Cloud result for Sea Trial’s (a)USN (b)ROKN

성능 확인에 중점을 두고 수행하고 있다.

### 2.5 한미 해군 시운전 목록 연관도 분석 결과

OO급 Batch-I 호위함을 기준으로 미 해군 함정 시운전 표준 종목을 Table 8과 같이 비교해보았다. OO급 Batch-I 호위함 기준 총 252종목 중 168종목인 66.6%가 유사하고 84종목인 33.3%가 유사하지 않은 것으로 분석되었다.

Table 5. Similarity between USN and ROKN AT List

Item	Total	Similar item	Other
Number	252	168	84
Rate(%)	100	66.6	33.3

다음 Table 6과 Fig. 3은 한미 해군 분야별 시운전 종목 수를 비교해보았다. 기본성능(GP), 주추진(MP), 전기(EP), 전자통신(EL), 무장(WP), 전투체계(CC) 분야는 한국 해군이 미 해군보다 세분화하여 시운전을 수행하는 것으로 나타났다. 특히, 전투체계(CC) 분야는 미 해군이 10종목 한국 해군이 47종목으로 한국 해군이 지나치게 세분화하여 평가하고 있는 것을 보여주고 있다. 보기(AU)와 의장(HF)은 미 해군이 더 세분화하여 평가하는 것으로 나타났다.

Table 6. Comparison between USN and ROKN AT List Base on SWBS(Ship Work Breakdown System)

Item	Total	GP 000	HS 100	MP 200	EP 300	EL 400	AU 500	HF 600	WP 700	CC 700	ETC
USN	176	8	-	7	21	30	41	45	8	10	6
Rate (%)	100	4.5	0	3.9	11.9	17.0	23.2	25.5	4.5	5.6	3.4
ROK N	252	16	1	23	37	51	21	32	24	47	0
Rate (%)	100	6.3	0.3	9.1	14.6	20.2	8.3	12.6	9.5	18.6	0

Table 7은 한국 해군과 유사종목 중 중복되는 시운전 종목을 분석해보았다. 2종목 이상 중복된 종목이 112종목이었다. 한국 해군 시운전 종목 252종목 대비 44%가 중복된 종목으로 세부적인 추가검토는 필요겠지만 해당

되는 미 해군 종목인 23종목으로 통합시 최대 89종목(전체 35%)이 축소가 가능하다. 특히 5종목 이상 중복되는 종목은 사통, 전투체계, 무장, 센서류, 취사장치, 탄약고 설치성, 살수장치, 발전기에 대한 종목은 한국 해군이 미 해군 대비 지나치게 세분화된 종목으로 시운전을 수행하여 대폭 통합이 필요하다.

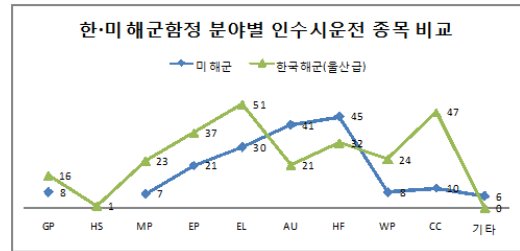


Fig. 3. Comparison between USN and ROKN AT List Based on SWBS(Ship Work Breakdown System)

Table 7. Overlapping Sea Trials between USN and ROKN AT List

Item	USN	ROKN
Search and fire control radar, EW, IFF, and TACAN parameter checks	1	13
HF/UHF/VHF communication equipment checks and demonstrations, ETC	2	16
Propulsion and electric plants hot/cold checks and testing of safety devices	1	7
Weapons handling equipment/systems	1	6
Magazine sprinkler and alarm systems test, ETC	4	20
Test controllable pitch propeller, ETC	8	32
Inspect OWS/OCM Systems, ETC	6	18
Total	23	112

### 3. 결론

본 연구의 목적은 빅데이터 기법을 활용하여 한미 해군의 시운전 종목명을 분석하여 시운전 목적과 주안점을 다각적으로 파악한 결과 다음과 같은 연구결과를 도출하였다.

첫째, 미 해군 시운전 종목 키워드 클렌징을 통해 1,190개의 키워드가 추출되었고 그중 최소 5회 이상의

빈도가 나타난 키워드는 45개로 주요 키워드는 System, Test, Inspect 등이 있었다. 이는 검사와 시연, 시험과 확인의 개념으로 시운전을 수행한다는 점을 볼 수 있으며, 단일 장비에 대한 시험의 개념으로 수행하는 한국 시운전보다 시스템에 초점을 두고 시운전을 진행한다는 것을 알 수 있다.

둘째, 한미 해군 시운전 목록 연관도 분석 결과 약 66.6%가 유사한 항목으로 분석되었으며 그중 2종목 이상 중복된 종목이 112종목이었으며, 한국 해군 시운전 종목 252종목 대비 44%가 중복된 종목으로 미 해군 시운전 종목으로 통합시 89종목(전체 35%)이 축소 가능하다는 것이 분석되었다.

함정사업은 복합 시스템으로 평가할 항목이 많고, 여러 개의 시운전이 동시다발적으로 수행되고 있어, 적기 전력화를 위해서는 시운전의 효율화가 필수적이다. 현재 한국 해군 함정 시운전과 같이 개별 장비의 기능, 성능 확인에 중점을 두고 수행하는 것은 시운전 대상 항목이 지나치게 많아 일정 및 평가 비용 증가로 필요한 예산이 필연적으로 증가하는 상황이다.[3] 향후 시운전의 통합 시스템적 평가 및 중복된 종목의 간소화를 통한 효율적이면서도 정확한 시운전에 대한 연구가 추가로 필요하다고 판단된다.

## References

- [1] Namgyu Kim, "Investigations on Techniques and Applications of Text Analytics" *The Journal of Korean Institute of Communication and Information Sciences*, Vol.42, No.2, pp.471-492, 2017.  
DOI: <http://dx.doi.org/10.7840/kics.2017.42.2.471>
- [2] Board of Inspection and Survey Instruction[Internet], INSURVINST, c2005 [cited 2005 APR 15], Available From:  
<http://www.navybmr.com/study%20material/INSURVINST%204730.1E.pdf> (accessed March. 15. 2020)
- [3] Jaeyong Kim, "A Study on Efficient Test & Evaluation Methods in Naval Ship Acquisitions" *Journal of the KIMST*, Vol.19, No.6, pp.703-711, 2016.  
DOI: <http://dx.doi.org/10.9766/KIMST.2016.19.6.703>
- [4] JongHa Kim, "A Plan for Improving the Naval Ship Acquisition Program Focused on the R&D Procedures", *Journal of the Korean Association of Defense Industry Studies*, Vol.18, No.2, pp.161-184, 2011.  
<https://www.earticle.net/Article/A163498>
- [5] JongHa Kim, "A Plan for Building up a Collaborative System between Navy and DAPA for Promoting Efficient Naval Ship Acquisition Program", *Korea Open Access Journals*, Vol.25, No.4, pp.187-213, 2010.  
DOI: <http://dx.doi.org/10.22883/jdps.2010.25.4.006>
- [6] S. S. Jeon. *A study on the process of weapon system management and developmental method*, Master's thesis, Korea University, Korea, pp.55-57, 2010.
- [7] J. S. Lee, *A Case Study on Application Methods and Effects based on the Type of Big Data at Corporations*, Ph.D dissertation, Hanshin University, Korea, pp.84-85, 2014.
- [8] Duckyoung Jeong, "A Technology Trend Analysis using Wordcloud" *Korean Institute of Intelligent Systems*, Vol.26, No.1, pp.17-18, 2016.  
<http://www.riss.kr/link?id=A101853005>
- [9] Myoung-gyun Lee, "Utilizing SNS data based on Opinion Mining using Big Data" *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, Vol.8, No.3, pp.859-877, 2018.  
DOI: <http://dx.doi.org/10.21742/AJMAHS.2018.03.11>
- [10] Hwang, Hae Jeong, "Exploration of User Experience Research Method with Big Data Analysis : Focusing on the Online Review Analysis of Echo" *The Journal of the Korea Contents Association*, Vol.16, No.8, pp.517-528, 2016.  
DOI: <https://doi.org/10.5392/JKCA.2016.16.08.517>
- [11] Joo-Seok Park , "A Comparative Study of Big Data, Open Data, and My Data" *The Korea Journal of Bigdata*, Vol.3, No.1, pp.41-46, 2018.  
DOI: <http://doi.org/10.36498/kbigdt.2018.3.1.41>

이 형 신(Hyeong-Sin Lee)

[정회원]



- 2010년 2월 : 홍익대학교 조선해양공학과 (공학사)
- 2015년 2월 ~ 현재 : 국방기술품질원 연구원

<관심분야>

조선공학, 품질관리

서 형 필(Hyeong-Pil Seo)

[정회원]



- 2003년 2월 : 해군사관학교 전자공학과 (공학사)
- 2013년 2월 : 연세대학교 일반대학원 전기전자공학부 (공학석사)
- 2014년 : 해군사관학교 무기체계공학과 조교수
- 2015년 9월 ~ 현재 : 국방기술품질원 선임연구원

<관심분야>

RCS 감소, 위상배열, 레이더

---

백 용 관(Yong-Kawn Beak)

[정회원]



- 2008년 3월 : 오사카대학 시스템공학과 (공학사)
- 2010년 3월 : 오사카대학 선박해양공학과 (공학석사)
- 2012년 7월 ~ 현재 : 국방기술품질원 선임연구원

<관심분야>

선체구조, 위험식별, 표준화

---

이 상 일(Sang-Il Lee)

[정회원]



- 1995년 2월 : 울산대학교 조선공학과
- 1995년 3월 ~ 2002년 6월 : 해군본부(조함단)
- 2002년 9월 ~ 2007년 9월 : 선박안전기술공단(現해양교통안전공단)

- 2007년 10월 ~ 현재 : 국방기술품질원 선임연구원

<관심분야>

선형설계, 구조설계, 자율운항