

## A Study on the Finding of Promising Export Items in Defense industry for Export Market Expansion-Focusing on Text Mining Analysis-

Seoyoon Yeo\*, Jong Hee Jeong\*\*, Seong Ho Kim\*\*

\*Researcher, Korea Researcher Institute for Defense Technology Planning and Advancement, Jinju, Korea

\*\*Lecturer, Department of Smart Distribution and Logistics, Gyeongsang National University, Jinju, Korea

\*\*Professor, Department of Smart Distribution and Logistics, Gyeongsang National University, Jinju, Korea

### [Abstract]

This paper aims to find promising export items for market expansion of defense export items. Germany, the UK, and France were selected as export target countries to obtain unstructured forecast data on weapons system acquisition plans for the next ten years by each country. Using the TF-IDF in text mining analysis, keywords that appeared frequently in data from three countries were derived. As a result of this paper, keywords for each country's major acquisition projects drawing. However, most of the derived keywords were related to mainstay weapon systems produced by domestic defense companies in each country. To discover promising export items from text mining, we proposed that the drawn keywords are distinguished as similar weapon systems. In addition, we assort the weapon systems that the three countries will get a plan to acquire commonly. As a result of this paper, it can be seen that the current promising export item is a weapon system related to the information system. Prioritizing overseas demands using key words can set clear market entry goals. In the case of domestic companies based on needs, it is possible to establish a specific entry strategy. Relevant organizations also can provide customized marketing support.

▶ **Key words:** Defense industry, Export market expansion, Promising export items, Text mining, Weapon systems

### [요 약]

경쟁이 심화되는 글로벌 방산수출환경은 다양한 비정형적 데이터의 체계적인 분석을 통해 수출목표국가에 대한 수출유망분야를 선정한 후 수출방향에 대한 전략적 맞춤형이 필요한 시점이다. 이에 본 연구는 방산수출품목 시장확대를 목적으로 수출목표국가를 독일, 영국, 프랑스로 선정하고 해당국가별 향후 10년간 무기체계 획득계획에 대한 비정형 예측 데이터를 확보하고자 하였다. 이후 텍스트 마이닝 분석 중 TF-IDF 기법을 활용하여 3개국의 데이터에서 자주 등장하는 핵심어를 도출하고자 하였다. 분석결과, 각 국의 주요 획득사업에 대한 핵심어를 도출할 수 있었고 이를 토대로 3국의 공통적 획득계획이 있는 핵심어를 유사한 의미로 분류하여 현 시점의 방산수출에 대한 유망수출품목을 발굴할 수 있었다.

▶ **주제어:** 방위산업, 수출시장확대, 수출품목발굴, 텍스트마이닝, 무기체계

• First Author: Seoyoon Yeo, Corresponding Author: Jong Hee Jeong, Co-Author: Seong Ho Kim

\*Seoyoon Yeo (syyeo@krit.re.kr), Korea Researcher Institute for Defense Technology Planning and Advancement

\*\*Jong Hee Jeong (jheesweet@gnu.ac.kr), Department of Smart Distribution and Logistics, Gyeongsang National University

\*\*Seong Ho Kim (shkim15@gnu.ac.kr), Department of Smart Distribution and Logistics, Gyeongsang National University

• Received: 2022. 09. 22, Revised: 2022. 10. 06, Accepted: 2022. 10. 13.

## I. Introduction

수출시장 확대를 위해 해외 유망수출 품목 및 유망기술 분야 등의 예측은 기술·품목 분류 또는 정형화된 코드 등을 활용한 데이터를 분석하는 것이 필수적이다. 그러나 국방 및 방위산업의 경우, 정형화된 데이터가 부족한 실정이다. 민수산업의 경우 산업분류코드(한국표준산업분류), HS CODE 등 표준화된 코드를 기반으로 매출액, 수출입 품목과 액수 등을 파악할 수 있는 정형화된 데이터가 있으나, 방산수출분야는 폐쇄적인 시장의 특성상 기업차원의 정보 접근이 매우 어렵다. 접근 가능한 정보조차도 국방 및 방위산업은 소수의 분석기관(SIPRI), 사설 방산전문분석 업체(Jane's, Forecast International 등), 각국의 국방백서, 뉴스 등 비정형적이고, 다양한 형태로 산재되어 있다.

기존의 유망수출품목을 설정한 후 적정한 수출목표국가를 탐색하고 지원방향을 결정하는 방식도 중요하다. 그러나 나날이 격변하는 글로벌 방위산업 환경과 무기체계의 수출 시장의 경쟁 심화를 고려해보았을 때, 다양한 비정형 데이터(문자, 단어, 문서 등)의 구체적이고 정확한 분석을 통해 각 국가별 수출유망분야를 선정한 후 국내기업의 수출방향에 대한 전략적 맞춤화를 실시하는 마케팅 고도화가 필요한 시점이다. 특히, 방산수출은 수출국 정부의 획득 정책 및 계획을 바탕으로 수요가 특정되거나 제한되어 있어 정부차원에서도 수출목표국가별 유망수출품목을 선정하여 집중적으로 지원하는 방식이 지원사업의 수출 성과 및 실효성을 높일 수 있다.

이러한 가운데, 지난 6월 NATO의 정상회의 결과는 시사하는 바가 크다. 첫째, 핀란드와 스웨덴의 NATO 가입 공식화와 더불어 러-우크라이나 전쟁 종식 이후에도 미국을 제외한 NATO의 회원국 중 특히 유럽국가들의 GDP의 2%에 해당하는 국방비 지출목표와 연간 국방비의 20% 이상에 해당하는 장비도입부문 지출목표를 설정하여 국방비 지출액 증가 추세가 일정기간 지속될 것으로 평가된다는 것이다. 둘째, 이번 회담에는 한국, 일본, 호주, 뉴질랜드 정상이 '글로벌 파트너' 자격으로 참여하여 높은 관심을 받은 점이다. 셋째, 미국은 이번 전쟁을 러시아산 무기의 세계시장 점유율을 줄이는 계기로 만들고자 전략적 상호운용성 강조에 대한 필요성을 역설하고 있다[1]. 이것은 곧 EU와 그 주변국에 대한 방산수출 시장규모가 확대될 것이고 전략적 협력 국가로 한국을 포함한 유럽 외 국가가 포함된 것으로 미루어 보아, 한국의 글로벌 방산수출 공급망 참여에 대한 청신호가 발생했음을 확인할 수 있다.

따라서 본 연구는 자연어(비정형 문장)의 군집화 및 통

계적 연산을 통해 문서간 중요도 등을 파악할 수 있는 텍스트 마이닝(text mining) 기법을 활용하여 수출목표국가별 주요 수출가능분야 키워드 선정을 통해 현시점의 유망수출품목을 발굴하여 방산수출 시장 확대 가능성을 제시하고자 한다. 본 연구결과는 국가별 유망수출품목을 파악할 수 있다. 또한 이러한 품목군에 대한 개발·생산 기술을 보유한 기업을 선제적으로 집중 육성하고 수출국가별 맞춤형 수출지원 기반 마련에 기여할 것으로 보인다. 이와 더불어 국가별로 상이한 방산정책과 심화되는 글로벌 경쟁하에 국내 방산기업들이 지속적 경쟁우위를 유지하기 위한 전략 수립 및 활용에 실무적 가이드라인을 제공할 수 있을 것이다.

## II. Preliminaries

### 1. Theoretical background and Related works

#### 1.1 Word embedding and TF-IDF

자연어는 수치화되어 있지 않아 컴퓨터 기반의 머신러닝, 딥러닝 기법을 원시 자료(raw data) 형태에 바로 적용할 수 없다. 따라서 자연어의 특징을 추출하여 벡터로 표현하는 과정이 필요하다. 이를 통해 분석 가능한 데이터 형태로 만드는 과정을 워드 임베딩이라고 한다[2]. 그 중 횡수 기반 임베딩의 대표적 기법인 TF-IDF (Term frequency-inverse document frequency)는 핵심어 추출 및 검색 결과 순위 결정을 위해 단어의 특정 문서 내 중요도를 산출하는 통계적 가중치 알고리즘으로 검색 순위 및 문서 내 핵심어 추출에 유용하다[3]. TF-IDF는 여러 문서로 이루어진 문서 군에서 어떤 단어가 특정 문서 내에서 얼마나 중요한 것인지를 나타내는 통계적인 수치로 문서의 핵심어를 추출하거나, 문서들 사이의 비슷한 정도를 구하는 등의 용도로 사용되는 텍스트 마이닝 기법이다. TF는 특정한 단어가 문서 내에서 얼마나 자주 등장하는지를 나타내는 값으로, 값이 높을수록 해당 문서에서 중요하다고 할 수 있다. 그러나 여러 문서의 집합 간 비교를 할 때 빈번하게 사용되는 것은 그 단어가 흔하다는 것을 의미한다. 따라서 해당 문서의 TF 값뿐만 아니라 문서 집합 간에 나타나는 빈도인 DF값의 역수, IDF를 적용함으로써 문서간에 자주 사용되는 단어에 대한 역가중치를 부여하여 문서별 중요 단어(핵심어)를 도출할 수 있다. TF-IDF는 0~1사이의 값으로 존재하는데 1에 가까울수록 해당 단어가 속한 문서에서 중요도가 높고, 해당 문서의 주제를 결정지을 수 있는 가능성이 크기 때문에 핵심어를 추출할 수

있는 척도라고 할 수 있다[4].

이를 수식으로 나타내면 다음과 같다. 여기서 N은 전체 문서의 수이다. TF는 특정 문서에서 특정 단어가 등장한 횟수이고, DF는 특정단어가 등장한 문서의 수이다.

$$TF-IDF = TF \times IDF(t,D)$$

$$TF(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$$

$$idf(t,D) = \log \frac{|D|}{|\{d \in D : t \in d\}|}$$

본 연구의 목적은 10년간의 국가별로 예상되는 획득 사업의 상세 내용 중 핵심어 도출과 도출된 핵심어를 토대로 한 유망 수출품목의 발굴이므로 횟수(Frequency) 기반의 TF-IDF 분석을 활용하고자 한다.

### 1.2 Antecedents in Text mining

문헌조사, AHP, 델파이 등의 접근방법은 결과를 도출하는데 많은 시간과 비용이 소요되며 전문가의 주관적 의견이 반영될 수 있다는 연구의 한계점이 존재한다. 정형화된 수치 데이터를 기반으로 분석할 수 있는 영역은 객관적인 분석이 가능하다는 점에서 이점이 있으나, 연구목적에 적합한 정량 데이터가 부족하며 연구 범위가 제한적이다. 최근에는 정성적 접근방식의 비효율성과 정량데이터의 부족으로 인한 연구 제한 등의 한계를 보완하고자 다양한 분야에서 텍스트 마이닝 등과 같은 비정형 데이터 분석기법을 도입한 연구가 증가하고 있다.

텍스트 마이닝이란 자연어(인간언어)로 구성된 비정형 텍스트 데이터(Unstructured Text Data)에서 숨겨진 패턴 또는 관계를 추출하여 의미 있고 활용 가치가 높은 정보 또는 지식을 찾아내는 분석 기법이다[5]. [6]에서는 국방기술품질원에서 운영하는 국방기술정보통합서비스에서 제공하는 국방강소벤처 Tech-Fi Net에 등록된 데이터를 대상으로 빈도기반의 키워드를 추출하고 단어간 중심성 분석을 통해 등록기업이 집중되어 있는 분야와 그렇지 않은 분야를 제시하여 향후 국방강소벤처 기업 기술개발시 방향을 제시할 수 있는 연구를 수행하였다. [7]에서는 통계청 경제총조사 기초 데이터를 활용하여 Word2Vec 알고리즘을 활용하여 유망 품목을 식별하고 이와 유사한 제품명들을 추출하여 이들의 매출액을 기준으로 시장규모를 추정하였다. 기존의 사전에 정해진 품목을 기준으로 시장규모를 추정했던 정량적 접근 방식에서 벗어난 텍스트 마이닝 기법을 통해 사전에 정해지지 않은 품목 혹은 제품의 시장규모를 추정하는 방법을 제안했는데 텍스트 마이닝 기법을 활용한 연구의 확장성을 보여주었으나, 문맥이 없

는 제품명간의 유사도를 분석하여 배열 순서에 따라 결과 값이 달라질 수 있는 한계가 있다.

다양한 분야에서 TF-IDF 가중치를 활용하여 특정 문서에서 자주 언급되는 단어를 도출하여 시장이나 기술 트렌드를 파악하는데 활용되고 있다. [8]에서는 부동산시장 동향분석을 위해 관련 기사에 TF-IDF 모델을 활용하여 주택시장의 변화특성을 분석하였으며, [9]에서는 경제 뉴스에 TF-IDF 기법을 적용하여 주식시장 트렌드를 분석하였다. 따라서 본 연구에서는 각 국별 예상 획득사업의 주요 핵심단어를 도출하여 국가별로 요구하는 유망 획득사업 및 품목 트렌드 차이를 비교·분석할 수 있도록 국가별 획득사업 설명 자료에 TF-IDF 가중치를 두어 획득사업의 핵심단어를 도출하고자 한다.

### 1.3 Antecedents in promising export items

다양한 산업의 유망 수출품목, 기술선정에 대한 선행연구는 존재하나 방위산업을 경영적 관점에서 바라보고 방산물자의 수출이나 품목선정에 관한 연구는 미흡한 실정이다[10]. 이에 방산수출에 관련된 연구는 민간수출분야의 연구를 바탕으로 방산수출의 특수성을 고려하여 연구 방향을 제시할 필요가 있다.

방위산업 분야의 유망수출품목 선정에 관한 연구로 [11]에서는 타 기관의 유망품목 선정 기준 등의 사례연구를 통해 품목 자체 경쟁력, 수출업체 역량, 수출파급효과 등 3가지 평가분야 및 하위 8가지 세부평가항목으로 구분하고 AHP 설문을 통해 항목별 상대적 가중치를 산정하여 유망수출품목 선정 모형을 제시하였다. 해당연구는 첫째 유망품목을 정의하고, 둘째 미래기술에 대한 예측을 토대로 유망품목후보군을 도출하며, 마지막으로 최종적인 유망품목을 선정하는 의사결정 과정의 순으로 이어지는데, 이러한 접근방법은 기업이 가진 역량을 바탕으로 유망수출품목을 정의하므로 구매자 별 맞춤형 니즈가 반영되기 어렵다는 한계점을 보여준다. [12]의 경우 현시비교우위지수(revealed comparative advantage: RCA) 분석을 통해 對세계비교우위가 있는 상위 50개 품목을 대상으로 선정하고, 對 비세그라드 4개국(체코, 슬로바키아, 폴란드, 헝가리)의 수출증감률을 시장성장으로, 수출점유율을 시장경쟁력으로 보아 일정 비율이상의 품목(HS CODE 4단위 기준)을 비세그라드 4국의 유망수출품목으로 선정하였다. 해당연구는 수출대상국을 선정한 후 시장진입이 유망한 품목을 선정한다는 점에서 목표시장 맞춤형 품목을 선정한다는 의의가 있기는 하지만, 민수분야에 국한된 HS CODE를 활용한 분석방법이라는 점에서 HS CODE 기준으로 분

류하기 어려운 방산분야 수출품목 선정에 적용하기에는 한계가 있다.

### III. The Proposed Scheme

#### 1. Research data and method

본 연구는 IHS Jane's(www.janes.com), Forecast international(www.forecastinternational.com), e-나라 지표에서 제공하는 방위산업 관련 Market Forecast 자료 등을 활용한다. 이를 토대로 2022년 2월에 발발한 러시아-우크라이나 침공으로 인해 22년 2월 이후 우크라이나에 무기체계를 직접 지원하거나 경제적 지원을 하고 있는 서방권 국가 중 방산 4대 강국이라 할 수 있는 중 미국을 제외한 독일, 영국, 프랑스의 향후 10년(2022~2031년)간의 국가별 무기체계 및 군수품 획득 예상소요를 데이터셋으로 활용하였다. 데이터셋의 기술적 특징은 다음과 같다.

첫째, 향후 10년간의 예측 데이터를 활용한 이유는 국방 획득의 경우 기본적으로 5개년 중기 계획을 수립하고 계획 후 사업화 및 예산 반영기간이 5년임을 고려함이다.

Table 1. The Ratio of Imports in Finished Weapon Systems to the Total US Defense Budget

Defense budget trend in America	2022
Import amount of finished W.S.(millions, USD)	2118.75
Defense Budget(millions, USD)	793,255
The ratio of import in finished W.S. to the total US D.B.	0.27%

Source: JANES Market Forecast & World Defense Budget

둘째, 본 연구에서 미국을 제외한 이유는 다음과 같다. 국방기술품질원의 2019년 세계 방산시장 연감에 따르면 미국의 국방비 지출은 7320억 USD(약 799조 3440억원)으로 전 세계 국방비 지출의 38%를 차지하며 2~5위를 나타내는 중국, 인도, 러시아, 사우디 등을 포함한 11위 국가까지 포함한 국방비 지출 총액보다 큰 금액을 지불하는 국가이자, 96개국에 무기를 납품하는 세계 최대 무기수출국이다. <표 1>은 2022년 기준 미국의 총 국방예산 대비 제 3국에서 현지합작, 공동연구개발을 제외한 무기체계 완제품을 직접 수입한 비중을 나타낸다.

이를 살펴보면, 총 국방예산에 대비하여 완성 무기체계 수입금액의 비중은 약 0.27%를 나타낸다. 즉, 99% 이상의 획득소요가 미국 내에서 이루어진다고 볼 수 있다. 앞서 언급한 것처럼, 미국은 세계 최고의 국방비 예산이 편성된

국가이기 때문에 미국의 방위산업시장은 시장매력도가 높다고 평가할 수 있지만, 완성무기체계를 활용한 시장진입은 <표 1>의 분석결과와 같이 상당한 어려움이 따르며, 방산수출에서 완성무기체계가 아닌 개별 부품까지 모두 고려하기에는 국가안보 논리의 문제로 인해 정확한 데이터의 수집이 아주 어렵다. 이러한 관점에서 본 연구의 목적이 방산 수출시장 확대를 위한 유망 방산수출품목 발굴인바, 미국은 본 연구를 위한 수출목표국가로 고려하는데 상당한 한계가 따른다.

Table 2. Summary of Raw Data

Classification	Value
The number of Export target	3
Total market size (millions, USD)	546,796
Estimated period of market size(2022~2031)	10 years
The number of text data	7,277
The number of effective words used in the analysis	91,735

셋째, 연구에 사용된 항목은 상세 사양, 요구배경, 발주처 등을 문장의 형태로 제공한 물품 명세와 품목군, 예상 획득금액이며, 분석에 사용된 텍스트 데이터의 개수는 <표 2>와 같이 7,277개, 총 시장규모는 5457억 9600만 달러, 텍스트 데이터를 단어 단위로 구분하고 불필요한 단어(관사, 전치사, 특수문자 등)를 제외한 분석에 활용된 유효 단어수는 91,735개이다. 3개 국가별로 사용된 텍스트 데이터 개수는 독일 1,891개, 영국 2,945개, 프랑스 2,441개이다. 국가별 Market Forecast 데이터셋은 각 국 정부 또는 군에서 획득계획을 발표하거나, 기존에 운용 중인 무기체계의 수명주기에 따른 도태시기를 예측하여 2022년 ~ 2031년간 해당 국가가 도입할 것으로 예상되는 획득사업에 대한 요구사항, 소요군, 연동체계, 교체시기 등을 평균 25개 내외의 단어로 기술된 텍스트 데이터를 포함한다. 이를 통해 본 연구는 무기체계 획득계획에 대한 텍스트를 분석하여 국가별로 잦은 빈도와 다른 국가에 비해 해당 국가에 특정하게 많이 언급되는 빈도를 기준으로 유망수출품목에 대한 핵심 키워드를 도출하고자 한다. 각 텍스트 데이터별 중복되는 단어와 불용어(관사, 전치사 등)를 제외하면 평균적으로 12~13개 내외의 단어가 핵심 키워드 도출에 활용된다.

텍스트 데이터를 분석하는 도구로는 R Studio Library 중 자연어처리 패키지인 Tidytext 패키지를 활용한다. Tidytext 패키지는 띄어쓰기 단위가 아닌 텍스트 포맷의 데이터를 글자 중 의미를 가진 단위를 총칭하는 token 단

위로 쪼개서 분석하기 때문에 데이터를 깔끔한 텍스트 형식으로 변환하여 필터링이 용이하고 처리가 쉬운 워크플로우를 제공한다는 장점이 있다[13].

## 2. Research model

본 연구에 활용된 전체적인 방법론과 절차를 <그림 1>로 나타내었다. 먼저 실험에 필요한 데이터를 수집하고 Text Mining 분석에 적합한 데이터로 가공하는 전처리 단계를 수행한다.

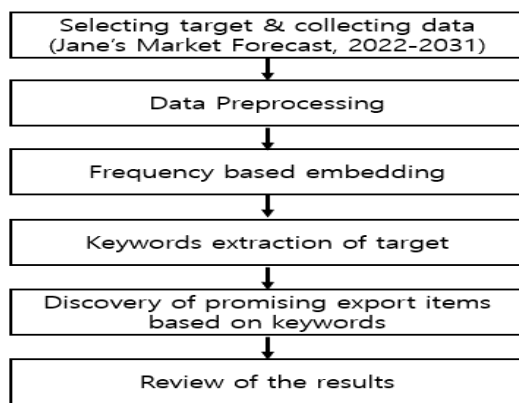


Fig. 1. Research process in this study

전처리 단계에서는 문장을 단어 단위로 나누고, 문장부호 및 특수기호 등과 불필요한 공백 등을 제거한다. 또한 영문 데이터를 분석하기 위해 관사, 전치사, 관용적 부사 등을 불용어로 분석에서 제외하며, 국가명과 같은 해당 국가를 상징하는 단어들을 제외하는 과정을 거친다. 이후, 국가별 유효단어로 정제된 데이터를 TF-IDF를 통해 국가별 문서내 중요하게 여겨지는 핵심어를 추출한다. TF-IDF 분석을 통해 국가별 문서에서 잦은 빈도로 다루어지지 않지만 타 국가별 문서에서는 추출되지 않아 해당국에서만 주요하게 다루어지는 단어를 도출할 수 있다. 또한 도출된 핵심어를 바탕으로 실제 국내상황에 맞추어 수출이 가능한 주요 품목들을 분류함으로써 목표 국가별 수출가능품목과 이에 맞춤형 수출전략 및 경쟁력 강화방안을 검토해보고자 한다.

## IV. Results

TF-IDF 기법을 통해 독일, 영국, 프랑스의 획득소요계획에 대한 텍스트 마이닝을 실시한 결과 도출된 핵심어는 TF-IDF 값이 0~1 사이인 경우 유효하므로 해당 기준에

의해 각 국가별 20개씩 도출되었고 크게 국방 획득관련 기관, 획득 무기체계, 방산기업, 보병장비 등으로 구분되었다(Appendix 참고). 그러나 해당 결과를 바탕으로 각 국가별 핵심어 20개씩 모두 분석하는데는 한계가 따른다. 첫째, 내용이 너무 광범위하다. 해당 결과는 본 연구의 목적인 수출시장 확대를 위한 유망수출품목을 발굴에는 적합하지 않은 글로벌 방산기업명, 연합방위군, 보병장비(방탄모 등)의 내용도 존재한다. 둘째, 가장 큰 방산 이슈 중 하나인 전 세계 방위산업의 글로벌 공급망 재편에 따라 신속한 공급망 참여가 필수적인 시점이다. 이에 따라 본 연구는 각 국가별 도출된 핵심어 중 국내에서 개발중이거나 완제품 양산역량을 갖추고 있어 신속한 수출이 가능한 유망수출품목을 중심으로 연구결과를 기술하고자 한다. <표 3>, <표 4>, <표 5>는 독일, 영국, 프랑스의 획득소요에 대한 핵심어의 출현빈도 및 TF-IDF 값을 나타낸다.

분석결과 <표 3, 4, 5>를 살펴보면 각 국별 7개의 핵심어가 도출되었고 TF-IDF 값이 1에 가까운 순서로 기재하였다. 또한 구체적인 유망수출품목을 발굴을 위해 본 연구는 <표 3, 4, 5>의 분석결과를 토대로 3 국간 공통 핵심어를 찾고 유사무기체계를 분류하여 유망수출품목을 분류하였다. 해당 결과는 <표 6>을 통해 확인 가능하다. 핵심어별 빈도는 <그림 2>와 같다.

<표 6>을 살펴보면 핵심어를 8개의 유사무기체계로 분류한 후 각 유사 무기체계별로 언급된 핵심어의 종류와 해당 핵심어가 도출된 국가가 나타나 있다. 핵심어를 토대로 유사무기체계를 분류해보면, 정보시스템, 전투기, 레이더, UAV, 미사일, 위성, 전함으로 구분할 수 있다.

Table 3. Cumulative Frequency of Keywords Required for Acquisition in Germany and TF-IDF

words	Meaning of Words	Frequency	TF-IDF
Luna	Unmanned Aerial Vehicle (UAV)	31	0.000665
MKS	Multi-Functional Battleship	29	0.000622
Tornado	Multi-Functional Fighter aircraft	47	0.000504
Motako	Mobile tactical communication	23	0.000493
Frankenthal	Minelayer	20	0.000429
Skeldar	UAV	20	0.000429
Attica	Thermal imager	19	0.000408

Table 4. Cumulative Frequency of Keywords Required for Acquisition in UK and TF-IDF

words	Meaning of Words	Frequency	TF-IDF
Insyte	Multi-Functional Radar	47	0.000608
Bowman	Tactical communication system	91	0.000588
Ceptor	Air-defense missiles	41	0.00053
Skynet	Satellites series	36	0.000465
Nautis	Modular command and weapon control system	30	0.000388
Bisa	Battlefield Information System Application	29	0.000375
Uat	Outfit UAT-ESM systems	29	0.000375

Table 5. Cumulative Frequency of Keywords Required for Acquisition in France and TF-IDF

words	Meaning of Words	Frequency	TF-IDF
Syracuse	Maritime satellite communications for naval	98	0.00077
Mica	Antiaircraft missile for naval	46	0.000723
Sperwer	Tactical UAV ground control station	42	0.00066
DRBV	Radar for naval	39	0.000613
ARBR	Electronic Support Measure for naval	38	0.000597
Senit	Combat systems for naval	34	0.000534
Rafale	Multi-functional fighter aircraft	65	0.000511

정보시스템의 경우 하위 단계로 지원, 통신, 제어, 통제 시스템으로 다시 구분된다. 정보시스템의 경우 언급된 핵심어의 개수는 영국 4개, 프랑스 3개, 독일 1개로 총 8번 도출되었다. 전투기는 독일 1개, 프랑스 1개로 2번 도출되었다. 레이더와 미사일, 위성은 영국과 프랑스에서만 도출되었고 UAV와 전함, 열화상 카메라는 독일만 도출되었다.

우선, 도출된 핵심어들을 살펴보면 대부분 자국 방산기업의 강점 분야를 중심으로 획득사업이 계획되고 있다는 것을 확인할 수 있다. 예를 들어 독일은 2022년 2월 국방 예산 배정 계획을 발표했는데 이때, 독일 내 방산기업들이 전투기, 레이더시스템, 함정 등의 공급을 제안했다는 것으로 알려져 있다. 프랑스에서 전투기 획득계획에 대한 핵심어로 Rafale이 도출됨을 확인할 수 있다. 이것은 방위산업의 특성상 무기체계 개발과 양산역량을 갖추에 있어 규모의 경제가 아주 중요한 요소로 고려되기 때문이라 사료된다.

다. 이에 본 연구는 방산수출에 있어서도 중요한 한 요소인 규모의 경제를 고려하여 수출가능품목 발굴을 위해 3개국 공통적으로 도출된 핵심어를 기반으로 유사어 분류 및 빈도분석을 진행하였다. 분석결과, 정보시스템에 관련된 무기체계 관련 핵심어는 전체 21개 단어 중 8번 도출되어 약 38%의 도출빈도를 보여준다.

Table 6. Classification of similar weapon systems on keywords in 3 countries

Weapon system	Classification	
Information system	electronic support	UAT(UK), ARBR(Fr)
	communication	Motako(Ger), Bowman(UK)
	control	Sperwer(Fr), Nautis(UK)
	command	Bisa(UK), Senit(Fr)
Fighter aircraft	Tornado(Ger), Rafale(Fr)	
Radar	Insyte(UK), DRBV(Fr)	
UAV	Luna, Skeldar (Ger)	
Missile	Ceptor(UK), Mica(Fr)	
Satellite	Syracuse(Fr), Skynet(UK)	
Combat ship	MKS, Frankenthal (Ger)	
Thermal imager	Attica(Ger)	

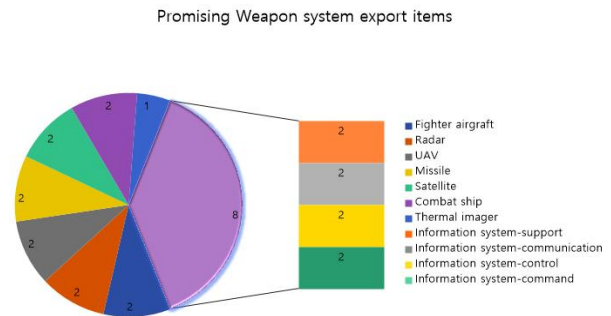


Fig. 2. Frequency of each keyword

정리하자면, 서유럽 국가들은 경쟁국과의 군사적 경쟁에서 우위를 차지하고자 하이테크 기술을 통해 전략 및 전술이 변화하는 미래전에 대비할 수 있는 정보시스템에 관련된 무기체계의 구체적인 획득계획이 존재한다고 볼 수 있다. 한국 내 방산기업들은 정보시스템에 관련된 완성품 무기체계의 양산이 가능할 뿐만 아니라 정보시스템 구축에 필수적인 부품이라 볼 수 있는 반도체 수급도 경쟁국에 비해 안정적인 조절 역시 가능하다. 이에 본 연구결과를 토대로 정보시스템에 관련된 무기체계들이 현시점에서의 유망수출품목이라 분석할 수 있다.

## V. Conclusions

제조·생산기술 수준의 차이가 빠르게 좁아지고, 방산수출의 특성상 불특정다수를 대상으로 한 방산수출 품목 발굴은 수출경쟁력 유지가 어렵기 때문에, 방산수출 분야 또한 고객맞춤화를 극대화한 마케팅 전략 수립이 상당히 요구되는 시점이다. 이에 본 연구는 방산수출 시장 확대를 목적으로 방산수출 관련 최신 이슈를 반영하여 수출목표 국가를 독일, 영국, 프랑스의 3개국으로 선정하고 해당국가별 향후 10년간 무기체계 획득소요계획에 대한 비정형 예측 데이터를 확보하여 텍스트 마이닝 기법 중 하나인 TF-IDF를 활용하였다. 이를 통해 3개국의 데이터에서 자주 등장하는 핵심어를 도출한 후 무기체계에 관련된 핵심어를 분류하고 그 중 국내양산역량을 고려하여 유망수출 품목을 발굴해내고자 하였다. 분석결과, 독일, 영국, 프랑스에 각각 7개의 수출품목을 분류할 수 있었다. 예를 들어 프랑스 전투기 소요계획에는 연관핵심어로 Rafale이 도출되는 등, 각각 자국 내 방산기업의 강점 분야를 중심으로 주요 획득사업을 계획하고 있음을 추론할 수 있다.

이에 보다 구체적인 수출품목 발굴을 위해 본 연구는 도출된 핵심어를 토대로 빈도분석을 실시하여 3국이 공통적으로 획득계획에 관련된 핵심어를 도출하여 방산수출의 주요 고려사항 중 하나인 규모의 경제를 달성할 수 있는 수출가능 무기체계의 우선순위를 비교해보고자 하였다. 분석결과, 3개국 전체 도출된 21개 단어 중 8개의 단어가 공통적으로 정보시스템에 관련된 무기체계 획득에 관련된 단어라는 것이 확인가능하다. 즉, 서유럽의 군사강대국들이 전장환경에서 실시간 상황을 공유할 수 있고 이를 통해 획득한 정보를 바탕으로 효율적이고 효과적인 의사결정을 내릴 수 있는 지휘통제 체제를 비롯한 정보시스템 구축에 큰 관심이 있고 대규모 예산 투자가 가능함을 나타낸다. 정리하자면, 완성무기체계에 대한 국내양산역량도 갖추고 있을 뿐만 아니라 정보시스템 구축에 필수적인 반도체 산업의 글로벌 경쟁우위가 존재하여 부품의 안정적인 공급이 가능하므로 방산수출에 있어 정보시스템에 관련된 무기체계들이 현재 가장 수출유망품목임을 제시가능하다.

본 연구결과를 바탕으로 국제적 이슈 등에 따른 방산분야의 공급 변화에 빠르게 대응할 수 있도록 대상국가를 선정하고, 대상국가에서 주요하게 다루는 핵심어 기반의 획득소요 파악을 바탕으로 현 시점에 집중 공략해야 하는 분야가 무엇인지 파악할 수 있다. 또한 목표시장의 공통 획득소요계획이 있는 무기체계 혹은 탑재장비 등을 선별할 수 있게 되어 향후 상세획득 계획 역시 추론 가능하다. 이

처럼 핵심어를 활용한 해외소요에 대한 우선순위 부여는 명확한 시장진입목표를 세울 수 있고 소요에 기반을 둔 국내기업의 경우 구체적인 진입전략 수립이 가능하며, 유관기관에서는 적합한 기업매칭 후 맞춤형 마케팅 지원이 가능함을 알 수 있다.

본 연구의 함의는 다음과 같다. 첫째, 가용할 수 있는 예산과 인력이 제한적인 환경하에 놓여 있는 국내 방산기업의 해외시장진출을 위해서 효율적으로 유망수출품목을 발굴함으로써, 국내기업의 신속하고 맞춤형된 진출전략의 수립에 대한 가이드라인을 제시할 수 있다. 둘째, 방산수출은 정부의 전폭적 지원이 필수적인 분야인 관점에서 본 연구 방법론을 통해 목표국가별 핵심어를 추출해내어 유관기관에서 우선적으로 지원해야 할 분야를 선정할 수 있으며, 이슈 발생시점 적기에 기업지원이 가능하다. 셋째, 비정형화된 데이터를 분석하여 수출목표국의 미래 핵심소요를 분석함으로써 기존 제공사, 향후 기회, 발주처 등을 분석할 수 있다. 이에 기존의 불특정 다수를 상대로 한 1차원적인 마케팅을 통해 소요처의 니즈를 파악하기 보다는, 사전 소요에 기업과 기술을 제안하는 형태의 고도화된 수출 마케팅전략 수립이 가능하다.

이러한 연구의 함의에도 불구하고 본 연구의 한계점은 다음과 같다. 앞서 언급했듯이, TF-IDF 기법을 활용한 분석결과와 경우 0~1사이에 도출된 값이 유효하다고 볼 수 있는데, 해당 기준을 적용할 경우 각 국가별 20개씩의 핵심어가 도출되어 연구 내용이 너무 광범위해짐에 따라 60개의 핵심어 모두를 분석하지 못한 점이다. 향후 단어의 비교범위를 넓혀 분석하고, 핵심어와 수출품목발굴에 대한 다차원적 관점에서 통계적 검증 등 추가 연구가 진행된다면 방산수출 확대를 위한 목표시장 선정과 수출품목 발굴에 있어 더욱 의미있는 결과를 도출해 낼 수 있을 것이다.

## REFERENCES

- [1] K. B. Park, , "Defense Industry Insight", Journal of the Defense Science and Technology Information, Vol. 109, pp.96-103, August, 2022.
- [2] G. Salton, C. Buckley and E. A. Fox, "Automatic Query Formulation in Information Retrieval", Journal of the American Society for Information Science, Vol. 34, Issue. 4, pp.262-280, July, 1983. DOI: 10.1002/asi.4630340406
- [3] Y. L. Jung, J. H. Kim and H. S. Yu, "A Study on Market Size Estimation Method by Product Group Using Word2Vec Algorithm", Journal of Intelligence and Information Systems, Vol.

- 26, No. 1, pp.1-21, March, 2020. DOI: 10.13088/jiis.2020.26.1.001
- [4] S. G. Lee and H. J. Kim, "Keyword Extraction from News Corpus using Modified TF-IDF", The Journal of Society for e-Business Studies, Vol. 14, No. 4, pp. 59-73, November, 2009. DOI: G704-000612.2009.14.4.010
- [5] H. J. Kim, N. O. Cho and K. S. Shin, "Text Mining-Based Emerging Trend Analysis for the Aviation Industry", Journal of Intelligence and Information Systems, Vol. 21, No. 1, pp. 65-82, March, 2015. DOI: 10.13088/jiis.2015.21.1.65
- [6] J. W. Park, I. R. Lee, J. W. Kwon, K. S. Byun, and S. Y. Cho, "Analysis Results in Technological Trends of Military Small Giant Venture Tech-Fi Net via Social Network Analysis", Journal of the Korea Academia-Industrial cooperation Society, Vol. 20, No. 12, pp.444-455, December, 2019. DOI: 10.5762/KAIS.2019.20.12.444
- [7] Y. L. Jung, J. H. Kim and H. S. Yu, "A Study on Market Size Estimation Method by Product Group Using Word2Vec Algorithm", Journal of Intelligence and Information Systems, Vol. 26, No. 1, pp.1-21, March, 2020. DOI: 10.13088/jiis.2020.26.1.001
- [8] H. J. Chun, "Analysis of Real Estate Market Trend Using Text Mining and Big Data", Journal of Digital Convergence, Vol. 17, No. 4, pp.49-55, April, 2019, DOI: 10.14400/JDC.2019.17.4.049
- [9] D. Madhavi. "Analyzing News Sentiments and their Impact on Stock Market Trends using POS and TF-IDF based approach." 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAET). IEEE, November, 2020. DOI: 10.1109/IICAET49801.2020.9257816
- [10] E. H. Joo and S. R. Shim, "A study on the Estimation of Weight of Purchasing Power Indicator for Export Market Selection of Defense Industry Products", Korea Trade Review, Vol. 44, No. 1, pp.193-205, February, 2019. DOI: 10.22659/KTRA.2019.44.1.193
- [11] J. H. Won, "A Study on the Selection Model of Promising Export Items Applicable to the Defense SMEs", Journal of the Korea Academia-Industrial cooperation Society, Vol. 21, No. 7, pp.321-330, July, 2020. DOI: 10.5762/KAIS.2020.21.7.321
- [12] T. D. Yeo and S. D. Kie, "Korea's Trade Strategy and Promising Export Products toward the Markets of Visegrád Group", E-Trade Review, Vol.18, No.1, pp.141-166, February, 2020. DOI: 10.17255/etr.18.1.202002.141
- [13] J. Silge and D. Robinson, "Tidytext: Text Mining and Analysis Using Tidy Data Principles in R", The Journal of Open Source Software, Vol. 1, No. 3, 37, pp.1-3, July, 2016. DOI: 10.21105/joss.00037

## Appendix.

Table 1A. Cumulative Frequency of Keywords Required for Acquisition in Germany and TF-IDF

words (Ger)	Frequency	TF-IDF	Ranking
eurofighter	75	0.000804575	1
ref	58	0.000622205	2
tornado	47	0.000504200	3
bundeswehr	37	0.000793847	4
heron	37	0.000396923	4
rim	35	0.000375468	6
corvettes	34	0.000364740	7
luna	31	0.000665115	8
mks	29	0.000622205	9
motako	23	0.000493472	10
frankenthal	20	0.000429106	11
lurssen	20	0.000429106	11
skeldar	20	0.000429106	11
attica	19	0.000407651	14
hafis	19	0.000407651	14
idz	19	0.000407651	14
sachsen	19	0.000407651	14
dsqs	18	0.000386196	18
heer	17	0.000364740	19
infosys	17	0.000364740	19

Table 2A. Cumulative Frequency of Keywords Required for Acquisition in UK and TF-IDF

words (UK)	Frequency	TF-IDF	Ranking
british	217	0.001402634	1
raf	145	0.000937244	2
gbp	120	0.000775650	3
babcock	96	0.000620520	4
bowman	91	0.000588201	5
eurofighter	53	0.000342578	6
insyte	47	0.000607592	7
astute	44	0.000568810	8
ceptor	41	0.000530027	9
skynet	36	0.000465390	10
dstl	33	0.000426607	11
nautis	30	0.000387825	12
bisa	29	0.000374897	13
uat	29	0.000374897	13
ccaaw	27	0.000349042	15
wildcat	27	0.000349042	15
albion	26	0.000336115	17
argyll	24	0.000310260	18
bgoaa	24	0.000310260	18
mastiff	23	0.000297332	20



Table 3A. Cumulative Frequency of Keywords Required for Acquisition in France and TF-IDF

words (Fr)	Frequency	TF-IDF	Ranking
syracuse	98	0.000770041	1
syst	74	0.001162919	2
csf	72	0.000565744	3
rafale	65	0.000510741	4
dcn	50	0.000785756	5
mica	46	0.000722895	6
spectra	45	0.000707180	7
sperwer	42	0.000660035	8
drbv	39	0.000612890	9
arbr	38	0.000597174	10
timent	36	0.000565744	11
senit	34	0.000534314	12
alouette	32	0.000502884	13
floreale	30	0.000471453	14
gaulle	30	0.000471453	14
drbn	29	0.000455738	16
entretien	29	0.000455738	16
fayette	29	0.000455738	16
indisponibilit	29	0.000455738	16
rubis	29	0.000455738	16

## Authors



Seoyoon Yeo received the M.A. degrees in Department of International Trade from Graduate School of International Studies in Pusan National University, Korea, in 2011. She works as a researcher at the Global

Defense Industry Promotion Team, Korea Researcher Institute for defense Technology planning and advancement. Her-research interests include Defense export promotion and strategy, Global Value Chain, Technology management.



Jong Hee Jeong received the M.S. and Ph.D degrees in Department of International Trade from Graduate School of International Studies in Pusan National University, Korea, in 2008 and 2014, respectively.

Dr. Jeong is an lecturer at Department of Smart Distribution and Logistics, Gyeongsang National University, Korea. Her-research interests include Global strategic management, Big data processing on Strategic decision making, Global supply chain management.



Seong Ho Kim received his bachelor's, master's and doctoral degrees from the Department of Trade, Pusan National University in 1997, 1999 and 2002. Dr. Kim joined the Department of Electronic

Commerce at Gyeongnam National University of Science and Technology in 2006 and has been with the Department of Smart Distribution and Logistics at Gyeongsang National University since 2021. He is currently studying international business strategy of big tech enterprise, the impact of information technology such as blockchain and big data on supply chain management and logistics.