

텍스트 마이닝을 활용한 자율운항선박 분야 주요 이슈 분석 : 국내 뉴스 데이터를 중심으로

이혜영¹⁾, 김진식¹⁾, 구병수¹⁾, 남문주¹⁾, 장국진¹⁾, 한성원^{1),2)}, 이주연^{1)*}, 정명석^{1)**}

1) 아주대학교 일반대학원 과학기술정책학과, 2) 해양수산과학기술진흥원 책임연구원

Analysis of major issues in the field of Maritime Autonomous Surface Ships using text mining: focusing on S.Korea news data

Hyeyeong Lee¹⁾, Jin Sick Kim¹⁾, Byung Soo Gu¹⁾, Moon Ju Nam¹⁾, Kook Jin Jang¹⁾,
Sung Won Han¹⁾, Joo Yeoun Lee^{1)*}, Myoung Sug Chung^{1)**}

1) Aju University, Department of Science and Technology Policy

Abstract: The purpose of this study is to identify the social issues discussed in Korea regarding Maritime Autonomous Surface Ships (MASS), the most advanced ICT field in the shipbuilding industry, and to suggest policy implications. In recent years, it has become important to reflect social issues of public interest in the policymaking process. For this reason, an increasing number of studies use media data and social media to identify public opinion. In this study, we collected 2,843 domestic media articles related to MASS from 2017 to 2022, when MASS was officially discussed at the International Maritime Organization, and analyzed them using text mining techniques. Through term frequency-inverse document frequency (TF-IDF) analysis, major keywords such as 'shipbuilding,' 'shipping,' 'US,' and 'HD Hyundai' were derived. For LDA topic modeling, we selected eight topics with the highest coherence score (-2.2) and analyzed the main news for each topic. According to the combined analysis of five years, the topics '1. Technology integration of the shipbuilding industry' and '3. Shipping industry in the post-COVID-19 era' received the most media attention,

Received: April 17, 2024 / **Revised:** May 28, 2024 / **Accepted:** May 31, 2024

*Corresponding Author: Joo Yeoun Lee / jooyeoun325@ajou.ac.kr

**Corresponding Author: Myoung Sug Jung / mschung333@gmail.com

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

each accounting for 16%. Conversely, the topic '5. MASS pilotage areas' received the least media attention, accounting for 8 percent. Based on the results of the study, the implications for policy, society, and international security are as follows. First, from a policy perspective, the government should consider the current situation of each industry sector and introduce MASS in stages and carefully, as they will affect the shipbuilding, port, and shipping industries, and a radical introduction may cause various adverse effects. Second, from a social perspective, while the positive aspects of MASS are often reported, there are also negative issues such as cybersecurity issues and the loss of seafarer jobs, which require institutional development and strategic commercialization timing. Third, from a security perspective, MASS are expected to change the paradigm of future maritime warfare, and South Korea is promoting the construction of a maritime unmanned system-based power, but it emphasizes the need for a clear plan and military leadership to secure and develop the technology. This study has academic and policy implications by shedding light on the multidimensional political and social issues of MASS through news data analysis, and suggesting implications from national, regional, strategic, and security perspectives beyond legal and institutional discussions.

Key Words: Text Mining, Topic Modeling, Maritime Autonomous Surface Ship, Policy Issues

1. 서론

최근 정보통신기술과 전통산업의 융합이 일상생활 전반으로 확장되어, 미래 사회를 이끌어갈 가장 핵심적인 성장 동력으로 주목받고 있다.[1] 조선 및 해양산업에서도 이러한 정보통신기술을 융합한 자율운항선박이 주목 받고 있으며 국내 연구기관과 기업에서도 자율운항선박 관련 안전 운항, 유지관리, 스마트 생산건조 등 다양한 분야에서 정보통신기술과 융합한 기술 개발을 추진 중에 있다.[2] 이러한 자율운항선박은 조선, 해운, 항만 정보통신, 기자재 업체 등 다양한 산업의 기술과 인프라가 융합한 선박으로 조선·해운 시장이 지속적으로 성장할 수 있는 기반 역할을 할 것으로 기대될 뿐만 아니라 중장기적인 관점에서 자율운항선박에 대한 수요 증가와 ICT 부문의 수요 증가로 관련 산업 또한 발전할 것으로 예상된다.[3]

이에 본 연구에서는 자율운항선박 관련 산업의 활성화를 위한 정책적 접근으로 최근 사회적으로 형성되고 있는 관련 이슈가 무엇인지, 도출된 주제들이 정책적으로 어떠한 시사점을 제공할 수 있는지

분석하였다. 이를 위해서 뉴스 데이터 2,843개를 수집하여 역빈도 분석(TF-IDF)으로 주요 키워드를 추출하였으며 토픽모델링을 실시하여 8개의 주요 토픽에 대해 분석하였다. 이러한 연구 결과를 바탕으로 주요 토픽별로 정리해 정책적 시사점을 제안하였다.

2. 선행연구 고찰

2.1 자율운항선박의 이론적 배경

2.1.1. 자율운항선박의 개념

국제연합(United Nations, 이하 "UN")의 전문기구이자 조선·해운 산업계에 지대한 영향을 미치는 국제해사기구(International Maritime Organization, 이하 "IMO") 내에서도 2017년 이전까지는 자율운항선박에 대한 명칭에 대해 스마트 선박(Smart ship), 디지털 선박(Digital ship), 무인 선박(Unmanned ship), 원격조종 선박(Remote ship), 자동화 선박(Automated ship) 등 다양한 용어로 혼용하였다.[4] 이러한 가운데 2017년 IMO MSC

제98차 회의 의제 20번 문서에서 우리나라를 포함한 8개국은 해상에서의 자율운항선박 운용과 관련된 IMO 규정 검토 작업을 신규로 제안하는 문서를 제출하였으며[5] 동 문서에서 처음으로 자율운항선박(Marine Autonomous Surface Ships, 이하 “MASS”)이라는 용어를 사용하였다. 이후에도 대상을 지칭하는 함축적인 의미는 같을지 모르나 자율운항선박을 정의하고 표현하기 위한 다양한 단어들 혼용되어 사용되어 왔고 그 정의가 명확하게 정립되어 있지 않은 상태로 논의가 진행되어 오다가 2018년 IMO MSC 제99차 회의의 논의 결과로써 “MASS definitions and concepts of different types and levels of autonomy, automation, operation and manning should be provisional, neutral in terms of technology and developed for the purpose of the exercise only.(다양한 자동화 수준에서 사람의 간섭 없이 독립적으로 운용될 수 있는 선박)”이라는 정의를 합의하였다.[5]

현재도 자율운항선박의 명칭과 개념은 산업계 및 연구단체 각국의 선급마다 그 수준이 조금 상이한 것으로 나타나고 있으나 최종 단계에 있어서 어떠한 형태를 취한다고 해도 그 운항 주체가 기존의 선박과 달리 인적요소의 배제를 언급하고 있는 특징이 있으며 최종단계에서는 인적요소의 어떠한 간섭도 없이 독립적으로 운용되는 선박이라는 공통점이 있

다.[4] 이를 정리하면 표 1과 같다.

2.1.2. 자율운항선박의 자율화 단계

자율운항선박의 정의에 합의한 IMO MSC 제99차 회의에서는 자율운항선박의 자율화 단계를 4단계로 분류하여 표 2와 같이 제시하였다.[5] 선박의 운항 주체가 누구인지 그리고 선박의 시스템(자동화, 무인화, 완전 자율화)에 따라 4단계로 분류하였지만 AI 시스템에 대해 과거의 기술개발 수준으로 반영한 것을 감안하면 관련 기술 발달과 함께 자율운항선박의 자율화 단계는 언제든지 바뀔 수 있다.

그러나 이렇게 개념 구분을 하는 것이 추후 규범의 적용과 기술개발에 도움이 될 수 있으므로 기술적으로 지나치게 세분화된 자율화 단계 구분법 보다는 국제적인 영향력 및 앞으로의 자율화 단계 사용가능성 등을 종합적으로 고려하여 IMO의 4단계 구분법을 따르는 것이 합리적이라 판단된다.

2.1.3. 자율운항선박 개발의 필요성

조선·해운산업 관점에서 자율운항선박을 개발함에 따른 장점으로는 첫 번째, 선박 건조 방식에 따라 친환경 선박으로 건조할 수 있는 장점을 지니고 있다. 전기추진 등의 방식을 사용하므로 대기환경오염원을 감소하여 대기환경을 보호하고, 선원이 승선하지 않으므로 인해 오수 및 중수와 같은 폐기물

〈Table 1〉 Compare similar terms for MASS

구분	영문명	정의	
자율운항선박	Maritime Autonomous Surface Ship	인적 여부에 관계없이 자율적 결정 시스템을 가지고 운항하는 선박	
유사 용어	스마트 선박	Smart Ship	이해당사자와 연결되어 정보 및 서비스를 제공하고, 자율적 또는 원격에 의한 조종으로 진단하고 관리함으로써 최적의 에너지 효율로 안전하게 운항하는 선박과 이를 위한 ICT 인프라
	디지털 선박	Digital Ship	디지털-IT 기반의 정보 처리를 토대로 안전하고 효율적 선박운항이 가능한 선박
	무인 선박	Unmanned Ship	선박과 육상에 설치된 무선장비와 통합제어 시스템에 의해 선박운항자의 별도 조종없이 선박조종 제어 시스템에 의해 운항하는 선박
	자동화 선박	Automated Ship	육상의 선박운항관리자 또는 컴퓨터가 입력한 동작신호를 받아 운항하는 선박
	원격조종 선박	Remote Ship	육상의 선박운항관리자의 직접적 통제를 지시받아 운항하는 선박

출처: 박혜리(2018)를 참고하여 재작성[6]

<Table 2> IMO's steps to autonomy

단계	상세 내용
1단계	Ship with automated processes and decision support: Seafarers are on board to operate and control shipboard systems and functions. Some operations may be automated (부분적 자율운항지원 자동화 단계 및 선원이 의사결정을 지원하는 기능의 선택)
2단계	Remotely controlled ship without seafarers on board:The ship is controlled and operated from another location. (선원이 승선, 원격으로 제어되는 선박, 시스템 고장 시 선원이 대응)
3단계	Remotely controlled ship without seafarers on board:The ship is controlled and operated from another location. (선원이 승선하지 않은 선박으로 원격으로 제어되는 선박, 시스템 고장을 대비하여 Stand-by 기능 구축)
4단계	Fully autonomous ship: The operating system of the ship is able to make decisions and determine actions by itself. (완전자율운항선박)

출처 : IMO(2017)자료를 참고하여 재작성

발생이 거의 없어 해양 환경보호에 기여할 수 있다.[7] 두 번째, 선원 비용의 절감을 통해 선박 운항비를 크게 절감할 수 있다.[4] 선원 비용은 선박 운항비에서 많은 부분을 차지하고 있는데 무인선박 운항을 통해 선박소유자는 선원 비용을 절감할 수 있다. 또한 생존 장비를 탑재하지 않아도 되며 선원이 거주 구역을 따로 설계하지 않아도 되므로 공간확보에 따른 비용 절감을 가져 올 수 있을 뿐만 아니라 연료유도 상당히 절감할 수 있다.[8] 셋째, 자율운항선박은 해상안전 및 해양환경오염을 경감하고 방지하는데 상당한 기여를 할 수 있다. 많은 해상사고가 직·간접적으로 선원의 부주의 및 고의·과실, 주의의무 태만에서부터 발생 되기 때문에 인적요소의 제거를 통해 해상사고를 사전에 예방 할 수 있다.[9] 마지막으로 자율운항선박은 선원의 근무 환경을 개선하는데 상당한 기여를 할 수 있다. 가족과 떨어져 고립된 생활을 해야하기 때문에 선원 생활을 기피하기도 하는데 자율운항선박이 상용화 될 경우 육상제어 통제실에서 선박의 운항 상태를 체크하며 정해진 출·퇴근 시간을 보장받고 가족과 함께 생활할 수 있는 새로운 근무 환경이 조성될 것으로 기대된다.[10]

2.1.4. 자율운항선박 관련 선행연구

국내의 경우 이현균(2022)은 자율운항선박의 상용화를 위한 연구·실증·시범 운항 등이 원활히 이루

어 질 수 있도록 관련 규제를 완화하는 등 자율운항선박의 개발 촉진과 상용화를 위한 법률의 필요성을 주장하며 관련 국내·외 법제 현황을 비교·분석하여 자율운항선박촉진법 제정안을 제시하였다.[11] 김민영 외(2019)는 자율운항선박이 국제해상충돌 예방 규칙을 준수하기 위해서 필요한 선박 범주 및 합성곱 신경망 기반의 선박 분류 기술을 제안하였다. 본 논문에서 정의된 선박 범주에 맞도록 기존 데이터 셋 및 인터넷 영상 검색 데이터의 정제를 통해서 40,300장 규모의 선박 범주 분류 데이터 셋을 구축하였다. 다음으로 최신 합성곱 신경망 모델을 구축된 선박 범주 분류 데이터 셋에 적용하여 선박 범주 분류 성능을 분석하였으며 실험 결과, 전이 학습을 통해 학습된 Inception-ResNet V2 모델은 14개 선박 범주를 91%의 높은 정확도로 분류함을 확인하였다.[12] 해외의 경우, Thieme et al(2019)은 현재 선박 충돌, 선박 구조 충돌 및 접지에 대한 선박 위험 모델이 얼마나 멀리까지 자율운항선박의 위험 평가에 적용 가능한지를 조사하였다. 시스템 엔지니어링 접근방식에서 도출된 9가지 기준에 대해 관련 선박 위험 모델을 평가하기 위해 사용하였으며 이러한 기준은 기술적 신뢰성, 소프트웨어 성능, 인간과 기계의 상호성, 작동 및 통신 등과 같은 자율운항선박의 운영에 필요한 고려 사항을 평가하는 것을 목적으로 하였다.[13] Robert Veal et al(2017)은 자율운항선박 도입에 따라 중

래의 선박들에 적용되는 국제해사협약을 그대로 적용하는게 적절한지에 대해 해석적 검토를 중심으로 연구하였다. 특히나 UN해양법협약, 국제해상안전협약, 선원의 훈련, 자격증명 및 당직근무 기준에 관한 국제협약, 국제해상충돌예방규칙을 중심으로 하여 자율운항선박에 대한 항해 안전성 및 감항성이 확보된다면 종래의 선박들과 같은 범주에서 현재의 규범 적용이 가능하다고 판단하였다.[14] 마지막으로 Kim et al(2024)은 자율운항선박 관련 SCI 논문 3,363개를 과학계량학적으로 분석하여 연구트렌드를 살펴보았다. 그 결과, 이미지 인식 연구가 우세한 빠르게 성장하는 분야에서 딥러닝을 통한 해양 데이터 수집 및 분석에 초점을 맞추고 있음을 밝혀냈다. 이를 바탕으로 정책 입안자와 연구자에게 깊은 통찰력을 제공하여 전기 추진 시스템의 해상 산업 통합에 대한 미래 정책과 연구 방향을 제시하였다.[15]

2.2 텍스트 마이닝의 이론적 배경

2.2.1 텍스트 마이닝의 개념

기술의 발전에 따라 이전에 처리할 수 없었던 많은 양의 자료에 대한 분석이 가능해졌으며 정형데이터 뿐만 아니라 비정형데이터를 활용한 연구 수행이 가능하게 되었다.[16] 비정형 데이터는 기존 정형데이터와는 다르게 텍스트, 영상, 이미지 등 구조를 갖추지 않은 데이터로 비정형데이터를 활용할 경우, 기존 정형데이터에서는 추출하기 힘든 다양한 인사이트(Insight)를 추출하는 것이 가능하다. 특히, 비정형 빅데이터 중 텍스트 데이터는 텍스트 마이닝(Text Mining) 기법을 통해 유의미한 정보를 추출하는 것이 가능하다. 일반적으로 텍스트 마이닝은 비정형데이터 수집, 데이터 전처리, 정보 추출, 정보 분석의 절차를 따르고 있다.[17] 국내에서는 1998년부터 관련 연구가 시작되었으며, 2010년 이후 빅데이터 기반의 연구가 활발해지면서 관련 연구가 증가하였다.[18] 텍스트 마이닝의 주요 기법인 토픽 모델링(Topic Modeling), 시맨틱웹(Semantic Web), 온톨로지 기법(Ontology), 감성 분석

(Opinion Mining), 등을 활용하여 이미 많은 연구에서 유의미한 정보를 도출하고 있다.[16]

2.2.2 뉴스 데이터 텍스트 마이닝 관련 연구

뉴스는 사회 전반에 발생하는 다양한 이슈들을 반영하며, 무엇보다 객관성있는 사실과 핵심 정보를 정확하게 전달하는데 목적이 있어 정책 의제를 도출할 때 중요한 역할을 한다.[18] 또한 뉴스는 전문성 있는 정보를 신속하게 전달하는 매체로서 미래의 신호를 탐색하기 때문에 특정 사회 이슈를 파악하는데 중요한 연구 데이터가 될 수 있다.[19]

이러한 뉴스 데이터는 서로 완전하게 독립한 형태가 아닌 시간의 흐름에 따라서 데이터 간 영향을 받는 시계열 데이터이다. 따라서, 과거의 뉴스가 현재 뉴스에 영향을 주기 때문에 트렌드, 이슈, 사회적 흐름을 파악하며 미래를 예측하는 데 핵심적인 역할을 할 수 있다. 특히 최근에는 빅데이터를 활용한 미래 예측 연구가 점차 증가하면서 데이터베이스화가 잘 된 뉴스를 이용한 연구들이 지속해서 증가하는 추세이다.[20]

송혜민 외(2022)는 정책 결정 과정에서 대중들이 관심을 갖는 사회 이슈의 정책 과정 반영의 중요성을 강조하며 코로나19 관련 관광 이슈를 도출하기 위해 정부 정책자료와 온라인 뉴스를 토픽 모델링하였다. 이후 결과를 비교·분석하여 관광 위기 상황에서의 사회 이슈 파악과 정책 수립에 대한 체계적 탐색을 위한 기반을 제공하였다.[21] 김태중(2022)의 연구에서는 세계적으로 큰 관심을 갖고 있는 디지털 전환 현상에 대해 거시적 차원에서 분석하여, 이에 대한 시사점을 제시하기 위해 뉴스 빅데이터를 활용하여 토픽모델링을 수행하였다.[22] 노희경(2022)은 뉴스 빅데이터 분석을 통해 메타버스를 관광 분야의 공공 및 비즈니스 영역에 어떻게 적용해야 하는지에 대한 방향성을 제시하기도 하였다.[23] 자율운항선박과 유사한 모빌리티 분야를 살펴보면 김길래(2021)는 자율주행자동차의 안전 위험을 파악하고 국내 안전규제체계를 검토하기 위해 뉴스 기사를 수집하고 토픽 모델링으로 12개의

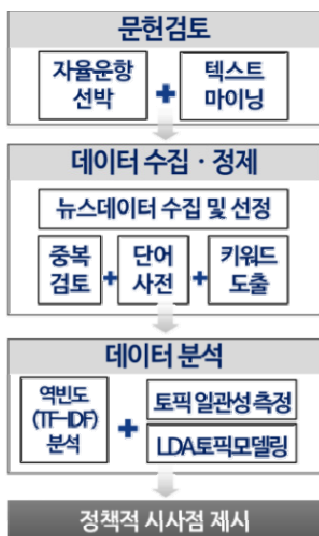
토픽을 도출, 해석하여 다양한 이슈를 파악하고 이를 바탕으로 안전한 자율주행차 개발 및 운영 환경 구축에 대한 시사점을 제시하였다.[24] 이 외에도 ESG, 탄소중립 등 최근 이슈화 되고 있는 사회적 문제에 대해 뉴스 데이터를 활용하여 분석하는 연구들이 증가하고 있다.[25] 이와같이 뉴스 데이터를 활용한 연구는 사회 전반에 걸쳐 복잡하고 불확실한 이슈에 대해 사회적 담론을 도출할 수 있으며 다양한 시사점을 제시할 수 있다.

3. 연구 방법

3.1 분석 방법

3.1.1. 연구개요

본 연구를 위해 그림 1과 같이 연구모형을 설계하였다. 먼저, 선행연구를 기반으로 자율운항선박의 개념과 특징을 살펴보았다. 그 이후 빅카인즈에서 자율운항선박 관련 뉴스데이터를 수집하였다. 수집한 데이터를 중복기사 제거, 용어 통일, 불용어 제거, 단어사전으로 필터링, 등의 과정을 거쳐 정제하였다. 정제된 데이터에 역빈도(TF-IDF)분석과 토픽 모델링을 적용하고 분석하여 정책적 시사점을 제시하였다.



[Figure 1] Research Models

3.1.2. 역빈도(TF-IDF) 분석

TF-IDF는 단어의 중요도 가중치이며, 값은 모든 문서에서 사용 될 수록 0에 가깝고 다수의 문항이 주어졌을 때, 특정 단어가 하나의 문서 내에서 얼마나 중요한지를 나타내는 통계 수치를 의미한다.[26] 단순히 빈도로 중요도를 측정하는 것이 아니라 문서 전체에서 단어의 분포도를 고려하여 가중치를 도출한다.[27] 따라서, TF-IDF 가중치를 사용하여 단순히 높은 빈도의 키워드보다 중요성이 높은 키워드 도출을 위한 방법으로 사용이 가능하며 수식은 다음과 같다.

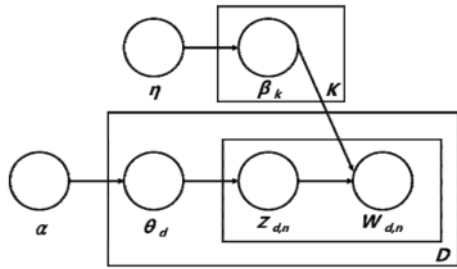
$$tfidf(t,d,D) = tf(t,d) \times idf(t,D)$$

[Figure 2] TF-IDF Arithmetic Expressions

3.1.3. LDA 토픽모델링

텍스트 마이닝은 비정형데이터의 일종인 텍스트 데이터로부터 의미 있는 정보를 추출하는 방법으로 토픽모델링이 대표적인 기법으로 사용된다.[28] 토픽모델링은 초기 잠재 의미 분석(Latent Semantic Analysis: LSA)와 확률론적 잠재 의미 분석(probabilistic LSA: pLSA)등 이 개발되었으나 pLSA가 보유한 학습 밖의 문서들에는 토픽 할당이 어렵다는 한계와 문서가 많아질 경우 과적합(overfitting)이 발생하였다. 이러한 문제를 해결하기 위해 제시된 것이 잠재 디리클레 할당(Latent Dirichlet Allocation: LDA)이다.[29]

LDA의 경우 문서가 여러 개의 토픽들로 구성되고 각 토픽들은 단어들로 표현될 수 있다고 전제하며 토픽과 단어의 확률분포에 대한 사전분포로서 디리클레(Dirichlet) 분포를 가정한다.[30] 이때 디리클레 분포를 가정하는 이유는 사후분포를 구하는 과정에서 디리클레 분포가 다항분포 함수와 곱해지면 다시 디리클레 분포 형태가 되는 ‘공액 사전 분포(conjugate prior distribution)가 발생하여 계산이 편리해지기 때문이다.[30] 이러한 전제하에 LDA 문서 생성의 과정은 아래 그림 3과 같다.



[Figure 3] LDA Topic Modeling

K는 토픽 수, D는 문서 수, N은 d번째 문서의 단어의 수를 나타낸다. α , β 는 하이퍼 파라미터 값으로 전체 문서에서 동일한 값을 갖는 특징이 있다. θ 는 문서별 토픽비율로 디리클레 분포를 따르며 α 값에 따라 형태가 결정되고 θ 는 문서별 토픽비율로 디리클레 분포를 따르며 α 값에 따라서 형태가 결정된다. Z는 단어에 할당된 토픽으로 다항분포를 따르며 문서별 토픽 비율인 θ 에 따라 결정된다. 각 단어에 토픽이 할당되면 토픽별 단어비율인 θ 에 따라서 단어 w가 결정된다.[31] 즉, ‘문서 내에서 토픽이 발생할 확률’에 따라 토픽을 선택하고 ‘토픽 내 단어의 출현확률’에 따라 하나의 단어를 선택하는 과정을 반복하며 문서를 생성하는 것이다.[32] 여기서 관찰이 가능한 단어(w) 외에 θ , ϕ , z의 경우 실제로 관측할 수 없는 잠재 변수이다. 따라서 LDA를 통한 토픽모델링은 역으로, 관찰된 자료인 문서 내 존재하는 단어들을 바탕으로 ‘문서 내에서 토픽이 발생할 확률’과 ‘토픽 내 단어의 출현확률’을 추정하는 과정이며 단어(w)의 집합인 문서(w)가 주어졌을 때 토픽을 추출하기 위해 사후분포를 계산하는 것이다.[33]

3.1.4. 일관성 점수(Coherence Score)

LDA를 이용한 토픽 분석에 있어 가장 어렵고 시행착오가 많은 부분이 토픽 수를 결정하는 일이다.[34] 따라서 토픽 수의 결정의 기준 지표로 일관성 점수(Coherence Score)를 사용하기도 하는데 본 연구에서는 U_Mass를 사용하였다. U_Mass은 아래와 같은 수식으로 계산한다.[35]

$$C_{UMass}(\omega_i, \omega_j) = \log \frac{D(\omega_i, \omega_j) + 1}{D(\omega_i)}$$

[Figure 4] U_Mass Arithmetic Expressions

수식에서 볼 수 있듯이, 토픽을 평가할 때, 토픽 내 단어의 동시 발생 확률의 측정에 기반하는데 U_Mass 방식의 측정값은 음수로 표현되는 확률의 로그 값이며, 0에 가까울수록 모델링이 잘 된 것으로 볼 수 있다.[35]

3.2 분석 대상

본 연구에서는 뉴스 빅데이터 분석 서비스인 ‘빅카인즈(www.bigkinds.or.kr)’를 활용하여 분석 데이터를 수집하였다. ‘빅카인즈’는 한국언론진흥재단에서 개발한 뉴스 분석 서비스로, 전국일간지, 지역일간지, 방송사, 경제지 등 다양한 언론사로부터 뉴스를 수집하여 빅데이터 분석 결과를 제공한다.[25] ‘자율운항선박’이 ‘무인 선박’ 등의 유사 단어와 혼용되어 사용하고 있음을 고려하여 박혜리(2018)의 자료를 참고하여 검색식을 작성하였다. 검색식으로는 ‘자율운항선박’ OR ‘스마트 선박’ OR ‘디지털 선박’ OR ‘무인 선박’ OR ‘자동화 선박’ OR ‘원격조종 선박’ 이다. 수집 기간은 IMO에서 자율운항선박의 개념을 처음 소개한 2018년부터 2022년까지의 5개년으로 지정하였으며 11개의 전국일간지를 대상으로 한 결과 2,921개의 기사에 대한 보도 일자, 언론사, 기고자, 제목, 분류, 인물, 위치, 특성추출, 본문 속성 중 특성추출 부분을 CSV형태로 수집할 수 있었다.

3.3 데이터 전처리

데이터의 신뢰성 확보를 위해 연구진 내부 협의를 통해 같은 내용을 신문마다 중복 게시한 기사와 본문에 키워드 ‘자율운항선박’이 존재하지만 기사의 핵심 내용은 자율운항선박과 관련이 없는 기사일 경우 분석 대상에서 제외하였다. 그리고 Python 3.12를 활용해 유의어, 지정어, 제외어, 포함어 등의 단

어사전을 만들어 전처리하였다. 제외어 사전은 자율 운항 선박과 직접적인 관련이 없는 키워드를 목록화 하였으며 뉴스 기사에서 많이 사용하는 ‘기자’ 등의 단어도 제외하였다. 유의어 사전의 경우 ‘해수부’를 ‘해양수산부’로 통일하였다. 표 3은 최종적으로 작성한 단어사전 예시이다.

<Table 3> Examples of lexicons

구분	예시
제외어 예시	관련, 기자, 단독, 속보, 언론, 뉴스, 독자, 정치, 제보, 이번, 한편, 가운데, 당시, 본격, 부분, 사회, 상황, 오늘, 오전, 오후, 이날, 일부, 지난달, 확인, 해당, 동안, 다음, 내용, 사실, 포토, 이후, 무단, 전제, 배포, 앵커, 영상, 화제, 사진, 가치, 적극, 반영, 저작권자, 세상, 노출, 제목, 작성일
유의어 예시	IMO : 국제해사기구, 국제 해사기구 울산 : 울산시, 울산특별시

최종적으로 총 2,843개의 기사를 선정하였으며 빅카인즈에서 제공하는 각 기사 별 특성추출 부분을 분석에 활용하였다. 특성추출 부분은 텍스트 랭크 알고리즘을 활용하여 검색 키워드와 한 문장내에서 동시에 출현하는 빈도를 계산해 뽑아낸 주요 키워드로 가중치순 상위 50개가 추출된다.[21] 이러한 텍스트 랭크 알고리즘은 자연어 처리 방식을 기반으로 하기 때문에 코더 간 신뢰도 측정이 필요가 없다는 장점이 있다.[37] 전처리가 완료된 각 단어를 대상으로 Python 기반의 NetMiner4.4 프로그램을 사용하여 역빈도(TF-IDF)분석과 LDA 토픽모델링을 실시하였다.

4. 연구 결과

4.1 역빈도(TF-IDF) 분석결과

본 연구의 대상으로 선정된 2,843개의 기사에 대해 추출된 명사 단어는 52,434개이며 TF-IDF 분석을 실시한 결과는 표 4와 같다. 5년간 데이터를 종합하여 주요 단어 등장 빈도를 알아보기 위해 상

위 50개 단어를 추출하였고 이를 통해 경향성을 살펴볼 수 있다.

<Table 4> TF-IDF Frequency Analysis Results

Rank	Keyword	TF-IDF Frequency	Frequency
1	조선업	578	1,037
2	HD현대	434	852
3	미국	432	440
4	경쟁력	309	317
5	해운업	305	489
6	스마트	299	354
7	부산	292	421
8	기술	288	349
9	중국	285	308
10	코로나	266	318
11	친환경	262	272
12	울산	258	429
13	경남	220	303
14	해양수산부	203	240
15	드론	198	221
16	실시간	194	194
17	일본	180	185
18	차세대	175	175
19	사업	173	209
20	연구	169	206
21	서울	169	175
22	기업	169	191
23	조선소	168	189
24	일자리	164	169
25	빅데이터	161	173
26	한화오션	155	173
27	전기	142	156
28	삼성중공업	142	142
29	정보	139	172
30	영국	138	141
31	경제	137	151
32	개발	137	148
33	인공지능	136	138
34	모빌리티	135	153
35	수소	133	208
36	자동	128	143
37	안전	125	186
38	반도체	124	127
39	연구원	122	155
40	센터	122	132
41	탄소	119	129
42	제조업	113	115
43	전문가	112	128
44	컨테이너선	110	111
45	운반선	108	134
46	추진	106	139
47	항공기	104	109
48	위원회	103	124
49	시스템	103	120
50	산업통상자원부	103	107

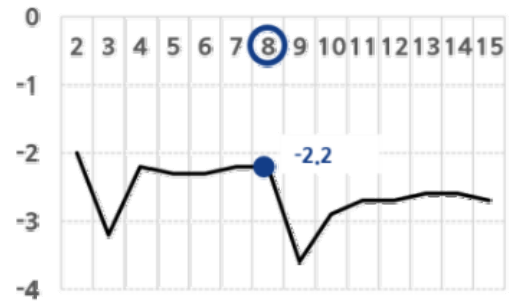
주요 단어를 살펴보면 ‘조선업’, ‘HD현대’, ‘미국’ 등이 있으며 이를 장소, 기관, 산업 등 3가지 분야로 분류하였다. 첫 번째 장소 분야의 경우 미국, 부산, 중국, 울산, 경남, 일본, 서울 영국 등이 상위단어로 나타났다. 해외의 경우 미국, 중국, 영국이 자율운항선박 관련 주요 국가이며 국내의 경우 부산, 울산, 경남, 서울에서 관련 연구 및 사업이 활발히 이루어지는 것으로 유추된다. 두 번째로 기관 분야에서는 HD현대, 해양수산부, 한화오션, 삼성중공업, 산업통상자원부 등이 뽑혔다.

국내 Big 3 조선사 모두 상위 50위에 속했으며 그 중 HD현대가 비중이 가장 높았다. 또한, 정부 부처 중에서는 해양수산부와 산업통상자원부가 자율운항선박 관련 주무 부처로 역할을 하고 있음을 유추해 볼 수 있다. 마지막으로 산업 분야에서는 조선업, 해운업, 경쟁력, 코로나, 차세대 등이 상위단어로 나타났다. 이는 자율운항선박이 포스트 코로나 시대에 조선업, 해운업의 차세대 신성장 동력으로써 경쟁력으로 갖추어야 할 분야로 인식했을 가능성이 높다. 또한 기술, 친환경, 수소, 빅데이터, 인공지능 등의 단어도 상위에 속했는데 이는 자율운항선박이 친환경인 요소를 지닐 뿐만 아니라 빅데이터, 인공지능 등 4차 산업혁명 기술을 융합하는 미래 선박임을 유추해 볼 수 있다.

4.2 LDA토픽 모델링 분석결과

4.2.1. 토픽 현황분석

최적화된 토픽 수를 결정하기 위해 W. Zhao 외 (2015)을 참고하여 α 값을 0.1, β 값을 0.01으로 설정하였으며 Iteration은 1,000으로 설정하였다.[36] 이후 토픽 수를 2부터 15까지로 설정하여 Coherence Score를 시뮬레이션 측정한 결과, 그림 5와 같이 가장 높은 점수(-2.2)가 나온 8개의 토픽으로 결정하였다.



[Figure 5] U_Mass measurement results

이후 각 토픽 별로 높은 빈도를 나타내는 상위 5개 키워드를 표 5와 같이 정리하고 토픽의 키워드 및 뉴스 원문을 검토하여 종합 토픽 명을 부여하였다. 각 토픽의 명칭은 연구 주제 관련 실무 경험 5년 이상인 내부 연구진 중심으로 심도 있는 논의를 통해 결정되었다. 초기의 토픽 분석 단계에서는 LDA 모델을 통해 도출된 각 토픽의 대표 키워드들을 바탕으로 관련 기사들을 철저히 검토하였다. 이후 연구진은 다음과 같은 기준에 따라 각 토픽의 명칭을 부여하였다. 먼저, 문맥적 의미를 살펴보았으며 각 키워드가 나타내는 문맥과 토픽 간 연관성을 주의 깊게 분석하였다. 키워드들이 토픽의 전반적인 주제와 어떻게 일치하는지를 고려하여, 가장 적합한 명칭을 선택하는 데 중점을 두었다. 다음으로 최종 명칭에 대한 결정은 전체 연구 팀의 합의를 통해 이루어졌다. 이 과정에서 모든 연구원들의 의견이 고려되었으며, 다양한 관점에서 토픽의 명칭을 검토하였다.

<Table 5> Topic-specific keywords

토픽	키워드
1	조선업, HD현대, 스마트, 한화오션, 조선소
2	부산, 안전, 해양수산부, 항만, 신항
3	해운업, 코로나, 중국, 경쟁력, 미국
4	미국, 중국, 드론, 영국, 일본
5	서해, 제주, 인천, 동해, 충남
6	경남, 부산, 울산, 창원, 조선업
7	울산, 수소, 기술, 전기, 모빌리티
8	HD현대, 미국, 조선업, 정기선, 아비커스

토픽 별 비율은 ①조선산업의 스마트화(16%), ②미래 항만산업과 자율운항선박(12%), ③포스트 코로나 시대의 해운업(16%), ④자율운항선박의 군사적 활용(14%), ⑤자율운항선박의 국내 시범운항 해역(8%), ⑥자율운항선박 관련 국내 주요도시(12%), ⑦자율운항선박의 핵심 클러스터 울산

(13%), ⑧자율운항선박 분야 국내 주요기업(10%)이며 토픽①과 ②의 비중이 가장 높았다.

정책적 시사점 도출을 위해서 각 토픽에 속하는 기사들의 내용을 정량적으로 검토한 후 토픽의 주요 주제 및 핵심 내용을 표 6에 정리하였다.

<Table 6> LDA Topic Modeling Analysis Results

구분	구분	설명
①	주요 키워드	조선업, HD현대, 스마트, 한화오션, 조선소
	토픽 명	조선산업의 스마트화(16%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 중앙일보(2021.10. 9.) “훈풍 부는 K-조선, 친환경 스마트 기술 앞세워 기세 잇는다.” 해석 : 정부는 친환경·스마트 선박의 개발·보급을 늘려 글로벌 주도권을 강화할 계획을 밝힘
②	주요 키워드	부산, 안전, 해양수산부, 항만, 신항
	토픽 명	미래 항만산업과 자율운항선박(12%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 한국일보(2020.11.17.) “2030년까지 37조 투입 부산항에 ‘한국형 스마트항만’ 가동”. 해석 : 해양수산부는 37조원을 투입해 자율운항선박과 연계할 수 있는 고부가가치 디지털 항만을 구축한다는 계획을 밝힘
③	주요 키워드	해운업, 코로나, 중국, 경쟁력, 미국
	토픽 명	포스트코로나 시대의 해운업(16%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 문화일보(2020. 6.29.) “한 중 일 교통 물류 분야 코로나19 대응에 협력...화상회의 개최”. 해석 : 비대면 선박검사를 비롯해 전염병 속에서도 국제 물류 운송이 이뤄질 수 있도록 검역, 방역 정보 등을 공유하기로 하였으며 해운물류 산업의 디지털화, 자율운항선박 개발 등에 협력하기로 함
④	주요 키워드	미국, 중국, 드론, 영국, 일본
	토픽 명	자율운항선박의 군사적 활용(14%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 서울일보(2022.05.26.) “미국이 원하는 대로 해줄 것” 세계 최초 中무인 항모 두고 ‘시끌’. 해석 : 미국, 중국, 일본, 영국 등 자율운항선박 관련 선도국가들은 자율운항선박 관련 기술을 상업적 이외에도 군사적으로도 활용
⑤	주요 키워드	서해, 제주, 인천, 동해, 충남
	토픽 명	자율운항선박의 시범운항 해역(8%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 동아일보(2021.03.26.) “스마트 선박 동해안 누빈다 울산 친환경 스마트 선박’ 관광사업 박차”. 해석 : 울산시에서 ‘정보통신기술(ICT)을 융합한 친환경 전기추진 스마트 선박의 계류지 위치 선정을 위한 공모를 진행
⑥	주요 키워드	경남, 부산, 울산, 창원, 조선업
	토픽 명	자율운항선박 관련 국내 주요도시(12%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 서울일보(2022.12. 1.) “울산 부산 경남, 조선업 기술경쟁력 강화”. 해석 : 지자체들이 자율운항선박 핵심기술 개발지원 등 제조산업인 조선산업에 기술경쟁력을 강화하는 추세
⑦	주요 키워드	울산, 수소, 기술, 전기, 모빌리티
	토픽 명	자율운항선박의 핵심 클러스터 울산(13%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 조선일보(2022.11. 3.) “울산 ‘자율운항 선박 실증센터’ 오늘 준공”. 해석 : 실증센터에 약 6년간 1,603억을 투자하여 실증 기반 마련
⑧	주요 키워드	HD현대, 미국, 조선업, 정기선, 아비커스
	토픽 명	자율운항선박 분야 국내 주요기업(10%)
	상세 내용	<ul style="list-style-type: none"> 주요 뉴스 : 국민일보(2022.01.09.) “현대중공업그룹, 미국선급협회 손잡고 자율운항선박 기술표준화 총력”. 해석 : 미국선급협회에서 HD현대의 자율운항 관련 기술에 대해 기본인증을 부여

4.2.2. 토픽 별 상세분석

본 연구에서는 각 토픽에 대한 상세 분석을 수행하기 위해 해당 토픽에 속한 뉴스 기사들을 전반적으로 검토하였다. 이 과정에서 상위 키워드를 주로 포함하는 뉴스 기사들을 주요 뉴스로 선정하고, 그 내용을 상세히 서술하였다. 첫 번째 토픽에서는 조선업, HD현대, 스마트, 한화오션, 조선소가 주요 키워드로 도출되어 '조선산업의 스마트화'로 명명하였다. 관련 뉴스를 살펴보면 문재인 전 대통령은 '21년 4월 확대경제장관회의를 주재한 자리에서 조선업계를 격려하는 한편 세계 1등 친환경 조선 강국으로 자리매김하겠다는 의지를 피력했다.[38] 발표 이후 해양수산부에서는 'K조선 재도약 전략'을 발표하였는데 친환경 및 자율운항 등 신기술을 접목한 선박기술 개발에 방점이 찍혀있었다.[39] 따라서 앞으로의 글로벌 조선산업에서 스마트·친환경 선박 및 건조공정에 대한 연구개발이 더욱 가속화 할 것으로 보이며 HD현대와 한화오션 등의 국내 조선소들은 현재 기술 우위를 유지하면서 자율운항선박 개발 등을 통해 가격경쟁력을 향상 시키는데 주력할 것으로 전망된다.

두 번째 토픽의 주요 키워드는 부산, 안전, 해양수산부, 항만, 신항이며 '미래 항만산업과 자율운항선박'으로 명명하였다. 자율운항선박의 완전 무인화가 이루어 질 경우 입출항 관리, 연료 충전, 최적 하역 등과 연계할 수 있는 고도화 된 항만 인프라가 필요할 것이다. 이에 따라 해양수산부는 광양항을 테스트 베드로 하여 시범운영을 시작해 오는 2030년까지 부산항 제2신항에 '한국형 스마트 항만'을 구축할 예정이다.[40] 관련 기사에 따르면, 10년 뒤 선보이게 되는 스마트 항만은 항만 내 자동화·디지털화 구축이 핵심이다. 자율운항 선박, 자율주행 트럭, 자동화 크레인 등을 이용해 선박과 부두의 접안, 배에서 항만으로 컨테이너를 싣거나 내리는 전 과정이 자동화되는 것을 의미한다. 또한 선사와 터미널 운영사, 세관·검역기관, 운송업체, 화주 등 항만 이용 주체가 항만의 상황과 화물정보를 실시간으

로 알 수 있도록 센서를 설치한다. 이를 통해 수집한 정보를 분석해 최적의 항만 운영계획을 세우고, 드론, 무인 수중 건설 로봇 등을 활용해 항만 인프라를 관리하게 된다.[41] 이와 같이 항만도 스마트화가 이루어지지 않을 경우 자율운항선박의 정상 운용이 안될 수 있으므로 이를 지원하는 항만 또한 고도화 될 것으로 예상된다.

세 번째 토픽에서는 해운업, 코로나, 중국, 경쟁력, 미국이 주요 키워드이며 이를 '포스트코로나 시대의 해운업'으로 명명하였다. 관련 기사를 살펴보면 코로나19 이후 시대를 준비하기 위해 한·중·일 3국은 앞으로 비대면 서비스 등 공동 협력 분야를 발굴할 예정이다. 또한 포스트코로나 시대를 준비하기 위해 비대면 선박검사, 스마트 기술에 기반한 해운물류 산업의 디지털화, 자율운항선박 개발, 원격교육 분야 개발 등을 위해 3개국 협력을 논의하였다.[42] 이와 같이 자율운항선박의 경우 자율주행차와 다르게 국내 운송 뿐만 아니라 국제해역을 향해하므로 국제해사협약 규정 검토 및 개정 또는 새로운 협약의 채택 등과 같이 국제사회의 협력을 통한 폭넓은 논의가 요구된다.[43] 따라서 단순히 국내 제도개선 및 규제혁신을 넘어 중국, 일본, 미국 등과 국제사회에서의 협력을 선제적으로 이끌어 나갈 필요가 있다.

네 번째 토픽은 '자율운항선박의 군사적 활용'으로 주요 키워드는 미국, 중국, 드론, 영국, 일본이다. 현재 미국의 경우는 군용 자율운항선박 수요 중심으로 기술개발을 진행 중에 있다.[44] 실제 분석결과에도 미해군에서는 자율운항선박으로 구축한 '유령함대'를 준비 중인 것으로 나타났다.[45] 관련 기사에 따르면 유령함대는 유·무인수상함과 수중드론이 혼재된 형태의 하이브리드 함대이다. 유령함대의 전략은 사람도 없고 적의 레이더에 잘 탐지가 되지 않는 '유령'과 같은 무인함과 드론을 선두에 세워 정보를 수집하고 방어망에 구멍을 내면, 이를 뒤따르는 유인 구축함 등이 공격에 나선다는 개념이다. 이러한 유령함대는 중국 해군을 겨냥해 만든 것으로 추정된다.[45] 미군보다 더 많은 전함을 보유한 중

국 해군에 대응하기 위해 건조 비용이 저렴하면서 유인 전투함 수준의 작전 능력을 갖춘 무인 함정을 개발한 것인데 태평양에 주둔하는 7함대사령부에 유령함대를 배치한 이유는 대만에서 유사한 상황이 발생 할 경우 즉시 출동시킬 수 있기 때문이다.[46] 이에 중국 또한 '22년 5월 무인 항행이 가능한 군함 '주하이윈(珠海云)'을 진수하였다.[47] 이 군함은 길이 88.5m, 폭 14m에 배수량 약 2천t으로 갑판이 넓어 수십대의 공중·해상·잠수용 드론을 탑재할 수 있다. 일본 또한 바닷속에서 항해하며 정보를 수집할 길이 10m 이상의 대형 수중 드론 개발을 추진 중에 있으며, 영국의 롤스로이스도 인공지능 지원으로 승조원 등 사람의 개입 없이 100일 동안 임무 수행이 가능한 대형 무인함을 설계 건조해 세계 해군을 상대로 판매할 계획을 발표하기도 하였다.[48] 이와 같이 강대국들이 경쟁적으로 해양 무인전력을 개발하는 현 시점에서 우리나라 또한 '해양 무인체계로의 중심 전환을 고려해 볼 필요가 있다.

다섯 번째 토픽에서는 서해, 제주, 인천, 동해, 충남이 주요 키워드로 도출되었으며 이를 '자율운항선박의 국내 시범운항 해역'으로 명명하였다. 현재 서해, 동해 등 전국의 해역에서 시범적으로 자율운항이 이루어지고 있는 것으로 나타났다. 중요 키워드를 중심으로 관련 뉴스를 살펴보면 삼성중공업이 국내 조선업계 최초로 서해·남해·동해를 잇는 도서 연안 해상 자율운항 실증에 성공한 사례가 보도되었다.[49] 삼성중공업은 목포해양대학교의 9,200t급 대형 실습선인 세계로호를 활용해 자율운항 실증에 나섰으며 전남 목포 서해상에서 출발해 남해 이어도와 제주도를 거쳐 동해 독도에 이르는 약 950km를 자율운항했다. 이와 같이 3면이 바다며 섬이 많은 국내의 해양 환경 특성상 복잡하고 다양한 충돌 위험 상황을 가정하여 자율운항 성능과 안정성을 확인하는 검증 작업을 계속할 필요가 있다.

여섯 번째 토픽의 주요 키워드는 경남, 부산, 울산, 창원, 조선업으로 토픽을 '자율운항선박 관련 국내 주요도시'로 명명하였다. 경남, 부산, 울산, 창원은 모두 조선소가 위치한 도시라는 공통점 이외에도

지역경제 통합을 위해 2022년 시행된 「지방자치법 제12장 특별지방자치단체」에 근거를 두고 '부산울산경남특별연합'을 출범하기도 하였다.[50] 기사 내용에 따르면 부울경 주력 전략사업으로 자동차 산업, 조선산업, 항공산업에 대해 집중 육성할 계획이다. 특히 조선산업의 경우 친환경 선박 기술개발과 디지털화에 집중 투자할 계획이며 생산성 향상을 위한 조선업 제조현장을 친환경, 디지털화한다고 밝혔다. 나아가 친환경 선박 기자재 국산화 선점을 위한 기반을 마련하고 조선업과 해운업을 연계하는 등의 협력 강화를 추진할 예정이다. 따라서 앞으로 자율운항선박 사업은 부울경 지역을 중심으로 기술개발이 이루어질 것으로 전망된다.

일곱 번째 토픽에서는 울산, 수소, 기술, 전기, 모빌리티가 주요 키워드로 나타나 '자율운항선박의 핵심 클러스터 울산'으로 토픽을 명명하였다. 여섯 번째 토픽에서 살펴본 바와 같이 부울경 지역을 중심으로 자율운항선박 기술개발이 이루어지고 있지만 특히 울산시의 관심이 높은 것으로 보인다. 관련 기사를 살펴보면 2022년 국내 최초로 울산 동구에 자율운항선박 성능실증센터가 준공되었다..[51] 이 센터에서는 해양수산부와 산업통상자원부가 추진 중인 '자율운항선박 기술개발사업'에서 개발되는 각종 장비와 시스템을 탑재해 평가·검증을 수행하고, 다양한 첨단 조선해양기자재의 성능 검증에도 활용된다. 또한 울산시는 국내 최초로 정보통신(ICT) 기술이 융합된 최첨단 전기 추진 스마트 선박인 '울산태화호'를 건조하였다.[52] 울산태화호는 총사업비 448억이 투입되었으며 선박 통합제어 시스템, 스마트 운항 보조 시스템, 이중 연료 엔진 시스템 등 32개에 국내에서 개발된 기술이 적용되었다. 이와 같이 울산은 자율운항선박 실증센터를 중심으로 기술개발과 실증을 위한 기반이 조성돼 있으며, 현대중공업과 현대미포조선 등 세계 최고 수준의 자율운항선박기술의 주요 기업이 위치하고 있어 앞으로 자율운항선박의 핵심 클러스터로 성장할 가능성이 높아 보인다.

여덟 번째 토픽의 주요 키워드는 HD현대, 미국,

조선업, 정기선, 아비커스이며 '자율운항선박 분야 국내 주요기업'으로 토픽을 명명했다. 앞서 역빈도(TF-IDF)분석 결과 HD현대의 키워드 빈도수가 높게 나온 것처럼 국내 조선사 중 자율운항선박 분야에 가장 적극적으로 투자하는 것으로 보인다. 토픽 관련기사에 따르면 정기선 HD현대 대표이사 사장은 2017년 부사장으로 승진하며 HD현대의 1호 사내벤처인 아비커스를 설립하였다.[53] 또한, 최초 60억 원을 출자해 아비커스를 자회사로 설립한 이후로도 5번의 유상증자를 통해 총 339억 원을 투입하는 등 아비커스에 지원이 필요할 때마다 출자자로 나서고 있다. 이러한 투자 덕분에 아비커스는 국내 최초로 12인승 크루즈 선박 완전 자율운항에 성공했고, 미국선급협회(ABS)와 선박 자율운항기술 단계별 기본 인증(AIP) 및 실증테스트 추진을 위한 양해각서를 체결하기도 하였다. 이와 같이 앞으로도 HD현대의 자율운항선박에 대한 적극적인 투자는 지속될 것으로 전망된다.

5. 결론 및 시사점

5.1 결론

본 연구는 자율운항선박 관련 뉴스 데이터를 활용하여 주요키워드와 토픽을 파악함으로써 자율운항선박의 주요 의제와 시사점을 도출하기 위해 수행되었다. 이를 위해 IMO에서 자율운항선박이 본격적으로 논의된 2018년부터 2022년까지 11개의 국내 주요일간지에서 보도된 2,843개의 뉴스에 대해 역빈도(TF-IDF)분석과 LDA 토픽모델링을 실시하였다. 역빈도(TF-IDF)분석 결과, 주요 단어들은 '조선업', 'HD현대', '미국' 등으로 이들을 장소, 기관, 산업의 세 분야로 분류하였다. 장소 분야에서는 미국, 중국, 일본, 영국 등이 주요 국가로 나타났으며 부산, 울산, 경남, 서울이 국내 주요 지역으로 나타났다. 기관 분야에서는 HD현대, 한화오션, 삼성중공업, 해양수산부 산업통상자원부 등이 주요 기관으로 강조되었다. 산업 분야에서는 조선업, 해운업, 경쟁

력, 코로나, 차세대 등이 주요 단어로, 자율운항선박이 포스트 코로나 시대의 신성장 동력으로 간주된다. 또한, 빅데이터, 인공지능 같은 4차 산업혁명 기술과 결합된 친환경 미래 선박으로 인식될 가능성이 높음을 시사하였다. LDA 토픽모델링 분석 결과 Coherecn Score 점수가 가장 높은 8개의 토픽으로 결정하여 ①조선산업의 스마트화, ②미래 항만산업과 자율운항선박, ③포스트코로나 시대의 해운업, ④자율운항선박의 군사적 활용, ⑤자율운항선박의 국내 시범운항 해역, ⑥자율운항선박 관련 국내 주요도시, ⑦자율운항선박의 핵심 클러스터 울산, ⑧자율운항선박 분야 국내 주요기업 등 8개의 토픽을 도출하였다. 5년 치를 통합하여 분석한 결과, '①조선산업의 스마트화' 토픽과 '②포스트코로나 시대의 해운업'이 각각 16% 비중을 차지하며 가장 언론의 관심이 높은 것으로 나타났다. 반면에 '⑤ 자율운항선박의 시범운항 해역' 토픽이 8% 비중을 차지하며 가장 언론의 관심이 낮은 것으로 나타났다. 이러한 연구결과를 중심으로 정책·사회·국제안보의 관점에서 시사점을 도출하였다.

첫째, 정부의 정책 추진 관점에서 각 분야 현황을 고려하여 단계적이며 신중한 추진이 필요하다. 자율운항선박은 역빈도(TF-IDF) 분석결과와 같이 빅데이터, 인공지능 같은 4차 산업혁명 기술과 결합된 친환경 미래 선박으로 인식될 가능성이 높을 뿐만 아니라 토픽2와 토픽3에서 나타난 것처럼 단순히 조선산업 이 외에도 항만산업, 그리고 해운 산업까지도 막대한 영향을 미칠 것으로 전망된다. 이러한 추진 과정에서 너무 급진적인 자율운항선박의 도입과 운용은 여러 부작용을 초래할 수 있다. 예를 들어, 항만 인프라의 미비로 인해 유인선박 운용과의 혼선으로 안전사고가 발생할 수 있고, 사고 발생 시 법적 책임 및 보험처리 문제, 윤리 문제 등이 예상되는 만큼 단계적이고 신중한 접근이 필요할 것이다.[47] 따라서 자율운항선박과 관련한 이해관계자 모두의 폭넓은 합의 및 참여를 포함할 수 있는 정책 추진을 제안한다.

둘째, 사회적 관점이다. 분석 결과 도출된 대부분

의 토픽은 투자, 개발과 같이 자율운항선박의 긍정적인 부분에 집중되어 보도되고 있으며 사이버 보안 등의 안전 문제, 선원 일자리 감소 문제 등과 같이 부정적인 토픽은 발견되지 않았다. 자율운항선박과 유사한 자율주행차의 뉴스 데이터를 분석한 김길래(2021)의 연구에 따르면 도로주행 테스트 중에 발생하는 위험, 개인정보와 사이버보안 위험, 그리고 사람과 기계와의 커뮤니케이션에서 발생할 수 있는 위험 등에 대한 자율주행차 관련 사회이슈를 도출하였다.[24] 이러한 안전문제 이슈는 자동차 뿐만 아니라 자율운항선박에서도 충분히 발생할 수 있는 이슈라고 판단된다. 특히, 자동차 사고와는 달리 선박이 충돌할 경우 수많은 사람의 사상을 초래하고 해양사고의 특성상 선박이 전복, 매몰되어 파괴될 경우 수억 내지 수천억 원의 물적 피해를 야기하며 기름이나 유해물질 배출로 인해 해양오염까지 초래하기 때문이다.[7] 자율운항선박의 경우도 실제로 상용화되기 전 다양한 제도의 정비가 선행되어야 하지만 이를 위해서는 장기간의 노력이 필요하므로 제도 정비 범위와 상용화 시점에 대한 전략적 선택이 필요할 것이다. 실제로 TF-IDF에서 주무 부처로 뽑힌 해양수산부는 '자율운항선박 선제적 규제혁신 로드맵'을 추진하고 있으나 무리한 규제혁신은 안전사고 등의 문제 발생 시 대응이 곤란하게 될 수 있다. 반면에 모든 제도를 정비하고 이해관계자 간 합의 후에 상용화를 추진하는 것 또한 장기간의 노력이 소요되어 현실적으로 어려운 것이 사실이다. 따라서 제도 정비 이전 혹은 이후 중의 이분법적 사고가 아니라 자율운항선박 상용화에 필요한 최소한의 제도를 마련한 이후 상용화를 추진하는 것이 바람직할 것이다. 나아가 자율운항선박 보다 앞서가고 있는 자율주행차의 사례를 살펴보면 사용자 안전 문제 이슈, 사이버 보안 이슈, 고용 감소 이슈 등의 사회문제를 어떻게 정책적으로 대응했는지 면밀히 검토한 후 자율운항선박 정책수립에 참고할 필요가 있다.

셋째, 안보적 관점이다. 토픽4에서 살펴본 바와 같이 자율운항선박은 경제적 측면에서 조선업, 해운업 등의 패러다임을 바꿀 뿐만 아니라 안보적 측면

에서 미래 해양전의 패러다임도 바꿀 것으로 보인다. 실제로, 역빈도(TF-IDF) 분석에서 주요 국가로 뽑힌 미국과 중국과 같은 자율운항선박 기술선도 국가들이 자율운항선박 기술을 상업적으로만 투자를 하는 것이 아니라 군사적으로도 무인 전력에 활용 중인 것으로 나타났다. 이러한 시점에서 우리나라 또한 '해양무인체계로의 중심 전환'을 검토해야 할 것이다. 현재 우리 해군에서도 변화하는 환경에 맞춰 해양무인체계 기반의 첨단 전력 건설을 추진하고 있다. 먼저, 해군 창설 100주년이 되는 2045년까지 해군비전 2045를 구현하기 위한 'SMART NAVY'를(Strong Maritime forces Accomplished with Revolutionary Technology) 추진 중에 있으며, 현재 3개로 편성된 함대사령부를 2040년까지 2개로 줄이는 대신 무인 수상함전대·무인 잠수정전대·무인 항공기전대로 구성된 해양 무인전력사령부를 창설할 계획이다.[54] 다만, 현재 우리 해군의 해양무인체계 기반 전력 건설은 다소 추상적 개념만이 제시되어 있는 수준이다.[55] 따라서 자율운항선박 기술을 기반으로 하여 해양무인 체계에 대한 명확한 개념을 정립하고, 군 주도로 관련 기술의 확보 및 개발이 필요할 것이다.

본 연구는 뉴스 데이터를 대상으로 자율운항선박의 현안과 동향을 분석함으로써, 자율운항선박에 대한 시사점을 제안하였고 정책적 방향 및 주요 논점을 제시했다는 정책적 의의가 있다. 또한, 선행연구에서는 자율운항선박 관련 법·제도를 중심으로 고찰되어 왔으나, 본 연구에서는 국가·지역·전략·안보 등의 다차원 관점에서 자율운항선박에 대한 사회적 이슈를 분석했다는 학술적 의의가 있다.

5.2 연구의 한계

본 연구의 한계와 제안 사항은 다음과 같다. 첫째, 본 연구는 국내 자율운항선박 관련 뉴스만을 대상으로 진행하였다. 그러나 현재 해외에서도 자율운항선박과 관련한 뉴스 및 실증데이터가 다양하게 생산되고 있으며 관심도도 증가하고 있다. 그렇기 때문에 추후 연구에서는 국제 뉴스와 비교하여 국내 뉴스와

어떠한 차이가 있는지를 분석한다면 더욱 의미있는 연구결과가 도출 될 것으로 기대된다.

둘째, 본 논문은 자율운항선박과 관련한 뉴스 데이터만을 한정하여 분석하였다. 추후 자율운항선박 관련 논문, 특허 데이터 등을 활용한 연구가 수행되기를 바라며 다양한 데이터를 활용한 폭넓은 변화의 쟁점에 대해 연구가 진행되길 기대해 본다.

셋째, 본 연구에서는 뉴스데이터의 수집 기간을 IMO에서 자율운항이 본격적으로 논의된 2018년부터 2022년까지의 5개년에 걸쳐 수집한 국내 주요 일간지의 뉴스데이터만을 대상으로 자율운항선박의 초기 이슈에 대한 분석을 탐색적으로 수행하였다. 다음 연구에서는 분석 기간을 더욱 확장하거나 경제지 등 여러 매체까지 확장하여 분석한다면 자율운항선박으로 인해 어떤 사회적 담론이 형성되었으며, 사회 전반에 어떠한 영향을 미쳤는지 심도 있는 논의를 진행할 수 있을 것이다. 또한, 시계열 분석을 추가하여 자율운항선박에 대한 사회적 인식과 논의가 시간이 지남에 따라 어떻게 변화하였는지, 특정 사건이나 정책 발표가 이러한 변화에 어떠한 영향을 미쳤는지를 구체적으로 분석하는 것도 가능할 것이다. 이러한 심화된 연구는 자율운항선박 기술의 사회적 수용과 관련 정책 수립에 있어 중요한 통찰을 제공할 것이다.

그럼에도 불구하고 본 연구는 자율운항선박과 관련한 뉴스기사를 최초로 분석하여 정책적 제언을 했다는 학술적 의의가 있다. 앞으로 자율운항선박 도입이 정책·사회·국제안보에 어떠한 영향을 미칠지 정확하게 예측하기는 어려울 것이다. 하지만 이러한 변화를 예상하고 어떻게 정책적으로 대응할 수 있는지 선제적으로 논의할 필요는 있다. 본 연구가 자율운항선박 관련 정책을 수립하는데 선행연구로 활용될 수 있기를 기대한다.

사 사

이번 연구는 과학기술정보통신부의 지원을 받아

수행되었다. 과학기술정책 전문인력 양성 및 지원사업 (N2020-0295).

References

1. 박성건, 황영찬, “4차 산업혁명의 스포츠 현장 적용을 위한 탐색적 연구: 소셜 빅데이터 활용 방안을 중심으로”, 한국체육학회지, 56(4): 397-413, 2017.
2. 중소조선연구원, “조선해양 ICT융합 기술 분야”, <https://www.rims.re.kr/sub02/sub03.php>, (검색일: 2024.03.15).
3. 김부기, 「선박의 이해」, 부산: 한국해양과학기술원, 2022.
4. 이상일, 「자율운항선박의 현황과 법적쟁점」, 부산: 다솜출판사, 2021.
5. International Maritime Organization (IMO), “MSC 98th Agenda Item 20, Maritime Autonomous Surface Ships (Proposal for a regulatory scoping exercise)”, 2017.
6. 박혜리, 「자율운항선박 도입 관련 대응정책 방향 연구」, 부산: 한국해양수산개발원, 2018.
7. Isaac Arnsdorf, 「Roll-Royce Drone Ships Challenge \$375 Billion Industry: Freight」, Bloomberg Online, 2014.
8. Ortiz de Rozas, J. M., “The Production of Unmanned Vessels and its Legal Implications in the Maritime industry” (Master's thesis), 2014.
9. Paul W. Pritchett, “Ghost Ships: Why the Law Should Embrace Unmanned Vessel Technology.” Tulane Maritime Law Journal, 40: 197, 2015.
10. 최정환, 이상일, “무인선박의 선박성에 관한 해석적 고찰”, 해양정책연구, 33(1): 171-191, 2018.
11. 이현균, “자율운항선박촉진법의 제정 필요성 및

- 제정 방향”, 과학기술법연구, 28(3): 121-163, 2022.
12. 김민영 · 정치윤. “자율운항선박의 국제해상충돌 예방규칙 준수를 위한 합성곱 신경망 기반의 선박 분류에 관한 연구”, 한국지능시스템학회 논문지, 29.4: 251-257, 2019.
 13. Ramos, M. A · Thieme, C. A, “Human-System Interaction in Autonomy Method-a Structured Approach to Risk Monitoring”, In Proceedings of the First International Workshop on Autonomous Systems Safety. NTNU, 2019.
 14. RINGBOM, Henrik Mikael VEAL Robert, “Unmanned ships and the international regulatory framework.” Journal of International Maritime Law, 23.2: 100-118, 2017.
 15. Kim J, Han S, Lee H, Koo B, Nam M, Jang K, Lee J, Chung M, “Trend Research on Maritime Autonomous Surface Ships (MASSs) Based on Shipboard Electronics: Focusing on Text Mining and Network Analysis”.Electronics, 13(10):1902, 2024.
 16. 정인석, 강지원, 이종덕, 박상민, “NTIS 데이터를 이용한 국내 자율주행 연구 동향 분석에 관한 연구”, 한국ITS학회 논문지, 22(2): 147-163, 2023.
 17. 오창석, 이용택, 고민수, “텍스트마이닝을 활용한 도로분야 ITS 정책이슈 탐색기법 정립”, 한국ITS학회 논문지, 15(6): 10-23, 2016.
 18. 황기연, 송재인, 이자영, 임이정, “텍스트 마이닝 기법을 활용한 자율주행자동차 인식분석연구”, 한국ITS학회 논문지, 2017.
 19. 나경식, 이지수, “신문 빅데이터를 바탕으로 본 국내 정보화의 경향과 도서관의 역할”, 한국콘텐츠학회논문지, 18(9): 14-33, 2018.
 20. 박대민, “장기 시계열 내용 분석을 위한 뉴스 빅데이터 분석의 활용 가능성”, 한국언론학보, 60(5): 353-407, 2016.
 21. 송혜민, 김홍범, “토픽모델링을 활용한 코로나19 관련 관광이슈 및 정부 정책 분석: 정부 정책자료와 온라인 뉴스 비교분석을 중심으로”, 관광학연구, 46(3): 161-182, 2022.
 22. 김태중, “뉴스 빅데이터를 활용한 코로나19 언론보도 분석: 토픽모델링 분석을 중심으로”, 한국콘텐츠학회논문지, 20(5): 457-466, 2020.
 23. 노희경, “뉴스 빅데이터를 활용한 관광분야 메타버스관련 이슈 분석”, 관광레저연구, 34(2): 151-166, 2022.
 24. 김길래, “자율주행자동차 안전위험 이슈 도출과 안전규제체계에 대한 연구”. 디지털콘텐츠학회 논문지, 22(3), 547-554, 2021.
 25. 김은정, 장석권, 이상용, “토픽모델링과 시계열 회귀분석을 활용한 헬스케어 분야의 뉴스 빅데이터 분석”, 경영정보학연구, 25(3): 163-177, 2023.
 26. 한성수, 양동우, “텍스트마이닝을 이용한 창업 관련 연구 동향 분석”, 벤처창업연구, 12(5): 1-12, 2017.
 27. 이예은, 장정현, “비정형 데이터 수집과 TF-IDF를 통한 이동학대 분석 및 키워드 추출”, 한국범죄심리연구, 18(4): 171-182, 2022.
 28. 박상민, 이환필, 소재현, 윤일수, “텍스트 임베딩을 이용한 자율주행자동차 교통사고 분석에 관한 연구”, 한국ITS학회 논문지, 20(1): 160-173, 2021.
 29. Blei D. M., Ng A. Y., and Jordan M. I., “Latent Dirichlet Allocation,” The Journal of Machine Learning Research, vol. 3, pp.993-1022, 2003.
 30. 유예림, “빅데이터 분석 기법을 활용한 2015 개정 교육과정 정책에 대한 언론보도 분석.”, 서울대학교 대학원 박사학위 논문, 2017.
 31. 이상엽, “토픽모델링을 활용한 교통경찰 민원 분석”, 한국ITS학회 논문지, 20(4): 57-70,

- 2021.
32. 유예립, “빅데이터 분석 기법을 활용한 2015 개정 교육과정 정책에 대한 언론보도 분석”, 서울대학교 대학원 박사학위 논문, 2017.
 33. 양혜진, 안정민, 이태현, “텍스트 빅데이터 분석을 통한 한국인의 불공정 경험 분석: 국민청원 게시판 데이터 분석 결과를 중심으로”, 조사연구, 22(1), 25-59, 2021.
 34. 양철원, 야마다 아키히코, “토픽모델링을 활용한 ‘산업연구’ 학술지 분석 - 과거와 미래에 대한 고찰”, 산업연구, 46(2), 29-57, 2022.
 35. Mimno, D., Wallach, H., Talley, E., Leenders, M., McCallum, A., “Optimizing Semantic Coherence in Topic Models”, In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 262-272, 2011.
 36. Zhao, W.; Chen, J.; Zen, W. “Best Practices in Building Topic Models with LDA for Mining Regulatory Textual Documents” CDER, Silver Spring, MD, USA, 2015
 37. 윤순옥 · 김민철, “LDA 기법을 이용한 미세먼지 이슈의 토픽모델링 분석”, 에너지공학, 29(2), 23-29, 2020.
 38. 중앙일보, “훈풍 부는 K-조선, 친환경 스마트 기술 앞세워 기세 이어간다”, 2021년 10월 9일.
 39. 해양수산부, “20230 항만정책 방향과 추진전략”, 2021.
 40. 한국일보, “2030년까지 37조 투입 부산항에 ‘한국형 스마트항만’ 가동”, 2020년 11월 17일.
 41. 해양수산부, “K조선 재도약 전략”, 2022.
 42. 문화일보, “한 중 일 교통 물류 분야 코로나19 대응에 협력...화상회의 개최”, 2020년 6월 29일.
 43. 최정환, 이상일, “무인선박의 선박성에 관한 해석적 고찰”, 해양정책연구, 33(1), 171-191, 2018.
 44. 과학기술정보통신부, “과학기술&ICT정책·기술 동향”, NO.191, 2021.
 45. 동아일보, “中 해군력 급속 팽창... 우크라이나 다음 분쟁은 서태평양?”, 2022년 6월 11일.
 46. 김성환, “미중 패권경쟁과 한국의 대응전략에 관한 연구”, 한국해양대학교 대학원 박사학위논문, 2023.
 47. 서울일보, “미국이 원하는 대로 해줄 것” 세계 최초 中무인 항모 두고 ‘시끌’, 2022년 5월 26일.
 48. 동아일보, “인공지능 선장이 이끄는 ‘新 대항해 시대’ 성큼”, 2021년 8월 8일.
 49. 한겨레, “자율운항 선박, 충돌 위기 29번 피하며 국내 연안 950km 항해”, 2022년 11월 24일.
 50. 한겨레, “전국 첫 특별지자체 ‘부산울산경남특별연합’ 출범”, 2022년 4월 19일.
 51. 동아일보, “울산 ‘자율운항선박 성능실증센터’ 문 열어”, 2022년 11월 4일.
 52. 경향신문, “울산시, 전기추진 스마트 선박 ‘대화호’ 건조...내년부터 해양관광사업에 투입”, 2022년 11월 29일.
 53. 조선일보, “‘사장 승진 1년’ 정기선, 현대중공업 흑자 눈앞... 자율운항·친환경도 순항”, 2022년 10월 12일.
 54. 정광호, 김동은, “미국 해양전략의 변화가 한국 해군에 주는 전략적 함의: 해양 기동전으로의 변화를 중심으로”, 한국군사학논집, 78(3): 57-89, 2022.
 55. 류재학, 허지용, 나윤후, “미래전 활용을 위한 해양무인체계 발전전략에 관한 연구”, Journal of the KNST, 6(2): 127-132, 2023.

[Appendix] 본 연구에서 인용한 뉴스 기사

매체명	기사 제목	기사 게시일
중앙일보	훈풍 부는 K-조선, 친환경 스마트 기술 앞세워 기세 이어간다.	2021년 10월 9일
한국일보	2030년까지 37조 투입 부산항에 ‘한국형 스마트항만’ 가동 “	2020년 11월 17일
문화일보	한 중 일 교통 물류 분야 코로나19 대응에 협력...화상회의 개최	2020년 6월 29일
동아일보	中 해군력 급속 팽창... 우크라이나 다음 분쟁은 서태평양?	2022년 6월 11일
서울일보	“미국이 원하는 대로 해줄 것”세계 최초 中무인 항모 두고 ‘시끌’	2022년 5월 26일
동아일보	인공지능 선장이 이끄는 ‘新 대항해시대’ 성큼	2021년 8월 8일
한겨레	자율운항 선박, 충돌 위기 29번 피하며 국내 연안 950km 항해	2022년 11월 24일
한겨레	전국 첫 특별지자체 ‘부산울산경남특별연합’ 출범	2022년 4월 19일
동아일보	울산 ‘자율운항선박 성능실증센터’ 문열어	2022년 11월 4일
경향신문	울산시, 전기추진 스마트 선박 ‘태화호’ 건조...내년부터 해양관광사업에 투입	2022년 11월 29일
조선일보	사장 승진 1년’ 정기선, 현대중공업 흑자 눈앞... 자율운항·친환경도 순항	2022년 10월 12일