

# 초소형모듈원자로 분야의 효율적인 두문자어 사전구축과 단어 임베딩을 이용한 지식 분석

최정우\*, 이병희\*\*

## Analyzing Knowledge through the Efficient Construction of an Initialism Dictionary and Word Embeddings in the Field of MMR

Jung-Woo Choi\*, Byeong-Hee Lee\*\*

---

본 연구는 2025년 한국과학기술정보연구원의(KISTI) 기본사업 과제로 수행한 것입니다(과제고유번호 K25L3M2C4)

---

### 요 약

본 연구는 MMR 관련 국내외 기술 동향과 연구개발 현황을 분석하여 우리나라의 기술 경쟁력 강화를 목적으로 한다. 이를 위해 본 연구는 R언어를 이용하여 우리나라 국가R&D과제 17건과 국제 논문 240건을 수집하여 처리하고, 국제 논문 전문 텍스트(Full-text) 데이터를 가지고 두문자어를 추출하여 사전을 구축하고, 전문 텍스트를 단어 임베딩하고 지식 분석을 수행하였다. 연구 결과 우리나라는 국가R&D 측면에서는 현재 기초연구 단계에 머물러 있으나, 미국은 민간 혁신과 시장 중심 접근으로 빠르게 상업화에 도달하고 있었다. 이를 통해 우리나라와 미국의 MMR 연구개발을 비교하고 토픽모델링과 토픽 실행계획을 마련하여 정책적 방향을 제시한다.

### Abstract

This study aims to enhance Korea's technological competitiveness by analyzing domestic and international technology trends and the current status of R&D related to MMR. To achieve this, the study utilizes the R programming language to collect and process data from 17 Korean national R&D projects and 240 international papers. It extracts initialisms from the full text of these papers to create a dictionary, embeds words for knowledge analysis. Results show that while our country is currently at the basic research stage in terms of national R&D, the United States is quickly reaching commercialization through private innovation and a market-oriented approach. By this research, we compares MMR R&D efforts in Korea and the United States, prepares topic modeling and action plans, and provides policy recommendations.

### Keywords

micro modular reactor, national R&D projects, acronym, word embedding, topic modeling

---

\* 한국과학기술정보연구원 NTIS센터 학생연구원, 과학기술연합대학원대학교 과학기술경영정책학과 석.박사 통합과정 · Received: Nov. 28, 2024, Revised: Feb. 27, 2025, Accepted: Mar. 02, 2025  
- ORCID: <https://orcid.org/0000-0002-7124-8789> · Corresponding Author: Byeong-Hee Lee  
NTIS Center, Korea Institute of Science and Technology Information  
Tel: +82-42-869-1724, Email: bhlee@kisti.re.kr

\*\* 한국과학기술정보연구원 NTIS센터 책임연구원, 과학기술연합대학원대학교 과학기술경영정책학과 교수(교신저자)  
- ORCID: <https://orcid.org/0000-0002-5379-9659>

## 1. 서 론

### 1.1 연구의 필요성

2022년 11월말 대규모언어모델(LLM, Large Language Models)인 챗GPT의 발표 이후 여러 LLM이 등장하면서 급격한 전기 에너지 상승 문제가 대두되고 있다. 이러한 급격한 전력 수요 증가는 소형 모듈원자로(SMR, Small Modular Reactor)와 초소형 모듈원자로(MMR, Micro Modular Reactor)와 같은 차세대 원전 기술의 주목을 받게 하는 요인이 되었다[1]-[3]. 특히 MMR은 전기출력이 10MWe(Megawatt electric)급으로 데이터 센터와 연구 시설 전력 소모가 큰 곳이나 특정 지역에 독립적이거나 분산형 전력원(DER, Distributed Energy Resource)이 요구되는 곳에서 안정적인 전력원으로 활용될 수 있어 중요한 변화의 기회로 등장하고 있다[4][5].

MMR은 소형 모듈 단위로 제작 및 설치가 가능하여 건설 기간 단축과 비용 절감 효과를 기대할 수 있다. 또한, 기존 원자로에 비해 규모가 작아 사고 발생 시 위험성이 낮고, 안전 시스템 구축 비용도 상대적으로 적게 든다. 즉 MMR은 초소형화, 모듈화 설계, 향상된 안전성, 다목적 활용 가능성 및 뛰어난 확장성의 장점을 지닌다. 이러한 MMR은 국가가 필요한 에너지를 안정적이고 경제적으로 확보할 수 있는 능력인 에너지 안보와 미래 에너지 시장에서 중요한 역할을 할 수 있게 하며 탄소중립 실현과 노후 석탄화력 대체 등의 과제 해결에 기여할 것으로 기대되고 있다.

전 세계적으로 MMR 기술 개발이 활발히 진행되고 있는 가운데 우리나라도 이 분야에서 기술 경쟁력 확보가 시급한 실정이며 최근 우리나라 에너지 믹스 정책과 맞물려 MMR의 상용화에 대한 관심이 증대되고 체계적인 연구 동향 분석의 필요성이 대두되고 있다. 우리나라는 2024년 2월 12대 국가전력 기술로 차세대 원자력을 포함시키며, 첨단산업의 발전에 따라 전 세계 전력수요 증가에 대응하고 탄소중립을 구현하기 위해 원자력 에너지의 중요성 강조하고 있다.

### 1.2 연구 목적

본 연구는 이렇게 최근 주목받고 있는 MMR에 대해 주요 핵심어와 기술 동향을 파악하고, 우리나라의 MMR 발전 방향과 경쟁력 강화를 모색하기 위해 MMR 신기술에 대한 기술적, 정책적 관점을 다각도로 접근하여 보다 종합적인 인사이트를 도출하기 위해 이들 분야의 지식 구조 연구가 부족한 상황에서 텍스트 마이닝과 효율적인 지식 사전 구축 방안을 제시하고자 한다.

과학기술 분야 논문에서는 전문용어와 기술적 개념을 간결하게 표현하고 복잡한 개념, 기술을 간결하게 표현하기 위해 약어의 일종인 두문자어(Initialism)를 통상적으로 사용한다. 과학기술 논문에서는 일반적으로 한 번 정의된 두문자어를 이후 본문에서 반복 사용해 가독성을 높인다. 따라서 과학기술 논문에서 두문자어의 의미를 정확히 이해하지 못하면 논문의 핵심 내용을 파악하기 어렵고, 검색 효율이 떨어질 뿐만 아니라 자동화된 데이터 분석에서 두문자어를 적절히 처리하지 않으면 잘못된 정보가 도출될 수 있다. 이런 상황에서 두문자어 사전을 활용하면 논문 및 데이터 검색의 정확성과 효율성을 크게 향상시킬 수 있다. 두문자어 사전은 자연어처리(NLP, Natural Language Processing) 모델에서 정확한 단어 의미 해석과 임베딩 학습에도 기여한다.

MMR은 최신 기술이자 제한 및 특화된 분야여서 논문, 특허, 연구보고서 등의 성과가 충분히 축적되지 않은 영역에서 객관적인 트렌드나 유용한 지식을 발견하기 위해서는 전통적인 연구 방식보다는 새로운 접근법과 다양한 데이터를 활용하거나 메타 분석이 연구 방법이 필요하다. 특히 논문은 새로운 발견이나 이론적 발전이 빠르게 반영되므로 최신 기술 동향 파악에 적합하고 특정 분야의 전문성 및 객관성을 구축하며 공개 접근성이 높다. 본 연구에서는 제한된 데이터로도 의미 있는 새로운 인사이트를 도출하고자 국제적인 논문의 전문(Full-text)을 이용하여 분석하고, 우리나라의 MMR 현황을 위해서 국가 연구개발(R&D) 과제정보를 이용하여 분석한다.

### 1.3 연구 방법

이를 위해 본 연구는 우리나라 MMR 연구개발 현황과 방향을 잘 파악하기 위해 국가연구개발 과제를 수집하여 분석하고, 국제 논문 240건 전체 내용 원문 pdf 파일을 수집하여 텍스트 파일로 변환한 후 337,054개 문장을 가지고 단어 임베딩과 클러스터링 등 지식구조를 분석하였다. 본 연구의 결과는 MMR 상용화 현황과 최근의 에너지 믹스에 대한 신속하고 효과적인 이해를 제공함으로써, 우리나라의 MMR 관련 국가 R&D 정책 방향 수립에 기여하고 데이터 기반의 분석 결과와 지식 사전 구축은 향후 과학기술 전문분야의 전략적 방향 설정에 기여할 것으로 기대된다.

## II. 관련 연구

### 2.1 MMR 개요 및 동향

MMR은 초소형 모듈형 원자로 기술로 기존 대형 원자력발전소 대비 높은 안전성과 유연한 배치를 제공하여 소규모 전력 수요 충족, 탈탄소화, 원격지 에너지 공급에 큰 잠재력을 지니고 있다[6]. 특히 최근 생성형 AI 기술 눈부신 발전에 따라 GPU나 TPU 고성능 연산과 고성능 하드웨어 작동 중 발열을 냉각하기 위해 폭증하는 AI 전력 수요에 OpenAI는 물론 MS, Google, Amazon 등의 거대 기업도 대규모 투자 및 건설에 뛰어 들고 있다.

하지만 MMR은 안전성 검증, 기술 규제와 표준은 아직 정립되지 않은 상태이며 초기 설계 및 테스트 비용과 관련하여 투자 경제성을 확인해야 하며 에너지 시장에서 신재생에너지 대비 상업적 경쟁력도 검증이 필요하다. 우리나라는 MMR 기술 개발의 중심 기관으로 한국원자력연구원과 한전KPS, 두산에너지빌리티 등 민간 부문 협력, 학계의 이론 연구와 산업 연계, 원자력 분야에 대한 국민적 인식과 사회적 수용성의 도전도 남아 있다.

표 1은 미국, 유럽, 중국, 호주, 우리나라와 참여한 회사가 진행한 MMR 테스트 결과로 대부분의 테스트는 전력 출력과 안정성/신뢰성 또는 안전성/확장성을 평가하는 데 집중되고 있으며 전력 출력

범위 1MW~1GW로 다양하다.

표 1. MMR 관련 국가(회사)별 동향  
Table 1. Trends by MMR-related countries(company)

Country(company)	Test results
USA (Westinghouse, NuScale Power)	* Tested 1 MW power output * Tested stability and reliability * Tested 50 MWe power output * Tested safety and scalability
Italy (ENEA)	* Tested 10 MW power output * Tested stability and reliability
France (AREVA)	* Tested 1 GW power output * Tested safety and scalability
UK (FIRST)	* Tested 10 MW power output * Tested stability and reliability
Switzerland (ABB, Urenco)	* Tested MIR1's safety and scalability * Tested 2 MW power output
China (CNNC)	* Tested first MMR * Tested stability and reliability
Australia (ANSTO)	* Tested first MMR * Tested stability and reliability
South Korea (KNS, Korea Hydro & Nuclear Power Co. Ltd.)	* Tested 1 GW power output * Tested safety and scalability

특히 미국의 USNC(Ultra Safe Nuclear Corporation)은 MMR의 설계 및 개발을 진행 중이며, 고온가스 로(HTR, High-Temperature Gas-cooled Reactor) 기반의 4세대 원자로 기술을 적용하고 있다. MMR은 안전성을 높이기 위해 혁신적인 세라믹 연료를 사용하며 이 연료는 녹지 않으며 방사성 물질의 환경 방출을 방지한다. 미국의 MMR 개발업체인 Last Energy는 2026년 첫 번째 초소형 원자로를 가동할 계획이며, 이를 위해 최근 4000만 달러의 투자를 유치했다. 이들은 MMR을 공장에서 사전 제작된 모듈로 조립하여 설치하는 방식을 채택하고 있다. 미국 에너지부는 MMR 기술 개발을 지원하고 있으며, 규제 프레임워크도 이러한 새로운 설계를 수용하기 위해 진화하고 있다.

우리나라 SK에코플랜트와 현대엔지니어링은 USNC와 협력하여 '수소 마이크로 허브' 구축을 위한 MOU를 체결했다. 이 프로젝트는 MMR에서 발생하는 전기와 고온 증기를 활용하여 수소를 생산하는 시스템을 개발하는 것을 목표로 하고 있다.

이들은 향후 5년간 공동 연구개발을 통해 가격 경쟁력 있는 수소 생산 체계를 구축할 계획이다. 이탈리아의 Ultra Safe Nuclear는 MMR의 안전성을 강조하며, 혁신적인 세라믹 연료 기술을 통해 방사성 물질의 방출을 차단하는 방법을 연구하고 있다. MMR은 공장 내 열 프로세스를 직접 제공하고, 소규모 커뮤니티에 전력을 공급하는 데 적합하다고 보고되었다.

## 2.2 두문자어와 단어 임베딩

과학기술 논문에서는 많은 전문용어와 두문자어가 사용된다. 언어학적으로 약어에 해당하는 acronym, abbreviation, initialism은 다음과 같은 차이가 있다. Acronym은 여러 단어의 첫 글자를 따서 만든 단어이다. 이 단어는 보통 하나의 단어처럼 발음되는 단계에 이른 상태이다. 예를 들어, NASA (National Aeronautics and Space Administration)나 NATO(North Atlantic Treaty Organization)가 있으며 나사, 나토로 발음되는 단계이다. Abbreviation은 단어나 구의 일부를 생략하여 만든 약어이다. 이는 보통 단어의 첫 글자나 몇 글자를 따서 만든다. 예를 들어, "Dr."(Doctor)나 "A.I."(Artificial Intelligence)가 있다. Initialism은 여러 단어의 첫 글자를 따서 만든 약어로 아직은 하나의 단어처럼 발음되지 않고 각 글자를 개별적으로 발음한다. 예를 들어, FBI (Federal Bureau of Investigation)나 AI(Artificial Intelligence)가 있다. 이러한 차이점은 약어가 어떻게 시기적으로 형성되고 발음되는지에 따라 구분된다[7].

일상적인 분야에서 두문자어와 관련된 웹 서비스로 AcronymFinder가 있는데 이는 인간 편집 정의를 포함한 세계 최대의 약어, 축약어 및 두문자어와 관련하여 광범위한 약어 사전을 구축하여 서비스하고 있다[8]. AcronymFinder는 정보 기술(IT)에 8.6만, 군사 및 정부 15.4만, 과학 및 의학 14.5만, 조직 및 학교 19.5만 이상의 정의를 포함하며 관료, 번역가, 의사, 무기 설계자 등 다양한 분야의 전문가들이 일상적으로 생성되는 약어를 해독하는 데 도움을 받을 수 있는 중요한 지식 자원이 되고 있다.

단어의 앞과 뒤에 나오는 단어 특히 뒤에 나오는

단어를 분석하는 언어 연구 분야도 있는데 이는 언어학, 자연어처리, 문헌 분석, 심리학, 마케팅 등 다양한 분야에서 단어 연쇄 분석, 키워드 연관어 또는 연어(Collocation) 분석에서 이루어지고 있다[9]. 이러한 연구는 단어들이 문맥 내에서 어떻게 결합하고 의미를 확장하는지, 그리고 그로 인해 발생하는 사회적, 경제적, 심리적 영향을 이해하는 데 큰 의미를 가지며 각 분야에서 이 분석은 언어의 패턴, 상호작용, 의미론적 관계 등을 파악하는 중요한 도구로 활용된다.

두문자어는 과학기술 및 전문 분야에서 정보 전달을 간소화하고 효율성을 높이지만 이들의 의미와 추출 정확성과 문맥 기반의 분석 없이는 명확하게 이해하기 어려운 한계도 존재한다. 이런 한계를 보완하기 위해서는 도메인 특화적으로 보정하고 필터링하는 시스템이 함께 설계되고 있다[10].

과학기술 분야의 논문에서 주요 개념을 대표하기 위해 보통 5개 정도의 키워드가 제시되고 있지만 두문자어는 긴 용어나 개념을 짧게 압축하여 간결하게 표현하며 논문의 가독성을 높이고 공간을 절약하는 데 도움을 준다. 특정 분야에서 사용되는 두문자어는 해당 분야의 전문 용어를 나타내며 이는 논문의 전문성을 높여준다. 두문자어는 새로운 개념에 대한 새로운 전문 용어를 정립하는 데 사용되며 반복적으로 사용되는 개념을 압축해 표현 가능한 장점이 있다. 이외에도 다국적 기업이나 국제 조직에서는 두문자어를 사용하여 언어 장벽을 극복하고 더 쉽게 소통할 수 있게 해 준다.

최근 딥러닝(DL, Deep Learning)에 기반한 단어 임베딩(Word embedding) 기술이 현대 자연어 처리에서 핵심적인 역할을 하고 있다. 이 기술은 각각의 단어를 수치화된 벡터로 변환하여 컴퓨터가 이해할 수 있는 형태로 표현한다. 단어들 간의 의미적 연관성을 수학적으로 계산할 수 있게 되어 비슷한 의미를 가진 단어들을 쉽게 찾아낼 수 있다. 이러한 단어 임베딩 기술은 단순히 단어 유사도 측정에만 그치지 않고, 유추 분석, 텍스트 분류, 문서 군집화, 검색 엔진 개선, 추천 시스템 구축, 기계 번역, 텍스트 생성 등 최근 폭넓은 응용 분야에서 활용되고 있다[11].

본 연구에서는 논문의 요약문에 나타난 두문자어를 추출하여 비교한 결과 다음과 같은 점을 알 수 있었다. AI라는 두문자어 단어의 뒤에 나오는 영어 단어를 조사해 보면 1,800개 이상이 등장하는데 AI라는 두문자어는 이제는 거의 보통 명사화된 상태로 진행됨을 알 수 있다. 이에 반해 비교적 최근 등장한 MMR이라는 두문자어는 17개, SMR이라는 두문자어는 70개 정도밖에 나오지 않아 AI에 비해 MMR이나 SMR은 아직 사회적으로 자주 쓰이지 않아 아직 보통 명사화되지 않은 상태임을 알 수 있다.

SMR과 MMR 단어 뒤에 오는 단어를 비교하면 표 2와 같다. SMR보다 MMR이 더 최근에 등장하여 아직은 MMR의 사용빈도가 떨어진다. SMR과 MMR 연구는 주로 기술적 측면에서 설계, 핵심 부품, 연료, 안전성 등이 중요한 이슈로 다뤄지고 있다. 또한, 상용화, 기술 개발과 배치에 대한 논의도 활발하게 이루어지고 있다.

표 2. SMR과 MMR 후속어  
Table 2. SMR and MMR following words

No.	SMR following word	n	MMR following word	n
1	design	59	core	5
2	analysis	24	fuel	5
3	technology	15	concept	4
4	plant	12	technology	3
5	domain	11	CSP	2
6	base	9	HTR	2
7	core	9	stable	2
8	deployment	9	combine	2
9	family	9	develop	2
10	protein	9	experience	2

본 연구에서는 MMR과 관련된 두문자어를 과학 기술 논문 전문(Full-text)에서 자동으로 찾아 두문자어와 전체 단어 쌍으로 사전을 구축하고자 한다. 본 연구에서는 MMR의 논문이 그리 많지 않은 분야이고 요약문을 가지고서는 데이터 부족 문제가 있어서 MMR 분야의 논문 240건의 전문을 가진 pdf 파일을 R언어의 pdftools 패키지를 이용해 텍스트로 변환하여 이용하고자 한다. 전문을 가지고 두문자어도 자동으로 찾아내고 두문자어의 전체 단어도 R언어를 이용해 프로그램으로 찾은 반자동으로 두문자

어와 전체 단어의 쌍을 맞추어 사전을 구축하고자 한다. MMR 분야는 다학제적 특성을 가지고 있어 물리학, 재료공학, 열역학, 원자력공학 등의 다양한 용어가 사용된다. 두문자어와 단어 임베딩을 통해 기술 용어의 의미를 맥락적으로 학습하고, 이를 표준화하여 분야 내 통일된 이해를 제공할 수 있다.

단어 임베딩은 자연어 처리(NLP)에서 단어를 수치 벡터로 표현하는 기법이다. 컴퓨터는 텍스트 데이터를 직접 처리할 수 없기 때문에, 단어를 고차원 공간의 벡터로 변환하고 압축 알고리즘에 해당하는 단어 차원 축소 기법을 이용하여 머신러닝 및 딥러닝 모델에서 활용할 수 있도록 한다. 이 과정에서 단어 간 의미적 관계와 유사성을 벡터 공간에 반영하고 코사인 유사도와 같은 기법을 통해 벡터 간 유사성을 계산할 수 있다.

단어 임베딩을 생성하는 방법으로는 Word2Vec, GloVe, FastText 등이 있으며 같은 단어라도 문맥에 따라 다른 벡터를 생성하는 맥락 기반 임베딩 기법도 있다. 단어 임베딩은 단어 간 의미적 관계를 벡터 간 거리로 직관적으로 표현하며, 희소 행렬에 비해 저장과 연산에 효율적이다. 하지만 단어 임베딩은 일반적인 사전 학습 모델을 사용하는 대신, MMR이나 의료 등 특정 도메인의 특성을 반영한 임베딩 학습이 필요하다. 또한 단어 임베딩을 통해 "MMR"과 "Micro Modular Reactor"은 형태는 다르지만 동일 의미를 갖도록 즉 두문자어를 학습할 필요가 있다.

### III. 데이터 수집과 연구 방법

본 연구는 데이터에 기반한 텍스트마이닝 분석 방법을 적용하고 R언어를 사용하여 구현한 연구 전체 흐름은 그림 1과 같다. 우리나라 MMR 관련 국가R&D과제를 가지고 데이터 분석을 하고, 국제 논문 240건을 가지고 두문자어 및 단어 임베딩 기법을 적용하여 데이터 분석을 한다. MMR 관련 국가 R&D과제 데이터를 수집하여 위해서는 NTIS에서 검색어로 '초소형모듈원자로MMR'를 이용하여 검색된 결과 176건의 과제 데이터를 수집하였고, 국제 논문 플랫폼인 Dimension에서 MMR 관련 논문 240개의 원문을 수집하였다.

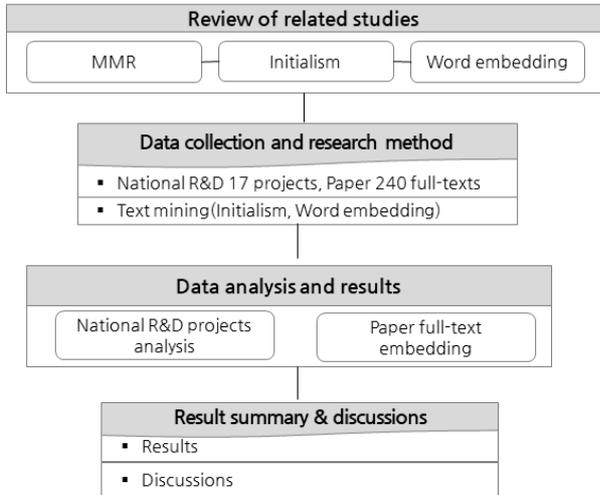


그림 1. 본 연구 전체 흐름도  
Fig. 1. Overall flow of the study

논문 PDF에서 텍스트를 추출하는 과정에서 쓸데 없는 특수 문자나 상용 문구가 포함되는 부분이 있어 이를 정규 표현식 등을 적용하여 제거 및 정제는 하였으나 일부 남아 있는 부분도 있었다. 텍스트에 잡음이 많으면 단어 임베딩의 품질도 떨어지기 때문에, 전처리 후 임베딩을 진행하는 것이 필요하며 텍스트를 많이 학습할 경우는 큰 문제는 되지 않게 나타났다.

#### IV. 데이터 분석 및 결과

##### 4.1 국가R&D과제 데이터 분석

2024년 2월 우리나라는 12대 국가전략기술에서 국가안보에 직결되는 거대과학 분야로 ‘차세대 원자력’을 선정하며 탄소중립을 주도할 세계 최고 수준의 소형모듈원자로 상용화를 포함시키며 세계 시장을 선도하는 차세대 원자력 강국 도약 목표로 하였다.

본 연구는 국가과학기술지식정보서비스(NTIS)에서 수집한 MMR 관련 2013~2023년 국가R&D과제 17건에 대한 정보를 살펴보면, 총연구비는 29억원에 과제당 평균 2.6억원에 달하며 부처는 모두 과학기술정보통신부(이전에는 미래창조과학부)이다. 과학기술표준의 적용분야분류도 모두 ‘에너지’로 나타났다. 과학기술표준분류의 연구분야에서는 표 3에서

보듯 MMR 관련 기술 개발의 주요 초점이 구조설계와 핵심기기, 핵연료 관련 기술에 있으며, 안전성 및 성능 평가도 중요한 과제를 보여준다.

표 3. MMR 관련 국가R&D과제의 연구분야분류  
Table 3. Classification of research fields for MMR-related national R&D projects

No.	S&T standards research category	Projects	%
1	Reactor system structural design/Analysis technology	5	29.4
2	Reactor system/core equipment technology	3	17.6
3	Nuclear fuel manufacturing technology	3	17.6
4	Nuclear fuel design/performance evaluation technology	2	11.8
5	Probabilistic safety assessment/risk information utilization technology	2	11.8
6	Fluid machinery	1	5.9
7	Special welding/joining technology	1	5.9

MMR 분야 국가R&D과제의 과제수행기관 측면에서는 한국과학기술원 9회(52.9%), 울산과학기술원 5회(29.4%), 중앙대학교 2회(11.8%), 한국재료연구원 1회(5.9%)로 수행하였다. 연구개발단계 측면에서는 ‘기초연구’가 94.1%, ‘개발연구’가 5.9%에 속하며, 연구수행주체 측면에서 대학이 94.1%, 출연연구소가 5.9%에 속해 주로 기초연구 분야에서 대학이 중심이 되어 연구가 되고 있음을 알 수 있다.

##### 4.2 MMR 논문 두문자어 사전 구축과 단어 임베딩

본 연구에서는 MMR 관련 국제 논문에서 MMR이라는 단어 뒤에 나오는 단어를 조사하여 정리하면 표 4과 같이 연어 패턴을 나타낸다. MMR이 소형화 및 모듈형 설계, 효율적 연료 사용, 안전성, 재생에너지와의 연계 가능성을 중점으로 연구되고 있음을 확인할 수 있다.

표 4. 논문 원문에 나타난 MMR 관련 후속어  
Table 4. MMR-related following words in full-text papers

Word	Ratio	Remarks
core	14.8%	Refers to the core part of the MMR reactor where nuclear fission occurs. Focuses on maximizing energy efficiency and safety
fuel	8.8%	Refers to nuclear fuel used in MMR, typically highly enriched uranium, with goals of high efficiency and minimal radioactive emissions
unit	6.6%	Refers to the modular reactor itself, designed in small units that allow flexible energy production and installation
energy	5.1%	Refers to the energy generated by MMR, which provides stable power supply and can serve as an independent power source
CSP	4.7%	CSP(Concentrated Solar Power) can be integrated with MMR, allowing for continuous energy supply through renewable energy
micro	4.6%	Emphasizes that MMR is a small modular reactor, making it adaptable to various environments
reactor	4.0%	Refers to the miniaturized reactor design of MMR, which allows for easy installation, management, and operation in diverse settings
design	3.8%	Refers to the design of MMR, focusing on safety, efficiency, and cost-effectiveness, with heat exchange systems as key elements
concept	3.1%	Refers to the modular approach of MMR, which allows flexible installation based on energy demand
HTR	2.4%	High temperature reactor, focusing on enhancing the high-temperature efficiency of MMR

본 연구에서는 MMR 관련 논문 원문을 텍스트로 변환하여 두문자어 추출하고 두문자어와 그 두문자어의 전체 단어(Full-word)를 갖는 사전을 구축하였다. 두문자어는 정규표현식(Regular expression)을 이

용하여 쉽고 효율적으로 찾을 수 있다. 예를 들어 단어의 경계에 있는 영어 대문자로 구성된 2-5글자를 찾아 그 빈도수를 찾고 두문자어 리스트를 만들어 국가명, 기관명 등과 같은 불필요한 단어는 제거하는 방식으로 진행하였다.

이런 과정을 통해 MMR 관련해서 추출한 두문자어는 2글자 단어 582개, 3글자 단어 1,950개, 4글자 단어 1,506개, 5글자 단어 764개이다. 이렇게 추출한 두문자어는 특정 키워드가 문맥 내에서 나타나는 방식을 확인할 수 있도록 하는 도구인 KWIC(Keywords-in-context) 분석을 통해 효율적으로 텍스트 데이터에서 지정한 단어를 찾아낸 뒤, 어떻게 나타나고 있는지 맥락적 의미를 확인할 수 있다. 본 연구에서는 R언어의 `quanteda` 패키지를 이용하여 `pattern`에 분석할 키워드를 지정하고, `window`를 통해 키워드 주변의 문맥 길이를 설정하면, 출력된 결과를 통해 키워드의 사용 맥락을 쉽게 찾아볼 수 있도록 구축하였다.

본 연구에서는 R언어의 `text2vec`, `doc2vec` 패키지를 사용하여 2014년 스탠포드 대학에서 개발한 단어 임베딩 방법인 GloVe(Global Vectors for Word Representation) 모델로 학습하고 문서 및 단어 벡터를 생성하고 난 후에 문서 분류, 단어 및 문장 유사도 측정 등 다양한 자연어 처리 작업을 수행할 수 있다. GloVe는 말뭉치 전체의 통계 정보를 반영하면서도 단어 간 유사도 측정을 효과적으로 할 수 있어, 다양한 자연어 처리 태스크에 유용하게 적용될 수 있다. `Text2Vec`과 `Doc2Vec`을 활용하면 MMR 논문 데이터를 보다 심층적으로 분석할 수 있으며 논문 간 유사성 기반 검색, 클러스터 또는 주제 탐색 등 다양한 활용이 가능하며, 연구와 정책 의사결정 모두에서 큰 가치를 제공할 수 있다.

먼저 단어 임베딩 결과를 통해 단어간은 물론 문서간 유사성 분석을 할 수 있다. 단어 간 의미적 유사성 및 문서 간 주제적 유사성을 계산할 수 있다. 새로운 용어(예: 새로운 핵연료 소재)와 기존 용어 간 의미적 관계 탐색할 수 있다. 특정 문서에서 중요한 단어 즉 키워드를 추출해 연구 방향을 제안할 수 있다.

표 5는 단어 임베딩을 하고 난 후 'MMR'과 유사한 단어와 그 유사도를 보여준다.

표 5. 단어 임베딩에 의한 MMR과 유사 단어와 유사도  
Table 5. Similar words and their similarity to MMR by word embedding

Word	Similarity	Word	Similarity
reactor	0.5885716	exchanger	0.4368933
modular	0.5536393	turbomachinery	0.4343008
MWth	0.5494913	evaluation	0.4226668
concept	0.5420359	CSP	0.4226212
design	0.5291978	power	0.4194072
plant	0.5181390	Small	0.4182448
SMR	0.5173926	operation	0.4180064
hybrid	0.5070701	conversion	0.4137839
Micro	0.4995372	units	0.4136602
core	0.4923092	microreactor	0.4127758
MWe	0.4832921	configuration	0.4119537
parameters	0.4651128	consideration	0.4118526
cooled	0.4586643	fueled	0.4110897
PWR	0.4474428	module	0.4103741
designed	0.4383875	nuclear	0.4102577

단어 임베딩을 통해 일반적인 사람이 생각하는 정도의 유사한 단어를 찾아준다. 예를 들어 MMR과 다른 SMR이 51.7% 정도의 유사한 단어임을 보여준다.

단어 임베딩을 통해 나온 결과를 가지고 클러스터링(Clustering) 알고리즘을 적용하여 문서 분류가 가능하다. 즉 논문 데이터의 구조를 파악하고 숨겨진 패턴을 찾을 수 있다. 클러스터링은 데이터 포인트의 유사성을 기준으로 그룹 및 클러스터를 형성하는 기법이며 논문 간의 유사성을 계산하여 비슷한 내용이나 특성을 가진 논문을 하나의 클러스터로 묶는다.

클러스터의 수(k)를 결정하기 위해 본 연구에서

는 그림 2와 같은 엘보우 기법을 적용하여 클러스터 수(k)가 증가함에 따라 클러스터 내 제곱합(SSE, Sum of Squared Errors)이 어떻게 변하는지를 관찰하여 팔꿈치(elbow)처럼 보이는 지점을 최적의 k 값으로 하는데 여기서는 k=5로 지정하였다.

다음으로 k=5일 때 문서를 kmeans 클러스터링 알고리즘을 적용하면 그림 3과 같이 얻을 수 있다. 크게 겹치는 영역이 없이 5개의 클러스터로 나뉘는 것을 볼 수 있다. 이렇게 하여 총 240개 논문 문서를 클러스터1은 44개, 클러스터2는 67개, 클러스터3은 7개, 클러스터4는 82개, 클러스터5는 40개로 나누고 5개의 각 클러스터에서 중심에 가장 가까운 문서를 찾기도 수행하였다.

클러스터링을 통해 문서의 분류를 예를 들어 k=5로 한다고 하고 5개의 중심어를 파악하고 이웃들을 찾아낼 수 있다. 이는 데이터가 구조화되어 있지 않은 경우에도 논문을 주제별 그룹화하여 연구 그룹의 특성, 기술 영역 또는 연구 범위를 파악하고 문서 간 관계를 시각적으로 표시할 수 있다.

본 연구에서는 구조적 토픽모델링(STM)을 통해 토픽 수를 정하고 각 토픽 내용을 살펴보았다. 그림 4는 STM 토픽 모델링을 위해 토픽 수를 결정하기 위한 진단 값이다. Held-Out Likelihood에서 일반적으로 이 값이 급격히 떨어지기 시작하는 지점(예: k=5 이후)이나 안정화되는 지점을 참고하고, Residuals도 k=5에서 급격히 떨어지는 첫 지점이고, Semantic Coherence에서도 k=5이후에 비교적 안정적인 값을 보여 그림 2의 클러스터링 시와 맞추어 k=5로 지정하고 토픽을 정하였다.

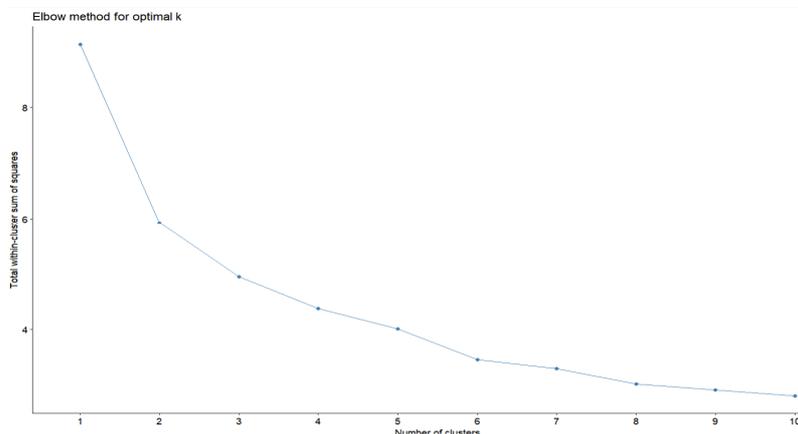


그림 2. 엘보우 기법에 따른 k값 결정  
Fig. 2. Determination of k according to Elbow method

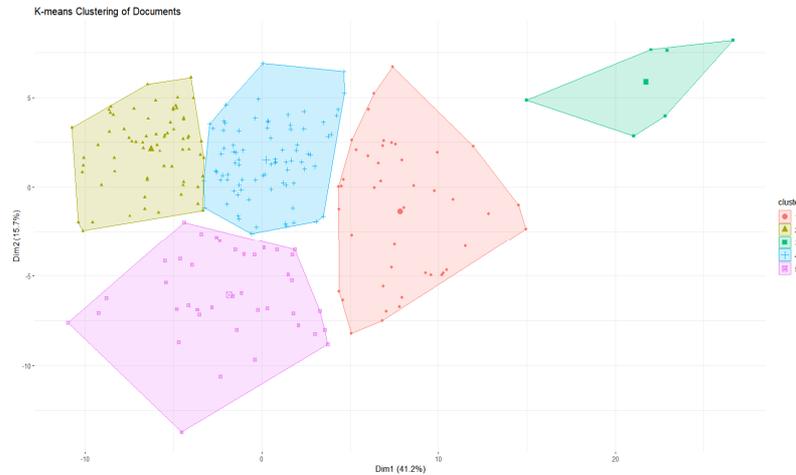


그림 3. K-means 클러스터링  
Fig. 3. K-means clustering of documents

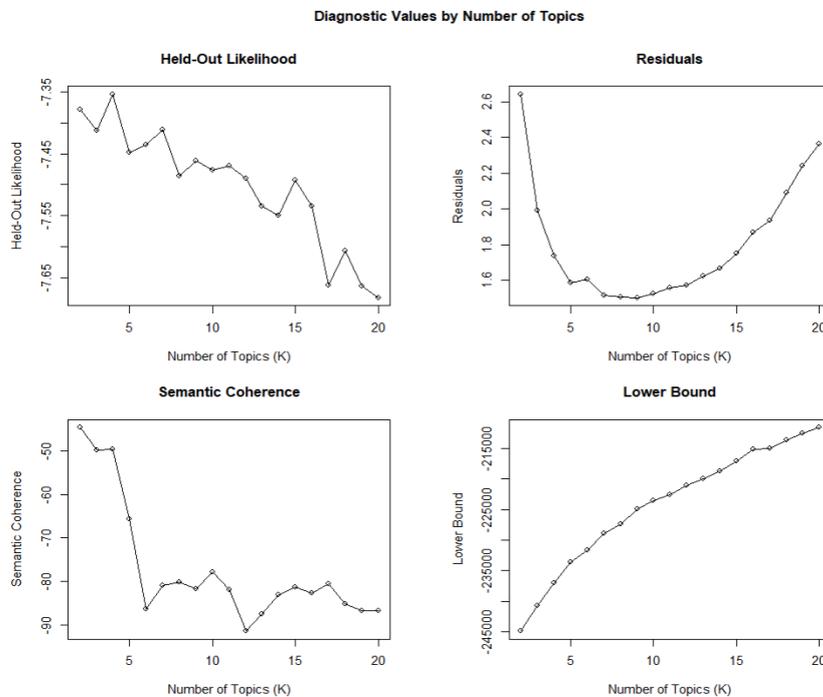


그림 4. 토픽 수 탐색  
Fig. 4. Exploring number of topics

토픽의 수를  $k=5$ 로 하여 STM 토픽모델링을 실행한 결과는 그림 5와 같다. 주요 5개 토픽은 열 생성 및 열 전달 설계, 열 생성과 전력 공급, Brayton 사이클을 기반 고효율 원자로, 효율적이고 안전한 연료 사용과 원자로 설계, 핵 에너지 기반 소형모듈형원자로로 나타났다.

본 연구는 기존 논문 유사도 비교 서비스가 주로 요약문(Abstract)을 활용해 유사도를 계산하는 것에 반해 본 연구는 전문 단어 임베딩하고 문서 기반으

로 유사도를 비교함으로써 논문의 전반적인 구조와 세부 내용을 더 정확히 반영할 수 있는 방법으로 활용될 수 있을 것으로 기대된다. 또한 기존 클러스터링 기법은 주로 수치화된 정형 데이터를 처리하였으나 비정형 데이터인 텍스트를 처리하기에는 한계가 존재하였다. 본 연구는 단어 임베딩 기법을 통해 비정형 텍스트 데이터를 벡터화하여 클러스터링에 활용할 수 있었다.

<p>Topic 1 Top Words: <b>Heat generation and heat transfer design</b>                  Highest Prob: nuclear, reactors, can, design, SMR, NPP, safety                  Lift: business, evacuation, grain, monitoring, PLCs, PRISM, safeguard                  Score: creep, 800H, Alloy, ice, PWT, 310S, SS</p> <p>Topic 2 Top Words: <b>Heat generation and power supply</b>                  Highest Prob: heat, power, reactor, transfer, temperature, model, nuclear                  Lift: fins, rNPPs, TAPIRSD, 0.6-m, 0.96, 0.97, 0.99                  Score: startup, transfer, wire, flow, pipe, sub-channel, wall</p> <p>Topic 3 Top Words: <b>High-efficiency reactor based on the Brayton cycle</b>                  Highest Prob: cycle, power, efficiency, control, Brayton, reactor, performance                  Lift: axis, BC-VCC, closed-cycle, controller, correction, helpful, isentropic                  Score: compressor, Brayton, off-design, recompression, inlet, SCO2, cycle</p> <p>Topic 4 Top Words: <b>Efficient and safe fuel use and reactor design</b>                  Highest Prob: fuel, reactor, core, design, power, control, reactivity                  Lift: #62, 0.00166, 0.19, 0.25, 0.258pcm, 0.4, 0.5525                  Score: core, enrichment, HPTWR, duplex, TRISO, reactivity, burnup</p> <p>Topic 5 Top Words: <b>Small modular reactor based on nuclear energy</b>                  Highest Prob: energy, nuclear, power, reactors, electricity, technologies, sources                  Lift: 2030, actions, charging, city, Conference, cost-effective, diversification                  Score: N-R, emissions, centres, GHG, charging, RESs, government</p>
---

\* Best score  
 \*\* Top word per score

그림 5. STM 토픽모델링 결과  
 Fig. 5. Results of STM topic modeling

### 4.3 결과 종합 및 토론

MMR은 우리나라 12대 국가전략기술 중 하나로 선정되어 국가안보와 탄소중립 목표를 동시에 달성할 핵심 기술로 상용화에 주력하고 있다. 그러나 우리나라가 MMR을 상용화하고 세계 시장을 주도하기 위해서는 다수의 기초연구를 수행하는 동시에 기초연구에서 개발연구로의 전환으로 상용화 가능성을 높이고 산업 전반에서 기술 생태계를 활성화, 글로벌 상용화를 위해서는 국제 표준 및 안전 규제 대응, 안전성 및 성능 평가가 필요하다. 국가R&D

과제 분석을 통해 우리나라의 MMR 연구가 기초연구 단계로 연구 초점은 구조설계, 핵심기기, 핵연료 제조 및 성능평가 등에 맞춰졌으며, 안전성 및 성능 평가도 중요한 과제임을 보여준다.

우리나라와 선진국과의 비교를 통해 기술 경쟁력 평가, 정책 개선, 글로벌 협력 방향 설정 등을 위한 MMR 기술 발전과 국가 경쟁력 강화에 중요한 기초 자료를 제공한다는 점에서 의의를 가진다. 또한 선진국의 사례를 바탕으로 국내 MMR R&D 정책 방향을 검토하고 국제 연구 협력 및 트렌드 정합성 확보를 위한 기반을 마련하고자 하였다.

우리나라와 미국의 MMR 관련 연구개발은 주체, 연구 동기, 시장 접근 방식에서 차이가 나면 이를 정리해 보면 표 6와 같다. 우리나라는 정부 주도, 미국은 민간 주도라는 핵심 차이를 보이며, 우리나라는 민간 참여 확대와 규제 개선, 상업화 로드맵 구축이 필요하다. 반면, 미국은 민간 혁신과 시장 중심 접근으로 빠르게 상업화에 도달하고 있음을 알 수 있다.

표 6. 우리나라와 미국의 MMR 연구개발 비교  
 Table 6. Comparison of MMR R&D between Korea and USA

Category	South Korea	United states
Key research entities	Institutions: KAERI, Doosan Enerbility, KEPCO KPS.	Private companies: TerraPower, NuScale Power, Oklo.
Research focus	Technology development and early-stage demonstration	Focus on commercialization and product launch.
Funding sources	Primarily reliant on government budgets.	Mainly funded by private investments.
Regulatory environment	Strict nuclear regulations, slowing technology adoption and demonstration.	Flexible regulations enabling simultaneous demonstration and approval.
Commercialization approach	Absence of a clear commercialization roadmap.	Aimed at market release and revenue generation.

우리나라 국가R&D과제 데이터 분석을 통해 우리나라 MMR 연구는 매우 활발하다는 사실을 발견했다. 코어 기술, 연료 시스템, 안전성, 다른 에너지 기술과의 통합 등 다양한 측면에서 개발이 진행되고 있으며, 이는 MMR이 미래 에너지 시스템의 중요한 부분으로 자리잡기 위한 노력의 일환으로 볼 수 있다. 이러한 개발 과정은 기존 원자력 발전 경험을 바탕으로 하면서도 새로운 혁신을 추구하는 방향으로 진행될 것으로 전망된다.

본 연구는 MMR 관련 국제 논문에서 언어 패턴을 분석하고 정규표현식을 사용해 4,802개 두문자어를 추출하고, KWIC 분석으로 문맥 내 의미를 파악하고 불필요한 단어를 제거해 효율적인 텍스트 분석을 제시하여 두문자어 사전 구축을 하고, GloVe 모델과 R언어의 text2vec 및 doc2vec 패키지를 활용해 단어 임베딩 및 문서 벡터 생성을 하고 각 클러스터에서 중심 문서를 식별하고 논문 주제 간 구조적 관계를 시각적으로 분석하였다. 과학기술 논문에서 두문자어는 정보를 효율적으로 전달하고, 전문성을 강화하며, 국제적 소통을 촉진하는 중요한 역할을 한다. 물론 두문자어가 특정 분야의 전문성을 강조할 수 있지만 두문자어를 모르는 사람들은 대화에서 소외감을 느끼게 하거나 세대 간 의사소통에서 문제가 될 수 있어 무분별한 두문자어 사용은 오히려 언어의 본질을 해치고 의사소통의 효율을 떨어뜨릴 수도 있어 상황에 맞게 적절히 사용하는 것이 중요하다.

이렇게 본 연구에서는 과학기술 논문의 전문에서 논문의 내용의 핵심에 해당하는 두문자어를 효과적으로 추출하고 이를 필터링하고 두문자어의 전문을 찾아서 MMR 분야의 도메인 특화 두문자어 사전을 구축하는 방법을 제안하였다. 물론 텍스트 내 여러 두문자어가 불규칙적으로 사용될 경우 정확한 추출 및 매핑이 어려울 수 있고, 필터링 과정에서 중요한 두문자어와 불필요한 두문자어를 구분하는 기준이 명확하지 않거나 주관적일 가능성이 있고, 도메인에서도 새로운 두문자어가 빠르게 등장하거나 기존 두문자어의 의미가 변화할 경우 사전이 금방 구식화될 우려가 있다.

이런 비판적인 측면에도 불구하고 다행히 과학기

술 논문에서 두문자어는 불규칙적으로 사용되지 않으며 불필요한 두문자어도 비교적 명확하여 MMR 도메인의 연구 효율성을 높이는 기초 자료와 분석 방법론을 제시했다는 점에서 가치를 가진다. 즉 MMR과 같은 전문적이고 특수한 분야처럼 두문자어 사용이 빈번한 분야에서 매우 유용하며, 논문 분석 자동화, 도메인 특화 자료 구축, 비정형 데이터 분석의 확장 가능성 등에서 실질적이고 유의미한 기여를 할 수 있을 것이다. 그림 5에서 나온 국제 논문의 토픽에 따라 표 7는 MMR 관련 토픽과 실행계획(Action plan)을 보여준다.

표 7. MMR 관련 토픽과 실행계획  
Table 7. MMR-related topics and action plan

No.	Topics	Action plan
1	Heat generation and heat transfer design	<ul style="list-style-type: none"> <li>* Research on optimizing heat transfer in MMR design.</li> <li>* Development and testing of temperature control systems.</li> <li>* Review of efficient pipe network design.</li> </ul>
2	Heat generation and power supply	<ul style="list-style-type: none"> <li>* Review the feasibility of applying new heat transfer technologies.</li> <li>* Develop and test a prototype of the MMR model design.</li> <li>* Perform technology improvement tasks to optimize power output.</li> </ul>
3	High-efficiency reactors based on the Brayton cycle	<ul style="list-style-type: none"> <li>* Efficiency analysis of systems based on the Brayton cycle.</li> <li>* Design and testing of control systems for MMR.</li> <li>* Performance evaluation studies under supercritical conditions.</li> </ul>
4	Efficient and safe fuel usage and reactor design	<ul style="list-style-type: none"> <li>* Research on optimizing fuel usage efficiency.</li> <li>* Reviewing improvements in the core design of MMR.</li> <li>* Developing temperature and reactivity control systems.</li> </ul>
5	Nuclear energy-based small modular reactors	<ul style="list-style-type: none"> <li>* Production and review of modular design prototypes.</li> <li>* Performance testing and analysis of key technologies in MMR.</li> <li>* Data collection and improvement efforts to optimize power and energy efficiency.</li> </ul>

토픽 1는 MMR 설계에서 열 전달 최적화 방안 연구, 온도 제어 시스템 개발 및 테스트, 효율적인 파이프 네트워크 설계 관련이며, 토픽 2는 새로운 열 전달 기술 적용 가능성 검토, MMR의 모델 설계 프로토타입 제작 및 실험, 전력 출력 최적화를 위한 기술 개선 작업 수행 관련이며, 토픽 3은 Brayton 사이클 기반 시스템의 효율성 분석, MMR의 제어 시스템 설계 및 테스트, 초임계 조건에서의 성능 평가 연구 관련이며, 토픽 4는 연료 사용 효율 최적화 연구, MMR의 코어 설계 개선 검토, 온도 및 반응도 제어 시스템 개발 관련이며, 토픽 5는 모듈형 설계 프로토타입 제작 및 검토, MMR의 핵심 기술 성능 시험 및 분석, 전력 및 에너지 효율 최적화를 위한 데이터 수집 및 개선 작업 수행 관련이 필요함을 알 수 있다.

본 연구는 기존의 논문 요약문에 제한하지 않고 논문 전문을 활용함으로써 유사도 분석의 깊이와 신뢰도를 높이고 학계의 데이터 분석 및 정보 검색 방법론을 한 단계 발전시킬 수 있다는 점에서 의의를 가진다. 논문의 서론, 본문, 실험, 결과, 토론 등 각 섹션에서 얻어진 다양한 정보가 분석에 포함되어, 유사도 분석의 정밀도가 높아 기존의 요약문 기반 접근법과 달리, 연구 배경과 맥락을 고려한 보다 종합적이고 심층적인 유사도 평가가 가능하다.

## V. 결론 및 향후 과제

MMR은 전력 수요 증가와 탄소중립 실현을 위한 중요한 기술로 부상하고 있다. MMR은 초소형화, 모듈화 설계, 향상된 안전성, 다목적 활용 가능성, 뛰어난 확장성 등의 장점을 가지고 있어 에너지 안보와 미래 에너지 시장에서 중요한 역할을 할 것으로 전망된다.

본 연구는 MMR 관련 국제 논문과 국가R&D 과제 데이터를 분석하여 기술 동향을 파악하고, 우리나라의 MMR 발전 방향과 경쟁력 강화를 위한 인사이트를 도출하고자 하였다. 특히 과학기술 논문의 전문에서 두문자어를 효과적으로 추출하고 필터링하여 MMR 분야의 도메인 특화 두문자어 사전을 구축하는 방법을 제안하였다. 또한 전문 텍스트를 단어 임베딩하고 지식 분석을 수행한 후, 우리나라와

미국의 MMR 연구개발 비교하고 토픽모델링과 토픽 실행계획을 마련하고 정책적 방향을 제시하고자 하였다. 이러한 연구 결과는 MMR 상용화 현황과 에너지 믹스에 대한 이해를 제공하여 국가 R&D 정책 방향 수립에 기여할 것으로 기대된다. 또한 데이터 기반의 분석 결과와 지식 사전 구축은 과학기술 전문분야의 전략적 방향 설정에 도움이 될 것이다.

본 연구를 통해 우리나라가 MMR을 12대 국가전략기술 중 하나로 선정하여 상용화에 주력하고 있지만, 국가R&D 측면에서는 현재 기초연구 단계에 머물러 있음을 알 수 있었다.

본 연구는 텍스트 데이터를 구조화하고 군집화하는 새로운 접근법을 제시하여, 기존의 정형 데이터 중심 클러스터링 기법을 비정형 데이터 처리로 확장하고 비정형 텍스트 데이터 처리의 정확성, 효율성, 유연성을 향상시켰다. 본 연구를 통해 텍스트 데이터 분석과 활용의 지평을 넓히기를 기대한다.

우리나라가 상용화와 세계 시장 주도를 위해서는 기초연구에서 개발연구로의 전환, 산업 전반의 기술 생태계 활성화, 국제 표준 및 안전 규제 대응, 안전성 및 성능 평가가 필요하다.

향후 우리나라가 MMR 상용화와 안정성과 경제성을 갖춘 최근의 에너지 믹스에 대한 신속하고 효과적인 이해를 제공하며, 우리나라의 MMR 관련 국가R&D 정책 방향 수립에 기여하길 기대한다.

## References

- [1] C. Zeliang Y. Mi A. Tokuhiko L. Lu, and A. Rezvoi, "Integral PWR-Type Small Modular Reactor Developmental Status, Design Characteristics and Passive Features: A Review", *Energies*, Vol. 13, No. 11, pp. 2898, Jun. 2020. <https://doi.org/10.3390/en13112898>.
- [2] B. Poudel, K. Joshi, and R. Gokaraju, "A Dynamic Model of Small Modular Reactor Based Nuclear Plant for Power System Studies", in *IEEE Transactions on Energy Conversion*, Vol. 35, No. 2, pp. 977-985, Jun. 2020. <https://doi.org/10.1109/TEC.2019.295670>.

- [3] J.-W. Choi and B.-H. Lee, "Analyzing SMR and Energy Mix Knowledge using Reddit Sentiment Analysis and National R&D Projects Need Mining", The Journal of Korean Institute of Information Technology, Vol. 22, No. 6, pp. 175-187, Jun. 2024. <https://doi.org/10.14801/jkiit.2024.22.6.175>
- [4] C. Deng, S. Zhu, Y. He, Y. Wu, K. He, J. Zhang, and G. Su, "Analysis of gas-cooled micro modular reactor (MMR) fuel", Journal of Nuclear Materials, Vol. 598, pp. 155191, Sep. 2024. <https://doi.org/10.1016/j.jnucmat.2024.155191>.
- [5] Y. J. Choi, S. Lee, S. Jang, I. W. Son, Y. Kim, J. I. Lee, and Y. H. Jeong, "Conceptual design of reactor system for hybrid micro modular reactor (H-MMR) using potassium heat pipe", Nuclear Engineering and Design, Vol. 370, pp. 110886, Dec. 2020. <https://doi.org/10.1016/j.nucengdes.2020.110886>.
- [6] M. A. Rahmanta, R. A. A. Hasibi, H. B. Tambunan, Ruly, A. Syamsuddin, I. A. Aditya, and B. Susanto, "Towards a Net Zero-Emission Electricity Generation System by Optimizing Renewable Energy Sources and Nuclear Power Plant", Energies, Vol. 17, No. 8, pp. 1958, Apr. 2024. <https://doi.org/10.3390/en17081958>.
- [7] Y. Koo, "ChatGPT-related Technology Neologism Extraction and Word Embedding in Paper Abstracts", The Journal of Studies in Language, Vol. 40, No. 2, pp. 109-125, 2024. <https://doi.org/10.18627/jslg.40.2.202408.109>
- [8] <https://www.acronymfinder.com/> [accessed: Nov. 28, 2024]
- [9] C. Lee and K. Kim, "Language Analysis in Mass Communication and Journalism : Using Korean Linguistic Inquiry and Word Count (K-LIWC)", Media Convergence Research, Vol. 24, pp. 88-115, 2017. <https://doi.org/10.22814/sgjcr.2017..24.88>.
- [10] M. Abulaish, M. Fazil, and M. J. Zaki, "Domain-Specific Keyword Extraction Using Joint

Modeling of Local and Global Contextual Semantics", ACM Transactions on Knowledge Discovery from Data, Vol. 16, No. 4, pp. 1-30, Jan. 2022. <https://doi.org/10.1145/3494560>.

- [11] M. Q. Khan, A. Shahid, M. I. Uddin, M. Roman, A. Alharbi, W. Alosaimi, J. Almalki, and S. M. Alshahrani, "Impact analysis of keyword extraction using contextual word embedding", PeerJ Computer Science, May 2022. <https://doi.org/10.7717/peerj-cs.967>.

### 저자소개

#### 최 정 우 (Jung-Woo Choi)



2015년 2월 : 연세대학교 상경대학  
경제학과(경제학사)

2019년 9월 ~ 현재 :  
과학기술연합대학원대학교  
과학기술경영정책 석.박사  
통합과정

관심분야 : 과학기술경영정책,

텍스트 마이닝, 에너지정책, R&D경제, 수소경제

#### 이 병 희 (Byeong-Hee Lee)



1992년 2월 : 충남대학교  
컴퓨터공학과(공학사)

2002년 2월 : 충남대학교  
컴퓨터공학과(공학박사)

2002년 9월 ~ 현재 :  
한국과학기술정보연구원  
NTIS센터 책임연구원

2012년 9월 ~ 현재 : 과학기술연합 대학원대학교  
과학기술경영정책학과 교수

관심분야 : 과학기술경영정책, 인공지능, 빅데이터,  
텍스트마이닝, 기술마이닝