

# 키워드의 도메인 속성을 고려한 고품질 키워드 정제기법

김민근\*, 정유철\*\*

## High-Quality Keyword Refinement Techniques Considering the Domain Attributes of Keywords

MinGeun Kim\*, Yuchul Jung\*\*

The work is supported by the Korea Basic Science Institute (Project No.A438200)

### 요약

키워드 추출은 연구 내용을 요약하고 관련 문헌을 검색하는 데 중요한 도구이다. 기술의 빠른 발전과 전문화로 연구 문서의 범위와 깊이가 확장되면서 키워드 추출은 연구의 핵심 주제 식별에 필수적 고려 요소이다. 그러나 기존 키워드 추출 방식만으로는 연구 내용의 전체적인 맥락과 복잡성을 완전히 이해하기 어려우며, 특히 전문 지식이 없는 사용자들에게는 추출된 키워드의 정확성과 관련성을 판단하기 어렵다. 이에 본 논문은 LLM(Large Language Model)을 활용한 프롬프트 기반 접근으로 키워드의 속성값을 고려한 정제방법론을 제안한다. 특히, 키워드에 속성값과 부연 설명을 추가하여 키워드의 정확성과 중요성을 의도, 역할, 구성 측면에서 분석하여 키워드 정제과정에서 어떻게 작용하는지 실험을 통해 검증하였다. 연구 결과, 이 방법은 기존 방식보다 키워드의 정확성과 관련성을 높여 연구 내용을 더욱 잘 파악하고 요약하는데 큰 효과를 보였다.

### Abstract

Keyword extraction is an important tool for summarizing research content and searching related literature. As technological advancements and specialization expand the scope and depth of research documents, keyword extraction becomes crucial for identifying core research themes. However, traditional keyword extraction methods have limitations in fully understanding the overall context and complexity of research content, and they can be particularly challenging for users without expert knowledge to assess the accuracy and relevance of the extracted keywords. This paper proposes a refined methodology considering the attributes of keywords using a Large Language Model(LLM)-based prompt approach. By adding attributes and explanations to keywords, we analyze their accuracy and significance in terms of intent, role, and structure, and conduct experiments to verify their impact on the keyword refinement process. The results show that this method improves keyword accuracy and relevance compared to traditional approaches, offering a better understanding and summarization of complex research content.

### Keywords

keyword extraction, attribute value, prompt engineering, LLM

\* 국립금오공과대학교 컴퓨터공학과 학사과정

- ORCID: <http://orcid.org/0009-0007-5871-5366>

\*\* 국립금오공과대학교 인공지능공학과 부교수(교신저자)

- ORCID: <http://orcid.org/0000-0002-8871-1979>

• Received: Jun. 02, 2024, Revised: Jul. 30, 2024, Accepted: Aug. 02, 2024

• Corresponding Author: Yuchul Jung

Dept. of Computer Engineering, Kumoh National Institute of Technology,  
61 Daehak-ro (yangho-dong), Gumi, Gyeongbuk, [39177] Korea

Tel.: +82-54-478-7536, Email: [enthusia77@gmail.com](mailto:enthusia77@gmail.com)

## I. 서론

현대 과학 연구와 기술 개발은 빠르게 발전하며 점점 더 전문화되고 복잡해지고 있다. 이러한 추세는 연구 문서와 논문에서 다루어지는 주제의 범위와 깊이를 확장시켜, 특정 분야에 대한 깊은 이해와 전문적인 지식을 필요로 하게 만든다. 키워드 추출은 연구 내용을 요약하고, 관련 문헌을 검색하며, 연구의 핵심 주제를 식별하는데 중요한 도구이다. 그러나 키워드만으로는 연구 내용의 전체적인 맥락이나 복잡성을 완전히 이해하는 데 한계가 있다. 특히, 전문적인 지식이 없는 이용자들에게는 추출된 키워드의 정확성과 관련성을 판단하기 어렵다.

이러한 배경을 바탕으로 본 연구는 서로 다른 기법으로 도출된 논문키워드들을 키워드의 관련성을 심도 있게 분석하고 체계적으로 정제할 수 있는 새로운 접근법을 제안한다. 이 접근법은 키워드의 속성값을 고려하되, 의도, 역할, 그리고 구성이라는 세 가지 핵심 요소에 초점을 맞춘다. 이를 통해 키워드의 다면적 특성을 명확하게 파악하여 키워드 정제에 활용하는 것을 의의로 둔다. 각 키워드에 속성값과 부연 설명을 추가함으로써, 키워드의 정확성과 중요성을 더욱 깊이 있게 파악하고 이해할 수 있으며 전문 지식이 없는 이용자들이 복잡한 연구 내용을 보다 명확하고 직관적으로 이해할 수 있게 함으로써, 연구 문서의 가독성과 접근성을 향상시킬 수 있다.

제안하는 프롬프트 기반 키워드 정제 시스템의 효율성을 검증하기 위해, 의도, 역할, 그리고 구성 요소의 포함 유무에 따라 키워드 분석의 정밀도와 깊이가 어떻게 달라지는지를 실험적으로 분석하였다. 이러한 분석을 통해, 제안하는 키워드 정제 시스템이 연구 초록 내에서 키워드가 갖는 의미와 중요성을 보다 명확하게 이해하고, 복잡한 데이터 집합에서 중요한 정보를 추출하고 분석하는 데 매우 유용함을 확인하였다. 이 방법론은 단순한 키워드 추출을 넘어서, 추출된 키워드에 대한 심도 있는 분석을 가능하게 하여 연구자들에게 보다 정교한 도구를 제공할 것으로 기대된다.

## II. 관련 연구

### 2.1 키워드 추출

키워드 추출은 정보 검색, 문서 요약, 내용 분류 등 다양한 자연어 처리 분야에서 중요한 역할을 한다. 이 과정에서 효과적인 키워드 추출은 텍스트의 핵심 내용을 식별하고, 문서의 주요 주제를 빠르게 파악하는 데 도움을 준다. 특히 최근에는 BERT (Bidirectional Encoder Representations from Transformers)[1]와 같은 딥러닝 기반 모델이 키워드 추출의 정확성과 효율성을 크게 향상시켰다.

BERT는 트랜스포머 구조를 활용한 언어 모델로, 문장 전체의 양방향 문맥을 고려하여 단어의 의미를 파악한다. 이는 각 단어 주변의 문맥적 정보를 포괄적으로 이해하고, 그 의미를 정확하게 해석하는데 기여한다. 추가적으로 분류 레이어를 사용하거나, CRF(Conditional Random Field)[2]와 같은 시퀀스 라벨링 기법을 결합하는 등 다양하게 응용되고 있다.

BERT 기반의 키워드 추출은 전통적인 통계적 접근법이나 초기 딥러닝 모델보다 더 높은 성과를 보여준다. 특히, 복잡한 문서에서 여러 주제가 교차하는 경우에도 각 주제와 관련된 키워드를 식별할 수 있다. 또한 BERT는 다양한 자연어 처리 작업에 적용해왔으며, 특히 키워드 추출에서는 학술 논문, 법률 문서, 뉴스 기사 등 다양한 장르의 텍스트에 적용하여 효과적인 결과를 얻었다.

### 2.2 키워드 생성

키워드 생성은 텍스트에서 의미 있는 정보를 추론하여 새로운 키워드를 도출하는 과정으로, 이는 특히 콘텐츠 제작, 검색 엔진 최적화, 자동 요약과 같은 응용 분야에서 중요하게 활용된다. 최근에는 LSM(Latent Semantic Modeling)[3], Seq2Seq (Sequence-to-Sequence)[4], BART(Bidirectional Auto-Regressive Transformer)[5], T5(Text-to-Text Transfer Transformer)[6]와 같은 다양한 딥러닝 모델들, 그리고 최근에는 대규모 언어 모델 기반의 프롬프트 기반 접근법들이 키워드 생성을 위해 활용되고 있다.

LSM은 텍스트 내 숨겨진 의미 구조를 파악하여 관련성이 높은 키워드를 생성하는 방법이다. 데이터에서 주요 개념과 그 관계를 찾아내어 보다 관련성 있는 키워드를 도출한다. Seq2Seq 모델은 입력된 텍스트 시퀀스로부터 새로운 텍스트 시퀀스를 생성하는 데 사용되며, 키워드 생성에 적용할 때는 주요 정보를 포착하여 정보적인 키워드를 생성한다.

BART와 T5는 변형된 텍스트를 입력받아 원래의 텍스트나 요약된 형태를 출력하는 데 사용되며, 키워드 생성에서는 문맥을 깊이 분석하고 텍스트의 중요한 요소를 식별하여 적절한 키워드를 제시한다. 또한, 대규모 언어 모델을 활용한 프롬프트 기반 접근법은 사용자가 제공한 프롬프트를 바탕으로 모델이 관련성 높은 키워드를 생성하며, 창의적이고 독창적인 키워드를 요구하는 상황에 특히 유용하다.

### 2.3 Prompt engineering

프롬프트 엔지니어링은 사용자가 정의한 프롬프트를 언어 모델에 제공하고, 모델이 생성한 출력을 기반으로 유용한 정보를 추출하는 과정이다. 언어 모델의 광범위한 지식과 이해 능력을 활용하여, 키워드에 대한 속성값과 부연 설명을 추가함으로써, 키워드의 중요성과 관련성을 더욱 명확하게 파악할 수 있게 한다.

제로샷(zero-shot) 프롬프트 엔지니어링은 언어 모델이 사전에 특정 작업에 대한 훈련을 받지 않고도 새로운 태스크에 대응할 수 있는 능력을 의미한다. 이 방법은 모델이 다양한 도메인의 지식을 활용하여, 주어진 프롬프트에 대해 적절한 출력을 생성할 수 있다. 제로샷 방식은 특히 데이터가 부족한 상황에서 유용하며, 다양한 작업과 환경에서의 범용성을 증가시킨다.

원샷(one-shot) 프롬프트 엔지니어링은 언어 모델에 단 하나의 예제를 제공하고, 그 예제를 통해 모델이 새로운 작업을 수행하게 하는 방법이다. 이 방법은 모델이 제공된 단일 사례를 학습의 기준점으로 사용하여 유사한 새로운 상황을 처리할 수 있게 한다. 원샷 학습은 빠른 적응력과 고도의 추론 능력을 요구하며, 제한된 데이터로도 효과적인 결과를 도출할 수 있다.

퓨샷(few-shot) 프롬프트 엔지니어링은 제로샷과 원샷의 개념을 확장하여, 소수의 예시를 활용해 모델이 새로운 작업을 수행하도록 한다. 이 방법은 몇 개에서 수십 개의 학습 데이터만으로도 모델이 새로운 상황에 효과적으로 적응하고 필요한 작업을 수행할 수 있도록 지원한다.

본 연구는 대규모 언어 모델의 제로샷 프롬프트 엔지니어링을 활용하여, 키워드 추출 과정을 한 단계 발전시키는 방법론을 제시한다. 기존 키워드 추출 생성기법들 중 BART와 RoBERTa(A Robustly Optimized BERT Pretraining Approach)[7] 모델을 기반으로 한 KBIR(Keyphrase Boundary Infilling with Replacement)[8] 모델을 활용한다. 이 모델들은 초기 키워드 추출 과정과 프롬프트를 통한 속성값 결합 과정을 통합하여, 최종적으로 연구 내용의 복잡성과 전문성을 효과적으로 반영하는 키워드를 생성한다. 이 과정은 연구 문서의 가독성과 접근성을 향상시키며, 특히 전문 지식이 없는 이용자들이 복잡한 연구 내용을 보다 명확하고 직관적으로 이해할 수 있게 한다.

키워드 추출과 프롬프트 엔지니어링의 통합 접근법은 과학 연구와 기술 개발 문서의 복잡성과 전문성에 대응하기 위한 혁신적인 방법을 제시한다. 이 접근법은 대규모 언어 모델의 능력을 활용하여, 연구 내용의 깊이 있는 이해와 신속한 정보 검색을 가능하게 하며, 연구자들과 일반 사용자 모두에게 가치 있는 도구를 제공한다.

### III. 프롬프트 기반 키워드 결합 시스템

본 절에서는 연구 초록에서 키워드의 관련성을 깊이 있게 분석하고 체계적으로 분류하는 새로운 접근법을 제안한다. 이 접근법은 의도, 역할, 및 구성의 세 가지 핵심 요소에 중점을 두고, 키워드의 다면적인 특성을 명확히 파악하기 위해 설계되었다.

제안하는 프롬프트 기반 키워드 결합 시스템 흐름은 그림 1과 같다. 분석 과정은 먼저 연구의 기본 목적과 분석 목표를 설정하는 ‘키워드 분석의 방향성’의 설정 유무에 따라 의도 요소를 적용하여 연구의 근본적인 목적과 분석 목표를 명확히 한다.

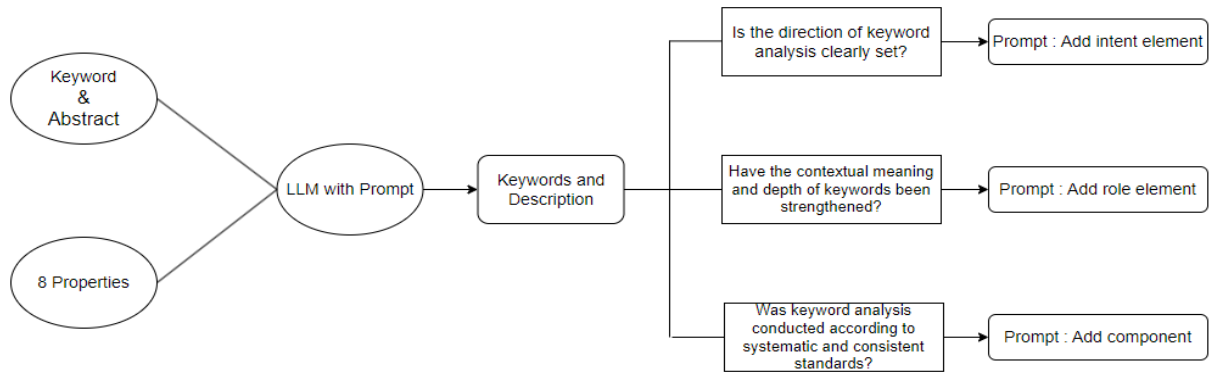


그림 1. 제안하는 프롬프트 기반 키워드 결합 시스템 흐름도  
 Fig. 1. Flowchart of the proposed prompt-based keyword combination system

다음으로, ‘키워드 분석 깊이’에 따라 각 키워드가 연구 내에서 수행하는 구체적인 역할을 심화시키며, 역할 요소를 통해 이를 분석한다. 마지막 단계에서는 ‘키워드 분석의 체계성’을 검토하여, 분석 과정의 표준화와 일관성을 구성 요소로 확보한다. 이러한 체계적 분석을 통해 키워드의 관련성과 중요성을 보다 명확하게 이해하고, 연구 내용의 심도와 체계성을 향상시킨다.

표 1. 키워드 속성 분류 유형의 예시  
 Table 1. Keyword attribute classification type

(Materials & Sample)	Keywords indicating the main substance, organism, or compound used in the study
(Research Technique & Equipmen)	Keywords related to experiments, measurements, analytical methods, numerical ratios, concepts, or specific equipment used in the study
(Physical/Chemical Propertie)	Keywords that describe the physical/chemical properties of the substance being studied
(Processes & Reactions)	Keywords related to specific processes or reactions observed (used) in research, or scientific reactions
(Applications & Uses)	Keywords that indicate how research results or findings can be used
(Time & Chronology)	Keywords that indicate a specific time or era covered by the study
(Topic)	Keywords representing topic information covered in the study
(Site & Facility & Environment)	Keywords that refer to a specific area or place where research or experiments were conducted

논문 내 키워드는 연구의 핵심 개념을 압축적으로 표현하며, 연구의 주요 내용과 방향성을 짚어내는 중요한 지표로 기능한다. 이는 특히 특정 기술이나 장비를 사용하는 연구에서 더욱 중요하다. 키워드의 속성을 체계적으로 분류함으로써, 해당 기술이나 장비와 관련된 연구의 주제와 방향성을 명확히 파악할 수 있다.

표 1은 과학기술분야의 키워드 속성 분류 유형 예시이다. 이 과정은 연구자들이 관련된 정보에 더 빠르고 효율적으로 접근하게 하며, 연구 주제 선정, 관련 문헌 검색, 그리고 연구 방향 설정 시간을 단축시키는 데 크게 기여한다. 따라서 키워드 속성의 정의와 검증 과정은 연구 분야의 최신 동향을 이해하고, 정보 검색을 용이하게 만드는 데 중요한 역할을 한다.

### 3.1 의도 요소

의도 요소의 포함은 키워드 분석 및 해석 과정에 있어 방향성을 설정하는 중요한 요소이다. 키워드를 결합 및 분류에 있어 연구 초록 내에서 각 키워드가 어떠한 의미를 지니며, 연구의 어떤 부분과 연결되는지를 명확하게 이해하는 데 핵심적인 역할을 한다. 의도 요소가 포함되지 않은 상태에서의 결합 및 분류는 표면적인 정보에 기반할 수 있으나, 의도 요소를 명시함으로써 연구의 본질적인 목적과 키워드의 깊은 의미를 더욱 정확하게 파악할 수 있다.

표 2은 의도 요소가 포함된 프롬프트 예시의 결과이다.

예를 들어, 'fluorescence chemosensor'라는 키워드는 의도 요소의 유무에 따라 다르게 해석될 수 있다. 의도 요소가 명확히 포함되지 않은 상황에서는, 이 키워드가 단지 연구에서 사용된 기술이나 장비로 분류될 수 있다. 반면, 의도 요소가 포함된 분석에서는, 이 기술이 특정 이온을 감지하는 데 어떻게 핵심적인 역할을 하는지, 연구의 전체적인 목표에 기여하는 방식을 명확히 이해할 수 있게 된다. 이는 해당 기술의 선택이 연구 결과와 직접적으로 관련되어 있음을 강조한다. 또한, 'excited-state intramolecular proton transfer (ESIPT)'와 같은 키워드의 경우, 의도 요소를 포함한 분석은 해당 과정이 연구 내에서 어떠한 중요한 역할을 수행하는지에 대한 심층적인 이해를 제공한다. 이는 연구 초록을 통한 해석 과정에서 연구 내용의 전체적인 맥락을 파악하고, 깊이 있는 의미를 이해하는 데 필수적이다. 결론적으로, 의도 요소의 명확한 포함은 키워드 분석 및 해석 과정에서 분석의 방향을 설정하는 데 결정적인 역할을 하며, 이는 연구 초록의 깊이와 정확성을 향상시키는 데 중요한 기여를 한다.

표 2. 의도 요소 포함된 프롬프트 예시  
Table 2. Example prompt with intent element

Points vary depending on whether intent element is included or not	
- "The purpose of this work is to": Clearly present work purpose "explain why based on information found in the abstract" : Request for explanation of reason for classification	
Prompt results that do not contain intent elements	Prompt result with intent element
1. Research Technique & Equipment: - 'fluorescence chemosensor' - This is a technique or tool used to detect or measure the presence of substances through fluorescence.	1. Research Technique & Equipment: - fluorescence chemosensor: This term implies a type of equipment or method being utilized to detect the presence of certain ions based on fluorescence.
2. Processes & Reactions: - 'excited-state intramolecular proton transfer (ESIPT)' - This refers to a specific process that occurs within the chemical structure (intramolecular) that is being inhibited by cation-induced interactions to cause fluorescence changes.	2. Processes & Reactions: - excited-state intramolecular proton transfer (ESIPT): This represents a specific process or reaction that occurs within the chemosensor when interacting with ions.

### 3.2 역할 요소

프롬프트에 역할 요소를 포함하는 접근법은 연구 초록 분석의 정밀도를 높이는 데 중요한 역할을 한다. 이는 특히 각 키워드가 연구 내에서 수행하는 구체적인 역할과 그 중요성을 명확히 드러내고자 할 때 유용하다. 의도 요소가 포함된 프롬프트가 제공하는 분석의 방향성에 깊이와 세밀함을 더해준다. 의도 요소가 연구의 전반적인 목적과 키워드가 연구 주제와 어떻게 연결되는지에 대한 분석을 제공한다. 역할 요소는 그 한 단계 더 나아가 각 키워드가 연구 내에서 실제로 어떠한 기능을 수행하는지에 대한 분석을 가능하게 한다. 이러한 접근은 키워드의 단순한 사전적 정의를 넘어서, 그것들이 연구 내용과 어떻게 상호작용하는지에 대한 깊이 있는 이해를 제공한다. 표 3은 역할 요소가 포함된 프롬프트 예시의 결과이다. 예를 들어, 역할 요소가 제외된 상황에서 'fluorescence'라는 키워드는 주로 그 물리적 혹은 화학적 특성으로 인식되어, 연구 내에서 이 현상이 어떠한 과정이나 반응을 나타내는지에 대한 구체적인 분석없이 분류될 수 있다. 반면, 역할 요소를 포함한 프롬프트를 사용하면, 'fluorescence'는 연구 내에서 fluorescence chemosensor의 핵심적인 기능으로 인식되며, 이 기능이 연구의 어떤 부분에 기여하는지, 왜 중요한지에 대한 더 깊은 분석을 제공한다. 이로써 키워드가 단순한 용어를 넘어서 연구 내용의 핵심 구성 요소로서의 역할을 분명히 할 수 있다. 더 나아가, 'chemosensor'라는 키워드 역시 역할 요소의 포함 유무에 따라 다르게 해석된다. 역할 요소가 없는 경우, 이 키워드는 연구에서 사용된 한 종류의 물질 또는 재료로 간주된다. 그러나 역할 요소를 포함하여 분석할 때, 'chemosensor'는 특정 이온에 반응하여 그 존재를 감지하는 중요한 도구로서의 역할이 강조된다. 이는 연구의 목적과 방향을 더 명확히 하며, 키워드와 연구 내용 간의 깊은 관련성을 드러내는 데 도움이 된다. 결론적으로, 역할 요소의 포함은 연구 초록의 복잡한 내용을 보다 정확하게 분석하고 해석하는데 있어 핵심적인 역할을 한다. 키워드가 연구 내에서 수행하는 구체적인 역할을 명확히 파악함으로써, 연구자들은 연구 내용에 대한 보다 심층적인 이해

를 얻을 수 있으며, 이는 연구 문서의 전반적인 가독성과 이해도를 향상시키는 데 기여한다.

표 3. 역할 요소 포함된 프롬프트 예시  
Table 3. Example prompt with role element

Points that vary depending on whether role elements are included or not	
- "As a research abstract analyst, you are responsible for" : Specify roles to be performed - "The validity of each classification must be made clear based on the contents of the abstract"와 "You should include which part of the abstract each keyword connects to, or, if not, why" : Emphasis on role performance	
Prompt results that do not contain role elements	Prompt results with role elements
1. <b>**fluorescence**</b> : Highly relevant because the key functional attribute of the chemosensor is its fluorescence behavior upon interaction with certain metal ions (Cu <sup>2+</sup> and Zn <sup>2+</sup> ). Fluorescence quenching and ratiometric change are the primary observations used to detect these ions.	1. <b>**fluorescence**</b> - This keyword refers to the light emitted by a substance when it has absorbed light or other electromagnetic radiation. It is a principle characteristic of the chemosensor being studied for its reaction to Cu <sup>2+</sup> and Zn <sup>2+</sup> ions. - Category: Physical/Chemical Properties
2. <b>**chemosensor**</b> : This keyword is directly relevant to the research because the abstract describes the development of a new bisthiazole chemosensor. It is essentially the main subject of the study, indicating the role of the new compound in sensing specific cations through fluorescence changes.	2. <b>**chemosensor**</b> - This term represents the main subject of the study, which is a device or substance specifically designed to detect the presence of particular elements or compounds—in this case, the bisthiazole chemosensor that detects Cu <sup>2+</sup> and Zn <sup>2+</sup> ions. - Category: Materials & Sample

### 3.3 구성 요소

프롬프트에 구성 요소를 포함시키는 것은 키워드 분석 과정에 명확한 지침과 구조를 제공한다. 구성 요소의 명시적인 포함은 분석할 키워드의 수를 제한하고, 각 키워드를 특정 형식에 따라 분류하도록 요구함으로써, 분석의 일관성과 체계성을 보장한다. 이러한 접근 방식은 키워드를 더 명확하게 분석하고, 그들의 연구 내용과의 연결점을 더 세밀하게 설명하는 데 도움이 된다. 표 4은 구성 요소가 포함된 프롬프트 예시의 결과이다. 예를 들어, 구성 요소가

포함되지 않은 프롬프트에서는 ‘fluorescence chemosensor’라는 키워드가 있을 때, 구성 요소를 제외한 상태에서는 이 키워드가 연구에서 사용된 특정 기술이나 장비로 간주된다. 그러나 구성 요소를 포함하여 분석할 경우, 이 키워드가 연구 맥락에서 차지하는 구체적인 위치와 중요성을 인지하고, 연구에서 수행하는 정확한 역할에 초점을 맞추게 된다. 이는 분석을 통해, ‘fluorescence chemosensor’가 단순한 장비가 아닌, 특정 이온과 상호작용할 때 빛을 발하는 방식으로 화학적 변화를 감지하는 중요한 도구로서의 역할을 강조한다.비슷하게, ‘ratiometric change’라는 키워드는 구성 요소의 포함 없이는 두 가지 다른 파장의 형광 강도 변화의 비율이라는 물리적/화학적 성질로 분류될 수 있지만, 구성 요소가 포함된 분석에서는 이 형광 강도의 비율적 변화가 어떻게 Zn<sup>2+</sup> 이온의 존재를 나타내는지, 즉 연구에서 이 현상이 갖는 의미와 기능을 좀 더 명확하게 파악한다.결론적으로, 구성 요소는 키워드를 단순한 카테고리에 할당하는 것을 넘어서, 각 키워드가 연구에서 어떻게 사용되고, 어떤 중요한 결과를 도출하는지에 대한 이해를 극대화하는 데 기여한다.

표 4. 구성 요소 포함된 프롬프트 예시  
Table 4. Example prompt with components

Points vary depending on whether components are included or not	
- "Keyword: Attribute Type - Reason" : Establish clear criteria for organizing analysis results	
Prompt results with no components included	Prompt results with components
1. <b>**fluorescence chemosensor**</b> - Research Technique & Equipment: This combined keyword reinforces the mentioned technique and equipment (chemosensor) specifically for the purpose of detecting changes in fluorescence.	1. fluorescence chemosensor: Research Technique & Equipment - It refers to a specific device or tool that measures or detects fluorescence to sense chemical changes.
2. <b>**ratiometric change**</b> - Physical/Chemical Properties: This might refer to a change in the ratios of fluorescence intensity at two different wavelengths. It is a property observed when the chemosensor interacts with Zn <sup>2+</sup> ions according to the abstract.	2. ratiometric change: Physical/Chemical Properties - This describes a change in the ratio of two properties (in this case, likely fluorescence intensities at different wavelengths) that is used to measure something, which in this context, is a physical property measurement.

이를 통해 연구 초록의 복잡한 내용을 보다 체계적이고 정확하게 분석하고, 연구 문서의 가독성과 이해도를 향상시키는 데 중요한 역할을 한다.

#### IV. 실험 및 분석

본 절에서는 텍스트 내 키워드 경계를 식별하고 적합한 키워드를 삽입하는 데 유용한 KBIR, 문서의 주요 내용을 요약하여 핵심 키워드를 도출하는 데 유용한 BART와 이 두 모델을 통합 사용함으로써 각 모델의 장점을 결합하여 키워드 추출 과정에 프롬프트 엔지니어링 기법이 어떻게 성능에 영향을 미치는지 평가한다. 정밀도(Precision)와 재현율(Recall)을 중심으로 한 성능 평가는, 프롬프트 엔지니어링의 다양한 요소 의도, 역할, 구성 요소의 적용을 통해 모델의 성능 변화를 관찰하는 데 초점을 맞췄다. 정밀도는 모델이 식별한 항목 중 실제로 관련이 있는 항목의 비율을 나타내며, 재현율은 실제로 관련있는 항목 중 모델이 정확히 식별한 항목의 비율을 의미한다.

제안기법의 적절성 평가를 위해, 다양한 하위 분야와 특정 전문 용어를 가지고 있는 지질·화학 분야의 SCI급 논문 100개의 초록들을 대상으로 키워드 추출과 분석 과정에 프롬프트 엔지니어링 기법을 적용하여 정제된 키워드들을 기준으로 성능 변화를 평가하였다. 표 5는 프롬프트 엔지니어링의 요소를 적용함으로써 얻어진 실험 결과이다.

KBIR 모델을 기본 설정으로 사용했을 때, 정밀도는 0.16, 재현율은 0.13으로 나타났다. 프롬프트 엔지니어링의 각 요소 의도, 역할, 구성을 순차적으로 적용하며 성능 변화를 관찰하였다. 의도 요소를 추가하면서 정밀도가 0.08로 감소하였으나, 이는 모델이 더 많은 관련 키워드를 식별하면서 동시에 불관련 키워드도 증가한 결과이다. 재현율은 0.21로 증가하였는데, 이는 모델이 더 많은 관련 키워드를 포착하였기 때문이다. 역할 요소를 적용한 후에는 정밀도가 0.09로 소폭 상승하고 재현율은 0.19로 더욱 증가하였다. 구성 요소 추가 시 정밀도는 0.17로 증가하고 재현율은 0.33으로 대폭 상승하였다. 이는 구성 요소가 분석의 체계성을 높여 모델이 보다 정확하게 관련 키워드를 식별하도록 도왔음을 보여준다.

BART 모델의 경우, 기본 설정에서 정밀도가 0.43, 재현율이 0.37로 KBIR 모델보다 높은 성능을 보였다. 의도 요소 추가 시 정밀도는 0.19로 감소하였으나, 재현율은 0.33으로 유지되었다. 역할 요소 적용 결과 정밀도는 변화가 없었으나, 재현율은 0.37로 상승하였다. 이는 모델이 더 많은 관련 키워드를 포착한 것을 알 수 있다. 구성 요소를 추가하면서 정밀도가 0.30으로 상승하고, 재현율은 0.47로 가장 높은 성능을 보였다. 이 결과는 BART 모델이 구성 요소의 체계적인 분석 프레임워크를 통해 키워드의 복잡성과 다양성을 더 잘 처리할 수 있음을 확인하였다.

통합 모델에서는 기본 설정에서 정밀도가 0.27, 재현율이 0.39로 시작한다. 이는 여러 프롬프트 엔지니어링 요소들이 적용되기 전의 초기 모델 성능을 나타낸다. 의도 요소를 추가한 후에는 정밀도가 0.15로 감소하였으나, 재현율은 0.42로 증가한다. 이는 의도 요소가 모델에게 더 많은 관련 키워드를 포착하도록 하지만, 동시에 관련성없는 키워드 추출도 증가시켜 정밀도를 저하시킨 것이다. 역할 요소를 적용하면 의도요소 적용시와 유사하게 관련성없는 키워드 추출이 증가하여 정밀도가 0.12로 더욱 낮아지지만, 재현율은 0.37로 비교적 높은 수준을 유지한다. 구성 요소를 추가했을 때 정밀도는 0.24로 상승하고, 재현율은 0.45로 높은 성능을 보이는데, 이는 구성 요소가 분석을 더 체계적으로 만들어 정밀도와 재현율이 모두 향상되었다.

P@5는 키워드 추출에서 가장 관련성이 높다고 판단되는 상위 5개의 키워드만을 고려하여 성능을 측정하는 지표이다. 기존 정밀도 대비 P@5 정밀도에서 수치가 향상되는 주된 이유는 모델이 불필요한 키워드의 추출을 줄이고 상위 5개 내에서 가장 관련성 높은 키워드에 집중하기 때문이다. 즉, 전체 키워드 추출에서는 많은 수의 불관련 키워드가 포함될 수 있지만, P@5 설정에서는 이를 제한함으로써 정밀도가 자연스럽게 높아지게 된다.

반면, 재현율이 감소하는 이유는 모델이 식별해야 할 전체 중요 키워드 중에서 오직 5개만을 선택하기 때문이다. 키워드의 총 수가 제한되면, 비록 상위 5개의 키워드가 매우 정확할지라도 전체 중요 키워드 중 일부만을 포착할 수 있어, 재현율이 낮아

진다. 이는 모델이 더 많은 정답 키워드를 식별할 가능성을 제한하므로, 키워드 추출의 전체적인 포괄성이 감소하게 된다.

본 연구 결과는 프롬프트 엔지니어링의 구성 요소가 언어 모델의 키워드 추출 과정에서 성능 향상 요소임을 보여준다. 구성 요소의 체계적인 적용은 모델이 보다 정확하게 관련 키워드를 식별하고 분류하는 데 기여한다. 다만, 각 모델의 특성에 따라 의도와 역할 요소의 추가가 정밀도와 재현율에 다른 영향을 미친다. KBIR 모델의 경우, 의도와 역할 요소의 추가로 정밀도와 재현율이 상승하는 경향을 보인다. 이는 각 요소가 모델의 성능을 개선하는 데에 긍정적인 효과를 가져온다. 반면, BART 모델은 의도와 역할 요소의 추가로 정밀도가 감소하는 경향을 보였지만, 재현율은 상승한다. 이는 BART 모델이 더 많은 관련 키워드를 추출하였지만, 불필요한 키워드도 증가하여 정밀도가 감소한 것이다.

표 5. 프롬프트 엔지니어링에 따른 실험 결과  
Table 5. Experimental results according to prompt engineering

	Precision (P@10)	Recall (R@5)	Precision (P@5)	Recall (R@5)
Basic KBIR	0.16	0.13	-	-
+ Intent element	0.08	0.21	0.25	0.21
+ Role element	0.09	0.19	0.26	0.19
+ Component	0.17	0.33	0.22	0.33
Basic BART	0.43	0.37	-	-
+ Intent element	0.19	0.33	0.26	0.27
+ Role element	0.19	0.37	0.24	0.24
+ Component	0.30	0.47	0.16	0.16
Basic KBIR + BART	0.27	0.39	-	-
+ Intent element	0.15	0.42	0.24	0.24
+ Role element	0.12	0.37	0.26	0.27
+ Component	0.24	0.45	0.21	0.21

따라서, 각 모델의 특성을 고려하여 추출 키워드의 깊이와 정확성을 향상시키고자 할 때, 해당 모델에 맞는 프롬프트 엔지니어링 기법의 적용을 고려해야 함을 시사한다.

제안하는 키워드 결합 시스템에서 얻은 답변의 판별 근거의 타당성을 파악하기 위해, OpenAI의 GPT-4-Turbo 모델을 활용하여, 속성값을 고려한 키워드 결합 결과가 연구 초록과 각각의 요소들에 얼마나 부합하며 어떤 상관관계를 가지는지에 대해 Likert 척도[10]로 평가했다.

기존의 키워드 결합 방법은 재현율과 정밀도를 기준으로 삼아 추출된 키워드들의 연구 초록과의 관련성을 평가했다. 그러나 이러한 방법은 추출된 키워드들이 연구의 주요 내용을 충분히 반영하고 있는지를 정량적으로 평가하기 어렵다. 이는 저자 키워드만이 절대적인 정답이 아니기 때문이다. 특정 연구의 핵심 내용은 다양한 요소에 의해 구성될 수 있으며, 이는 종종 저자 키워드로만 충분히 표현되지 않을 수 있다. 따라서 이전 방법보다 더욱 타당한 평가를 위해 Likert 척도를 활용하여 키워드 추출의 타당성을 평가하였다.

표 6의 실험에서는 3가지 주요 요소가 키워드 추출의 타당성에 얼마나 개선이 되었는지 확인하였다. 먼저, 의도 요소를 추가한 경우, 기본 KBIR, BART 및 통합 모델에서의 평균 타당성이 0.34점 높은 점수를 보였다.

표 6. 결합 키워드 타당성 평가 Likert 척도 평균 점수  
Table 6. Combined keyword validity assessment likert scale mean score

	Score avg.
Basic KBIR	3.27
+ Intent element	3.65
+ Role element	3.82
+ Component	3.85
Basic BART	3.65
+ Intent element	3.83
+ Role element	3.94
+ Component	3.98
Basic KBIR + BART	3.34
+ Intent element	3.8
+ Role element	3.84
+ Component	3.93



이는 키워드 추출 과정에서 의도에 맞게 추출된 키워드들이 연구 초록의 핵심 내용을 보다 정확하게 반영했음을 알 수 있다.

역할 요소를 추가한 경우, 3개의 기본 모델에서 의도 요소가 추가됐을 때의 타당성보다 평균 0.1점의 높은 점수를 보였다. 이는 역할 요소가 연구 내용의 역할과 책임을 명확히 이해하고 반영함에 따라 의도 요소보다 키워드 추출의 타당성이 향상된 것을 시사한다.

마지막으로, 구성 요소를 추가한 경우, 역할 요소가 추가됐을 때의 타당성보다 평균 0.05점 높은 점수를 보였다. 이는 구성 요소가 추가됨으로써 추출된 키워드 내용이 보다 체계적이고 구조화되어 있어, 연구 초록의 주요 내용을 더 명확하게 전달함을 시사한다.

이러한 실험 결과는 키워드 추출의 정량적 평가에서 Likert 척도를 활용하는 것이 이전 방법보다 더욱 타당하고 신뢰할 만하다는 것을 보여준다. 또한, 의도, 역할 및 구성 요소와 같은 주요 요소가 키워드 추출의 품질을 향상시키는 데 중요한 역할을 한다는 것을 확인했다.

## V. 결 론

본 연구에서는 데이터 정제 및 키워드 결합 시스템에서의 신뢰성 향상을 위해 프롬프트 엔지니어링 기법을 적용하여 이를 Likert 척도를 활용하여 평가하였다. 이전에는 재현율과 정밀도를 주요 기준으로 삼아 키워드 추출의 품질을 평가했으나, 이는 연구의 핵심 내용을 충분히 반영하기 어려웠다. 그에 따라 Likert 척도를 도입하여 보다 신뢰성 있는 평가를 실시하였고, 이로써 키워드 추출의 타당성을 보다 정량적으로 확인할 수 있었다.

실험을 통해, 의도, 역할 및 구성 요소와 같은 주요 요소가 키워드 추출의 품질을 향상시킬 수 있음을 확인할 수 있다. 특히 BART 모델은 이러한 요소들을 통합할 때 좋은 성능을 보여주었고, 복잡한 데이터에서 키워드의 정확성과 관련성을 높이는 데 효과적임을 볼 수 있다. 이러한 결과는 프롬프트 엔지니어링 기법이 단순히 키워드 추출뿐만 아니라,

데이터의 정제 및 결합 과정에 있어서도 중요한 역할을 할 수 있다. 이는 연구 문서의 복잡성과 전문성을 보다 잘 이해하고, 연구 내용의 체계성과 깊이를 향상시키는 데 기여할 것으로 기대된다. 또한, 본 연구 방법론은 다양한 학문 분야뿐만 아니라 산업적 응용에서도 키워드 추출의 정확성과 효율성을 높이는 데 도움이 된다.

향후 연구에서는 이 방법론을 더 다양한 언어와 문헌 유형에 적용하여 그 유효성을 검증할 예정이다. 이를 통해 더 정교하고 다양한 키워드 추출 기법을 개발하고, 기존의 방법론을 보완하는 새로운 접근 방법을 모색할 것이다. 더 나아가, 인공지능 기술의 발전과 함께 이러한 접근법이 연구자들에게 보다 효과적인 도구를 제공하여, 연구 내용의 분석과 해석을 더욱 풍부하고 다층적으로 할 수 있게 할 것으로 기대된다.

## References

- [1] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", arXiv preprint, arXiv:1810.04805, Oct. 2018. <https://doi.org/10.48550/arXiv.1810.04805>.
- [2] J. Lafferty, A. McCallum, and F. C. N. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data", ICML '01: Proceedings of the Eighteenth International Conference on Machine Learning, MA, USA, pp. 282-289, Jun. 2001.
- [3] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by Latent Semantic Analysis", Journal of the American Society for Information Science, Vol. 41, No. 6, pp. 391-407, Sep. 1990. [https://doi.org/10.1002/\(SICI\)1097-4571\(199009\)41:6<391::AID-ASII>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9).
- [4] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to Sequence Learning with Neural Networks", NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing

- Systems, Montreal, Canada, Vol. 2, pp. 3104-3112, Dec. 2014.
- [5] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, and L. Zettlemoyer, "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension", arXiv preprint, arXiv:1910.13461, Oct. 2019. <https://doi.org/10.48550/arXiv.1910.13461>.
- [6] C. Raffel, et al., "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer", arXiv preprint, arXiv:1910.10683, Oct. 2019. <https://doi.org/10.48550/arXiv.1910.10683>.
- [7] Y. Liu, et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach", arXiv preprint, arXiv:1907.11692, Jul. 2019. <https://doi.org/10.48550/arXiv.1907.11692>.
- [8] M. Kulkarni, D. Mahata, R. Arora, and R. Bhowmik, "Learning Rich Representation of Keyphrases from Text", arXiv preprint, arXiv:2112.08547, Dec. 2021. <https://doi.org/10.48550/arXiv.2112.08547>.
- [9] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, "A Neural Probabilistic Language Model", Journal of Machine Learning Research, Vol. 3, pp. 1137-1155, Mar. 2003.
- [10] R. Likert, "A Technique for the Measurement of Attitudes", Archives of Psychology, No. 140, pp. 1-55, 1932.
- [11] T. B. Brown, et al., "Language Models are Few-Shot Learners", arXiv preprint, arXiv:2005.14165, May 2020. <https://doi.org/10.48550/arXiv.2005.14165>.
- [12] T. Mikolov, et al., "Efficient Estimation of Word Representations in Vector Space", arXiv preprint, arXiv:1301.3781, Jan. 2013. <https://doi.org/10.48550/arXiv.1301.3781>.
- [13] Z. Dai, Z. Yang, Y. Yang, J. Carbonell, Q. V. Le, and R. Salakhutdinov, "Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context", arXiv preprint, arXiv:1901.02860, Jan. 2019. <https://doi.org/10.48550/arXiv.1901.02860>.
- [14] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation", Journal of Machine Learning Research, Vol. 3, pp. 993-1022, Jan. 2003.
- [15] C. D. Manning, P. Raghavan, and H. Schütze, "Introduction to Information Retrieval", Cambridge University Press, 2008. <https://doi.org/10.1017/CBO9780511809071>.
- [16] A. Vaswani, et al., "Attention Is All You Need", arXiv preprint, arXiv:1706.03762, Jun. 2017. <https://doi.org/10.48550/arXiv.1706.03762>.
- [17] T. B. Brown, et al., "Language Models are Few-Shot Learners", arXiv preprint, arXiv:2005.14165, May 2020. <https://doi.org/10.48550/arXiv.2005.14165>.
- [18] V. Sanh, et al., "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter", arXiv preprint, arXiv:1910.01108, Oct. 2019. <https://doi.org/10.48550/arXiv.1910.01108>.
- [19] Z. Yang, et al., "XLNet: Generalized Autoregressive Pretraining for Language Understanding", arXiv preprint, arXiv:1906.08237, Jun. 2019. <https://arxiv.org/abs/1906.08237>.
- [20] L. McMillan, et al., "Natural Language Processing with Transformers", O'Reilly Media, ISBN: 978-1098103000, Jan. 2022.
- [21] C. Olah, "Understanding LSTM Networks", 2015. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> [accessed: May 11, 2024]
- [22] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning", MIT Press, ISBN: 978-0262035613, 2016.

## 저자소개

김민근 (MinGeun Kim)



2019년 3월 ~ 현재 : 국립금오공과  
대학교 컴퓨터공학과 학사과정  
관심분야 : 메타추출 성능고도화,  
LLM(fine-tuning)

정유철 (Yuchul Jung)



2011년 2월 : 한국과학기술원  
(KAIST) 전산학과(공학박사)  
2013년 7월 : 한국전자통신연구원  
(ETRI) 선임연구원  
2017년 8월 : 한국과학기술정보  
연구원(KISTI) 선임연구원  
2022년 2월 : 국립금오공과대학교

컴퓨터공학과 조교수

2022년 2월 ~ 현재 : 국립금오공과대학교 인공지능공학과  
부교수

관심분야 : 거대언어모델, 자연어처리, 지식그래프,  
한국어 음성 인식/합성, AI응용