

A Performance Context Ontology Centered Integrated Ontology for Effective Expert Search

In-Ha Jun*, Hyeon-Jeong Mun**, Young-Ji Kim**, Su-Hyun Lee***, Yong-Tae Woo***

*Student, Dept. of Computer Engineering, Changwon National University, Gyeongsangnamdo, Korea

**Principal Research Engineer, Hibrainnet Co., Gyeongsangnamdo, Korea

***Professor, Dept. of Computer Engineering, Changwon National University, Gyeongsangnamdo, Korea

[Abstract]

In this study, we propose a Performance Context Ontology (PCO) that expresses expert competency information embedded in recruitment or academic domains in the same performance context structure and enables effective competency matching even between cross domains. First, we design an ontology schema by deriving a conceptual framework of performance contexts that can be commonly applied across cross domains using LLM. Then, ontology instances are constructed by extracting performance context concepts from unstructured data collected from each domain. To verify the effectiveness of the proposed ontology, an integrated ontology for recruitment and academic information was constructed to confirm whether competency-based matching is appropriately performed even in cross domains. In the future, we expect the proposed ontology to serve as a foundational technology for developing an expert search system to effectively identify key experts required by institutions.

▶ **Key words:** Performance Context Ontology, Ontology Design, Ontology Learning, Large Language Model, A.I., Expert Search

[요 약]

본 연구에서는 채용이나 학술 도메인에 내포된 전문가의 역량 정보를 동일한 수행 맥락 구조로 표현하고, 이종 도메인 간에서도 효과적으로 역량 매칭이 가능한 수행 맥락 온톨로지(Performance Context Ontology, PCO)를 제안한다. 먼저, LLM을 이용하여 이종 도메인에서 공통적으로 적용할 수 있는 수행 맥락 개념 체계를 도출하여 온톨로지 스키마를 설계한다. 그리고 각 도메인에서 수집한 비정형 데이터로부터 수행 맥락 개념을 추출하여 온톨로지 인스턴스를 구축한다. 제안 온톨로지의 효과성을 검증하기 위하여 채용 및 학술 정보에 대한 통합 온톨로지를 구축하여 이종 도메인에서도 역량에 부합하는 매칭이 적합하게 수행되는지 확인하였다. 향후, 제안 온톨로지는 기관에서 요구하는 핵심 전문가를 효과적으로 탐색하기 위한 전문가 검색 시스템 개발을 위한 기반 기술로 활용될 수 있으리라 기대한다.

▶ **주제어:** 수행 맥락 온톨로지, 온톨로지 설계, 온톨로지 학습, 대형 언어 모델, 인공지능, 전문가 검색

- First Author: In-Ha Jun, Corresponding Author: Yong-Tae Woo
- *In-Ha Jun (ihjun@hibrain.net), Dept. of Computer Engineering, Changwon National University
- **Hyeon-Jeong Mun (hjmun@hibrain.net), Hibrainnet Co.
- **Young-Ji Kim (yjkim@hibrain.net), Hibrainnet Co.
- ***Su-Hyun Lee (sleep@changwon.ac.kr), Dept. of Computer Engineering, Changwon National University
- ***Yong-Tae Woo (ytwoo@changwon.ac.kr), Dept. of Computer Engineering, Changwon National University
- Received: 2026. 05. 12, Revised: 2026. 05. 25, Accepted: 2026. 06. 09.

I. Introduction

현대 사회에서 인공지능과 데이터 기반 기술이 빠르게 확산되고 있다. 최근에는 산업 전반의 기술 구조가 급격하게 변화하는 인공지능의 전환 시대를 맞이하고 있다. 이에 따라, 첨단 기술을 연구하는 기관에서는 핵심 전문가의 확보가 기업의 성패를 좌우하는 중요한 요인이 되고 있다. 특히, 기관에서 필요로 하는 핵심 역량을 갖춘 전문가를 효과적으로 탐색할 수 있는 기법은 기관의 기술 경쟁력을 선도적으로 확보할 수 있는 중요한 전략의 하나이다[1].

일반적으로 기관에서 전문가 채용을 위해 등록하는 채용공고는 주로 기술 용어로 작성된 직무가 다양한 문서 형식으로 제시된다. 또한, 전문가의 역량을 표현하는 프로필 정보는 학력이나 경력 그리고 논문 같은 연구실적 정보로 구성된다. 이에 따라, 직무 중심의 채용공고와 개인별 역량 중심의 프로필 정보간의 불일치로 인해 전문가 탐색 과정이 비효율적으로 진행되고 있다. 또한, 스카우트의 수작업에 의한 전문가 탐색 과정으로 인해 많은 비용과 시간이 소요되는 문제점이 발생한다[1-3].

인공지능 기반의 전문가 검색 기법은 대량의 텍스트에 의해 사전에 학습된 LLM(Large Language Model)을 활용하여 문서간의 의미적인 연관성을 분석하기 위한 기법이다. 최근에는 이종 도메인 간의 문맥 정보를 기반으로 전문가 검색을 위한 Hi-BERT 모델이 제시되었다[4]. 이 모델은 학습된 언어 표현을 바탕으로 문서 간의 의미적 유사성과 문맥 정보를 효과적으로 반영할 수 있다. 하지만, 이 모델은 특정 도메인 지식에 대한 맥락적인 관계를 분석하기 위해 대량의 문서를 학습해야 하는 문제점이 있다.

온톨로지 기반의 전문가 검색 기법은 도메인 지식을 형식화된 개념 구조로 표현하고, 이를 기반으로 용어의 맥락 정보를 반영하여 의미적인 검색을 수행하는 기법이다[5]. 이러한 온톨로지 모델을 전문가 검색에 활용하는 연구가 활발하게 진행되고 있다[6-10]. 하지만, 이종 도메인에서 독립적으로 구축된 온톨로지는 서로 다른 개념 체계로 설계되어 상호간의 정보 매칭이 어려운 문제점이 있다. 이에 따라, 이종 도메인간에 수행 정보를 동일한 개념 체계로 구조화하여 정보 매칭에 활용할 수 있는 온톨로지 모델 개발이 필요하다.

본 연구에서는 채용과 학술 도메인에서 추출한 전문가 역량 정보를 동일한 맥락 구조로 표현하기 위한 수행 맥락 온톨로지(Performance Context Ontology, PCO) 모델을 제안하고 구축 방법을 제시한다. 제안 기법은 이종 도메인 간에 정보 매칭이 가능한 통합 온톨로지를 구축하기 위한

기법이다. 이 기법에서는 LLM을 이용하여 이종 도메인에서 적용 가능한 수행 맥락 개념 체계(Performance Context Conceptualization)를 구성하여 통합 온톨로지 로 형식화하는 방법을 제시한다. 또한 비정형 데이터로부터 수행 맥락 개념을 추출하여 인스턴스를 생성한다. 그리고 이종 도메인 간에 발생하는 용어간의 이질성을 해소하기 위한 방법을 제시한다.

제안 기법에 대한 효과성을 검증하기 위하여 채용공고와 학술논문을 대상으로 통합 온톨로지를 구축하였다. 하이버레인넷 사이트에 등록된 AI 분야의 채용공고와 Scopus 논문을 수집하여 실험하였다. 실험 결과, 본 연구에서 제안한 방법에 의해 이종 도메인에서 추출한 수행 정보를 동일한 맥락 구조로 일관성 있게 구성하였다. LLM을 이용하여 이종 도메인 간에 통합 온톨로지를 효과적으로 구축할 수 있었다. 또한, 통합 온톨로지에서도 직무 역량과 학술연구 활동 정보가 효과적으로 매칭되는지 여부를 검증하였다. 향후, 제안 기법은 전문가 검색 시스템 개발을 위한 핵심 기법으로 활용될 수 있으리라 기대한다.

II. Preliminaries

1. Related works

1.1 Ontology-based Search Method

온톨로지 기반 검색은 도메인 지식을 명시적으로 정의하여 의미적인 검색을 지원하기 위한 기법이다[5]. 그동안 채용 도메인에서 직무 정보를 효과적으로 표현하기 위한 온톨로지 모델이 제시되었다. ESCO, O*NET, NCS 기반 모델은 직업 분류, 직업 요구 스킬과 역량, 직무 능력 단위 정보를 표현하는 모델이다[6-8]. 하지만 기존의 직무 모델은 방법, 도구, 역량 등의 개념이 복합적인 스킬 개념으로 표현되거나, 직무에 포함된 스킬과 역량이 어떤 역할로 사용되는지를 구분하기 어려운 문제점이 있다.

학술 도메인에서도 연구 정보를 표현하기 위한 모델들이 제시되었다. 연구 절차 기록을 위한 EXPO, PROV[11,12], 연구 기여 정보를 표현하기 위한 ORKG, CRediT[9,10], 연구 활동의 구조적 표현을 위한 SKOO, TaDiRAH[13,14] 등과 같은 모델이 제시되었다. 하지만 기존의 연구 활동 온톨로지 모델은 주로 특정 도메인의 연구 절차, 연구 기록, 산출물, 기여 역할을 표현하는 데 초점을 두어, 연구자의 수행 역량을 직무 정보와 연결 가능한 공통 의미 구조로 표현하는 데 한계가 있다. 또한 기존 모델들은 각 도메인의 목적에 따라 서로 다른 개념과 속성

구조를 정의하여 도메인 간의 수행 정보를 동일한 기준에서 직접 비교하고 매칭하기 어려운 문제점이 있다.

1.2 Ontology Design Methodology

온톨로지 설계를 위한 원칙, 개발 절차, 재사용, 통합 및 검증 방법에 대한 연구도 활발하게 진행되고 있다. Gruber는 온톨로지 설계를 위한 명료성, 일관성, 확장성 등의 원칙을 제시하였다[15]. METHONTOLOGY는 대표적인 온톨로지 구축 방법론의 하나로 소프트웨어공학의 개발 생명주기 원칙을 온톨로지 구축에 적용하였다[16]. 이 방법은 온톨로지 구축 단계를 명세화, 지식획득, 개념화, 형식화, 구현, 통합, 평가 단계로 구분하였다. NeOn Methodology는 온톨로지 재사용과 통합을 지원하는 개발 방법론이다[17]. OntoClean에서는 온톨로지 계층 구조의 논리적 설계를 위한 방법론을 제시하였다[18].

1.3 Ontology Learning Methods

온톨로지 학습 기법은 데이터로부터 개념, 관계, 분류 체계와 같은 구성 요소를 추출하여 온톨로지를 효과적으로 구축하기 위한 기법이다[19]. 기존의 온톨로지 학습 기법은 자연어 처리, 통계 분석, 기계 학습 등과 같은 기법을 결합하여 용어 추출, 개념 관계 그리고 분류 체계를 구축한다[20]. 최근에는 LLM을 이용한 용어 분류, 분류 체계, 비분류 관계 추출 등과 같은 온톨로지 학습 기법이 제시되었다[21]. 이러한 학습 기법은 다양한 도메인에서 온톨로지 자동 생성과 의미 매핑 자동화로 확장되고 있다[22,23]. 하지만, 기존의 학습 기법은 이종 도메인의 수행 정보를 공통 구조로 통합하여 표현하기 어렵다.

III. The Proposed Ontology

본 연구에서는 채용 도메인의 직무 정보와 학술 도메인의 연구 활동 정보를 동일한 맥락 구조로 표현하여 이종 도메인 간에 역량 기반의 매칭을 위한 통합 온톨로지 구축 방법을 제안한다. 제안 방법은 METHONTOLOGY 방법론에 기반한 통합 온톨로지 구축 기법이다. 먼저, LLM을 이용하여 이종 도메인의 공통 수행 맥락 정보를 개념화하고 온톨로지 스키마를 설계한다. 설계된 스키마를 기반으로 비정형 데이터에서 수행 맥락 구조를 자동으로 추출하여 인스턴스를 생성한다. 그림 1은 본 연구에서 제안한 통합 온톨로지에 대한 개념도이다.

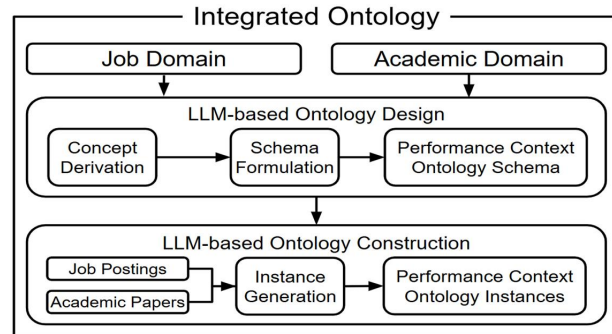


Fig. 1. PCO-Centered Integrated Ontology Concepts

1. Performance Context Ontology Model

본 연구에서는 전문가가 수행하는 활동 분야, 목적, 대상, 방법, 결과 등에 대한 개념과 개념간의 의미적 관계를 표현하는 수행 맥락 온톨로지 모델을 제안한다. 예를 들어, ‘답러닝 기반 의료 영상 분석 모델 개발’에서 수행 맥락 정보는 ‘의료’ 분야에서 ‘답러닝’ 방법으로 ‘영상 분석 모델’을 ‘개발’하는 활동으로 표현할 수 있다. 본 단계에서는 이러한 수행 맥락에 의해 전문가의 역량을 의미적인 구조로 표현하기 위한 온톨로지 모델을 제안한다.

온톨로지 설계는 채용 직무 정보 및 학술 도메인의 연구 활동 정보에서 도출한 수행 정보를 개념화하여 온톨로지 스키마로 형식화하는 과정이다. 먼저, 수행 정보 개념화는 LLM을 이용하여 이종 도메인에 공통적으로 적용되는 수행 맥락 개념(Performance Context Concept) 체계를 도출하는 과정이다. 여기서 수행 맥락 개념은 전문가가 직무와 연구 과정에서 수행하는 활동 역량을 범위와 방법 등으로 분류하고, 활동 역량을 세분화한 개념이다. 그리고 온톨로지 스키마 형식화는 도출된 개념 체계를 기반으로 온톨로지를 설계하고 스키마를 정의하는 과정이다.

1.1 LLM-based Conceptualization Method

본 연구에서는 LLM을 이용하여 도메인 전문가의 개입 없이 채용과 학술 도메인의 수행 맥락을 자동적으로 추출하여 개념화하기 위한 새로운 방법을 제안한다. 제안 방법은 LLM 기반의 온톨로지 학습 기법을 이용하여 이종 도메인에서 공통적인 수행 맥락 개념, 수행 맥락 개념간의 관계, 개념 유형을 자동으로 도출하는 방법이다. 제안 방법에서는 이종 도메인에서 도출되는 공통 개념 체계를 이용하여 채용 직무와 연구 활동 정보를 동일한 맥락 구조로 표현한다. 그림 2는 LLM 기반의 수행 맥락을 도출하여 개념화하는 과정이다.

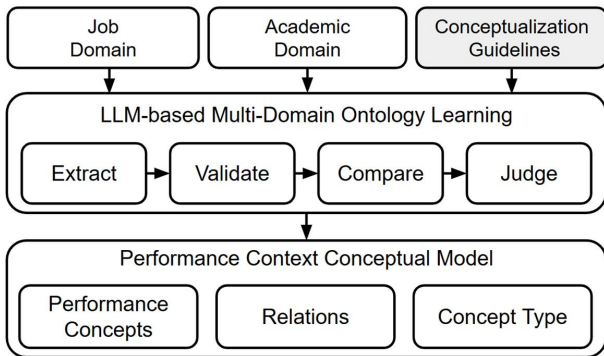


Fig. 2. Conceptualization Process for Deriving Performance Context Concepts

먼저, 채용과 학술 도메인에서 공유하는 수행 맥락 개념을 도출한다. 이 과정에서 LLM이 어휘 분포로 인해 편향된 개념을 도출하는 것을 방지하기 위해 SKOO(Scientific Knowledge Objects Ontology)와 같은 학술 온톨로지 명세와 NCS(National Competency Standards)와 같은 직무 모델 명세를 활용한다. LLM은 직무 및 논문 데이터를 분석하여 유효한 수행 정보 개념만 도출할 수 있도록 개념 원칙과 후보 채택 기준을 사전에 프롬프트 형식으로 정의한다. 온톨로지 학습 과정에서는 개념 후보 추출(Extract), 적합성 검증(Validate), 의미 비교(Compare), 채택 판정(Judge) 단계에서 정의된 지침에 따라 후보 개념을 추출하고, 수행 맥락 개념을 판정한다. 이를 통해, 이종 도메인이 공유 가능한 수행 맥락 개념과 개념간 관계를 도출할 수 있다. 그림 3은 활동 역량을 중심으로 개념과 개념간의 관계를 표현한 다이어그램이다.

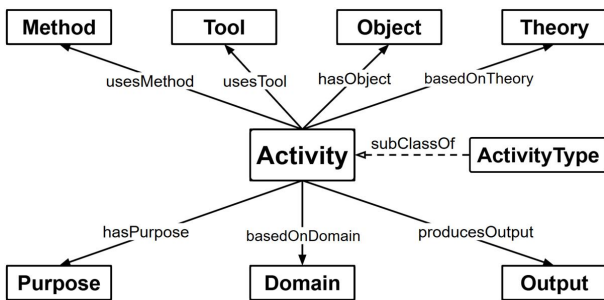


Fig. 3. Diagram of the Activity-Centered Performance Context Concept Structure

1.2 Derivation of Performance Activity Types

본 연구에서는 채용공고의 직무 내용과 학술논문의 연구 활동에 대한 매핑을 위해 활동 역량을 세분화하였다. 예를 들어, '자율주행에서 보행자를 감지할 수 있는 객체 인식 딥러닝 모델을 개발하였다'와 '자율주행 시스템의 신뢰성을 확보하기 위해 딥러닝 모델의 정확도를 분석하였

다'에서 '딥러닝'이라는 Method는 동일하지만, Purpose에 따라 '모델 개발'과 '정확도 분석'의 Activity로 구별된다. 이러한 Activity를 세분화함으로써 전문가의 역량을 표현하는 수행 맥락에 대한 정확도를 높일 수 있다.

여기서 Activity는 연구, 개발, 분석, 평가, 검증, 분류 등과 활동 역량을 표현하는 술어로, 유사 술어를 클러스터링한 후 의미를 정렬하여 개별 활동 유형으로 세분화한다. 이러한 활동 유형 체계는 적용 도메인의 특성과 범위에 따라 추가적으로 확장될 수 있다. 그림 4는 이종 도메인에서 수행 활동 유형(Activity Type)을 도출하는 과정이다.

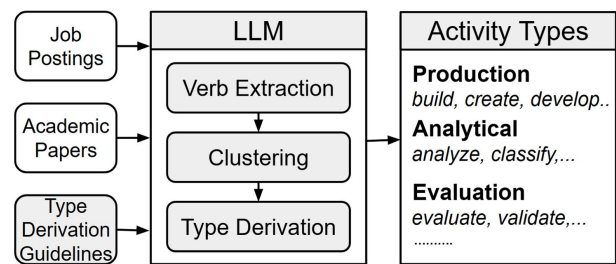


Fig. 4. Performance Activity Type Derivation Process

1.3 Integrated Ontology Schema Design and Formalization

본 연구에서 제안한 통합 온톨로지는 채용 및 학술 도메인을 표현하기 위한 도메인 온톨로지와 이종 도메인간 공유를 위한 수행 맥락 온톨로지 구성된다. 먼저, 도메인 온톨로지는 각 도메인의 고유 개념과 속성으로 정의된다. 수행 맥락 온톨로지는 이종 도메인에서 추출한 공통 수행 맥락 개념을 정의한다. 이러한 구조는 도메인별로 독립성을 보장하면서 이종 도메인간 수행 맥락의 효과적인 연결과 확장에 유연하다. 그림 5는 도메인 온톨로지와 수행 맥락 온톨로지의 개념과 온톨로지간 관계를 표현한 것이다.

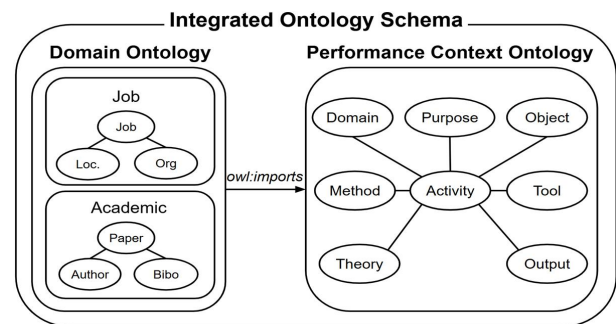


Fig. 5. PCO-Centered Integrated Ontology Schema

2. Performance Context Ontology Centered Integrated Ontology Construction

본 단계는 도메인별로 수집한 데이터를 대상으로 설계된 수행 맥락 온톨로지 스키마를 기반으로 수행 맥락 인스턴스를 생성하고 온톨로지를 구축하는 과정이다. 그리고, 이종 도메인 간의 의미적 용어 연결은 추출된 용어의 의미와 관계를 확장하여 이종 도메인의 수행 정보를 동일한 개념 맥락으로 연결하는 과정을 말한다.

2.1 LLM-based Performance Context Instance Construction

본 단계는 도메인별로 수집한 비정형 데이터로부터 온톨로지 스키마 기반의 수행 맥락 구조를 추출하여 인스턴스를 생성하는 과정이다. 수행 맥락 인스턴스는 직무와 연구 수행 정보를 수행 맥락 온톨로지 스키마에 유효한 의미적 관계로 구조화하여 표현한 것이다. 그림 6은 수행 맥락 온톨로지 스키마 정의에 따라 이종 도메인의 수행 맥락 인스턴스를 생성하는 과정이다.

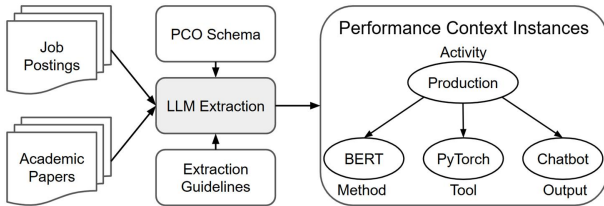


Fig. 6. LLM-based Generation of Performance Context Instances

먼저, LLM을 이용하여 채용공고와 학술논문 데이터를 분석하여 공통된 수행 맥락 온톨로지 인스턴스를 생성한다. 이 과정에서 LLM의 할루시네이션을 최소화하고 수행 맥락 추출 결과에 대한 유효성을 보장하기 위하여 수행 맥락 온톨로지 스키마와 인스턴스 생성 규칙을 제공한다. 예를 들어, ‘딥러닝 모델’ 용어가 수행 활동 유형에 따라 Object 또는 Method 개념에 매핑될 수 있도록 각 개념 및 개념간 생성 규칙을 명시한다. 추출된 수행 맥락 데이터는 RDF 트리플로 변환하여 온톨로지로 구축된다.

2.2 Cross-domain Semantic Linking

본 단계는 이종 도메인 간의 개념 매칭의 정확도를 개선하기 위하여 외부 지식베이스인 Wikidata를 활용하여 용어의 의미와 관계를 강화하는 과정이다. 온톨로지 구축 과정에서 추출된 용어에 대하여 Wikidata에서 항목을 검색하고, LLM을 이용하여 해당 용어의 적합성과 관계를 판정하여 온톨로지를 강화한다. 여기서 LLM은 검색된 용어와 설명이

실제로 동일 개념을 지칭하는지를 판별하고, 지식연결(Knowledge Linking)로 용어의 유사 개념, 상위 관계, 한국어 레이블, 약어 등과 같은 의미적인 관계를 생성한다.

예를 들어, ‘Deep Learning’ 용어가 Wikidata와 매핑 되면 한국어 라벨은 딥러닝, 약어는 DL, 상위 개념은 Machine Learning 등과 같은 용어 관계를 생성한다. LLM 판정 결과에 따라 용어간 관계는 SKOS(Simple Knowledge Organization System)와 OWL 트리플로 생성되어 수행 맥락 온톨로지에 자동으로 삽입된다. 그림 7은 수행 맥락 온톨로지 구축 과정에서 용어의 의미와 관계를 강화하는 과정을 보여준다.

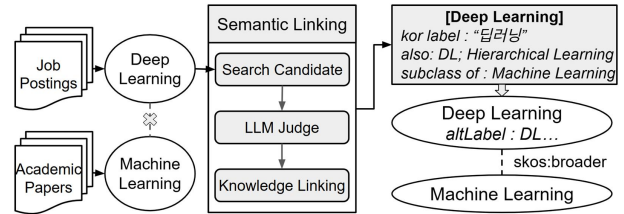


Fig. 7. Cross Domain Semantic Linking Process

IV. Experimental Results

본 연구에서 제안한 통합 온톨로지의 효과성을 검증하기 위하여 하이브레인넷 사이트에 공고된 AI 분야의 채용 공고 364건과 Scopus 논문 4,185편을 대상으로 실험을 진행하였다. 채용 및 학술 도메인에서 수집한 데이터를 기반으로 LLM을 이용하여 수행 맥락 온톨로지를 설계하였다. 그리고 채용과 학술 도메인을 연결한 통합 온톨로지를 구축하였다. 구축한 통합 온톨로지의 효과성을 검증하기 위하여 기존의 채용 및 학술 도메인 온톨로지와 제안 모델의 성능을 비교하였다.

1. Ontology Design and Construction Results

채용과 학술 도메인을 대상으로 LLM 기반의 온톨로지 학습을 통해 수행 맥락 온톨로지를 설계하고 스키마를 생성하였다. 그리고 수행 맥락 온톨로지 스키마를 기반으로 채용과 학술 데이터에서 수행 맥락 구조를 추출하여 RDF 인스턴스를 생성하고 통합 온톨로지를 구축하였다.

1.1 Derivation Results of Performance Context Concepts and Activity Types

체계적인 온톨로지 설계를 위하여 채용과 학술 도메인에서 개념 후보를 추출하여 공통 개념 체계를 도출하였다.

본 연구에서 개념 도출 단계별로 사용한 LLM은 Claude Sonnet 4.6 버전을 사용하였다. 개념 채택 단계는 도메인 데이터에서 수행 맥락에 해당하는 개념 후보 추출 단계, 추출된 후보가 개념 정의 기준에 부합하는지를 점검하는 적합성 검증 단계, 서로 다른 도메인에서 추출된 개념 후보들 사이의 의미적인 유사성 비교 단계로 구성된다.

LLM 프롬프트는 단계별 역할을 수행하기 위해 수행 정보 표현 여부, 개념 독립성, 의미 중복 여부, 관계 적절성 등의 개념 검증 기준과 수행 지침을 명시하여 구성하였다. 그리고 도출된 개념 체계의 검증과 평가를 위하여 온톨로지 설계 검증 도구인 역량 질문(Competency Question)을 이용하여 검증하였다. LLM을 이용한 온톨로지 학습에서는 8개의 수행 맥락 개념이 도출되었다. 표 1은 수행 맥락 개념에 대한 예이다.

Table 1. Example of Performance Context Concepts

Concept	Description	Example
Activity	What is performed	develop, evaluate
Domain	Field of performance	AI, NLP
Method	How it is performed	RAG, LLM
Tool	Tool used	python, sparql
Purpose	Why it is performed	accuracy gain
Object	Target of performance	image data
Output	Result of performance	classifier
Theory	Underlying theory	transformer

채용과 학술 도메인간의 수행 맥락 매칭의 정확도를 향상하기 위하여 LLM을 이용하여 수행 활동을 세분화하는 실험을 진행하였다. 여기서 수행 활동 세분화는 채용 직무와 연구 활동 정보에서 술어를 추출하여 클러스터링한 후, 다양한 활동 유형 체계로 세분화하는 과정이다. 표 2는 수행 활동에 대한 세분화된 유형 체계의 예이다.

Table 2. Example of Performance Activity Types

ActivityType	Activity Verb Pool
Production	build, create, develop...
Analysis	analyze, classify, compare...
Evaluation	evaluate, validate, assess...
Design	design, specify...
Application	apply, execute, operate...
Refinement	improve, optimize, fine-tune...

본 실험을 통해 도출된 8개의 수행 맥락 개념은 클래스, 19개의 관계는 객체 속성 그리고 11개의 수행 활동 유형은 서브 클래스로 각각 정의되어 형식화된다.

1.2 Integrated Ontology Schema Results

본 실험에서는 도출된 수행 맥락 개념과 수행 유형 체계

를 기반으로 통합 온톨로지 스키마를 생성하였다. 통합 온톨로지는 수행 맥락 온톨로지와 도메인 온톨로지 로 구성된다. 스키마는 OWL 2 DL(Description Logic)을 이용하여 생성하였다. 통합 온톨로지는 FOAF(Friend Of A Friend), BIBO(Bibliographic Ontology), SKOS와 같은 외부 표준 온톨로지를 사용하여 설계되어 시스템간에 호환이 가능하다. 그림 8은 통합 온톨로지 스키마의 일부 예이다. 통합 온톨로지 스키마는 수행 맥락 온톨로지 클래스, 채용과 학술 도메인 클래스, 그리고 외부 표준 클래스 간의 관계를 포함하여 OWL 2 DL로 정의한다.

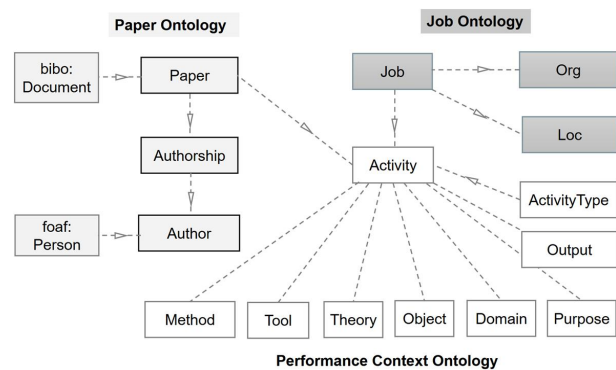


Fig. 8. Integrated Ontology Schema

1.3 LLM-based Instance Construction Results

본 실험에서는 LLM 모델 Claude Haiku 4.5를 이용하여 AI 분야의 채용공고와 학술논문 데이터에서 그림 3의 구조와 같은 형태의 수행 맥락 구조를 추출하였다. 그림 9는 논문에서 온톨로지 스키마를 기반으로 수행 맥락을 추출하여 인스턴스를 생성한 결과이다. 논문에서 추출된 수행 동사 'developed', 'evaluated'는 수행 활동 유형에 의해 Production과 Evaluation 수행 활동 유형으로 매핑된다. 문맥상 'linear regression'과 'Google Colab'은 방법(Method)과 도구(Tool)로 매핑된다.

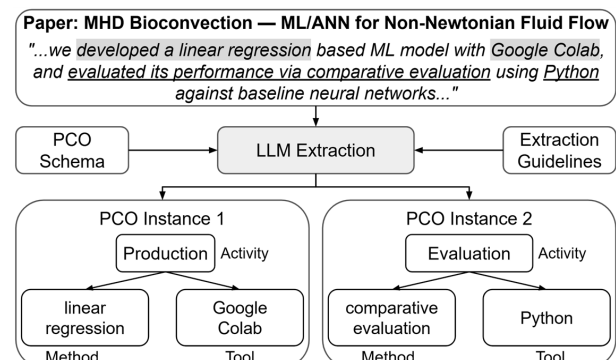


Fig. 9. Example of Performance Context Instance Extraction from a Paper

1.4 Cross Domain Semantic Linking Results

본 실험에서는 Wikidata를 이용하여 이종 도메인 간의 용어 관계를 강화하는 실험을 진행하였다. 먼저, 수행 맥락 온톨로지 구축 과정에서 추출된 용어 중에서 4,250개가 Wikidata와 매핑되었다. 그리고, LLM 모델 Claude Haiku 4.5를 이용하여 매핑된 정보를 분석하고 용어간의 관계 유형을 추가하였다. 그림 10은 용어의 의미와 관계 정보가 수행 맥락 온톨로지에 추가된 결과의 예이다.

```
Keyword:Large Language Model
a skos:Concept, owl:class ;
owl:sameAs wd:Q115305900
skos:prefLabel "대규모 언어 모델"@ko
skos:altLabel "LLM"@en,
skos:definition "language model built with very..
skos:border "language model"
skos:narrow "generative AI"
skos:related "transformer"
```

Fig. 10. Example of Semantic Linking Results

그림 10에서처럼 Wikidata를 이용하여 ‘Large Language Model’ 용어에 대한 동의어(altLabel), 정의(definition) 정보와 용어 관계 정보를 효과적으로 분석할 수 있다. 의미 연결 강화 실험을 통하여 7,102개의 Wikidata 지식정보가 추가로 연결되었다. 그리고 수행 맥락 온톨로지의 트리플은 58% 증가하였다. 이러한 트리플 증가는 용어간의 의미 관계가 효과적으로 확장됨을 의미한다. 또한, 이러한 확장을 통하여 제안 모델이 이종 도메인 간의 용어 이질성을 완화하고 매칭의 정확도를 향상할 수 있음을 보여준다.

2. Comparative Experimental Results between Domain-Specific Ontology Models and PCO

본 연구에서 제안한 모델은 향후 채용공고에서 요구하는 전문가의 역량과 논문에서 도출한 전문가의 연구 수행 정보를 비교하여, 해당 채용공고에 적합한 전문가 후보를 효과적으로 탐색하기 위한 목적으로 활용하기 위함이다. 따라서, 제안한 온톨로지 모델에 대한 성능은 클래스, 관계, 인스턴스 개수와 같은 단순한 기술 지표에 의해 평가하기 어려운 측면이 있다.

2.1 Retrieval Performance Evaluation Method for the Proposed Model

제안 모델에 대한 성능 평가는 제안 모델이 채용공고에서 요구하는 직무와 의미적으로 일치하는 학술논문을 정확하게 검색하여 전문가 후보를 효과적으로 탐색할 수 있

는지를 검증하였다. 이를 위해, 본 실험에서는 채용공고로 논문을 검색하는 태스크 기반(Task-based) 평가 방식을 사용하였다. 예를 들어, ‘생성형 AI 시스템 구축 및 개발 경험이 있는 연구원 모집’이라는 채용공고에 적합한 학술 논문을 온톨로지 추론을 통해 검색한 결과를 평가하였다.

제안 모델의 효율성을 검증하기 위하여 기존 도메인 온톨로지 모델과 검색 성능을 비교하는 실험을 수행하였다. 실험에서는 채용 도메인 온톨로지에 포함된 30건의 채용 직무 정보를 질의 형태로 변환하고, 이를 기반으로 학술 도메인 온톨로지서 1,000편의 학술논문을 대상으로 검색을 수행하였다. 실험 평가를 위한 정답지는 LLM 모델 Claude Sonnet 4.6을 평가자(LLM-as-a-Judge)로 이용하였다. 채용공고별로 의미적 유사도가 높은 논문 후보 집합을 구축한 후, IT 도메인 전문가들이 후보 집합을 검증하여 최종적인 정답지를 구축하였다. 그리고, 제안 모델과 기존 모델이 채용공고와 의미적으로 일치하는 논문을 검색한 결과를 정답지와 비교 분석하였다.

2.2 Comparative Retrieval Results between Existing and Proposed Models

본 실험에서 성능 비교를 위해 사용한 평가 지표는 전문가 검색 분야에서 일반적으로 사용되는 MRR(Mean Reciprocal Rank), Precision, Recall, F1을 사용하였다. MRR은 질의에 대한 첫 번째 정답 문서가 검색 결과 목록에서 얼마나 상위에 위치하는지를 평가하는 지표이다. Precision은 상위 N개의 검색 결과 중에서 정답이 차지하는 비율, Recall은 전체 정답 중에서 상위 N개가 검색 결과에 포함된 비율을 평가하는 지표이다. F1은 Precision과 Recall의 조화 평균으로 정밀도와 재현율의 균형을 평가한다.

그림 11은 제안 모델의 검색 성능과 기존의 채용 및 학술 도메인 온톨로지 모델 간의 검색 성능을 여러 평가 지표로 비교 실험한 결과이다. 제안 방법의 검색 결과가 채용공고와 학술논문 도메인에서 기존의 도메인 온톨로지 간의 검색 방식보다 우수한 성능을 보였다. 실험 결과에서 MRR 지표의 결과는 제안 모델이 도메인 매핑 기반의 기존 방식에 비해 의미적으로 일치하는 첫 번째 결과를 더 높은 순위에 노출할 수 있음을 보여준다. 특히 Precision@10 지표의 결과는 상위 10개의 검색 결과에서 관련 논문의 비율이 높아, 제안 모델이 검색 정확도를 향상시킬 수 있음을 의미한다. 이는 본 연구에서 제안한 검색 방식이 도메인 온톨로지 간의 매핑을 이용한 기존 검색 방식보다 더 정확한 결과를 제공할 수 있음을 의미한다.

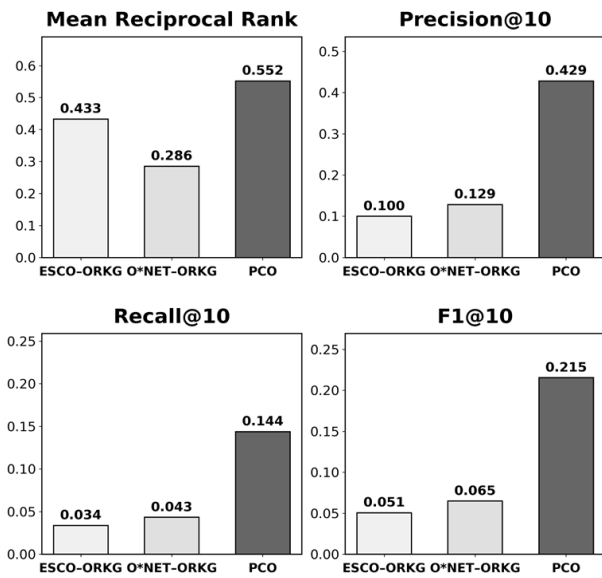


Fig. 11. Comparative Experimental Results between Existing Domain Ontology and PCO

2.3 Discussion on Retrieval Performance of the Proposed Model

제안 모델과 기존 모델간의 성능을 비교한 결과, 기존 독립적인 모델 간의 검색은 채용 및 학술 도메인 온톨로지의 구조와 개념 체계가 상이하고, 어휘의 표현 차이로 인해 의미 기반 검색이 어려웠다. 이로 인해, 검색 과정에서 여러 단계의 복잡한 추론 경로가 요구되어 연산 시간이 증가하고, 단순 키워드 기반 검색에 의존하게 되어 정확도와 신뢰도가 저하되었다. 하지만, 제안 모델은 채용공고와 학술논문이 동일한 수행 맥락 구조로 인스턴스화되고, 용어 간의 이질성도 수행 맥락 온톨로지에서도 완화되어 수행 맥락 기반의 의미적 추론이 가능하였다. 결과적으로 제안 모델은 기존의 개별 도메인 온톨로지 간의 검색보다 우수한 성능을 보였다.

V. Conclusions

본 연구에서는 채용과 학술 도메인에서 추출한 수행 정보를 동일한 맥락 구조로 표현한 수행 맥락 온톨로지 모델을 설계하였다. 그리고 온톨로지 스키마를 구성하고 채용 공고와 학술논문 데이터를 인스턴스로 생성하여 이종 도메인 간에 통합 온톨로지를 구축하였다. 먼저, 온톨로지 설계 단계에서는 LLM을 이용하여 이종 도메인에서 추출한 수행 정보를 동일한 구조로 표현하는 공통 수행 맥락 개념 체계를 도출하였다. 구축 단계에서는 LLM을 이용하여 채용공고와 학술논문 데이터에서 수행 맥락 인스턴스

를 생성하였다. 그리고 이종 도메인 간의 수행 정보를 동일한 의미 구조로 연결한 통합 온톨로지를 구축하였다.

본 연구에서 제안한 수행 맥락 온톨로지의 효과성을 검증하기 위해 채용공고와 학술논문 데이터를 수집하여 온톨로지를 구축하였다. 구축한 온톨로지는 기존 도메인 온톨로지와 비교 실험을 통해 이종 도메인 간의 검색 성능을 평가하였다. 실험 결과, 제안한 수행 맥락 온톨로지 기반의 검색 방식이 도메인 온톨로지 간의 매핑을 이용한 기존 검색 방식보다 더 정확한 검색 결과를 제공하였다. 이는 제안 모델이 기존의 도메인 온톨로지간 검색에서 탐색하기 어려운 수행 역량 기반으로 검색이 가능함을 의미한다.

결과적으로, 본 연구에서 제안한 수행 맥락 온톨로지는 채용공고의 직무 정보와 학술논문의 연구 활동 정보를 동일한 맥락 구조로 표현하여 이종 도메인 간의 매칭 효과성을 향상시켰다. 또한 제안한 수행 맥락 온톨로지는 기관에서 요구하는 직무 역량과 부합하는 전문가를 효과적으로 탐색할 수 있는 전문가 검색 시스템 개발에서 기반 기술로 사용할 수 있으리라 기대한다.

REFERENCES

- [1] K. Balog, Y. Fang, M. de Rijke, P. Serdyukov, and L. Si, "Expertise Retrieval", *Foundations and Trends in Information Retrieval*, Vol. 6, No. 2-3, pp. 127-256, July 2012. DOI: 10.1561/1500000024
- [2] O. Husain, N. Salim, R. A. Alias, S. Abdelsalam, and A. Hassan, "Expert Finding Systems: A Systematic Review", *Applied Sciences*, Vol. 9, No. 20, Article No. 4250, Oct. 2019. DOI: 10.3390/app9204250
- [3] K. Balog, L. Azzopardi, and M. de Rijke, "Formal Models for Expert Finding in Enterprise Corpora", *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2006)*, pp. 43-50, Aug. 2006. DOI: 10.1145/1148170.1148181
- [4] S. K. Song, H. J. Mun, Y. J. Kim, and Y. T. Woo, "Hi-BERT: A Semantic Relation Learning Model for Recruitment Ontology and Academic Papers", *Journal of The Korea Society of Computer and Information*, Vol. 30, No. 6, pp. 21-29, June 2025. DOI: 10.9708/jksoci.2025.30.06.021
- [5] M. N. Uddin, T. H. Duong, K. J. Oh, J. G. Jung, and G. S. Jo, "Experts Search and Rank with Social Network: An Ontology-Based Approach", *International Journal of Software Engineering and Knowledge Engineering*, Vol. 23, No. 1, pp. 31-50, Jan. 2013. DOI: 10.1142/S0218194013400032
- [6] M. H. T. de Boer, R. M. Bakker, and M. Burghoorn, "Creating Dynamically Evolving Ontologies: A Use Case from the Labour

- Market Domain," Proceedings of the AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), CEUR Workshop Proceedings, Vol. 3433, pp. 1-9, Mar. 2023.
- [7] J. Beverley, S. Smith, M. A. Diller, W. D. Duncan, J. Zheng, J. W. Judkins, W. R. Hogan, R. McGill, D. Dooley, and Y. He, "The Occupation Ontology (OccO): Building a Bridge between Global Occupational Standards," Proceedings of the 9th Joint Ontology Workshops (JOWO 2023): Ontologies for Services and Society (OSS 2023), CEUR Workshop Proceedings, Vol. 3637, pp. 1-12, July 2023.
- [8] Y. G. Jin and W. G. Lee, "A Study on Ontology Design for NCS 'Application SW Engineering' Supporting Intelligent Knowledge Management and Search Reasoning", Journal of the Korea Convergence Society, Vol. 8, No. 9, pp. 17-23, Sep. 2017. DOI: 10.15207/JKCS.2017.8.9.017
- [9] M. Y. Jaradeh, A. Oelen, K. E. Farfar, M. Prinz, J. D'Souza, G. Kismihók, M. Stocker, and S. Auer, "Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge", Proceedings of the 10th International Conference on Knowledge Capture (K-CAP 2019), pp. 243-246, Nov. 2019. DOI: 10.1145/3360901.3364435
- [10] National Information Standards Organization, "ANSI/NISO Z39.104-2022, CRediT, Contributor Roles Taxonomy", NISO, <https://www.niso.org/publications/z39104-2022-credit>
- [11] L. N. Soldatova and R. D. King, "An Ontology of Scientific Experiments", Journal of the Royal Society Interface, Vol. 3, No. 11, pp. 795-803, Dec. 2006. DOI: 10.1098/rsif.2006.0134
- [12] T. Lebo, S. Sahoo, and D. McGuinness, "PROV-O: The PROV Ontology", W3C Recommendation, <https://www.w3.org/TR/prov-o/>, Apr. 2013
- [13] V. Daponte and G. Falquet, "An Ontology for the Formalization and Visualization of Scientific Knowledge", arXiv, <https://arxiv.org/abs/2107.04347>, July 2021
- [14] L. Borek, Q. Dombrowski, J. Perkins, and C. Schöch, "TaDiRAH: A Case Study in Pragmatic Classification", Digital Humanities Quarterly, Vol. 10, No. 1, 2016.
- [15] T. R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing", International Journal of Human-Computer Studies, Vol. 43, No. 5-6, pp. 907-928, Nov. 1995. DOI: 10.1006/ijhc.1995.1081
- [16] M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "METHONTOLOGY: From Ontological Art Towards Ontological Engineering", Proceedings of the AAAI Spring Symposium on Ontological Engineering, pp. 33-40, Mar. 1997.
- [17] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The NeOn Methodology for Ontology Engineering", Ontology Engineering in a Networked World, Springer, pp. 9-34, 2012. DOI: 10.1007/978-3-642-24794-1_2
- [18] N. Guarino and C. Welty, "Evaluating Ontological Decisions with OntoClean", Communications of the ACM, Vol. 45, No. 2, pp. 61-65, Feb. 2002. DOI: 10.1145/503124.503150
- [19] A. Maedche and S. Staab, "Ontology Learning for the Semantic Web", IEEE Intelligent Systems, Vol. 16, No. 2, pp. 72-79, Mar.-Apr. 2001. DOI: 10.1109/5254.920602
- [20] P. Cimiano, "Ontology Learning and Population from Text: Algorithms, Evaluation and Applications", Springer, 347 pp., 2006. DOI: 10.1007/978-0-387-39252-3
- [21] H. B. Giglou, J. D'Souza, and S. Auer, "LLMs4OL: Large Language Models for Ontology Learning", The Semantic Web - ISWC 2023, pp. 408-427, Nov. 2023. DOI: 10.1007/978-3-031-47240-4_22
- [22] T. Aggarwal, A. Salatino, F. Osborne, and E. Motta, "Large Language Models for Scholarly Ontology Generation: An Extensive Analysis in the Engineering Field", Information Processing & Management, Vol. 63, No. 1, Article No. 104262, Jan. 2026. DOI: 10.1016/j.ipm.2025.104262
- [23] H. B. Giglou, J. D'Souza, F. Engel, and S. Auer, "LLMs4OM: Matching Ontologies with Large Language Models", The Semantic Web: ESWC 2024 Satellite Events, pp. 25-35, May 2024. DOI: 10.1007/978-3-031-78952-6_3

Authors



In-Ha Jun received the M.S. degree in Computer Engineering from Changwon National University, Korea, in 2007. Mr. Jun is a Senior Research Engineer in Hibrainnet Co..

He is interested in Expert Search System, Ontology, and Generative AI application services.



Hyeon-Jeong Mun received the M.S. and Ph.D. degrees in Computer Engineering from Changwon National University, Korea, in 1996 and 2003, respectively. Dr. Mun is a Principal Research Engineer at Hibrainnet Co..

She is interested in Ontology, Data Modeling, and Data Analysis.



Young-Ji Kim received the M.S. and Ph.D. degrees in Computer Engineering from Changwon National University, Korea, in 1997 and 2004, respectively. Dr. Kim is a Principal Research Engineer at Hibrainnet Co..

She is interested in Generative AI application services, Big data analysis, and hyper-personalization.



Su-Hyun Lee received the B.S. in Computer Science from Kwangwoon University, Korea in 1987. He received the M.S. and Ph.D. degrees in Computer Science from Korea Advanced Institute of Science and

Technology(KAIST), Korea, in 1989, 1994, respectively. Dr. Lee is a Professor in the Department of Computer Engineering, Changwon National University since 1996. He is interested in computer algorithm, programming languages, compiler, and blockchain.



Yong-Tae Woo received the B.S., M.S. and Ph.D. degrees in Computer Science and Engineering from Kyungpook National University, Korea, in 1982, 1984 and 1995, respectively.

Dr. Woo is a Professor Emeritus in the Department of Computer Engineering, Changwon National University since 1987. He is also CEO of Hibrain.net Co. He is interested in Data Modeling, Internet Business, and Big Data Analysis areas.