

## 인공지능 노출 정도에 따른 고용 추세 분석: K자형 고용 양극화\*

이예슬\*\* · 황현준\*\*\*

### 요약

기술 발전이 고용에 미치는 영향은 자동화에 의한 대체 또는 새로운 업무 도입에 따른 고용 증가 등 여전한 논쟁의 대상이다. 특히 인공지능 기술 발전과 고용에 대한 실증 논의는 더욱 부족한 실정이다. 이에 본 연구는 자연어처리 기법(SBERT)과 특허를 이용하여 직업별 인공지능 노출 점수를 계산하고 평균 점수를 기준으로 상위 집단과 하위 집단으로 구분하여 집단별 고용 추세를 분석한다. 자연어처리 기법을 통해 한국 특허와 미국 직업의 업무 설명을 연계하는 인공지능 노출 점수 계산 방식과 한미 표준직업분류 연계 방식을 제시하고 이를 국내 고용 통계에 적용하여 추세를 분석한다. 2013년 이후 국내 인공지능 출원 특허와 통계청 지역별고용조사를 분석한 결과 한국의 고용은 시간이 지남에 따라 평균 이상의 인공지능 노출 집단에서 우상향하고, 평균 이하 집단에서는 우하향하는 K자형 양극화 양상을 보인다.

주제어 : 인공지능, 고용, 양극화, 특허, 자연어처리

## Analyzing employment trends in response to AI exposure: K-shaped labor polarization in Korea\*

Lee, Yeseul\*\* · Hwang, Hyeonjun\*\*\*

### Abstract

The impact of technological advancements on employment is a matter of ongoing debate, with discussions on the effects of AI technology development on employment being particularly scarce. This study employs the natural language processing technique (SBERT) and patents to calculate an occupation-based AI exposure score and to analyze employment trends by group. It proposes a method for calculating the AI exposure score based on the similarity between Korean patent information and US job descriptions and linking SOC(U.S.) and KSCO(Korea). The analysis of domestic AI patent applications and regional employment data in the KOSIS Database since 2013 reveals a K-shaped polarization pattern in Korean employment trends among groups with above and below average levels of AI exposure.

Keywords : AI, employment, polarization, patent, NLP

Received Mar 16, 2023; Revised Apr 21, 2023; Accepted May 30, 2023

\* This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(RS-2023-00242528).

\*\* Master Student, Graduate School of Data Science, Kyungpook National University (yeseull@knu.ac.kr, <https://orcid.org/0009-0000-7241-1560>)

\*\*\* Corresponding Author, Assistant Professor, Graduate School of Data Science, Kyungpook National University (hhwang@knu.ac.kr, <https://orcid.org/0000-0002-2192-0290>)

## I. 서론

한국지능정보사회진흥원(National Information Society Agency, NIA)의 <2022 국가지능정보화백서>에 따르면 디지털 전환 가속화에 따라 국내 인공지능 기술 시스템 도입이 활발하게 이루어지고 있고, 관련 시장 규모는 2025년 1조 9700억 원 규모로 성장할 것으로 예상된다. Lee and Nam(2022)은 문헌 연구를 통해, 인공지능 활용과 관련한 논문 수는 2015년부터 2021년까지 꾸준히 증가하고 있으며, 활용 분야 또한 제조, 의료, 건설, 교육 등으로 다양하다고 밝혔다. 이러한 산업 현장에서의 인공지능 기술 도입 그리고 그에 대한 산업 규모에 대한 논의와 달리 인공지능과 같은 기술의 발전이 고용과 어떤 관계에 있는지에 대한 실증적 연구는 충분하게 이루어지지 않고 있다(Hur, 2019).

어떤 산업에서 급속한 기술 발전은 노동시장에서 인간의 노동력을 대체하고 새로운 일자리를 창출하기도 하는 복합적인 형태를 보인다(WEF, 2019). 구체적으로 노동시장에서 인간의 노동력을 요구하는 일자리는 기술 발전에 의한 자동화(Automation)로 대체되거나 새로운 업무의 도입(Creation of New Tasks)에 따라 창출되는 두 가지 양상을 보인다(Acemoglu & Restrepo, 2018). 특히 4차 산업혁명의 핵심이라 할 수 있는 인공지능 기술은 로봇, 소프트웨어에 이어 사회 구조적 변화를 가져올 것으로 주목받는 기술 분야로, 인공지능이 노동을 대체한다는 비판론과 노동을 보완한다는 낙관론이 대립하고 있다(Seo, 2019).

좀 더 자세하게, Acemoglu and Restrepo(2018)은 기술 발전에 의한 노동시장의 변화는 업무 단위에서 일어난다는 '업무 기반 모형(Task-based)'을 제시하였다. Acemoglu and Restrepo(2019)는 인간의 노동력을 대체하는 자동화와 달리 새로운 업무의 도입은 고용에 긍정적인 영향을 미칠 수 있지만, 자동화가 가속화되고 새로운 업무의 도입은 둔화하면서 노동 수요가 침체된다고 주장하였다. Huang and Rust(2018)은 서비스 부문에서 인공지능이 수행할 수 있는 업무가 늘어남

에 따라 기술에 의한 인간 노동력의 완전 대체 가능성을 강조하였다. Kim(2019)는 문헌 연구를 통해, 인공지능 기술혁신은 노동자에게 불리한 시장 배분을 가져오기 때문에 불평등 문제를 완화할 적극적인 재분배 정책이 필요함을 밝혔다.

반면, Hur(2019)는 기술 진보로 인한 업무 대체가 사회 전반의 일자리를 감소시킨다는 증거는 찾을 수 없으며, 인공지능이 노동시장에 미치는 영향 분석은 장기적 관점에서 볼 필요가 있다고 주장하였다. Autor and Salomons(2018)은 기술혁신이 노동자를 대체한다고 하지만 전체 노동 수요를 줄이지는 않았으며, 생산성 증가로 시간당 임금이 증가했음을 확인하였다. Eom and Lee(2020)은 인공지능 기술의 업무 대체는 주로 비숙련, 반숙련 직무에서 진행되며, 인간과 인공지능이 협력함으로써 생산성이 향상될 수 있다고 보았다.

기술 발전에 따른 노동시장의 변화를 실증적으로 분석하기 위해서 기술적인 측면에서 선행 연구들은 기술의 특징을 구체화하고 이를 업무(Task)나 요구되는 능력(Ability)에 연계하였다. 대표적으로 기술의 특징을 파악하고 이를 직업과 연계하는 과정에는 주로 전문가 인터뷰와 설문조사를 활용한 Fery and Osborne(2017)의 연구가 있다. Fery and Osborne(2017)은 전문가 인터뷰로 미국 직업 사전인 O\*NET의 702개 직업 중 자동화 가능성이 있는 70개의 직업을 선정하고, 컴퓨터화에 의한 대체 확률을 직업별로 계산하는 직업 기반(Occupation-based) 접근법을 시도하였다.

유사하게 Arntz, et al.(2016)은 Fery and Osborne(2017)의 직업별 자동화 가능성 지표를 OECD PIACC(The Programme for the International Assessment of Adult Competencies) 자료와 연계하여 업무 기반 접근에 기초한 직업의 자동화 정도를 추정하였다. Brynjolfsson, et al.(2018)은 전문가 평가를 통해 머신러닝 적합성 지시문(Suitability for Machine Learning Rubric)을 만들고, O\*NET의 18,156개 업무와 연계하는 작업을 통해 인공지능 기술의 영향을 받는 업무에 대해 분석하였다. Webb(2019)는 특히 정보

와 O\*NET의 업무 텍스트 간의 일치 정도로 각각의 업무가 인공지능에 노출된 정도를 파악하고, 이를 집계하여 직업의 인공지능 노출 점수를 계산하는 업무 기반 인공지능 노출 정도 계산법을 제시하였다.

또한 Felten, et al.(2019, 2021)은 O\*NET의 52개 능력(Abilities) 지표를 활용하는 능력 기반(Ability-based) 접근법을 고안하였다. 아마존 mTurk(Amazon's Mechanical Turk) 웹 서비스의 깃워커(Gig Workers) 설문조사를 통해 EFF(Electronic Frontier Foundation)가 정의한 10개 인공지능 기술을 O\*NET의 능력 지표와 연계하여 직업별 인공지능 노출 점수를 계산하였다.

국내 연구는 크게 Fery and Osborne(2017)의 자동화 가능성 지표를 사용한 연구, Felten, et al.(2019, 2021)의 인공지능-능력 연계표를 사용한 연구, 한국의 직업별 업무수행능력 조사에 기반한 연구로 나누어진다.

먼저 Kim(2015)는 미국 직업을 대상으로 한 Fery and Osborne(2017)의 지표를 한국고용정보원의 2012 한국직업사전, 통계청의 2014년 하반기 지역별 고용조사와 연계하여 국내 직업의 대체 가능성을 분석하였다. 다음으로 Oh, et al.(2016)은 통계청의 2008년, 2015년 상반기 지역별고용조사와 한국고용정보원의 2007년, 2013년 대졸자직업이동경로조사 원자료에 Fery and Osborne(2017)의 지표를 적용하였다. Lee, et al.(2017)은 한국표준직업분류 세분류에 Fery and Osborne(2017)의 직업별 자동화율을 연계하여 소분류별 대체율을 계산하였다. Park, et al.(2016)은 전문가 의견 조사로 2016년, 2020년, 2025년의 기술 발전에 따른 일자리 대체 가능성 자료를 구축하여 세분류 단위 업무능력 대체 비율을 산출하였다. Lee, et al.(2021)은 Park, et al.(2016)의 지표를 한국노동패널조사 원자료의 소득정보에 적용하여 2018년 기준 직업별 직무대체율을 계산하였다. Cheon, et al.(2022)는 Felten, et al.(2019, 2021)의 결과를 한국고용정보원의 2020년 재직자 조사의 능력 지표와 연계하여 직업별 인공지능 노출 점수를 구하였다.

한편 기술 발전의 정도를 분석에 반영하기 위해 특허를 이용하는 연구들이 존재한다. Kim and Bae(2017)은 유망 기술을 예측하는 모형을 구축하면서 특허 분석을 활용하였고, 세계지식재산기구(World Intellectual Property Organization, WIPO)는 2019년 보고서에서 특허 분석을 통해 인공지능 기술 동향을 파악하였다. Ha, et al.(2022)는 연도별 특허 수 추세를 분석하여 기술 발전에 따른 유망직종을 예측하였다. Lee, et al.(2022)는 특허 출원 수와 토픽모델링을 통한 특허 트렌드 분석으로 한국의 인공지능 경쟁력을 평가하였다.

기술 발전에 따른 노동시장의 변화를 실증적으로 분석한 연구들은 인공지능이라는 기술적 변화보다는 노동력을 대체할 수 있는 자동화에 초점이 맞춰진 경향이 크다. 또한 미국에서 진행된 선행연구에서 도출된 지표 등을 국내에 그대로 적용한 경우가 대부분이다. 이러한 한계로 인해 인공지능이라는 기술변화가 국내 노동시장에 어떤 영향을 미쳤는지 실증적으로 분석한 연구들은 부족하다. 이러한 한계를 극복하기 위해 본 연구는 국내 특허 정보를 이용해 기술(인공지능) 발전 정도(인공지능 노출 점수)를 계산하는 방식과 미국의 노동시장 분석 연구를 국내 노동시장에 연계할 수 있도록 한미 직업분류 연계 방식을 제시하여 인공지능이라는 기술적 변화가 국내 고용시장에 어떻게 반영되었는지 실증적으로 분석하고자 한다.

특히 본 연구에서 제안하는 인공지능 노출 점수 계산, 한미 직업 연계를 이용해 실제 통계자료를 통해 직업별 인공지능 노출 정도에 따라 고용 양상이 어떻게 나타나는지를 확인해보고자 한다. 선행연구에 따르면, 기술 발전에 의한 노동시장의 변화는 자동화로 인한 인간 노동력 대체와 새로운 업무 도입에 따른 일자리 창출로 나타난다. 이러한 경향에 있어서 국내의 고용시장이 노동력 대체 또는 일자리 창출 혹은 두 양상이 혼합된 상태로 나타나는지 등을 확인함으로써 실증 자료에 기반한 고용시장 경향을 파악한다.

이를 위해 본 연구에서는 먼저 자연어처리 기법과 특허를 이용하여 직업별 인공지능 노출 점수를 계산한

다. 본 연구가 제시하는 인공지능 노출 점수는 국내 특허 정보와 미국 O\*NET의 업무 설명 텍스트 간의 유사도 측정 결과로 산출된다. 다음으로 한미 양국 간 표준 직업분류 연계 또한 자연어처리 방법을 이용해 연계 방식을 제안하고 이를 통해 기술 노출 측정법을 국내 노동시장에 적용하고, 인공지능 노출 점수가 높은 집단과 낮은 집단을 분류한다. 마지막으로, 인공지능 노출 정도에 따라 다르게 나타나는 고용 추세를 확인하기 위해 국내 직업별 취업자 수 자료와 결합한다.

본 연구의 연구목적은 크게 세 가지다. 첫째, 기술 발전의 정도를 반영하는 직업의 인공지능 노출 정도 계산에 있어서 상대적으로 높은 비용이 요구되는 전문가 평가나 설문조사 대신 공공데이터 기반의 특허 정보를 이용하는 자연어처리 기반 접근법을 제시한다. 둘째, 해외 지표를 그대로 국내 자료에 사용한 국내 선행연구와 달리 국내 특허 정보를 이용하여 한국의 인공지능 기술 수준을 반영하고, 한국과 미국의 표준직업분류를 연계하여 미국 직무 기반으로 계산된 결과를 국내 연구와 연계할 수 있는 방안을 제시하고 이를 통계 자료에 연계하여 국내 노동시장을 분석할 수 있도록 한다. 셋째, 인공지능 노출 점수로 집단을 나누고 집단별 고용 추세를 분석하여 인공지능 기술의 발전과 노동시장의 관계에 대한 실증 분석을 수행한다.

## II. 분석 자료

본 연구는 KIPRIS Plus 특허 데이터, 미국 O\*NET의 업무 설명(텍스트)로 구성된 직업 관련 데이터, 통계청의 지역별고용조사 등 3가지의 자료를 이용한다.

### 1. KIPRIS Plus 특허 데이터

KIPRIS Plus(Korea Intellectual Property Rights Information Service Plus)는 특허청이 개방 중인 모든 특허 정보를 실시간으로 활용할 수 있는 정보 제공 서비스이다. KIPRIS Plus의 공보 데이터는 크게 특허/

실용, 디자인, 상표, 정정공보로 나뉘는데, 본 연구에서는 한국의 특허와 실용신안과 관련된 '특허·실용 공개·등록공보'의 Open API 크롤링으로 데이터를 수집하였다. 특허법상 출원일에서부터 1년 6개월이 지난 시점에 기술 내용을 대중에게 공개하기 때문에 2020년까지의 자료를 수집할 수 있었다. 따라서 분석은 2000년부터 2020년까지 한국에서 출원된 특허 2,630,628건을 사용한다.

Open API 크롤링으로 구축한 특허 데이터셋은 출원일자, 출원번호, 국문제목, 영문제목, IPC 코드, 초록으로 구성된다. 그중 국제특허분류인 IPC(International Patent Classification)는 발명의 기술분야를 나타내는 국제적으로 통일된 특허분류체계다. 특허청은 특허 문헌의 분류와 검색의 편의를 위해 모든 국내 출원 특허에 대해 IPC 코드를 부여하고 있다. IPC 코드는 특허의 성격을 분류하기 때문에 우리는 인공지능 기술 관련 IPC 분류를 통해서 어떤 특허의 인공지능 기술 관련 여부를 판단한다.

## 2. 직업 관련 데이터

### 1) 2018 미국표준직업분류

2018년 작성된 미국표준직업분류(Standard Occupational Classification, SOC-2018)는 여섯 자리 숫자 분류체계를 가지고 있으며, 두 번째 자리까지는 대분류(Major Group), 세 번째 자리는 중분류(Minor Group), 네 번째와 다섯 번째 자리는 소분류(Broad Occupation), 여섯 번째 자리는 세분류(Detailed Occupation)로 구분된다. 대분류 23개, 중분류 98개, 소분류 459개, 세분류 867개로 구성되어있다. 예를 들어, '대분류 51-0000 Production Occupations'는 '중분류 51-8000 Plant and System Operators'를 포함하며, '소분류 51-8010 Power Plant Operators, Distributors, and Dispatchers'와 '세분류 51-8013 Power Plant Operators'가 하위 항목으로 존재한다. 세분류 코드 끝자리 숫자가 9인 경우는 기타 직종(All

Other)으로 분류된다.

### 2) O\*NET의 업무 텍스트

미국 노동부(United States Department of Labor) 산하 고용훈련실(Employment Training Administration)에서 제공하는 O\*NET 데이터베이스는 미국 직업 사전(Dictionary of Occupational Titles)의 후속으로, 2019년 SOC-2018을 기반으로 하는 'O\*NET-SOC-2019'라는 분류체계를 구축하여 1,016개 직업에 대한 설명을 제공하고 있다. O\*NET 데이터베이스의 직업 코드는 '51-8013.03 Biomass Plant Technicians'와 같이 SOC-2018의 여섯 자리 분류체계에 소수점 아래 두 자리가 추가된 형태다. O\*NET은 직업 코드, 직업 제목과 함께 업무(Tasks), 직업 내 업무 중요도, 직능(Skills), 사용 도구 등 직업 특수적 정보를 제공한다. '대분류 55-0000 Military Specific Occupations'와 기타 직종에 대해서는 직업 제목만 서술하기 때문에 이를 제외한 923개 직업의 업무 텍스트와 직업 내 업무 중요도 수치를 사용한다.

### 3) 7차 한국표준직업분류

통계청이 2017년 7월 개정·고시한 제7차 한국표준직업분류(Korean Standard Classification of

Occupations, KSCO-7)는 대분류 10개, 중분류 52개, 소분류 156개, 세분류 450개, 세세분류 1,231개로 구성되어있다. 분류번호는 숫자와 알파벳 A로 표시하며 대분류 1자리, 중분류 2자리, 소분류 3자리, 세분류 4자리, 세세분류 5자리로 표시된다(통계청, 2017). 통계청이 제공하는 영문 분류 항목표에서 '대분류 A 군인'을 제외한 직종의 직업 코드와 영문 제목을 사용한다.

### 3. 통계청 지역별고용조사

통계청의 지역별고용조사는 KSCO 소분류 단위 취업자 수를 제공하는 통계로 2008년부터 조사가 시작되었고, 국가통계포털 KOSIS(Korean Statistical Information Service)에서 확인할 수 있다. 2013년부터 반기 조사로 실시되었으며, 상반기 조사의 경우 4월, 하반기 조사의 경우 10월의 15일이 포함된 일주일을 기준으로 표본가구 내에 상주하는 만 15세 이상 가구를 조사 대상으로 한다. KSCO-7이 적용된 2013년 이후의 하반기 자료를 활용하며, '전국 직업/성별 취업자(소분류)'의 소분류 단위 취업자 수와 ' 시도/성별 경제활동인구 총괄'의 15세 이상 인구수를 사용한다. 연도별 15세 이상 인구수와 전체 취업자 수는 <표 1>에서 확인할 수 있다.

〈표 1〉 연도별 15세 이상 인구수와 취업자 수  
 (Table 1) 15 years old and over population, and employed persons by year

Year	2008	2009	2010	2011	2012	2013
Population, 15 years old and over (Thousand Persons)	39,933	40,441	40,933	41,487	41,937	42,444
Employed persons (Thousand Persons)	24,045	24,038	24,270	24,604	25,257	25,798
KSCO	6th	6th	6th	6th	6th	7th
Year	2014	2015	2016	2017	2018	2019
Population, 15 years old and over (Thousand Persons)	42,929	43,357	43,704	44,015	44,262	44,601
Employed persons (Thousand Persons)	26,247	26,519	26,746	27,026	27,090	27,509
KSCO	7th	7th	7th	7th	7th	7th

source: KOSIS database



### Ⅲ. 분석 방법

본 연구는 국내 인공지능 특허 제목과 미국 O\*NET 업무 텍스트를 활용한 인공지능 노출 점수 계산과 SOC-KSCO 연계에서 딥러닝 자연어처리 기법인 SBERT와 문장 간 유사도 측정을 사용한다. 업무 단위에서 계산한 인공지능 노출 점수를 SOC 소분류 단위 점수로 환산하고, 소분류 기준 SOC-KSCO 연계표를 통해 KSCO 소분류 단위 인공지능 노출 점수를 산출한다. 인공지능 노출 점수에 따라 상위 집단과 하위 집단으로 분류한 다음, 최종적으로 통계청 지역별고용조사사의 소분류 단위 취업자 수를 사용하여 집단별 고용 추세를 분석한다. 전체적인 연구 흐름은 <그림 1> 분석 프레임워크와 같다.

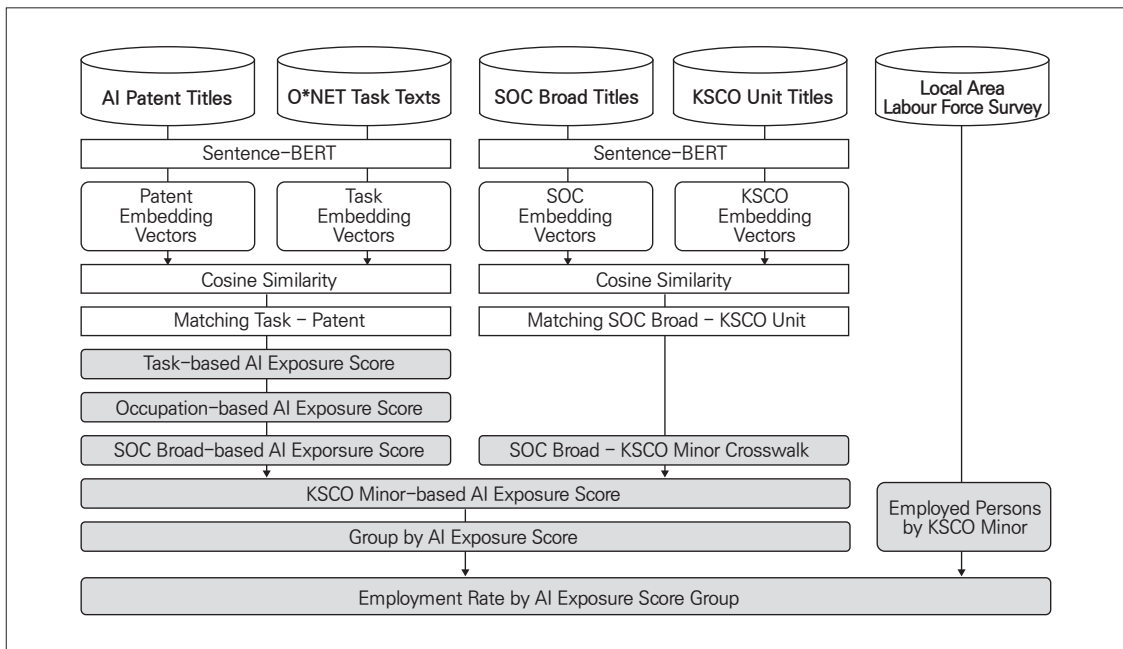
본 장의 구성은 다음과 같다. 우선 인공지능 특허 분류 방식을 설명하고 딥러닝 언어 모델인 SBERT를 소개한다. 그리고 우리가 제안하는 인공지능 노출 점수 계산 모형을 구체적으로 다룬다. 이어서 한국과 미국의

표준직업분류 연계 방식과 인공지능 노출 점수에 따른 고용 추세 분석 방법을 소개한다.

#### 1. 인공지능 노출 점수 계산

##### 1) 인공지능 특허 분류

특허청은 2018년 ‘4차 산업혁명 관련 7대 기술분야 특허분류 체계’를 발표하였다. 7개 기술분야는 인공지능, 빅데이터, 사물인터넷, 3D 프린팅, 자율주행차, 클라우드, 지능형 로봇이며, 각 기술분야에 해당하는 CPC (Cooperative Patent Classification, 협력적 특허분류) 코드와 IPC 코드가 있다. KIPRIS Plus에서 수집한 21년간의 특허 데이터에는 IPC 코드만 부여되어 있으므로 인공지능 분야 IPC 코드로 특허를 분류한다. 다만, 특허청에서는 2022년 4차 산업혁명 관련 기술분야를 16개로 확대하여 신(新)특허분류 체계를 발표하였는데, 기존 코드를 활용하는 것이 아닌 새로운 코드를 만든 것



<그림 1> 분석 프레임워크  
<Fig. 1> Analysis framework

이므로 본 연구의 분석 자료에는 적용할 수 없다. 따라서 통계청의 2018년 분류체계를 따른다.

IPC 코드는 특허의 기술분야를 명시하는 것을 목적으로 하기 때문에 한 건의 특허에 여러 개의 IPC 코드가 할당될 수 있다. 분석에 사용한 IPC 코드는 부록의 <표 1>과 같은데 개별 특허의 IPC 코드를 모두 조회하여 부록 <표 1>의 인공지능 관련 IPC 코드를 한 개 이상 포함한다면 인공지능 특허로 분류하였다.

## 2) Sentence-BERT (SBERT)

Devlin, et al.(2018)에서 제시한 BERT는 트랜스포머를 기반으로 하는 사전 훈련된 언어 모델(Pre-trained Language Model)이다. 언어 모델은 문장과 같은 단어의 시퀀스에서 주어진 단어들로 빈칸에 들어갈 단어를 예측하는 모델을 말한다. BERT는 마스크드 언어 모델(Masked Language Model, MLM)과 다음 문장 예측(Next Sentence Prediction, NSP)으로 학습한다. Devlin, et al.(2018)은 입력된 단어들 중 15%를 랜덤으로 마스크 하였는데, 이 중 80%는 [MASK]로 처리하여 단어를 가리고, 10%는 다른 단어로 변환, 나머지 10%는 그대로 둔 다음 모델이 실제 단어를 예측하도록 훈련하였다. 이러한 MLM은 인공 신경망이 문맥을 양방향으로 학습할 수 있도록 한다. NSP는 두 개의 문장을 주고 이들이 이어지는 문장인지 아닌지 판단하는 훈련 방법이다. BERT는 위키피디아(Wikipedia)의 25억 개 단어와 북스코퍼스(BookCorpus)<sup>1)</sup>의 8억 개 단어라는 레이블이 없는 방대한 데이터로 사전 훈련되어 사용자의 필요에 따라 하이퍼 파라미터를 재조정하는 파인 튜닝(Fine-Tuning)과 추가 훈련을 할 수 있기 때문에 자연어처리 분야에서 향상된 성능을 보인다.

Reimers and Gurevych(2019)이 처음 제안한 SBERT는 BERT의 문장 임베딩(Sentence Embedding) 성능을 우수하게 개선한 모델이다. BERT는 문장을

입력하면 문장을 구성하는 각 단어의 벡터를 출력하기 때문에 문장 단위 분석을 하기 위해서는 문장 대표 벡터를 구해야 한다. BERT에서 대표 문장 벡터를 얻는 방법은 BERT가 각 문장의 시작마다 출력하는 CLS(Special Classification Token) 값을 그대로 사용하는 방법과 단어 벡터들을 풀링(Pooling) 하는 방법 두 가지가 있다. 풀링에는 벡터들의 평균을 구하는 평균 풀링(Mean Pooling)과 벡터들의 최대값을 뽑아내는 최대 풀링(Max Pooling)이 있다. SBERT는 평균 풀링으로 문장 대표 벡터를 구하고 자연어 추론(Natural Language Inferencing, NLI)과 문장 간 유사도 측정(Semantic Textual Similarity, STS)으로 추가 훈련을 한 모델이다.

머신러닝 라이브러리를 개발하는 기업인 허깅페이스(Hugging Face)는 트랜스포머를 기반으로 하는 다양한 자연어 처리 모델을 제공하고 있다. 허깅페이스 Transformer 패키지의 SentenceTransformer 모듈로 분석을 진행하였고, 평균 풀링으로 문장 임베딩을 하고 100가지 언어에 대해 NLI와 STS로 추가 훈련을 한 'xlm-r-100langs-bert-base-nli-stsb-mean-tokens' 모델을 사용하였다.

## 3) SBERT를 활용한 인공지능 노출 점수 계산

본 연구는 Acemoglu and Restrepo(2018)의 업무 기반 모형(Task-based Model)을 확장한 Webb(2019)의 자료 선정 방식과 기술 노출 점수 계산 방법을 Ha, et al.(2022)의 코사인 유사도를 활용한 직업-특허 코드 연결과 결합하여 업무-특허 연계를 시도하였고, 그 과정에서 SBERT를 사용하여 자연어 처리 성능을 높였다.

Webb(2019)는 문장 내에서 단어 간의 문법적 관계를 파악하는 의존 구문 분석(Dependency Parsing) 기능을 제공하는 자연어 처리 모델 spaCy로 O\*NET의 업무 텍스트와 특허 제목에서 동사-명사 쌍을 추출하

1) 문장 유사도 모델을 훈련하기 위해 11,038권의 책에서 수집한 말뭉치로, 출판되지 않아 무료로 사용 가능한 책을 대상으로 하였다. 2만 단어 이상인 책만 포함하였고, 로맨스, 판타지, sf 등 16가지 다른 장르로 구성되어있다. (Zhu, et al., 2015, 21쪽).

였다. WordNet<sup>2)</sup>의 계층적 어휘 구조 데이터베이스로 동의어를 처리하고, 업무 텍스트의 동사-명사 쌍이 특히 제목의 동사-명사 쌍과 겹치는 경우 해당 업무가 기술에 노출된 것으로 보았다. 개별 업무에 대해서 기술 노출 정도를 측정하고, O\*NET의 직업 내 업무 중요도를 가중치로 업무 단위 노출 점수를 집계해 직업별 기술 노출 점수를 도출하였다. Webb(2019)의 연구는 특허와 업무 텍스트로 직업의 기술 노출 정도를 정량화했다는 데 의의가 있다. 하지만 업무의 인공지능 노출 여부 판별에서 사용된 자연어처리 방식이 문장의 구조나 의미를 충분히 반영하기에는 단순한 방식을 이용하고 있기 때문에 자연어처리 부분의 정확도를 개선할 필요가 있다.

Ha, et al.(2022)는 O\*NET의 직업 설명과 특허의 CPC 코드 설명을 워드 임베딩(Word Embedding) 기법으로 벡터화하고 코사인 유사도로 직업과 유사한 특허 코드를 연결하였다. Word2Vec, Doc2Vec, FastText, BERT로 출력한 임베딩 벡터를 군집화하고, 직업 설명 벡터 군집은 O\*NET의 31개 커리어 그룹(Career Cluster)과 비교하고 특허 설명 벡터는 CPC 체계의 136개 그룹과 비교하는 방법으로 정확도를 측정하였다.

Ha, et al.(2022)의 연구에서는 BERT의 정확도가 가장 낮게 나왔는데 이는 전처리 과정에서 WordNet 토큰나이저(Tokenizer)를 사용했기 때문인 것으로 보인다. 주어진 자료를 일정한 의미 단위로 분할하는 토큰화(Tokenization)는 자연어 처리 과정에서 중요한 부분이다. 토큰나이저는 토큰화를 수행하는 모듈로 다양한 방법으로 작동한다. 내부 단어 토큰화(Subword Tokenization)는 입력 문장을 단어보다 더 작은 단위인 내부 단어로 분할하여 사전에 없는 단어나 자주 등장하지 않는 단어에 대응할 수 있다. BERT는 기본적으로 내부 단어 토큰나이저인 WordPiece 토큰나이저를 사용하는데, WordNet 토큰나이저로 전처리한 데이터

를 입력으로 넣으면 높은 성능을 기대하기 어렵다.

또한, Ha, et al.(2022)가 사용한 O\*NET 직업 설명 텍스트는 한두 문장 정도의 간략한 설명으로 이루어져 있어서 업무에 대한 구체적인 설명과는 텍스트의 성격이 다르다. 국제표준직업분류(International Standard Classification of Occupation, ISCO)는 직무(Job)를 개별 종사자들이 수행하는 일련의 업무와 과업(Tasks and Duties)으로 설명하고, 직업(Occupation)을 유사한 직무의 집합으로 정의한다(통계청, 2017). 즉, 업무는 직업을 표현하는 중요한 요소이며, 직업별 업무 텍스트가 가진 정보를 직업 설명이 대체할 수는 없다.

본 연구에서 한국이 아닌 미국의 직업 자료를 사용하는 이유도 위의 설명과 이어진다. 먼저, KSCO 해설서는 세계분류에 포함되는 직업의 예시와 주요 업무를 서술하고 있지만, 대부분 세 문장에서 다섯 문장 정도로 간략히 설명되어 있고, 이것마저 없는 세계분류도 존재한다. 그리고 한국고용정보원의 2020 한국직업사전은 개별 직업에서 수행하는 업무를 수행직무에서 서술하고 있지만 다음과 같은 이유에서 본 분석에 활용하기 부적합하다. 같은 업무라도 직업에 따라 중요도 차이가 있는데 한국직업사전에서는 직업 내 업무 중요도 정보를 확인할 수 없다. 한국직업사전의 수행직무는 작업순서에 따라 작성되기 때문에 나열된 문장들의 순서로 중요도를 판단하기는 힘들다. 중요도 지표는 후술할 인공지능 노출 점수 계산에서 가중치의 역할을 하기 때문에 직업의 특성을 명확히 설명하며 직업 내 업무 중요도 정보를 제공하는 O\*NET의 자료를 사용한다.

언급된 두 연구를 보완하는 본 연구가 제시하는 인공지능 노출 정도 계산 방식은 다음과 같다. 먼저 SBERT를 활용하여 KIPRIS Plus에서 수집한 (한국) 특허 영문 제목과 (미국) O\*NET 업무 텍스트 각각의 문장 대표 벡터를 구한다. 예를 들어, 특허 영문 제목의 문장 대표 벡터  $p$ 는 'the system for auditing information

2) 영어의 명사, 동사, 형용사, 부사를 'synset'이라는 유의어 집합으로 정리한 어휘 데이터베이스다. 'synset'은 개념적 의미(conceptual-semantic)와 어휘적 관계(lexical relation)로 연결된다. (<https://wordnet.princeton.edu/>).



system and the method for auditing information system thereof와 같은 특허 제목을 768차원 벡터로 표현한 것이다. 분석에 사용한 SBERT 모델은 BERT-base를 기반으로 하는데, BERT-base로 출력된 단어의 임베딩 벡터는 768차원이다. SBERT는 BERT-base의 단어 벡터들을 평균 풀링하여 문장 임베딩 벡터를 구하기 때문에 마찬가지로 768차원의 크기를 가진다. 다시 말해서, SBERT는 문장 길이와 상관없이 모든 문장을 같은 크기의 벡터로 만들어주기 때문에 문장 길이 차이로 인해 발생할 수 있는 문제를 방지할 수 있다.

업무 텍스트의 문장 대표 벡터는  $q_{i,j}$ 로 표기하며  $p$ 와 마찬가지로 768차원이다.  $i$ 는 직업( $i = 1, \dots, I$ ),  $j$ 는 업무를 의미하며  $i$ 직업은  $N_i$ 개의 업무를 가진다 ( $j = 1, \dots, N_i$ ). 즉,  $q_{i,j}$ 는  $i$ 직업의  $j$ 업무 텍스트의 문장 대표 벡터다. O\*NET은 개별 직업이 수행하는 업무를 설명하는데, 한 직업 내에 여러 개의 업무가 있고 하나의 업무 텍스트는 예를 들어 'Prepare detailed reports on audit findings.'와 같이 문장 형태로 서술되어있다.

인공지능 노출 점수 계산에 사용되는 유사도 값  $s_{i,j}$ 는 식(1)에 제시된 바와 같이  $p$ 와  $q_{i,j}$ 의 최대 코사인 유사도 값을 의미한다.  $q_{i,j}$ 와 모든  $p$ 의 코사인 유사도를 계산하고 그 중 최댓값을  $s_{i,j}$ 라고 하고 그때의 특허를  $i$ 직업의  $j$ 업무와 가장 유사한 특허로 연결한다. 즉,  $s_{i,j}$ 는  $i$ 직업 내  $j$ 업무와 해당 업무에 연결된 특허 사이의 유사도 값이다.

$$s_{i,j} = \max \frac{p \cdot q_{i,j}}{\|p\| \|q_{i,j}\|} \quad (1)$$

O\*NET에서 제공하는 직업 내 업무 중요도 (*importance*)를 변환하여 가중치로 사용한다. 가중치  $w_{i,j}$ 는  $i$ 직업에서  $j$ 업무가 가지는 중요도  $importance_{i,j}$ 를 직업의 업무 중요도 총합으로 나눈 값이다.

$$w_{i,j} = \frac{importance_{i,j}}{\sum_j importance_{i,j}} \quad (2)$$

여기서  $i$ 직업의 모든 가중치를 합하면 1이 된다.

$$\sum_j w_{i,j} = 1 \quad (3)$$

업무 단위 인공지능 노출 점수  $E_{i,j}$ 는  $i$ 직업의  $j$ 업무가 인공지능에 노출된 정도를 나타내는 수치로 가중치  $w_{i,j}$ 와 유사도 값  $s_{i,j}$ 를 곱한 값이다.

$$E_{i,j} = w_{i,j} s_{i,j} \quad (4)$$

업무 단위 인공지능 노출 점수  $E_{i,j}$ 를 집계하면 직업 단위 인공지능 노출 점수  $E_i$ 를 구할 수 있다.

$$E_i = \sum_j E_{i,j} \quad (5)$$

다음으로 직업 기준으로 계산된 정보를 고용 정보로 연계하기 위해 SOC-2018 소분류(Broad Occupation) 단위 인공지능 노출 점수를 계산한다. 앞서 계산한 직업 단위 점수  $E_i$ 는 SOC-2018의 최하위 분류인 세분류 코드에 소수점 아래 두 자리 숫자가 더해진 O\*NET-SOC-2019 체계의 코드를 기반으로 한다. SOC-2018은 459개 소분류를 가지는데, 각각의 소분류를  $l$ 이라고 하자( $l \in \{L_1, L_2, \dots, L_{459}\}$ ). 소분류 단위 인공지능 노출 점수  $E_i^l$ 은 소분류  $l$ 에 속하는  $i$ 직업의 인공지능 노출 점수 평균으로 정의한다.  $E_i^l$  계산에 사용한 지시함수(indicator function)는 그 형태가  $1[A]$ 일 때  $A$ 가 일어나면 1, 아니면 0을 갖는 함수를 의미한다.

$$E_i^l = \frac{1}{\sum 1[i|i \in l]} \sum_{i \in l} E_i \quad (6)$$

## 2. 고용 추세 분석

### 1) SOC-KSCO 연계

미국의 O\*NET-SOC-2019 체계의 인공지능 노출 점수를 국내 노동시장에 적용하기 위해서 SOC와 KSCO를 연계한다. 한미 직업분류 연계에도 SBERT와 코사인 유사도를 사용한다.

기존의 SOC-KSCO 연계는 ISCO를 통한 간접 연계 방식을 이용하였다. 예를 들어. 통계청은 KSCO-7

과 ISCO-08 연계표를 제공하고, 미국 노동통계국(Bureau of Labor Statistics)은 SOC-2010와 ISCO-08 연계표를 제공한다. SOC-2018과 SOC-2010 연계표도 고시되어 있어 SOC-2018과 KSCO-7의 연계가 가능하다. Ha and Kwon(2021)은 위와 같은 방법으로 SOC와 KSCO를 연계하였다. 하지만 본 연구에서는 딥러닝 언어 모델인 SBERT를 사용하여 SOC와 KSCO를 연계함으로써 두 국가의 직업분류 항목이 가지는 단어의 의미를 자연어 처리 기법을 이용해 최대한 반영할 수 있는 방법을 시도한다.

본 연구의 SOC-2018과 KSCO-7의 연계는 SOC-2018의 소분류 항목과 KSCO-7의 세분류 항목을 연결하고, 연결된 세분류를 상위 단위인 KSCO-7 소분류로 변환하는 순서로 진행된다. SOC-2018 소분류는 459개, KSCO-7 소분류는 150개로 항목 수에서 차이가 나고, SOC는 최하위 단위가 세분류이고 KSCO는 최하위 단위가 세세분류이므로 두 분류 체계의 소분류를 같다고 보기는 어렵다. 예를 들어, 세부 항목을 조회하면, KSCO-7은 '소분류 241 의료 진료 전문가' 이하 '세세분류 24111 내과 전문 의사', '세세분류 24113 외과 전문 의사', '세분류 2412 치과의사' 체계를 가지고 있는 반면, SOC-2018은 소분류 단위에서 '29-1020 Dentists(치과의사)', '29-1210 Physicians(내과의사)', '29-1240 Surgeons(외과의사)'를 정의하고 있는 것을 알 수 있다. 따라서 SOC-2018 소분류 항목을 KSCO-7의 세분류 단위에서 연계한 뒤 해당 세분류 항목을 포함하는 KSCO-7 소분류로 변환하는 방법을 사용한다.

SBERT를 활용한 직업 연계 방법은 업무-특허 연계와 동일하다. 통계청에서 제공하는 KSCO-7의 영문 세분류 항목명과 SOC-2018 소분류 항목명을 벡터로 변환한 다음, SOC-2018 소분류 항목 기준 코사인 유사도가 가장 높은 KSCO-7 세분류 항목을 선택한다. 세분류 단위에서의 연계가 끝나면 KSCO-7 체계에 따라 세분류를 소분류로 변환한다.

## 2) 고용 추세 분석

고용은 소분류 단위 인공지능 노출 점수  $E_i^l$ 의 평균인  $\bar{E}$ 를 기준으로 군집을 나눠 분석한다. 먼저, SOC-2018 소분류 단위 인공지능 노출 점수  $E_i^l$ 을 연계된 KSCO-7 소분류( $m$ ) 항목에 할당하고 이를  $E^m$ 이라 한다.  $E^m$ 값이  $\bar{E}$ 이상인 소분류 항목을 평균 이상 그룹으로,  $E^m$ 이  $\bar{E}$ 이하인 소분류 항목을 평균 이하 그룹으로 구분한다. SOC-2018 소분류를 KSCO-7 세분류 단위에서 연계한 것을 KSCO-7 소분류 단위로 변환하였기 때문에 SOC-2018 소분류와 KSCO-7 소분류가 일대일로 연결되지 않는다. 따라서 평균 이상 집단과 평균 이하 집단에 중복으로 등장하는 소분류 항목이 존재한다. 이런 경우에는 중복되는 소분류 항목의 인공지능 노출 점수  $E^m$ 을 모두 조회한 뒤  $\bar{E}$ 이하인 점수가 과반이면 평균 이하 집단으로 분류하고, 그 외는 평균 이상 집단으로 처리한다.

한국과 미국의 표준직업분류 체계 차이로 인해 SOC-2018 소분류와 KSCO-7 소분류를 바로 연결할 수 없어 KSCO-7 세분류와의 연계 작업을 거쳐야 했다. 이 과정에서 일대일 연계가 이루어지지 못해 하나의 KSCO-7 소분류 항목이 여러 개의  $E^m$ 값을 가지게 되었고 평균 이상 집단과 평균 이하 집단 간 중복이 발생하였다. 이는 통계청에서 제공하는 직업별 취업자 수 정보의 최하위 단위가 KSCO-7 소분류이기 때문이다.

SOC-2018 소분류와 KSCO-7 세분류 연계 과정을 거친 것을 고려하여 KSCO-7 소분류 단위 점수를 새롭게 계산하기보다는 KSCO-7 세분류 항목에 연계된 점수를 그대로 사용하는 방법을 택하였다. 따라서 평균 이상 집단과 평균 이하 집단 간 중복 항목 처리 또한 해당 소분류 항목이 가진  $E^m$ 을 모두 조회하여  $\bar{E}$ 이하인  $E^m$ 이 과반인 경우는 평균 이하 집단으로 분류하고 그렇지 않으면 평균 이상 집단으로 분류하는 방법을 사용하였다.

통계청의 고용률은 만 15세 이상 인구 중 취업자가 차지하는 비율을 말한다. 본 연구의 집단별 고용률은 통계청의 고용률 공식을 이용한다. 앞서 정의한 인공지

능 노출 점수 평균 이상 집단을  $M^H(H: E^m \geq \bar{E})$ , 평균 이하 집단을  $M^L(L: E^m \leq \bar{E})$ 로 표기한다. 집단  $k \in \{H, L\}$ 에 속하는 소분류  $m$ 의  $t$ 년도 취업자 수를  $m_t^k$ ,  $t$ 년도 만 15세 이상 인구수를  $Y_t$ 라고 했을 때  $k$ 집단의  $t$ 년도 고용률  $R_t^k$ 는 다음과 같다.

$$R_t^k = \frac{\sum_{m \in k} m_t^k}{Y_t} \quad (7)$$

## IV. 분석 결과

### 1. 인공지능 노출 점수

#### 1) 인공지능 특허

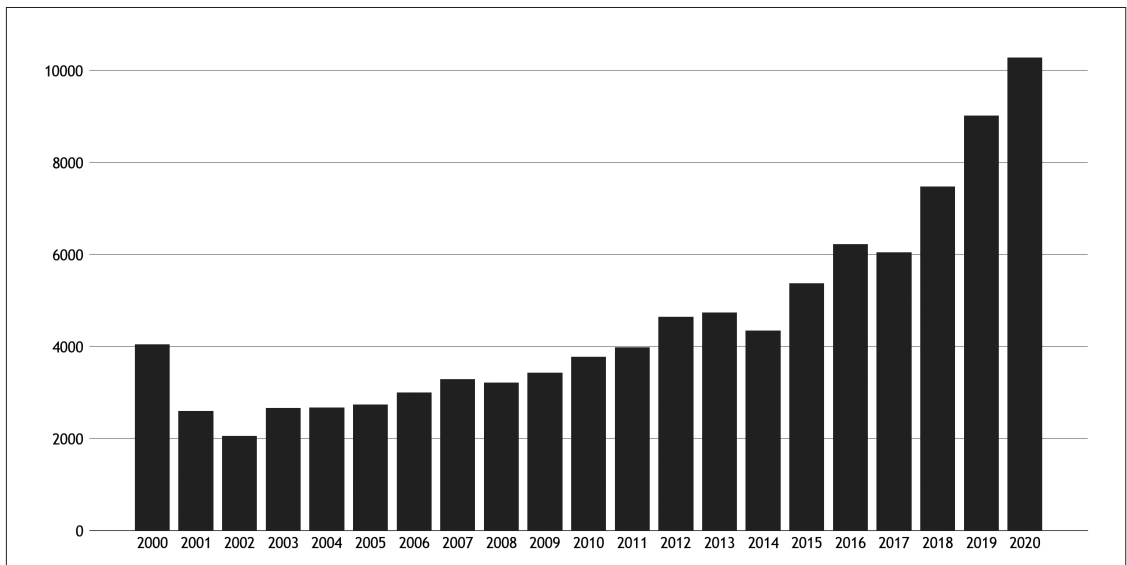
특허청의 인공지능 특허 분류 체계의 IPC 코드와 KIPRIS Plus 특허 데이터의 IPC 코드를 연결하여 분류한 인공지능 특허는 95,602건이다. <그림 2>를 보면, 국내 인공지능 특허의 출원은 2000년부터 2002년까지는 감소하지만 이후 꾸준히 증가하는 것을 확인할

수 있다. WIPO(2019)는 인공지능 기술 관련 세계 특허 출원이 2013년부터 급격하게 늘기 시작했으며, 인공지능 특허에서 가장 큰 비중을 차지하는 머신러닝의 경우 2013년부터 2016년까지 매년 연평균 28% 증가했다고 밝혔다. 한국의 특허를 분석한 본 연구의 결과에서는 인공지능 특허 출원 수가 같은 기간 매년 연평균 10% 증가했음을 확인하였다.

#### 2) O\*NET-SOC-2019 직업 단위 인공지능 노출 점수

O\*NET에서 직업 제목 외 추가적인 정보를 제공하는 직업 923개 중 업무 텍스트와 직업 내 업무 중요도를 모두 가진 직업은 823개다. 개별 직업을 설명하고 있는 업무 텍스트와 직업 내 업무 중요도 수치 자료의 형태는 <표 2>에서 제시한다. <표 2>는 '13-2011.00 Accountants and Auditors'가 수행하는 26개 업무 중 5개를 나타낸다.

분석에 사용된 823개 직업에 대한 전체 업무 텍스트 수는 18,291개이며, 대다수 직업이 15개 이상 25개 이하의 업무 텍스트를 가지고 있다. 직업 내 업무 텍스트



<그림 2> 연도별 인공지능 특허 출원 수 (KIPRIS)  
<Fig. 2> Patent application of AI by year (KIPRIS)

〈표 2〉 Accountants and Auditors의 업무 텍스트와 중요도  
 〈Table 2〉 Task text and importance of Accountants and Auditors

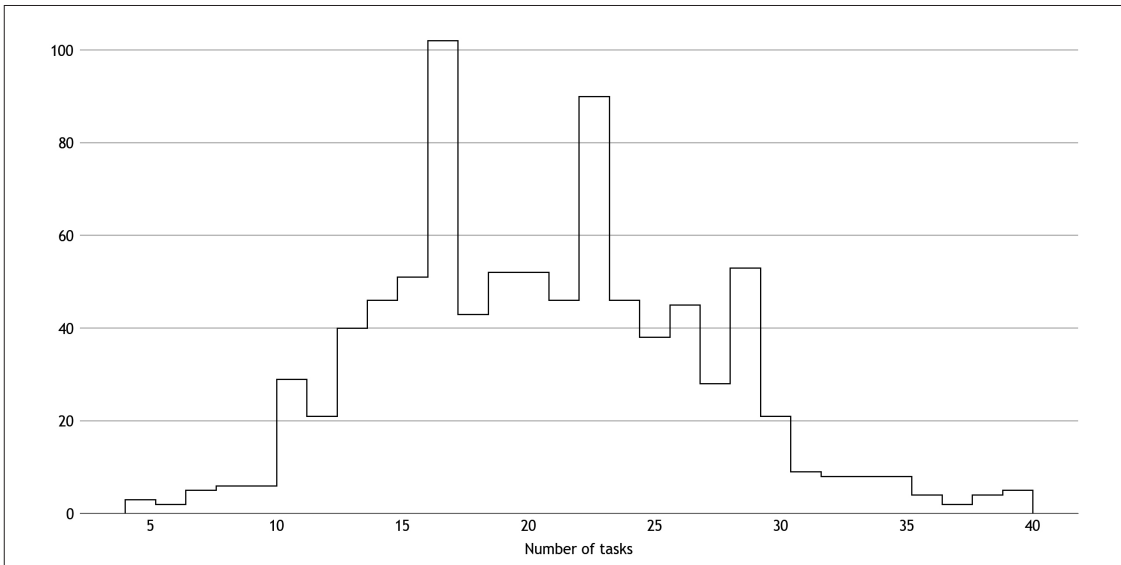
Task text	Importance
Prepare detailed reports on audit findings.	90
Report to management about asset utilization and audit results, and recommend changes in operations and financial activities.	89
Collect and analyze data to detect deficient controls, duplicated effort, extravagance, fraud, or non-compliance with laws, regulations, and management policies.	88
Inspect account books and accounting systems for efficiency, effectiveness, and use of accepted accounting procedures to record transactions.	88
Supervise auditing of establishments, and determine scope of investigation required.	88

source: O\*NET database

수의 분포는 〈그림 3〉에서 확인할 수 있다. 직업에 따라 차이 나는 업무 텍스트의 수는 분석 방법에서 설명한 업무 내 중요도 수치를 활용한 가중치를 통해 보정하였다. 업무 단위 인공지능 점수는 업무 텍스트와 특히 제목 간의 유사도 값에 가중치를 곱하여 계산되고, 이를 집계한 직업 단위 인공지능 노출 점수에도 가중치

가 반영되므로 업무 텍스트 수의 차이로 발생할 수 있는 문제는 통제되었다.

2000년부터 2020년까지 한국에 출원된 인공지능 특허 95,602건과 O\*NET의 18,291개 업무 텍스트를 이용하여 업무 단위 인공지능 노출 점수를 계산하였다. SBERT와 코사인 유사도 측정을 통한 업무-특허 연계



〈그림 3〉 직업 내 업무 텍스트 수 분포 (O\*NET)  
 〈Fig. 3〉 Distribution of the number of task text across occupations (O\*NET)

와 이를 기반으로 계산한 업무 단위 인공지능 노출 점수의 예시는 <표 3>과 같다. <표 3>은 '29-1071.00 Physician Assistants'의 12개 업무 중 5개 업무에 대한 연계 특허와 업무 단위 인공지능 노출 점수를 보여 주고 있다.

예를 들어 <표 3>에서 Physician Assistants가 수행하는 12개 업무의 중요도 합이 1065이므로 중요도 수치가 99인 업무 'Make tentative diagnoses and decisions about management and treatment of patients.'의 가중치  $w_{i,j}$ 는 0.093이다. 95,602건의 인공지능 특허 중 이 업무와 가장 높은 코사인 유사도 값을 가지는 특허 'method and system for managing patient'가 연결되었고, 이때의 유사도 값 0.786이  $s_{i,j}$ 가 되었다. Physician Assistants 첫 번째 업무의 인공지능 노출 점수  $E_{i,j}$ 는  $w_{i,j}$ 와  $s_{i,j}$ 를 곱한 값으로 0.073이다.

업무-특허 연계와 업무 단위 인공지능 노출 점수 계산에 사용된 코사인 유사도는 1부터 -1까지의 값을

가질 수 있다. 본 연구에서는 특허 제목 벡터  $p$ 와 업무 텍스트 벡터  $q_{i,j}$ 사이의 유사도 값 중 가장 큰 값을 선택하였기 때문에  $s_{i,j}$ 에 음수 값은 할당되지 않았다. 가중치  $w_{i,j}$ 는  $i$ 직업 내  $j$ 업무 중요도를  $i$ 직업이 가진 전체 업무 수를 고려하여 환산한 값이므로, 개별 직업의 업무 가중치 합은 1이다. 따라서 직업 단위 인공지능 노출 점수  $E_i$ 는 0과 1 사이의 값을 가진다.

<표 4>와 <표 5>에서 각각 업무 텍스트와 특허 제목 간의 유사도 상위 항목과 하위 항목을 정리하였다. 업무 텍스트 'Perform wedding'과 특허 제목 'Wedding service system'의 유사도가 0.935로 가장 높았지만, 이 업무는 직업 'Judges, Magistrate Judges, and Magistrates'에서 중요도가 35밖에 되지 않아 인공지능 노출 점수에 영향을 주지 못했다. 반면, 'Manage safety programs at power generation facilities.'는 'Biomass Power Plant Managers'에서 가장 높은 중요도인 91을 가지는 업무이기 때문에 해당 직업의 인공지능 노출 점수 계산 시 크게 반영되었다.

**<표 3> Physician Assistants 업무 단위의 인공지능 노출 점수**  
**<Table 3> Task-based AI exposure score of Physician Assistants**

Tasks	Weight ( $w_{i,j}$ )	Patents	Similarity ( $s_{i,j}$ )	Score ( $E_{i,j}$ )
Make tentative diagnoses and decisions about management and treatment of patients.	0.093	method and system for managing patient	0.786	0.073
Interpret diagnostic test results for deviations from normal.	0.092	apparatus and method for detecting abnormal behavior using data pattern analysis	0.789	0.073
Prescribe therapy or medication with physician approval.	0.092	method for guiding of taking medicine and apparatus for executing the method	0.819	0.075
Obtain, compile, and record patient medical data, including health history, progress notes, and results of physical examination.	0.092	method combined with medical photography and stethoscopic record for supplying patient status information	0.855	0.079
Examine patients to obtain information about their physical condition.	0.092	method for diagnosing user using body information	0.775	0.071



〈표 5〉의 ‘Bailiffs’는 전체 15개 업무 중 중요도 85 ‘Guard lodging of sequestered jury.’와 중요도 76 ‘Stop people from entering courtroom while judge charges jury.’가 매우 낮은 유사도 값을 가져 인공지능 노출 점수가 낮게 계산되었다. 유사도 값은 하나의 업무와 모든 인공지능 특허를 비교했을 때 가장 높게

측정되는 값을 가져온 것인데, 〈표 5〉의 업무는 〈표 4〉의 업무보다 낮은 값을 가짐과 동시에 상이한 특허와 연결되었음을 확인할 수 있다.

Webb(2019)의 기술 노출 점수 측정 방법에 코사인 유사도를 도입하면서, 0과 0.002 사이의 값을 가지는 Webb(2019)의 결과보다 점수대가 높아졌다. 직업 단

〈표 4〉 업무 텍스트 - 특허 제목 유사도 상위 5개

〈Table 4〉 Similarity between task text and patent title, top 5

SOC-2018 Occupation	Tasks	Patents	Similarity ( $S_{i,j}$ )
Judges, Magistrate Judges, and Magistrates	Perform wedding ceremonies.	wedding service system	0.935
Paving, Surfacing, and Tamping Equipment Operators	Control traffic.	traffic control method	0.933
Telephone Operators	Update directory information.	system and method for updating directory	0.930
Transportation Planners	Produce environmental documents, such as environmental assessments or environmental impact statements.	system for collecting and evaluating environmental information	0.930
Biomass Power Plant Managers	Manage safety programs at power generation facilities.	the management method of safety evaluation procedure innuclear power plants facilities	0.929

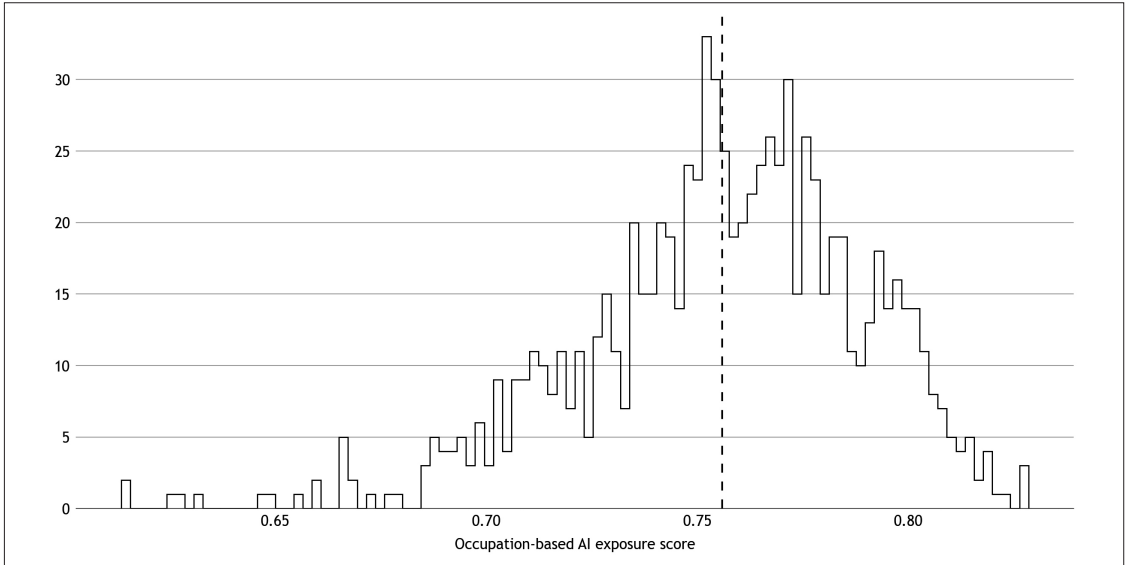
〈표 5〉 업무 텍스트 - 특허 제목 유사도 하위 5개

〈Table 5〉 Similarity between task text and patent title, bottom 5

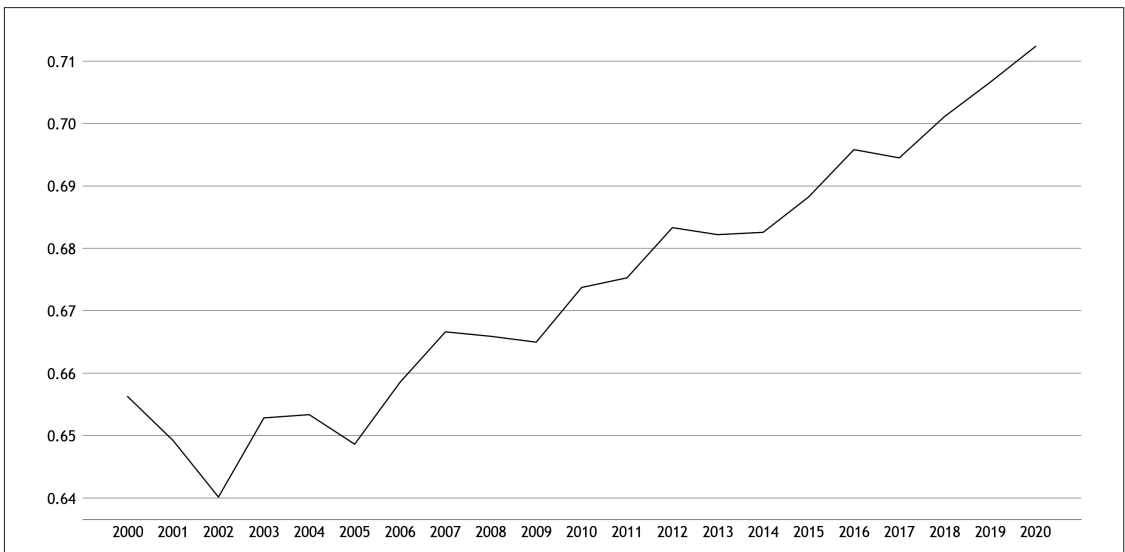
SOC-2018 Occupation	Tasks	Patents	Similarity ( $S_{i,j}$ )
Log Graders and Scalers	Saw felled trees into lengths.	wood furniture production system	0.464
Bailiffs	Stop people from entering courtroom while judge charges jury.	blocking bar for parking lot	0.477
Fishing and Hunting Workers	Kill or stun trapped quarry, using clubs, poisons, guns, or drowning methods.	die capable of being opened, manufacturing method thereof, and mold	0.488
Actors	Read from scripts or books to narrate action or to inform or entertain audiences, utilizing few or no stage props.	toy performance apparatus and method using chatting	0.494
Bailiffs	Guard lodging of sequestered jury.	system for checking the number of persons rided in a bus	0.497

위 인공지능 노출 점수  $E_i$ (식 (5) 참조)의 분포를 나타낸 <그림 4>를 보면 점수가 오른쪽으로 치우쳐져 있음을 알 수 있는데, 이는 미국 특허 자료를 사용한 Webb(2019)

와 비교했을 때 한국 특허를 사용한 연구에서 나타나는 특징이다. <그림 4>의 세로 점선은 직업 단위 인공지능 노출 점수의 평균으로, 그 값은 0.756이다.



〈그림 4〉 직업 단위 인공지능 노출 점수 분포  
 〈Fig. 4〉 Distribution of occupation-based AI exposure score



〈그림 5〉 연도별 직업 단위 인공지능 노출 점수 평균  
 〈Fig. 5〉 Average of occupation-based AI exposure score by year

직업 단위 인공지능 노출 점수에 어떤 변화가 있는지 알아보기 위해 연도별 인공지능 노출 점수를 계산하였다. 2000년부터 2020년까지의 한국에 출원된 인공지능 특허 전체를 사용한  $E_i$ 와 달리  $E_{i,t}$ 는  $t$ 년도에 출원된 한국의 인공지능 특허로 계산한 직업 단위 인공지능 노출 점수다. <그림 5>는  $E_{i,t}$ 의 평균값으로 본 2000년부터 2020년까지의 인공지능 노출 점수 변화다. 연도별 인공지능 출원 특허 수가 분석 기간 초기에 감소했다가 이후로 상승세가 이어지는 것과 비슷한 양상을 보인다.

## 2. 고용 추세 분석

### 1) SOC-KSCO 연계

SOC-2018 소분류는 ‘대분류 55-0000 Military Specific Occupations’를 제외하면 총 436개이고, O\*NET에서 업무 텍스트와 직업 내 업무 중요도를 제공하는 직업의 소분류는 430개다. SOC-2018의 430개 소분류를 KSCO-7의 450개 세분류와 연계한 결과 230개의 세분류가 연결된 것을 확인하였다. 연계된 세분류를 소분류로 변환하면, 최종적으로 KSCO-7의

〈표 6〉 인공지능 노출 점수 상위 항목 SOC-2018-KSCO-7 연계표  
 〈Table 6〉 SOC-2018-KSCO-7 crosswalk of top AI exposure score group

SOC-2018 Broad		KSCO-7 Minor group	
Code	Title	Code	Title
51-8030	Water and Wastewater Treatment Plant and System Operators	881	Water Treatment Plant Operators
29-1070	Physician Assistants	246	Health and Medical Service Related Workers
29-1030	Dietitians and Nutritionists	244	Dietitians
51-8020	Stationary Engineers and Boiler Operators	841	Metal Casting and Metal Processing Related Operators
15-1220	Computer and Information Research Scientists	221	Computer Hardware and Telecommunication Engineering Professionals

〈표 7〉 인공지능 노출 점수 하위 집단 SOC-2018-KSCO-7 연계표  
 〈Table 7〉 SOC-2018-KSCO-7 crosswalk of bottom AI exposure score group

SOC-2018 Broad		KSCO-7 Minor group	
Code	Title	Code	Title
33-3010	Bailiffs, Correctional Officers, and Jailers	131	Research, Education and Legal Services Related Managers
45-4020	Logging Workers	620	Forestry Related Workers
47-5050	Rock Splitters, Quarry	784	Mining and Civil Engineering Related Technical Workers
51-3020	Butchers and Other Meat, Poultry, and Fish Processing Workers	710	Food Processing Related Trades Workers
39-4020	Funeral Attendants	423	Wedding Ceremony and Funeral Service Workers

156개 소분류 중 119개 항목이 SOC-2018 소분류와 연계된다.

〈표 6〉과 〈표 7〉은 각각 SCO-2018 소분류 단위 인공지능 노출 점수 상위 항목과 하위 항목의 SOC-KSCO 연계표다.

미국 노동통계국의 SOC-2018 설명서와 통계청의 KSCO-7 해설서의 항목별 업무 설명 내용을 보면, 분류 항목명만을 사용하였는데도 직업적 특성을 유의미하게 반영하고 있는 것으로 판단된다. '의사의 감독하에 의료 서비스를 제공'하는 '29-1070 Physician Assistants'를 '의사가 수행하는 것보다 범위와 복잡성에서 한정된 진단 및 치료 업무를 지원·수행'하는 '246 보건 의료 관련 종사자'와 연결하였다. 〈표 7〉의 세 번째, 네 번째 항목의 경우 SOC-2018 기준으로는 '47-0000 Construction and Extraction Occupations'와 '51-0000 Production Occupations'라는 다른 대분류에 속하는 소분류이지만, KSCO-7 기준으로는 두 항목 모두 '대분류 7 기능원 및 관련 기능 종사자'에 포함된다.

〈표 6〉의 '244 영양사'의 경우 중요도 수치가 91로 가장 높은 업무 'Assess nutritional needs, diet restrictions, and current health plans to develop and implement dietary-care plans and provide nutritional counseling'가 특히 'method for provision of health information based diet recommendation and brokerage platform(건강정보기반의 식단 추천 및 조리영양사 중개 플랫폼 제공 방법)'과 연결되고, 그 다음으로 높은 업무 'Evaluate laboratory tests in preparing nutrition recommendations.'가 특히 'an application that allows to check the results of nutrition detection(영양소 검출 결과를 확인할 수 있는 어플리케이션)'과 연결되었다. 업무 텍스트와 관련 인공지능 특허를 직접 비교했을 때, 인공지능 노출 정도가 높은 직종임을 알 수 있다.

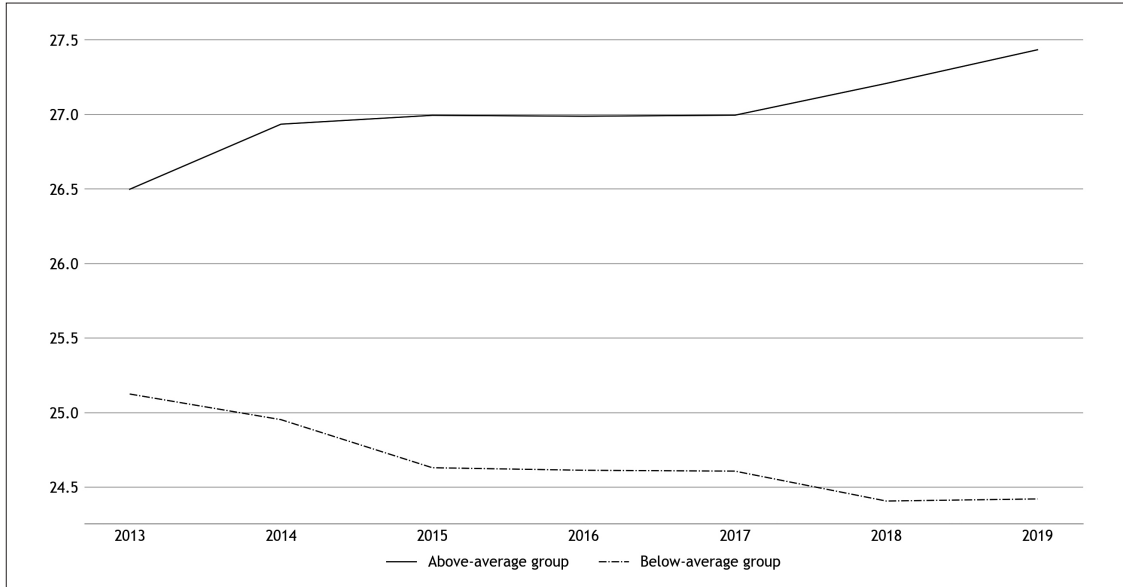
〈표 7〉의 '33-3010 Bailiffs, Correctional Officers, and Jailers'는 '세분류 1313 법률 경찰 소방 및 교도 관리자'와 연결되어 소분류 단위에서 '131 연구 교

육 및 법률 관리자'와 연계되었다. SOC-2018 소분류 항목의 세분류로는 '33-3011 Bailiffs'와 '33-3012 Correctional Officers and Jailers'가 있다. 'Bailiffs'는 앞의 〈표 5〉에서 업무 단위 인공지능 점수가 낮음을 확인하였다. 'Correctional Officers and Jailers'의 경우, 중요도 95인 업무 'Inspect conditions of locks, window bars, grills, doors, and gates at correctional facilities to ensure security and help prevent escapes.'와 가장 높은 유사도를 가지는 특허가 'head mounted display apparatus and method for controlling the same(헤드 마운티드 디스플레이 장치 및 이의 제어방법)'인 것으로 볼 때 수행하는 업무가 인공지능 기술과 큰 관련이 없다는 것을 알 수 있다.

엄밀히 따지면, '1313 법률 경찰 소방 및 교도 관리자'는 관리자 직군이고, '33-3010 Bailiffs, Correctional Officers, and Jailers'는 실무를 담당하는 직군이기에 때문에 정확한 연계라고 보기 어려울 수 있다. 이는 자연 어처리 기반 연계 모형이 직업의 직능수준(Skill Level)을 판별하는 데는 한계가 있다고 해석된다. 하지만 본 연구는 단어 간의 의미를 유기적으로 결합함으로써 기존의 ISCO 활용 연계와는 다른 접근을 제시하였고, 분석 결과 국가 간 분류 체계의 차이에도 불구하고, 딥러닝 언어 모델을 활용한 표준직업분류 연계가 가능함을 확인하였다.

## 2) 인공지능 노출 점수 집단별 고용 추세 분석

고용추세 분석을 위해 KSCO-7이 적용된 2013년부터 코로나19 영향이 미치기 전인 2019년까지의 통계청 지역별고용조사를 사용한다. 통계청은 2018년 시행된 KSCO-7을 2013년 이후의 '전국 직업/성별 취업자(소분류)'에 소급 적용하였다. KSCO-6과 KSCO-7을 연계하면 2008년부터의 자료를 사용할 수 있지만, 세세분류 단위 기준인 통계청의 연계표를 소분류 단위 고용 추세 분석에 적용하는 것에는 한계가 있으므로 2013년 이후의 자료를 사용하였다. 그리고 2020년은 코로나19의 영향으로 전체 고용률이 전년 대비 1.3%p



〈그림 6〉 인공지능 노출 점수 집단별 고용률  
 〈Fig. 6〉 Employment rate by AI exposure score group

하락한 해다. 특허와 직업을 연계하여 인공지능 기술 발전이 노동시장에 미치는 영향을 분석하고자 하는 연구의 목적에 맞지 않다고 판단하여 2020년 고용 자료는 제외하였다.

집단별 고용 추세 분석 결과 〈그림 6〉과 같이 한국 고용시장에서 인공지능에 따른 기술 발전은 평균적인 인공지능 노출 정도에 따라 K자형 양상으로 양극화되어 영향을 미치고 있다. 소분류 단위 인공지능 노출 점수의 평균  $\bar{E}$ 를 기준으로 소분류 항목을 분류하였을

때, SOC-2018 소분류와 연계된 119개의 KSCO-7 소분류 중 평균 이상 집단  $M^H$ 와 평균 이하 집단  $M^L$ 에 각각 61개, 58개 항목이 포함되었다. 분석 기간 동안 인공지능 노출 점수 평균 이상 집단의 고용률이 평균 이하 집단의 고용률보다 매해 높았고, 평균 이하 집단의 고용률은 감소하는 데에 반해 평균 이상 집단의 고용률은 증가하여 두 집단의 고용률 격차는 벌어졌다.

평균 이상 집단  $M^H$ 의 평균 고용률은 27.01%, 평균 이하 집단  $M^L$ 의 평균 고용률은 24.68%로 두 집단의

〈표 8〉 인공지능 노출 점수 집단별 고용률  
 〈Table 8〉 Employment rate by AI exposure score group

Group	Number of KSCO-7 Minor group	Average employment rate(%)	$R_t^H - R_t^L$ (%P)						
			2013	2014	2015	2016	2017	2018	2019
$M^H$	61	27.01	1.37	1.98	2.36	2.38	2.39	2.81	3.01
$M^L$	58	24.68							

source: Author's calculations based on the KOSIS database



평균 고용률 차이는 2.33%p다. 두 집단의 연도별 고용률 차이는 <표 8>에서 확인할 수 있다. 2013년 평균 이상 집단의 고용률은 26.49%, 평균 이하 집단의 고용률은 25.12%로 두 집단 간 고용률 차이는 1.37%p였다. 2019년 평균 이상 집단의 고용률은 27.43%로 증가하고, 평균 이하 집단의 고용률은 24.42%로 감소하여 차이는 3.01%p로 늘어났다.

## V. 결론

본 연구는 인공지능 노출 점수 평균을 기준으로 KSCO 소분류 항목을 분류하여 분석한 집단별 고용 추세에서 한국의 경우 인공지능 노출 점수 평균 이상 집단의 고용은 증가하는 반면, 평균 이하 집단의 고용은 감소하는 K자형 양극화가 나타남을 확인하였다. 우리의 결과는 한국의 고용 환경에서 인공지능 활용과 같은 기술 진보는 자동화와 새로운 업무의 도입이라는 상반되는 양상이 직업에 따른 양극화로 나타난다는 새로운 해석을 제시한다.

이러한 새로운 해석은 본 연구가 시도한 딥러닝 언어 모델 SBERT 기반의 업무-특허 연결을 통한 인공지능 노출 점수 계산과 SOC-KSCO 연계라는 기여들로부터 도출되었다고 볼 수 있다. 기존 연구 방식을 활용 또는 변형하고 거기에 새로운 기술을 적용함으로써 현상에 대한 새로운 시각을 제시했다는 점에서 기존의 기술 진보와 고용에 관한 연구와는 차별점이 있다.

NIA(2023)은 2023년 12대 디지털 트렌드 중 하나로 양극화, N극화 현상을 지목하였다. 인공지능의 발전을 비롯한 기술혁신 속에서 디지털 역량 격차가 심해지고 이로 인한 불평등이 사회적 문제로 대두된다는 전망을 한 것이다. 본 연구는 2013년부터 2019년까지의 고용 추세 분석을 통해 일자리 분야에서 기술 발전으로 인한 양극화가 진행되고 있음을 확인하였다. 분석 기간 동안 인공지능 노출 점수 평균 이상 집단의 고용률이 항상 높았던 것은 각 집단에 포함된 소분류 항목 수에서 차이가 나기 때문일 수 있다. 하지만 항목 수의 차이

는 크지 않았고, 그 점을 고려하더라도 평균 이상 집단의 고용은 증가 추세를, 평균 이하 집단의 고용은 감소 추세를 보이는 것으로 양극화가 이미 진행 중임을 알 수 있다. 따라서, 인공지능 노출 점수 평균 이하 직업군에 대한 고용 정책 마련이 필요함을 시사한다.

서론에서 밝혔듯이 인공지능 기술의 발전과 노동시장의 관계는 큰 관심의 대상인 것에 비해 실증 연구는 부족하다. 특히 국내 선행연구의 경우 해외 연구의 지표를 그대로 사용했다는 한계가 존재한다. 본 연구는 국내 출원 특허 자료를 사용하여 한국의 인공지능 기술 발전 정도를 반영한 직업별 인공지능 노출 점수 계산을 시도하였다. 모형 제시에서 그치지 않고, 한미 양국 간 표준 직업분류 연계와 통계청 지역별고용조사를 이용한 고용 추세 분석으로 한국의 노동시장 상황을 파악하였다.

인공지능은 현재도 활발히 발전하고 있는 분야이기 때문에 7년간의 고용 추세 분석 결과가 미래에도 이어진다고 단정 지을 수 없다. 하지만 본 연구는 인공지능과 고용의 관계에 대한 실증 분석 방법을 제시하였다. 전문가 인터뷰 또는 설문조사를 사용한 기존의 방법론에 비해 적은 시간과 비용을 요구하기 때문에 빠르게 발전하는 기술의 영향력을 평가할 수 있는 접근법을 제시했다는 의의도 있다. 이를 기반으로 추후 연구가 이어간다면 인공지능을 비롯한 새로운 기술에 대응하는 정책적 의사 결정에 도움이 될 것으로 기대한다.

본 연구의 한계는 다음과 같다. 먼저, 미국 직업 중심으로 만들어진 O\*NET의 업무 텍스트를 사용했다는 점이다. 분석 방법에서 서술했듯이, 통계청의 KSCO 해설서나 한국고용정보원의 한국직업사전을 활용하는 방법도 고려해보았지만, 본 연구의 모형은 직업별로 상세한 업무와 각 업무의 중요도가 정리되어있는 자료가 필수적이었기 때문에 O\*NET을 사용하였다.

다음으로, SOC-KSCO 연계가 일대일로 이루어지지 못했다는 점이다. 통계청이 제공하는 직업별 취업자 수의 최하위 단위가 소분류이기 때문에 SOC-2018 소분류와 KSCO-7 소분류를 연결할 필요가 있었다. SOC-2018과 KSCO-7의 체계 차이로 인해 SOC-2018 소분

류와 KSCO-7 세분류를 연계한 뒤 KSCO-7 소분류로 변환하는 작업을 거쳐야 했고, 그 과정에서 SOC-2018 소분류와 KSCO-7 소분류가 일대일로 연결되지 못하였다. 이 때문에 하나의 KSCO-7 소분류 항목이 여러 개의 인공지능 노출 점수를 가지게 되었다.

마지막으로, 딥러닝 언어 모델이 가지는 한계가 있다. SBERT를 사용하여 자연어처리에서 선행연구가 가지는 문제점을 보완하였지만, 여전히 개선이 필요하다. 분석 결과에서도 밝혔듯이 SOC-KSCO 연계 시, 직업의 직능수준을 구분하는 성능에서 아쉬움이 있다. 이는 언어 모델이 직업분류라는 특수한 분야에 대해서 완벽히 학습을 하지 못했기 때문인 것으로 보인다. 다양한 직업 텍스트로 추가 훈련을 하는 등의 방법을 사용하면 향후 연구에서 개선된 성능을 보일 것으로 기대된다.

앞서 밝힌 한계점들은 국내 직업 데이터베이스 보완, 좀 더 발전된 딥러닝 언어 모델의 도입, 추가 훈련을 통한 본 분석에 특화된 언어 모델 구축 등으로 해결될 수 있을 것이라 기대한다. 본 연구는 국내 인공지능 출원 특허를 사용하여 한국의 특성을 반영하고자 하였고, 딥러닝 자연어처리 기술을 이용한 국가 간 표준직업분류 연계를 처음 시도하였다. ISCO를 매개로 하는 보편적인 방법 대신, 문맥을 파악하여 벡터화하는 딥러닝 언어 모델의 특성을 살려 의미 중심 직업분류 연계를 진행하였다는 데에 의의가 있다.

본 연구가 인공지능 기술과 고용 간의 인과관계를 밝힌 것은 아니기 때문에 추가 분석이 필요하고, 이는 향후 연구 주제로 남겨둔다. 하지만 기존의 연구와는 차별된 접근으로 한국 노동시장의 고용 추세는 자동화로 인한 인간 노동 대체 혹은 새로운 업무 도입으로 인한 일자리 증가 중 한쪽으로 치우치지 않고 인공지능 노출 정도에 따른 K자형 양극화로 나타남을 확인하였다. 인공지능 기술에 대한 사회적 관심이 높고 산업적 활용도 늘어날 것이라는 전망 속에서 본 연구의 결과는 인공지능 노출 정도에 따라 집단을 분리한 고용 정책 설계가 필요함을 보여준다.

## References

- Acemoglu, D. & Restrepo, P. (2018). "The race between man and machine: Implications of technology for growth, factor shares, and employment." *American Economic Review*, 108(6), 1488-1542.
- Acemoglu, D. & Restrepo, P. (2019). "Automation and new tasks: How technology displaces and reinstates labor." *Journal of Economic Perspectives*, 33(2), 3-30.
- Arntz, M., Gregory, T. & Zierahn, U. (2016). *The risk of automation for jobs in OECD countries: A comparative analysis*. Paris: Organisation for Economic Co-operation and Development.
- Autor, D. & Salomons, A. (2018). "Is automation labor-displacing? Productivity growth, employment, and the labor share." *National Bureau of Economic Research*, No. w24871.
- Brynjolfsson, E., Mitchell, T. & Rock, D. (2018). *What can machines learn, and what does it mean for occupations and the economy?* Paper presented at the 130th Annual Meeting of the American Economic Association, Jan 5-7.
- Cheon, B., Jeong, J. & Chang, J. (2022). "The Employment and Wage Effects of Artificial Intelligence (AI)." *Journal of Economics Studies*, 40(1), 133-156.
- {전병유·정준호·장지연 (2022). 인공지능(AI)의 고용과 임금 효과. <경제연구>, 40권 1호, 133-156.}
- Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. (2018). "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805*.
- Eom, H. & Lee, M. (2020). "A Study on Labor Market Changes from Artificial Intelligence (AI) in the Intelligence Information Society." *Information Society & Media*, 21(2), 1-20.

- {엄효진·이명진 (2020). 인공지능 (AI) 기반 지능정보사회 시대의 노동시장 변화: 경제사회학적 접근을 중심으로. <정보사회와 미디어>, 21권 2호, 1-20.}
- Felten, E. W., Raj, M. & Seamans, R. (2019). "The occupational impact of artificial intelligence: Labor, skills, and polarization." *NYU Stern School of Business, Available at SSRN 3368605*.
- Felten, E., Raj, M. & Seamans, R. (2021). "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses." *Strategic Management Journal*, 42(12), 2195-2217.
- Frey, C. B. & Osborne, M. A. (2017). "The future of employment: How susceptible are jobs to computerisation?" *Technological forecasting and social change*, 114(1), 254-280.
- Ha, J. & Kwon, H. (2021). "The Types of Job Mobility for Young Workers According to Automation Risk." *Korean Journal of Industrial Relations*, 31(2), 61-88.
- {허재영·권현지 (2021). 자동화 위험도에 따른 청년 근로자 직업이동 유형화. <산업관계연구>, 31권 4호, 61-88.}
- Ha, T., Lee, M., Yun, B. & Coh, B. (2022). "Job Forecasting Based on the Patent Information: A Word Embedding-Based Approach." *IEEE Access*, 10, 7223-7233.
- Huang, M. H. & Rust, R. T. (2018). "Artificial intelligence in service." *Journal of service research*, 21(2), 155-172.
- Hur, J. (2019). "AI and the Future of Work: Concerns, Theory and Reality." *The Korean Economic Forum*, 12(3), 59-92.
- {허재준 (2019). 인공지능과 노동의 미래: 우려와 이론과 사실. <한국경제포럼>, 12권 3호, 59-92.}
- Kim, G. & Bae, J. (2017). "A novel approach to forecast promising technology through patent analysis." *Technological Forecasting and Social Change*, 117, 228-237.
- Kim, S. (2015). *Changes in the Labor Market and Response to Technological Advancement*. Seoul: Korea Labor Institute.
- {김세움 (2015). <기술진보에 따른 노동시장 변화와 대응>. 서울: 한국노동연구원.}
- Kim, Y. (2019). "AI, Employment, Economic Growth, and Inequality: Recent Literature Survey and Policy Implications." *The Korean Economic Forum*, 12(3), 1-34.
- {김영식 (2019). AI 와 고용, 경제성장, 불평등: 최근 문헌 개관 과 정책 함의. <한국경제포럼>, 12권 3호, 1-34.}
- Lee, C., Cho, H. & Chang, Y. (2021). "Exploring the Effect of Job Replacement by Technological Innovation: Focusing on Social Cost Estimation and Social Perception Analysis." *Journal of Social Science*, 60(1), 233-271.
- {이채정·조희찬·장운정 (2021). 기술혁신에 따른 직무대체의 사회적 영향 탐색: 사회적 비용 추계와 사회적 인식 검토를 중심으로. <사회과학연구>, 60권 1호, 233-271.}
- Lee, H., Qiao, X., Shin, S., Kim, G. & Oh, S. (2022). "Analysis of Korea's Artificial Intelligence Competitiveness Based on Patent Data: Focusing on Patent Index and Topic Modeling." *Informatization Policy*, 29(4), 43-66.
- {이현상·차오신·신선영·김규리·오세환 (2022). 특허데이터 기반 한국의 인공지능 경쟁력 분석: 특허지표 및 토픽 모델링을 중심으로. <정보화정책>, 29권 4호, 43-66.}
- Lee, S., Cheong, J., Kim, S., Hong, H., Cheong, S., Lee, J., Lee, Y., Kim, J., Hwang, G., Kim, E., Hong, S., Kong, J. & Lee, H. (2017). *Prospects for Mid-Long-Term Manpower Demand Reflecting Technological Innovation*. Sejong: Korea Employment Information Service.
- {이시균·정재현·김수현·홍현균·정순기·이진명·이용호·

- 김재진·황규성·김은·홍성민·공정승·이혜연 (2017). <기술혁신을 반영한 중장기 인력수요 전망(2016-2030)>. 세종: 한국고용정보원.
- Lee, Z. & Nam, H. (2022). "Applications, AI-Related Management, and AI Application Risk." *Informatization Policy*, 29(2), 3-36.
- {이준기·남효경 (2022). 인공지능의 활용, 프로젝트 관리 그리고 활용 리스크에 대한 문헌 연구. <정보화정책>, 29권 2호, 3-36.}
- NIA (2022). *National Intelligent Informatization White Paper*. Seoul: National Information Society Agency.
- {한국지능정보사회진흥원 (2022). <국가지능정보화백서>. 서울: 한국지능정보사회진흥원}
- NIA (2023). *IT & Future Strategy report: NIA Forecasts 12 Digital Trends in 2023*. Seoul: National Information Society Agency.
- {한국지능정보사회진흥원 (2023). <IT & Future Strategy 보고서: NIA가 전망한 2023년 12대 디지털 트렌드>. 서울: 한국지능정보사회진흥원.}
- Oh, H., Zoo, H. & Choi, D. (2016). *Career Future and Human Resource Development Strategy*. Seoul: Korea Research Institute for Vocational Education & Training.
- {오호영·주회정·최대선 (2016). <직업의 미래와 인적자원 개발 전략>. 서울: 한국직업능력개발연구원.}
- OMB (2018). *Standard Occupational Classification Manual*, Washington, D.C.: Office of Management and Budget.
- Park, G., Cheon, Y., Hong, S. & Son, Y. (2016). *A Study on the Job Impact of Technological Change*. Sejong: Korea Employment Information Service.
- {박가열·천영민·홍성민·손양수 (2016). <기술변화에 따른 일자리 영향 연구>. 세종: 한국고용정보원.}
- Reimers, N. & Gurevych, I. (2019). "Sentence-bert: Sentence embeddings using siamese bert-networks." *arXiv preprint arXiv:1908.10084*.
- Seo, H. (2019). "A Preliminary Discussion on Policy Decision Making of AI in The Fourth Industrial Revolution." *Informatization Policy*, 26(3), 3-35.
- {서형준 (2019). 4차 산업혁명시대 인공지능 정책의사결정에 대한 탐색적 논의. <정보화정책>, 26권 3호, 3-35.}
- Statistics Korea (2017). *Korean Standard Classification of Occupations*, Seoul: Statistics Korea.
- {통계청 (2017). <한국표준직업분류>. 서울: 통계청.}
- Statistics Korea (2018). *Patent Classification System for the Seven Major Technologies Related to the Fourth Industrial Revolution*. Seoul: Statistics Korea.
- {통계청 (2018). <4차산업혁명 관련 7대 기술분야 특허분류 체계>. 서울: 통계청.}
- Webb, M. (2019). "The impact of artificial intelligence on the labor market." Available at SSRN 3482150.
- WEF (2019). *Towards a Reskilling Revolution Industry-Led Action for the Future of Work*. Geneva: World Economic Forum.
- WIPO (2019). *WIPO Technology Trends 2019: Artificial Intelligence*. Geneva: World Intellectual Property Organization.
- Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A. & Fidler, S. (2015). *Aligning books and movies: Towards story-like visual explanations by watching movies and reading books*. Proceedings of the IEEE international conference on computer vision, Dec 11-18.

**[부록 1] 인공지능 IPC 코드 (IPC code of Artificial Intelligence)**

IPC 서브클래스	제목	IPC 서브그룹
A63F	<ul style="list-style-type: none"> <li>카드게임, 보드게임 또는 톨렛게임</li> <li>작은 움직이는 물체를 사용하는 실내용 게임</li> <li>비디오 게임</li> <li>그 밖에 분류되지 않는 게임</li> </ul>	전체
A63H	<ul style="list-style-type: none"> <li>완구(예. 팽이, 인형, 후프 또는 블록)</li> </ul>	전체
B60W	<ul style="list-style-type: none"> <li>다른 종류 또는 가른 기능의 차량용 부품의 관련 제어</li> <li>하이브리드 차량에 특별히 적합한 제어 시스템</li> <li>특정의 단일의 부품의 제어에 관한 것은 아닌, 특정의 목적을 위한 도로상의 차량의 운전 제어 시스템</li> </ul>	B60W 30/14
G01C	<ul style="list-style-type: none"> <li>거리, 고저 또는 방위 측정, 측량, 항법</li> <li>자이로스코프 기구</li> <li>사진측량 또는 영상 측량</li> </ul>	G01C 21
G06F	<ul style="list-style-type: none"> <li>전기에 의한 디지털 데이터처리</li> </ul>	G06F 19/24, G06F 17/21, G06F 17/22, G06F 17/24, G06F 17/25, G06F 17/26, G06F 17/27, G06F 17/28, G06F 17/289
G06K	<ul style="list-style-type: none"> <li>그래픽 데이터의 판독</li> <li>데이터의 표현</li> <li>기록 매체</li> <li>기록 매체 처리</li> </ul>	G06K 9/6286, G06K 9/00
G06N	<ul style="list-style-type: none"> <li>특정 컴퓨터 모델에 기반한 컴퓨팅 장치</li> </ul>	G06N 3/082, G06N 3/084, G06N 3/086, G06N 3/088, G06N 5/043, G06N 5/045, G06N 5/046, G06N 5/048, G06N 99/005, G06N 3/02
G06Q	<ul style="list-style-type: none"> <li>관리, 상업, 재무, 관리 또는 감독 목적을 위해 특별히 적용된 정보 통신 기술</li> <li>달리 제공되지 않은 행정, 상업, 재무, 관리 또는 감독 목적에 특별히 적합한 시스템 또는 방법</li> </ul>	G06Q 40/00, G06Q 40/02, G06Q 40/04, G06Q 40/06, G06Q 40/08, G06Q 50/24, G06Q 50/04, G06Q 50/30, G06Q 50/26, G06Q 50/20, G06Q 40/12, G06Q 50/18, G06Q 40/12
G06T	<ul style="list-style-type: none"> <li>이미지 데이터 처리 또는 발생, 일반</li> </ul>	G06T 7/50
G08G	<ul style="list-style-type: none"> <li>측정치, 제어신호 또는 유사신호를 위한 전송방식</li> </ul>	전체
G10L	<ul style="list-style-type: none"> <li>음성분석 또는 합성</li> <li>음성 인식</li> <li>음성(speech) 또는 음성(voice) 처리</li> <li>음성(speech) 또는 오디오(audio) 부호화 또는 복호화</li> </ul>	G10L 15/00, G10L 15/01, G10L 15/02, G10L 15/04, G10L 15/05, G10L 15/06, G10L 15/065, G10L 15/07, G10L 15/08, G10L 15/10, G10L 15/12, G10L 15/14, G10L 15/16, G10L 15/18, G10L 15/183, G10L 15/187, G10L 15/19, G10L 15/193, G10L 15/197, G10L 15/20, G10L 15/22, G10L 15/24, G10L 15/25, G10L 15/26, G10L 15/28, G10L 15/30, G10L 15/32, G10L 15/34
G16H	<ul style="list-style-type: none"> <li>헬스케어 인포매틱스</li> <li>의료 또는 건강 관리 데이터의 취급 또는 처리에 특히 적합한 정보통신 기술</li> </ul>	전체

source: Statistics Korea (2018).