

연구보고서
2025-02

인공지능 시대의 숙련

장지연 · 심지환 · 정준호 · 이철승 · 전병유

목 차

요 약	i
제1장 서 론	(장지연) 1
제2장 숙련의 네트워크적 특성과 전이가능성 (Transferability)	(장지연 · 심지환) 3
제1절 서 론	3
1. 문제제기	3
2. 네트워크 분석	4
3. 숙련을 식별하고 분류하기 : 숙련사전	5
4. 본 장의 구성	6
제2절 숙련의 AI 노출도	7
1. 숙련의 AI 노출도 측정 방법	7
2. 측정 과정 및 결과 예시	10
3. 숙련군별 AI 노출도 점수	14
제3절 숙련의 네트워크적 특성	15
1. 기존 연구 검토 : 네트워크 분석을 활용한 숙련연구	15
2. 숙련 네트워크 분석 : IT 일자리를 중심으로	18
3. 네트워크 특성 지표와 AI 노출도의 상관관계	22
제4절 숙련의 직업 간 전이가능성(Transferability)	30
1. 직종 간 숙련 전이 확률 측정	30
2. 직업 간 숙련 전이가능성 측정 결과	33
제5절 요약과 결론	38

제3장 숙련-직업 간 네트워크의 ‘임계적’ 양극화

..... (정준호)	40
제1절 서론	40
제2절 선행 연구	42
제3절 실증 분석	44
1. 분석 방법	44
2. 자료	51
3. 분석 결과	53
제4절 소결	82

제4장 숙련 네트워크에서 브로커리지의 역할 : 노동시장 매개

구조 분석	(이철승)	86
제1절 연구 배경과 목적		86
제2절 이론적 배경		87
1. 브로커리지 이론		87
2. 브로커리지를 이용한 기존의 연구들		89
3. 중심성 지표들의 한계와 브로커리지의 독자적 기여		92
4. 브로커리지를 통한 경제 횡단의 전략적 가치와 보상 구조 (임금) 분석		93
5. Gould and Fernandez의 브로커리지 유형		94
6. Burt의 구조적 홀 이론과 제약 지수		96
제3절 자료 및 방법론		99
제4절 분석 결과		100
1. 기술통계 분석 결과		100
2. 회귀분석 결과		104
제5절 결론 : 연구 기여 및 정책 함의		110

제5장 AI 숙련 네트워크와 AI 숙련-역량 정합성 평가	
..... (전병유)	112
제1절 연구 배경과 목적	112
제2절 이론적 배경 및 기존 연구	114
1. 숙련의 네트워크적 구조	114
2. AI 확산과 숙련 수요 구조의 재편	116
3. 숙련 수요-공급 간 정합성 분석의 진화	118
제3절 자료와 방법론	119
1. AI 직업 숙련	119
2. AI 교과 역량	129
3. 숙련 네트워크 지표	137
4. 숙련 네트워크 지표의 정책적 의미	142
제4절 AI 숙련 네트워크 구조	147
제5절 AI 숙련과 AI 역량의 연계와 정합성	154
제6장 요약과 결론	(장지연) 172
참고문헌	181
부록	189

표 목 차

〈표 2- 1〉 한국고용직업분류 IT 일자리	18
〈표 3- 1〉 사전 숙련 유형 구분과 루뱅 방식 숙련 유형 구분 간 불일치	63
〈표 3- 2〉 숙련 유형이 로그 임금에 미친 효과 : 사전적 유형 구분의 경우	66
〈표 3- 3〉 숙련 유형이 로그 임금에 미친 효과 : 루뱅 유형 구분의 경우	68
〈표 3- 4〉 직업 간 연관성이 직업 간 로그 임금 차이에 미친 효과 ...	80
〈표 4- 1〉 중심성 지수들과 임금(로그) 간 관계에 대한 OLS 회귀 분석	105
〈표 4- 2〉 브로커리지(Brokerage) 지수들과 임금(로그) 간 관계에 대한 OLS 회귀분석	107
〈표 4- 3〉 구조적 공백(Structural Hole) 지수들과 임금(로그) 간 관계에 대한 OLS 회귀분석	109
〈표 5- 1〉 AI 숙련 분류별 빈도수 상위 10개 AI 키워드	125
〈표 5- 2〉 직업별 AI 숙련 키워드(상위 30개 직업)	128
〈표 5- 3〉 대학별 AI 강좌 분포	130
〈표 5- 4〉 대학별 AI 강좌 전공/교양 분포	134
〈표 5- 5〉 AI 교과 역량 분류와 상위 10개 역량 키워드(강좌당 복수 역량)	136
〈표 5- 6〉 숙련 네트워크 지표의 지표 특성과 정책적 함의	144
〈표 5- 7〉 AI 숙련의 지표별 상위 및 하위 10개	145

〈표 5- 8〉 파급 중심성과 맥락 중심성에 따른 숙련 분포와 그 의미 해석	153
〈표 5- 9〉 OJP와 대학 강좌에서의 숙련 빈도와 비율 비교	155
〈표 5-10〉 AI 네트워크 지표별 상위 10개 숙련 키워드의 AI 관련 강좌 반영 비율	157
〈표 5-11〉 AI 교육 강좌에서 상위 20개 숙련(역량에 매칭된 숙련) ...	159

그림목차

[그림 2- 1] 숙련의 AI 노출도 점수 분포	9
[그림 2- 2] 숙련군별 AI 노출도 점수	14
[그림 2- 3] IT 분야 숙련 네트워크(2024년)	20
[그림 2- 4] IT 분야 숙련 네트워크(2021년)	22
[그림 2- 5] 숙련군별 숙련 확산도	25
[그림 2- 6] 숙련군별 페이지랭크 중심성	28
[그림 2- 7] 직업 간 숙련 전이가능성 시각화	34
[그림 3- 1] 숙련 간 보완성 지수 분포	54
[그림 3- 2] 임계치(0.6) 기반 숙련 간 보완성 지수의 네트워크 구조	55
[그림 3- 3] 숙련 간 보완성 지수 변화에 따른 모듈성 지수(Q)의 변화	57
[그림 3- 4] 숙련 보완성 지수에 기반한 MST 네트워크 구조	58
[그림 3- 5] 숙련 유형 구분을 투영한 임계치(0.6) 기반 숙련 간 보완성 네트워크 구조	60
[그림 3- 6] 숙련 유형 구분을 투영한 숙련 간 보완성 MST 네트워크 구조	61
[그림 3- 7] 교육연수를 투영한 숙련 간 보완성 네트워크 구조	65
[그림 3- 8] 직업 간 연관성 지수 분포	69
[그림 3- 9] 직업 간 연관성 지수 변화에 따른 모듈성 지수(Q)의 변화	70
[그림 3-10] 직업 간 연관성 MST 네트워크 구조	72
[그림 3-11] 직업 유형 구분을 투영한 직업 간 연관성 MST 네트워크 구조	74

[그림 3-12] 숙련 유형 구분을 투영한 임계치(0.65) 기반 직업 간 연관성 네트워크 구조	76
[그림 3-13] 루뱅 방식의 인지 숙련 비중을 투영한 직업 간 연관성 네트워크 구조	78
[그림 4- 1] 숙련 네트워크 공동체의 브로커리지 구조	89
[그림 4- 2] Gould and Fernandez' Brokerage Types(Gould and Fernandez, 1989)	96
[그림 4- 3] 숙련 공동체에서 브로커리지 역량과 제약	99
[그림 4- 4] 산업 및 직종 간 브로커리지 히트맵	101
[그림 4- 5] 각 브로커리지 하위 유형의 출현 빈도	102
[그림 4- 6] 브로커리지 지수(대표자 + 문지기) 기준 상위 40개 숙련의 네트워크 시각화	103
[그림 4- 7] 각 브로커리지 유형의 OLS 회귀 계수 유의성 및 크기 비교	110
[그림 5- 1] AI 온라인 구인광고에서의 AI 숙련 키워드의 분포(OJPD당 복수 키워드 반영)	124
[그림 5- 2] AI 구인광고에서 나타나는 숙련 키워드 수	126
[그림 5- 3] 교과 역량 대분류별 강좌 수(강좌당 복수 역량 반영)	135
[그림 5- 4] AI 숙련 네트워크	148
[그림 5- 5] AI 숙련의 위계 구조(파급 중심성, LRC)와 교육 수준 ..	150
[그림 5- 6] 파급 중심성과 맥락 중심성에 따른 숙련 키워드의 분류 ..	153
[그림 5- 7] AI 직업 숙련과 AI 강좌 역량 간의 층적 다이어그램 (Alluvial Diagram)과 히트맵(Heatmap)	161
[그림 5- 8] AI 직업 숙련과 AI 강좌 역량 간의 층적 다이어그램과 히트맵 : 연결 중심성 가중 정합성	164
[그림 5- 9] AI 직업 숙련과 AI 강좌 역량 간의 층적 다이어그램과 히트맵 : 아이젠벡터 중심성 가중 정합성	166

[그림 5-10] AI 직업 숙련과 AI 강좌 역량 간의 총적 다이어그램과
히트맵 : 과급 중심성 가중 정합성 168

요약

본 보고서는 생성형 AI의 급속한 발전과 확산으로 전통적인 업무 방식이 근본적으로 변화하는 상황에서, AI 시대 노동시장에서 요구되는 숙련의 본질과 구조적 변화를 네트워크 분석 방법론을 통해 체계적으로 규명하고자 했다. 숙련을 개별적이고 고립된 능력 단위가 아닌 관계적·위계적 개념으로 이해할 때, 보다 풍부한 인사이트를 얻을 수 있다는 판단하에, 숙련 간 상호작용, 숙련-직업 간 관계, 그리고 노동시장 전체의 구조적 특성을 다층적으로 분석했다.

제2장에서는 한국숙련사전 6,558개 숙련에 대한 AI 노출도를 측정하고 네트워크 분석을 수행했다. ‘관리및지원’과 ‘정보기술’ 숙련군이 가장 높은 AI 노출도를 보인 반면, ‘개인돌봄서비스’와 ‘소프트스킬’은 가장 낮은 노출도를 나타냈다. 숙련 네트워크 분석에서는 2021년 ‘빅데이터 컴퓨팅’ 중심에서 2024년 ‘통합 보안 관리’ 중심으로의 전환이 관찰되어 데이터 중심 시대에서 보안 중심 시대로의 패러다임 전환을 확인했다. ‘고객및클라이언트지원’과 ‘관리및지원’ 숙련군이 네트워크 중심성에서 최상위를 차지했으며, 흥미롭게도 네트워크 중심성과 AI 노출도 간 강한 상관관계가 나타나지 않아 AI 대체 가능성이 높은 숙련이 현재 노동시장에서 광범위하게 요구되는 역설적 상황이 확인되었다. 직업 간 전이가능성 분석에서는 AI/ML 기술이 범용 기술로서 IT 영역을 넘어 금융, 제조, 전자 등 다양한 산업으로 확산되는 현상이 드러났다.

제3장에서는 KNOW 재직자조사 자료를 활용하여 숙련 보완성과 직업 간 연관성을 네트워크 관점에서 분석함으로써 한국 노동시장의 숙련 및 직업 양극화 구조를 탐색했다. 한국의 숙련 체계는 표면적으로 연속적이지만 내부적으로는 인지·사회 숙련과 신체 숙련으로 이원화된 ‘중심-교량-주변부’의 삼중 구조를 형성하고 있었다. 보완성 지수가 임계

치를 넘어서면 일부 숙련 클러스터가 결속되어 외부와 단절되는 ‘임계적 양극화’ 상태로 전이되는 현상이 관찰되었다. 397개 직업의 연관성 네트워크 분석 결과, 화이트칼라(고인지 숙련)와 블루칼라(중·저인지 숙련) 직업군이 분리된 구조가 나타났으며, 임계치 이상의 네트워크에서는 고숙련 직종의 중심 집중과 주변부 직군의 분리가 더욱 뚜렷해졌다. 임금 분석에서는 인지 숙련 비중이 양(+)의 효과를, 신체 숙련 비중이 음(-)의 효과를 보여 ‘인지 숙련 프리미엄-신체 숙련 디스카운트’ 구조가 확인되었다. 직업 간 연관성이 높을수록 경쟁이 강화되어 임금 격차가 확대되는 ‘연관성 속의 경쟁 구조’가 발견되었으나, 고숙련 내부에서는 유사성이 조정 기능을 수행하는 다층적 구조가 존재했다.

제4장에서는 IT 일자리 온라인 구인공고를 분석하여 브로커리지와 구조적 공백 개념을 통해 숙련의 전략적 위치를 측정했다. 엑셀, 파워포인트, 문서작성, 프로젝트 관리, 데이터 분석 등 범용성 높은 기술들이 가장 강한 브로커리지 역할을 수행하며 직종 간 연결을 주도했다. 브로커리지 유형 중 조정자, 대표자, 문지기 역할의 숙련이 통계적으로 유의미하게 높은 임금과 연관되었으며, 구조적 공백 분석에서는 제약 지수만이 유의미한 음의 효과를 보여 다른 브로커에 의해 통제되는 숙련이 낮은 보상을 받음을 확인했다. 이는 기존 중심성 지표가 포착하지 못한 산업 간 융합과 신기술 확산의 촉매로서 브로커 숙련의 중요성을 보여 준다.

제5장에서는 온라인 구인공고와 국내 대학 및 K-MOOC AI 교육과정 477개를 분석하여 숙련의 수요-공급 간 구조적 정합성을 평가했다. 산업 현장에서는 AI 프레임워크, AI/ML 모델, 데이터 분석 등 범용 기술 역량 수요가 높았고, 교육과정은 머신러닝, 딥러닝, 데이터 분석 중심의 기초 기술 교육 비중이 높았다. 그러나 생성형 AI, 강화학습, MLOps 등 최신 특화 기술은 교육과정에 거의 반영되지 않았다. 120개 숙련 키워드 중 47.5%만이 교육과정에 반영되었으며, 범용 기술은 높은 일치를 보인 반면 생성형 AI, 빅데이터 플랫폼, 책임 있는 AI 등은 반영이 미흡했다. 네트워크 기반 지표를 활용한 분석에서 Python, 데이터 마이닝, 생성형

AI, 지능형 인프라 등이 중심 허브 역할을 하며 통합적 교육 설계의 필요성을 보여주었다. 중심성을 고려한 가중 적합성 분석에서는 AI 프레임워크, 자율 시스템, 책임 있는 AI 등이 전략적 중요 숙련으로 확인되었다.

본 연구는 AI 시대 숙련 구조 변화를 네트워크 분석을 통해 다층적으로 조명함으로써, 숙련을 관계적·구조적 개념으로 재정의했다. 인지 숙련 중심의 위계적 결속과 신체 숙련의 주변화라는 이중 구조, 범용 기술의 브로커리지 역할과 임금에 미치는 차별적 효과, 그리고 교육-노동 시장 간 구조적 불일치를 가시화함으로써, AI 기술 확산이 초래하는 숙련 재편이 노동시장 전체의 네트워크 위상과 연결 구조를 근본적으로 변화시키는 과정임을 밝혔다. 이는 향후 인력 개발 정책이 개별 숙련 육성을 넘어 숙련 간 연결성의 다양성 유지, 브로커 숙련에 대한 전략적 투자, 교육과정의 구조적 재설계를 통해 포용적이고 균형 잡힌 숙련 생태계를 조성해야 함을 시사한다.

제 1 장 서 론

인공지능이 사람이 하던 과업을 대체하는 시대를 맞이하여 가장 많은 사람들이 궁금해하는 질문은 ‘내 일자리는 어떻게 될까’이다. 그런데 이것은 사실 엄밀하게 들여다보면, ‘내가 가진 숙련(skill)을 계속 활용할 수 있을까’라는 질문으로 귀결된다. 생성형 AI의 급속한 발전과 확산으로 인해 전통적인 업무 방식이 근본적으로 변화하고 있는 상황에서, AI 시대에 일터에서 요구되는 숙련은 무엇인지, 나는 그리고 내 자녀는 무엇을 배워야 하는지에 대한 질문들이 대두되고 있다. 단순히 기술적 변화에 대한 우려를 넘어서, 미래 노동시장에서 경쟁력을 유지하기 위해 필요한 핵심 역량이 무엇인지에 대한 체계적인 분석과 전망이 절실히 요구되는 시점이다. 본 연구는 이러한 시대적 질문과 사회적 관심에서 출발하여, AI 시대 숙련의 변화 양상과 미래 인재 개발 방향을 모색하고자 한다.

본 연구는 2024년 발표된 ‘인공지능 발전의 고용효과’ 연구의 후속편이라고 할 수 있다. 이전 연구는 인공지능이 사람이 하던 일 중에서 어떤 과업을 대신하고, 따라서 어떤 일자리가 사라지게 될지에 대해 분석하였다. 여기서 발견한 중요한 함의는 일자리 자체의 소멸 못지않게 중요한 것이 일하는 방식의 변화라는 점이었다. 이에 본 연구는 과업(task)이나 일자리(job)에 미치는 영향 대신에 숙련(skill)에 미치는 영향을 고민하고자 한다. 일자리가 새로 생기거나 소멸되거나, 혹은 일하는 방식이 바뀌거나, 모든 경우에 우리가 주목해야 할 것은 ‘우리에게 요구되는 숙련은 어떻게 변화할 것인가’이기

때문이다.

OECD가 제시하는 바에 따르면, 숙련은 인적자본을 나타내는 포괄적 지표이기도 하고, 동시에 데이터와 사물을 조작하는 특정한 전문능력의 집합을 의미하는 용어로 사용된다. 인적자본의 포괄적 지표로서 숙련은 인지 능력, 신체 능력, 사회정서적 능력, 메타인지 능력(예: 정보처리기술, 손재주, 팀워크, 자기조직화)과 같은 광범위한 능력 집합뿐만 아니라 특정 직무나 과업을 수행하는 능력(예: 회계나 헤어컬러링)을 포함한다. 이처럼 숙련은 단순한 기술적 능력을 넘어서 개인의 종합적인 역량과 잠재력을 아우르는 다층적 개념으로 이해할 수 있다(OECD, 2023).

본 보고서는 숙련 변화를 연구하는 방법으로 네트워크 분석을 선택하였다. 숙련을 개별적으로 고립된 능력 단위로 간주하는 것보다, 관계적 개념이나 위계적 개념으로 이해할 때 보다 풍부한 인사이트를 얻을 수 있을 것으로 판단했기 때문이다. 네트워크 분석(Network Analysis)은 개체 간의 관계적 특성을 분석하는 데 적합한 기법으로, 복잡한 관계와 상호작용을 가진 데이터를 효과적으로 분석하고 시각화할 수 있다는 장점을 가진다. 이러한 분석을 통해 개별 숙련의 중요성을 파악함과 동시에, 숙련 간 상호작용과 숙련-직업 간 관계를 체계적으로 시각화하고자 한다. 이는 AI 시대 숙련의 변화 양상을 보다 입체적으로 이해하고, 미래 인재 개발 전략을 수립하는 데 필요한 핵심적 통찰을 제공할 것으로 기대된다.

제 2 장

숙련의 네트워크적 특성과 전이가능성(Transferability)

제1절 서론

1. 문제제기

AI 기술의 급속한 발전이 노동시장과 일터에 미치는 영향에 대한 논의는 그동안 주로 사람이 수행하던 과업의 일부를 기계가 대체한다는 관점에서 이루어져 왔다. 그러나 AI의 영향은 단순한 과업 대체를 넘어서 일하는 방식의 근본적인 변화, 나아가 노동과정과 작업조직의 전면적인 재편을 수반한다는 점에 주목해야 한다. 이러한 변화의 결과는 특정 직업의 소멸과 새로운 직업의 창출로 나타날 수도 있지만, 동시에 동일한 과업을 수행하는 과정에서 인간이 AI를 활용하여 생산성을 향상시키는 증강(augmentation) 형태로 전개될 가능성 또한 크다.

이러한 논의는 숙련(skill)의 관점에서 보다 구체적으로 접근할 수 있다. AI 기술의 발전은 필요한 숙련의 변화, 즉 숙련 수요의 구조적 전환을 의미한다. AI 기술이 고도화됨에 따라 점점 더 중요해지는 숙련이 있는 반면, 더 이상 필요하지 않게 되는 숙련도 존재한다. 이는 우리가 현재 보유하고 있는 숙련과 미래에 요구되는 숙련 사이의 괴리를 야기할 수 있으며, 이러한 불일치는 개별 근로자뿐만 아니라 조직과 사회 전체의 생산성에 중대한 영

향을 미칠 수 있다. 현재의 숙련 분포에 대한 체계적인 분석과 이해를 기반으로 미래 숙련 분포에 대한 전망이 요구된다.

숙련에 대한 이해는 다음과 같은 정책적 함의로 이어진다. 근로자 개인은 변화하는 환경에 적응하여 업스킬링(upskilling)과 리스킬링(reskilling)을 수행해야 하는데, 이를 위해서는 무엇보다도 미래에 필요한 숙련 수요가 무엇인지에 대해 정확히 파악하는 것에서 출발할 필요가 있다. 정책적으로는 기술 발전으로 인한 구조적 실업을 최소화하면서 동시에 원활한 노동 이동을 가능하게 하는 방법을 모색해야 한다. 이는 단순히 개별 근로자의 적응 노력에만 의존할 것이 아니라, 숙련의 구조와 변화 양상에 대한 깊이 있는 이해를 바탕으로 한 체계적이고 선제적인 접근을 요구한다. 본 연구가 이러한 정책적 과제 해결에 기여할 수 있기를 기대한다.

2. 네트워크 분석

본 연구는 숙련 구조가 ‘네트워크’ 형태를 띠고 있다는 가정에서 출발한다. 개별 숙련들은 서로 고립된 채로 존재하는 것이 아니라 복잡한 연결 관계를 형성하고 있으며, 이러한 연결 구조는 중심성이나 밀도 등의 네트워크 특성으로 파악할 수 있다. 또한 네트워크 내에서 밀접하게 연결된 하위 그룹을 식별할 수 있어, 숙련 A는 숙련 A'와 긴밀하게 연결되어 있고 숙련 B와는 상대적으로 거리가 멀다는 관계적 특성을 구체적으로 분석할 수 있다.

이러한 네트워크 관점을 통해 특정 분야에서 핵심적인 역할을 하는 숙련을 식별할 수 있으며, 관련 숙련들의 묶음(skill bundle)을 시각화할 수 있다. 동시에 다양한 분야에서 두루 필요로 하는 범용적 숙련도 구별해낼 수 있다. 특히 일터에서 생산 공정이나 서비스 과정에 AI가 투입되는 상황에서 숙련 네트워크는 기존과는 다른 양상을 보일 것으로 예상된다. 예컨대, STEM 숙련, 소프트 스킬(soft skill), 사회적 숙련(social skill) 등 다양한 숙련 집합들은 AI 숙련과의 연계성에 따라 네트워크 내에서의 중요도와 위치가 근본적으로 달라질 것이다.

숙련들은 직업을 매개로 하여 서로 연결된다. 특정 직업에서 숙련 A, 숙련 B, 숙련 C가 동시에 요구된다면, 이들 숙련은 해당 직업을 통해 서로 연결되

어 있다고 볼 수 있다. 이러한 관점에서 숙련 네트워크뿐만 아니라 숙련-직업 네트워크의 구조를 파악할 수 있으며, 이는 숙련이 실제 노동시장에서 어떻게 활용되고 있는지를 이해하는 데 중요한 단서를 제공한다. 직업과의 연계성을 기반으로 한 숙련-직업 네트워크 분석은 개별 숙련의 경제적 가치와 숙련이 노동 이동에 미치는 영향을 살펴보는 데 유용한 분석틀을 제공한다.

숙련 네트워크 구조 분석은 여러 차원에서 중요한 의미를 갖는다. 첫째, 숙련 네트워크 내에서 개별 숙련의 상대적 중요성을 객관적으로 파악할 수 있다. 둘째, 숙련 간의 유사성과 상호보완성, 그리고 개별 숙련의 확산 속도 등을 정량적으로 측정하고 비교할 수 있다. 셋째, 복잡한 숙련 간 상호작용과 숙련-직업 간 관계를 직관적으로 시각화함으로써 숙련 생태계의 전체적 구조를 이해할 수 있도록 돕는다.

3. 숙련을 식별하고 분류하기 : 숙련사전

숙련 네트워크 분석을 위해 맨 먼저 해야 할 일은 숙련을 식별(identify)하는 작업이다. 이는 개별 숙련에 명확한 명칭을 부여하고 체계적으로 분류하는 과정으로, 모든 후속 연구를 위한 사전적 준비 단계에 해당한다. 숙련의 정의와 분류가 없이는 숙련 간 관계나 네트워크 구조를 분석할 수 없기 때문이다.

숙련(skill)은 직업에서 요구되는 과업을 수행할 수 있는 능력으로 정의된다. 전통적으로 직업 숙련에 관한 정보와 분류체계는 전문가들의 직업 조사에 기반해 왔다. 가장 널리 알려진 숙련분류체계는 미국 노동통계국(BLS)의 O*NET이며, 유럽연합의 ESCO(European Skills, Competences, and Occupations)도 개념적 숙련분류체계의 대표적 사례이다. 그러나 전문가 의존적인 숙련 정보와 분류체계는 AI와 같이 급변하는 기술 환경에서 시의성 있는 숙련 수요-공급 정보를 제공하는 데 근본적인 한계가 있다. 전문가 조사와 분류 작업에는 상당한 시간이 소요되며, 새로운 기술과 숙련이 빠르게 등장하는 현실을 따라잡기 어렵기 때문이다.

이에 반해 온라인 구인공고(Online Job Postings, OJPs)를 활용한 숙련분

류체계는 실시간으로 변화하는 노동시장의 숙련 수요를 세밀하게 포착할 수 있는 새로운 가능성을 제시한다. 구인공고는 기업들이 실제로 필요로 하는 숙련을 직접적으로 표현하는 자료이므로, 이를 체계적으로 분석하면 현실적이고 시의성 있는 숙련분류체계를 구축할 수 있다. 미국의 LightCast(구 Burning Glass)는 실제 온라인 구인공고 분석을 통해 구축한 경험적 숙련분류체계로 연구에도 널리 활용되고 있다.

우리나라에는 개념적 숙련분류체계나 경험적 숙련분류체계가 모두 부재하였던바, 2024년에 한국노동연구원 연구팀이 미국의 LightCast사의 방식을 벤치마킹하여 우리나라 온라인 구인공고에 등장하는 숙련을 6,500여 개의 숙련명으로 목록화하였다(장지연 외, 2024). 이 과정은 다음과 같은 단계를 거쳤다. 첫째, 구인공고 샘플링 및 숙련 용어 추출 단계에서는 업종 분포를 고려한 대표성 있는 샘플을 확보하고 API를 활용하여 숙련 관련 용어를 추출하였다. 둘째, 숙련용어 정제 단계에서는 중복 단어를 제거하고 정규식을 활용하여 불용어, 공백, 특수문자를 체계적으로 정리하였다. 셋째, 숙련용어 통합 및 분류 단계에서는 LLM을 활용한 초별 분류와 소프트 스킬 식별을 위한 질적 코딩을 수행하였으며, 숙련 용어마다 정의를 생성하여 벡터화한 후 군집분석을 통해 클러스터를 생성하고, 이 정보를 활용한 퍼지 매칭으로 용어를 정제하였다.

이러한 과정을 통해 구축된 한국 숙련사전은 대분류 29개, 중분류 116개, 소분류 445개로 구성된 포괄적인 숙련분류체계를 제시한다. 특히 숙련 단어의 다양성이 높은 IT 정보기술 분야에서는 80개의 소분류를 적용하여 2,250건의 숙련을 더욱 세밀하게 분류하였다. 또한 전통적 의미의 기술적 숙련뿐만 아니라 역량 프레임워크에 기반하여 337건의 소프트 숙련을 체계적으로 포착함으로써, 현대 노동시장에서 점차 중요해지고 있는 비기술적 숙련까지 포괄하는 종합적인 분류체계를 완성하였다(부표 2-1 참고).

4. 본 장의 구성

본 장은 다음과 같이 구성된다. 제2절에서는 이렇게 식별된 각각의 숙련에 대해서 AI 노출도를 계산한다. 이는 개별 숙련이 얼마나 AI로 대체 가능

한지를 추론해 보는 작업으로, 어떤 특징을 갖는 숙련이 AI 노출도가 높은지 분석하기 위한 준비 작업이다. AI 노출도 측정을 통해 기술 변화가 각 숙련에 미치는 영향의 정도를 정량적으로 파악할 수 있으며, 이는 향후 숙련 수요 변화를 예측하는 기초 자료가 된다.

제3절에서는 숙련의 특징을 포착하기 위해 네트워크 분석을 채택한다. 이는 개별 숙련을 고립된 요소로 보지 않고 숙련 간 관계적 특성과 위계적 특성으로 이해하고자 하는 접근이다. 네트워크 분석에 기반하여 중요한 숙련을 식별하고 이를 시각화함으로써 숙련 생태계의 전체적 구조를 파악할 수 있다. 또한 네트워크상에서의 주요 지표들과 AI 노출도 간의 관계를 분석하여 어떤 네트워크적 특성을 가진 숙련이 AI의 영향을 더 많이 받는지 탐구한다.

제4절에서는 네트워크 분석을 활용하여 직업 간 숙련 전이가능성(Skill Transferability)을 측정한다. 특정 숙련을 기반으로 어떤 다른 직업으로 이동할 수 있을지를 분석하는 것이다. 개별 숙련은 하나의 직업에서만 유용한 것이 아니라 여러 직업에서 공통적으로 활용될 수 있으며, 이러한 숙련의 범용성을 바탕으로 한 직종 간 노동 이동 가능성을 체계적으로 분석한다. 이는 노동시장의 유연성과 근로자들의 적응력을 이해하는 데 중요한 통찰을 제공할 것이다.

제2절 숙련의 AI 노출도

1. 숙련의 AI 노출도 측정 방법

작년 보고서(장지연 외, 2024)에서는 직업을 과업(tasks)의 합으로 정의하고 AI의 대체 가능성을 논의하였다. 이러한 접근법의 이론적 기반은 Eloundou et al.(2023)의 연구에서 제시되었는데, 이들은 대규모 언어 모델(LLM)을 활용하여 노출도를 측정할 수 있음을 실증적으로 보여주었다. 구체적으로 대규모 언어 모델인 ChatGPT-4를 활용하여 개별 직무의 수행 시간

을 AI가 얼마나 단축시킬 수 있는지를 평가하여 AI 노출도를 측정하였다. 이들은 O*NET이 제시하는 특정 작업(tasks) 및 세부 작업 활동(detailed work activities)을 ChatGPT-4에 제시하고, 해당 작업을 생성형 AI가 수행할 경우 소요 시간이 50% 이상 단축되는지, 또는 결과물의 품질이 저하되는지 등의 질문을 통해 대체 가능성을 평가하였다. 특히 사람의 평가와도 비교하여 그 유사성을 확인함으로써 AI 기반 평가의 신뢰성을 검증하였다. 이후 Gmyrek et al.(2023)을 비롯하여 여러 연구에서 이러한 방법론이 활용되고 있다.

본 연구에서는 유사한 방법론을 활용하되, 작업(task) 단위가 아닌 숙련(skill) 단위를 기준으로 노출도를 측정하고자 한다. 이는 숙련이 여러 직업에 걸쳐 전이가능한 특성을 가지고 있어, 숙련 단위의 분석이 노동시장 전반의 변화를 이해하는 데 더욱 유용한 통찰을 제공할 수 있기 때문이다. 또한 숙련 중심의 접근은 개별 근로자의 역량 개발과 재교육 방향을 설정하는 데 있어서도 더욱 실용적인 가이드라인을 제시할 수 있다.

본 연구에서는 한국 숙련사전에 등재된 전체 숙련에 대해 AI 기술 도입 시 예상되는 작업시간 절감 효과를 정량적으로 분석하였다. 이를 위해 OpenAI의 GPT-4를 기반으로 전문가 패널 시뮬레이션 모델인 [SkillAIAssessment] 클래스를 개발하여 체계적인 평가를 수행하였다. 이 접근법은 단순한 일회성 평가를 넘어서 전문가 간 토론과 의견 수렴 과정을 시뮬레이션함으로써 보다 신뢰성 있는 결과를 도출하고자 하였다 (SkillAIAssessment 코드는 부표 2-2 참고).

이 시뮬레이션에서는 각 숙련에 대해 5명의 가상 전문가로 구성된 패널이 3라운드에 걸친 델파이 방식의 합의 과정을 통해 AI 기반 시간 절감률을 산정하였다. 델파이 기법은 전문가들의 의견을 반복적으로 수집하고 피드백을 제공하여 합의에 도달하는 구조화된 의사소통 기법으로, 복잡하고 불확실한 문제에 대한 전문가 판단을 체계적으로 집약하는 데 효과적이다.

0~100점의 점수 척도를 사용하여 AI 노출도를 다음과 같이 구분하였다.

- 0~20점 : AI 적용이 어려운 영역으로 제한적 효과만 기대할 수 있는 숙련
- 20~50점 : 부분적 AI 보조가 가능한 영역으로, AI가 보완적 역할을 수행할 수 있는 숙련
- 50~80점 : 대폭적인 AI 지원 효과를 기대할 수 있는 영역으로, AI 도입

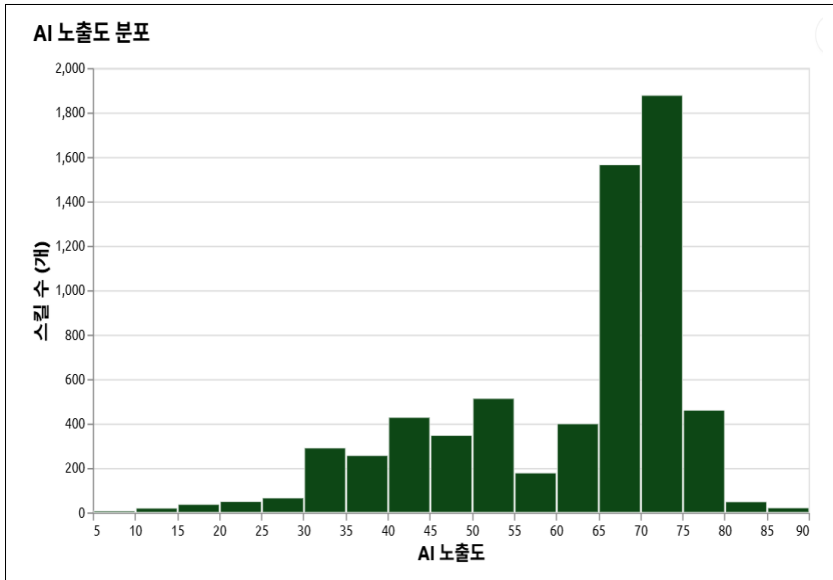
을 통해 상당한 효율성 향상이 가능한 숙련

- 80~100점 : 전면적 자동화가 가능한 영역으로, AI가 인간의 역할을 대부분 대체할 수 있는 숙련

각 라운드에서 전문가들은 이전 라운드의 의견을 참고하여 점수를 조정하며, 최종 3라운드의 평균값을 합의 점수로 산정하였다. 이러한 반복적 과정을 통해 극단적 의견들이 조정되고, 보다 균형 잡힌 전문가 판단이 도출될 수 있었다. 또한 각 전문가는 점수와 함께 200자 이내의 상세한 코멘트를 제공하여 평가 근거를 명확히 하고, 다른 전문가들이 이를 참고할 수 있도록 하였다.

위와 같은 방법으로 6,558개 숙련에 대하여 AI 노출도를 측정하였으며, 측정된 AI 노출도는 최대 89점, 최소 7점으로 나타났다. 전체 숙련에 대한 AI 노출도 값의 분포는 [그림 2-1]과 같다.

[그림 2-1] 숙련의 AI 노출도 점수 분포



자료 : 저자 작성.

2. 측정 과정 및 결과 예시

가. 노출도 상위 숙련 예시

AI 기술을 활용했을 때 작업시간이 크게 단축될 것으로 예상되는 상위 10개 숙련을 분석한 결과는 대부분 반복적이고 규칙성이 높으며, 명확한 절차와 기준에 따라 수행되는 업무들로 구성되어 있어 AI 시스템으로의 자동화가 용이한 특성을 보인다.

가장 높은 AI 노출도를 보이는 숙련은 “빌드 및 테스트 및 배포 자동화하기”로 89%의 시간 절감 효과가 예상된다. 정보기술 분야에서는 “순환 중복 검사(CRC) 수행하기”(89%), “Prettier 코드 포맷터 활용하기”(87%), “로그아웃 기능 구현하기”(86%) 등이 상위권에 위치하고 있다. 이들은 모두 명확한 규칙과 알고리즘에 기반하여 수행되는 업무로, AI가 인간보다 더 빠르고 정확하게 처리할 수 있는 영역들이다.

물류 및 사무 숙련에서도 높은 자동화 가능성을 보인다. “입출고 전산 입력하기”(88%), “송장 출력하기”(87%), “팩스 송수신 업무 처리하기”(86%), “전산 처리 업무 수행하기”(86%) 등은 정해진 양식과 규칙에 따라 정보를 처리하는 반복적 업무들이다. 제조업에서는 “자동 광학 검사 수행하기”(88%)가 높은 순위에 위치하며, 이는 AI의 컴퓨터 비전 기술이 인간의 시각적 판단을 대체할 수 있는 대표적인 사례이다. 또한 “전자 서명 업무 관리하기”(87%)와 같은 법무 및 컴플라이언스 분야의 업무도 높은 자동화 가능성을 보이는 것으로 나타났다.

나. 노출도 상위 숙련의 점수 산정 과정

AI 노출도 점수는 5명의 가상 전문가가 3라운드에 걸친 델파이 방식의 합의 과정을 통해 도출되었다. 각 라운드에서 전문가들은 이전 라운드의 의견을 참고하여 점수를 조정하며, 최종 3라운드의 평균값을 합의 점수로 산정하는 방식으로 진행되었다.

예를 들어 “입출고 전산 입력하기” 숙련의 경우, 1라운드에서 전문가들은

75~90점의 폭넓은 범위의 점수를 제시하였다. 전문가1은 85점을 주며 “반복적이고 규칙적인 성격이 강해 자동화가 매우 용이하다”고 평가했고, 전문가4는 90점으로 “현재 기술로 완벽하게 자동화할 수 있으며, 거의 모든 입력 과정을 AI가 처리할 수 있다”고 의견을 제시했다. 반면 전문가3은 75점으로 “자동화 도구가 대부분의 일을 대신할 수 있지만, 간혹 예외 사항 처리가 필요할 수 있다”며 상대적으로 보수적인 평가를 하였다.

2라운드에서는 전문가들이 서로의 의견을 반영하여 점수를 조정하는 과정을 거쳤다. 전문가1은 85점에서 88점으로 상향 조정하며 “동료의 의견을 듣고 보니, 예외적인 경우를 제외하고는 거의 완벽하게 자동화 가능하다”고 설명했다. 전문가3도 75점에서 80점으로 점수를 높이며 “예외 사항 처리의 필요성은 인정하지만, 그 비중이 크지 않다는 점을 감안하여 점수를 상향 조정한다”고 밝혔다.

최종 3라운드에서는 전문가들의 의견이 85~90점 범위로 수렴하였다. 전문가1은 90점으로 “토론을 통해 거의 모든 경우의 자동화가 가능하다는 결론에 도달했다”고 평가했으며, 전문가3도 85점으로 “대부분의 의견이 자동화의 높은 가능성을 지지하므로, 저도 이에 동의하여 점수를 높인다”고 하였다. 최종적으로 3라운드 평균인 88점이 합의 점수로 도출되었다.

“자동 광학 검사 수행하기”의 경우에도 유사한 패턴을 보였다. 1라운드에서 80~90점의 범위였던 점수가 라운드를 거듭할수록 86~90점으로 수렴하면서 최종 88점의 합의 점수를 얻었다. 전문가들은 “AI의 정밀도와 반복성을 고려할 때 거의 완벽한 자동화가 가능하다”는 데 의견을 모았다.

이러한 델파이 방식의 합의 과정을 통해 극단적 의견들이 조정되고, 보다 균형 잡힌 전문가 판단이 도출된다. 특히 각 전문가가 점수와 함께 200자 이내의 상세한 근거를 제시하게 함으로써 평가의 신뢰성을 확보하고자 하였다.

다. 노출도 하위 숙련 예시

AI 기술로 자동화하기 어려운 하위 10개 숙련을 분석한 결과, 이들은 모두 10% 이하의 매우 낮은 시간 절감 효과를 보이는 것으로 나타났다. 이러

한 숙련들은 인간 고유의 감정, 창의성, 신체적 경험, 그리고 복잡한 인간관계가 핵심적인 역할을 하는 영역들로 구성되어 있다.

가장 낮은 AI 노출도를 보이는 숙련들 중 “수영 강습하기”(7%)와 “첼로 연주하기”(7%)는 각각 신체적 기술 전수와 예술적 표현이라는 인간 고유 영역을 대표한다. 수영 강습의 경우 “AI는 수영 강습의 기술 분석에서 일정 부분 도움을 줄 수 있으나, 전체적인 자동화는 어렵다”고 평가되었으며, 첼로 연주하는 “현재 기술로는 감성적 요소를 재현하기 어렵다”는 판단을 받았다. “재즈 연주하기”(10%) 역시 “창의성과 즉흥성을 완벽히 재현하기 어렵다”는 평가를 받아 예술 분야에서 인간 고유의 창조적 능력이 여전히 중요함을 보여준다.

특히 주목할 점은 소프트 스킬 영역에서 낮은 AI 노출도 숙련들이 대부분 집중되어 있다는 것이다. “자기 성찰 및 만족도 관리하기”(7%), “사명감 가지기”(8%), “자기 동기부여하기”(8%), “인내심 발휘하기”(8%) 등은 모두 개인의 내적 역량과 관련된 숙련들이다. 이들은 “매우 개인적이고 주관적인 과정이므로 AI 자동화는 매우 제한적”이라는 공통된 평가를 받았다.

“강인한 정신력 발휘하기”(9%)와 “장기적으로 헌신하기”(10%)의 경우, 전문가들은 “개인의 내적 요소로 AI의 자동화 가능성은 매우 제한적”이며 “주로 인간의 독특한 감정과 가치에 기반한다”고 평가하였다. “자기 성찰하기”(10%)도 “깊은 인간의 이해가 필요하나, AI는 정보 제공에 한 해 소극적 역할만 가능”하다는 분석을 받았다.

이러한 결과는 AI 기술이 아무리 발전하더라도 인간의 감정적 지능, 자아 인식, 창의적 표현, 그리고 신체적 기술 전수와 같은 영역에서는 여전히 인간 고유의 역할이 중요함을 시사한다. 특히 이들 숙련의 공통점은 주관적 경험, 감정적 교감, 창조적 사고, 그리고 복잡한 인간관계에서의 판단이 핵심 요소라는 점이다. 따라서 이러한 영역에서는 AI가 보조적 역할을 하더라도 인간 중심의 접근이 계속 유지될 것으로 예상된다.

라. 노출도 하위 숙련의 점수 산정 과정

AI 노출도가 낮은 숙련들의 점수 산정 과정을 살펴보면, 높은 노출도 숙

련들과는 정반대의 패턴을 보인다. 전문가들은 초기부터 매우 낮은 점수를 제시했으며, 라운드를 거듭할수록 오히려 점수가 하향 조정되거나 낮은 수준에서 유지되는 경향을 보였다.

“수영 강습하기”의 경우 1라운드에서 전문가들은 5~10점의 낮은 점수를 제시했다. 전문가1은 5점을 주며 “수영 강습은 신체 조정 능력과 개인별 맞춤 교육이 중요하여 AI 자동화가 어렵다”고 평가했고, 전문가2는 10점으로 “AI로 수영 동작 판단이 가능하지만, 직접적인 교정과 안전 관리는 인간의 개입이 필수적”이라고 의견을 제시했다. 2라운드에서 전문가1은 5점에서 6점으로 소폭 상승시키며 “AI가 수영 기술의 일부 분석을 도울 수 있음을 인정한다”고 했지만, 여전히 매우 낮은 수준을 유지했다. 최종 3라운드에서는 모든 전문가가 7~8점으로 수렴하여 평균 7점의 합의 점수를 얻었다.

“인내심 발휘하기”는 더욱 극명한 패턴을 보였다. 1라운드에서 전문가4는 5점으로 “AI는 인내심과 같은 인간의 감정적 특성을 자동화할 수 있는 수준에 이르지 못했다”고 평가했으며, 가장 높은 점수를 준 전문가3도 20점에 그쳤다. 2라운드에서는 전문가들의 점수가 대부분 하향 조정되었는데, 전문가3은 20점에서 15점으로, 전문가2는 15점에서 10점으로 낮췄다. 전문가1은 “동료의 의견을 반영하여 조정하였으나 여전히 자동화는 어렵다”며 10점에서 8점으로 더욱 낮췄다. 최종 3라운드에서는 7~10점 범위로 수렴하여 평균 8점의 합의에 도달했다.

“장기적으로 헌신하기”의 경우도 유사한 양상을 보였다. 1라운드에서 전문가4는 5점으로 “장기적 헌신은 인간만의 고유한 특성이며 AI로는 접근하기 어렵다”고 평가했고, 가장 높은 점수를 제시한 전문가2도 20점에 머물렀다. 2라운드에서 전문가2는 20점에서 18점으로, 전문가3은 15점에서 12점으로 하향 조정하며 “AI의 보조 사용은 가능하지만 본질적인 이해나 대체는 어렵다”고 밝혔다. 최종 3라운드에서는 6~15점으로 수렴하여 평균 10점의 합의 점수를 얻었다.

이러한 낮은 AI 노출도 숙련들의 공통점은 전문가들이 라운드를 거듭할수록 AI의 한계를 더욱 명확히 인식하게 되어 점수를 낮추거나 낮은 수준에서 유지했다는 것이다. 특히 “인간의 감정적 특성”, “개인의 동기와 가치관”, “신체적 안전 관리” 등이 핵심 제약 요인으로 지속적으로 언급되었으며, AI

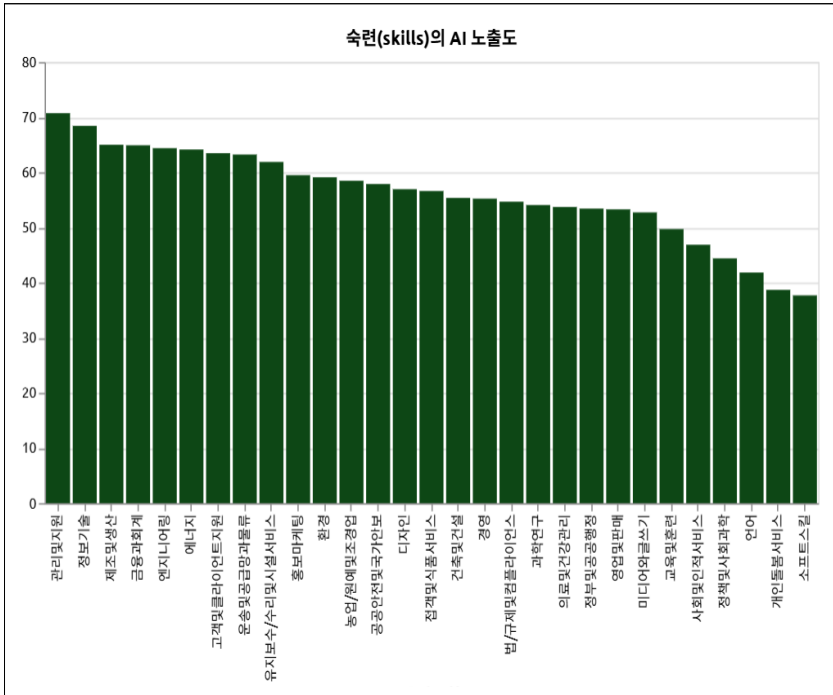
의 역할은 “보조적 지원”이나 “부분적 분석”에 국한될 것이라는 데 전문가들의 의견이 모아졌다.

3. 숙련군별 AI 노출도 점수

[그림 2-2]는 숙련군별 AI 노출도의 평균값을 보여준다. 관리및지원, 정보기술 숙련의 경우 높은 AI 노출도를 나타내며, 인적서비스나 돌봄서비스, 소프트스킬 등에서 낮은 AI 노출도를 보인다.

‘관리및지원’ 숙련의 경우 총 58개의 세부 숙련으로 구성되어 있으며, ‘Excel 및 PowerPoint 고급 활용하기’, ‘Outlook 일정 및 메일 관리하기’, ‘회의록 작성 및 배포하기’, ‘포장명세서 작성 및 관리하기’ 등과 같이 문서 작성, 데이터 처리, 일정 관리 등 디지털 도구를 활용한 업무가 주를 이룬다.

[그림 2-2] 숙련군별 AI 노출도 점수



자료 : 저자 작성.

이러한 숙련들은 생성형 AI의 문서 자동화, 데이터 분석, 업무 효율화 기능과 직접적으로 연관되어 높은 AI 노출도를 나타낸다.

‘정보기술 숙련’은 2,075개로 가장 많은 세부 숙련을 포함하고 있으며, ‘빅데이터 분석하기’, ‘Word2Vec 모델 활용하기’, ‘AIOps(인공지능 IT 운영) 구현하기’, ‘사용자 스토리 작성하기’ 등 AI 기술 자체를 다루는 숙련들이 다수 포함되어 있다. 특히 머신러닝, 데이터 분석, 시스템 자동화 관련 숙련들이 AI 기술의 핵심 영역과 직접 겹치면서 최고 수준의 AI 노출도를 보인다.

반면 ‘개인돌봄서비스 숙련’은 45개의 숙련으로 구성되어 있으며, ‘보육하기’, ‘헤어 펌 시술하기’, ‘미용 서비스 제공하기’, ‘전문 코칭 서비스 제공하기’ 등과 같이 직접적인 인간 접촉과 감정적 교감, 섬세한 손기술이 요구되는 업무들이 주를 이룬다. 이러한 특성으로 인해 AI가 대체하기 어려운 영역으로 분류되어 낮은 AI 노출도를 나타낸다.

‘소프트스킬’의 경우 194개 숙련으로 구성되어 있으며, ‘대인관계 능력 발휘하기’, ‘핵심 파악하기’, ‘도전 정신 발휘하기’, ‘아이디어 확산 능력 발휘하기’ 등과 같이 인간 고유의 창의성, 감성, 사회적 상호작용 능력을 다루는 영역이다. 이러한 숙련들은 인간의 정서적 지능과 사회적 맥락 이해가 핵심이 되어 현재 AI 기술로는 완전히 구현하기 어려운 특성을 지니므로 가장 낮은 AI 노출도를 보인다.

제3절 숙련의 네트워크적 특성

1. 기존 연구 검토 : 네트워크 분석을 활용한 숙련연구

숙련(Skill) 및 노동 시장 연구 분야에서 사회 네트워크 분석(Social Network Analysis, SNA) 방법론은 숙련 간의 복잡한 상호 작용, 구조적 중요성, 그리고 노동 시장의 동역학적 변화를 이해하는 데 핵심적인 도구가 될 수 있다. 기존 연구들은 숙련을 단순한 목록이 아닌 상호 연결된 시스템으로 모델링함으로써 ‘Skills Space(숙련 공간)’라는 네트워크 형태로 구축하였

으며, 특정 직업에서 두 숙련이 함께 발생하는 빈도를 기반으로 숙련 간 상호보성을 측정하고 중심성 지표(연결 중심성, 매개 중심성, 근접 중심성, 고유벡터 중심성)를 통해 각 숙련의 구조적 중요도를 정량화하였다. 이러한 네트워크 기반 접근법은 노동 시장의 양극화 현상을 사회-인지적 숙련과 감각-신체적 숙련이라는 두 군집으로 설명하고, 숙련 집합 간의 네트워크 근접도를 활용하여 직업 전환 경로를 예측하는 데 활용되는 등 네트워크 분석이 실제 노동 이동을 설명하는 데 효과적임을 입증하였다. 또한 숙련의 가치가 시장의 수요와 공급뿐만 아니라 네트워크 내에서의 상호보성에 의해 결정된다는 점을 밝혀내어, 숙련을 구조적 관점에서 이해하는 데 강력한 분석틀을 제공하고 있다.

이러한 방법론의 실제 적용 사례로, Maghsoudi(2024)는 컴퓨터 과학 직무에서 요구되는 숙련 집합을 밝혀내기 위해 사회 네트워크 분석을 활용한 연구를 수행하였다. 이 연구는 LinkedIn에서 수집한 5,763개의 숙련 목록과 7,777개의 구인 광고를 분석하여 1,315개의 고유한 숙련을 노드로, 구인 광고에서 공통적으로 나타나는 빈도를 관계의 가중치로 설정한 숙련 통신 네트워크를 구축하였다. 분석 결과 컴퓨터 과학 직무의 숙련은 제너럴리스트(40%), 소프트웨어 개발(25%), 인프라 및 보안(33%), 임베디드 시스템(2%)이라는 네 가지 뚜렷한 커뮤니티로 분류되었으며, 모든 중심성 지표에서 커뮤니티케이션, 영어, SQL이 상위 3위를 차지하여 기술적 전문성과 함께 소프트스킬의 중요성을 확인하였다. 특히 이 연구는 네트워크 분석을 통해 제너럴리스트 숙련의 수요가 감소(87.3%→82.6%)하고 전문화된 숙련에 대한 수요가 증가하는 노동 시장의 변화 동향을 포착함으로써, 네트워크 분석이 시장 변화의 선행 지표로서 활용될 수 있는 가능성을 보여주었다.

한편 Henning et al.(2023)은 네트워크 분석의 적용 범위를 지역 경제 차원으로 확장하여 숙련 관련성과 지역 경제 성과 간의 관계를 탐구하였다. 이들은 스웨덴의 600만 건 구인 광고 데이터를 기반으로 직업과 산업의 교차 조합으로 정의된 직무 간의 숙련 관련성을 Doc2Vec과 코사인 유사도를 활용하여 측정하고, 이를 통해 지역의 숙련 일관성을 정량화하였다. 흥미롭게도 연구 결과는 역설적인 패턴을 보여주었는데, 숙련 일관성이 낮을수록(즉, 다양성이 높을수록) 지역의 평균 임금과 생산성이 더 높게 나타났으며,

이는 고등 교육을 받은 근로자 비율과 밀접한 관련이 있었다. 반면 근로자들이 숙련 관련성이 높은 직무로 전환하는 비율이 높은 지역일수록 더 높은 평균 임금을 누리는 것으로 나타나, 지역 내 다양성과 개별 전환의 관련성이라는 두 차원이 모두 경제 성과에 중요한 역할을 한다는 점을 밝혀냈다. 이 연구는 네트워크 분석이 개별 숙련이나 직무 수준을 넘어서 지역 경제 정책 수립에도 활용될 수 있는 분석틀임을 실증적으로 보여주었다.

Huang et al.(2021)은 네트워크 분석을 산업 간 직업 전환 경로 분석에 적용하여 현대 산업 근로자들의 이직 가능한 숙련과 경력 전환 가능성을 탐구하였다. 이들은 O*NET 데이터를 활용하여 현대 및 관광, 보건 과학, 정보 기술, 운송 및 물류, 경영 및 행정 등 5개 경력 군집 간의 숙련 공유 네트워크를 구축하고, 현대 근로자들이 보유한 소프트 스킬(능동적 경청, 말하기, 비판적 사고 등)이 다른 산업에서도 가치 있는 이직 가능한 숙련임을 확인하였다. 특히 현대 산업과 보건 과학 군집 간에 가장 강력한 숙련 연결고리가 존재한다는 점을 밝혀내어, 현대 근로자들이 IT 직업보다는 보건 의료 직업으로의 전환이 더 수월할 것임을 시사하였다. 이 연구는 자동화 위험이 높은 현대 산업(평균 85% 컴퓨터화 가능성) 근로자들에게 실질적인 재교육 경로를 제시함으로써, 네트워크 분석이 개별 근로자의 경력 개발과 정책적 재교육 프로그램 설계에 구체적인 방향성을 제공할 수 있음을 보여주었다.

마지막으로 이현경(2025)은 국내 AI 숙련 연구에서 네트워크 분석 활용 시 직면하는 현실적 한계와 방법론적 과제를 제기하였다. 이 연구는 OECD, Stanford HAI, CSET 등 주요 해외 연구기관들이 Lightcast나 LinkedIn과 같은 글로벌 플랫폼 데이터를 활용하여 AI 노동력과 숙련을 측정하는 방식을 비교 분석하면서, 이러한 해외 플랫폼 기반 접근법이 국내 현실에 적용될 때의 한계를 지적하였다. 특히 국내 AI 관련 구인공고의 37%가 주요 3대 플랫폼(사람인, 잡코리아, AI 커리어) 중 어디에도 게시되지 않는다는 실증적 발견을 통해, 구인 광고 기반 네트워크 분석 연구에서 데이터 출처의 편향성과 선택 편향 문제가 연구 결론의 일반화에 심각한 제약을 가할 수 있음을 경고하였다. 이는 네트워크 분석의 강력함에도 불구하고 데이터의 대표성과 포괄성이 분석 결과의 신뢰성을 좌우하는 핵심 요소임을 시사하며, 향후 국내 숙련 네트워크 연구에서 다각화된 데이터 수집 전략과 국가 차원의

체계적 조사의 필요성을 제시하였다.

2. 숙련 네트워크 분석 : IT 일자리를 중심으로

가. 2024년 IT 직무 숙련 네트워크의 구조적 특성

이 절에서는 2021년과 2024년 IT 일자리 온라인 구인공고에 나타난 숙련을 네트워크로 시각화하였다. 이때 IT 일자리는 한국고용직업분류(KECO)의 131~136 일자리를 의미한다(표 2-1). 시각화 도구로는 Gephi를 활용하였으며, frequency 1000 이상의 엣지만 필터링하여 표시하였고 PageRank 값을 반영하여 작성하였다.

[그림 2-3]과 같이 2024년 IT 일자리 구인공고에서 추출한 숙련들을 네트워크로 시각화한 결과, IT 노동시장의 숙련 구조는 네 개의 뚜렷한 클러스터

〈표 2-1〉 한국고용직업분류 IT 일자리

코드	직업명
1311	컴퓨터 하드웨어 기술자 및 연구원
1312	통신공학 기술자 및 연구원
1320	컴퓨터시스템 전문가
1331	시스템 소프트웨어 개발자
1332	응용 소프트웨어 개발자
1333	웹 개발자
1339	기타 컴퓨터 전문가 및 소프트웨어 전문가
1341	데이터 전문가
1342	네트워크 시스템 개발자
1343	정보시스템 운영자
1344	웹 운영자
1349	기타 데이터 및 네트워크 전문가
1350	정보보안 전문가
1360	통신·방송송출 장비 기사

로 구성되어 있음을 확인할 수 있다. 이러한 클러스터 구조는 현대 IT 산업의 기술적 특성과 시장 수요를 반영하며, 각 클러스터 간의 연결 패턴은 IT 전문가에게 요구되는 복합적 역량의 성격을 보여준다.

가장 주목할 만한 특징은 보안 관리 클러스터(초록색)가 전체 네트워크에서 압도적인 중심성을 차지한다는 점이다. “통합 보안 관리 및 통합 로그 관리 시스템 구축 및 유지보수하기”를 핵심 노드로 하는 이 클러스터는 보안 시스템 통합 관리, 로그 관리 및 모니터링, 보안 인프라 구축 및 운영, 시스템 보안 운영 등의 숙련들로 구성되어 있다. 이 클러스터가 네트워크의 중심에 위치하면서 다른 모든 클러스터와 강한 연결을 보이는 것은 보안이 현재 IT 시장에서 최우선 과제가 되었으며, 모든 IT 영역에서 필수 요소로 자리 잡았음을 의미한다.

서비스 개발 및 운영 클러스터(빨간색)는 네트워크 중앙부에 위치하며 “회원 서비스 전산 개발 업무 수행하기”, “SI 환경 개발환경 수행하기”, “빅데이터 컴퓨팅하기” 등을 핵심 노드로 한다. 이 클러스터는 웹/모바일 서비스 개발, 시스템 통합 개발, 데이터베이스 관리, 빅데이터 처리 및 분석, 클라우드 서비스 운영 등 전통적인 IT 개발 및 운영의 핵심 영역을 포괄한다. 특히 이 클러스터가 다른 모든 클러스터와 광범위하게 연결되어 있다는 점은 개발 역량이 여전히 IT 전문가의 기본 소양으로 기능하고 있음을 보여준다.

하드웨어 및 인프라 클러스터(주황색)는 네트워크의 좌측 상단에 위치하며 “BSP(Board Support Package) 개발하기”를 핵심 노드로 한다. 이 클러스터는 임베디드 시스템 개발, 하드웨어 지원 패키지 개발, 저수준 시스템 프로그래밍, 펌웨어 개발 등의 전문 기술 영역을 포함하며, IoT와 스마트 디바이스의 확산과 관련된 기술 수요를 반영한다. 상대적으로 특화된 위치에 있으면서도 다른 클러스터들과의 연결을 유지하고 있어 하드웨어 전문성이 소프트웨어 영역과 통합적으로 요구되고 있음을 시사한다.

전문 기술 클러스터(파란색)는 네트워크 하단에 위치하며 “머신러닝을 활용한 분석 고도화하기”를 핵심 노드로 한다. 이 클러스터는 머신러닝 및 AI 기술, 고급 데이터 분석, 알고리즘 최적화, 예측 모델링 등의 첨단 기술을 포함한다. 다른 클러스터들에 비해 상대적으로 독립적인 위치에 있다는 점은 이러한 기술들이 높은 전문성을 요구하는 특화된 영역임을 나타낸다.

라는 세 가지 핵심 특징으로 요약된다.

가장 두드러진 변화는 네트워크 중심 기술의 완전한 전환이다. 2021년에는 “빅데이터 컴퓨팅”이 네트워크의 중심에 위치하며 최대 중심성을 보였으나, 2024년에는 “통합 보안 관리 및 통합 로그 관리 시스템 구축 및 유지보수하기”가 가장 큰 중심 노드로 변화했다. 이는 데이터 중심 시대에서 보안 중심 시대로의 패러다임 전환을 보여주는 변화로, 디지털 전환 가속화와 함께 사이버보안 위협이 급증한 현실을 반영한다.

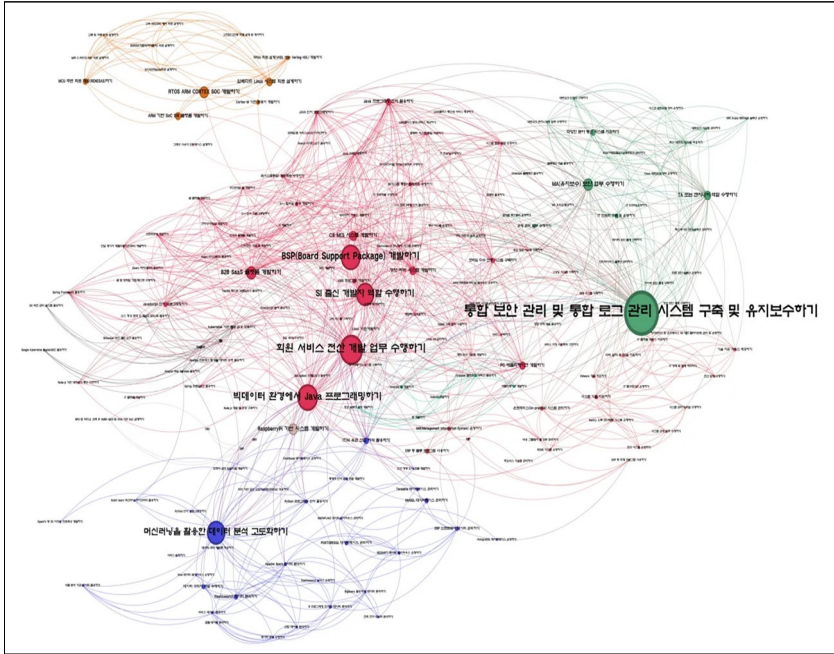
기술 생태계의 구조적 재편도 주목할 만한 변화이다. 2021년에는 빅데이터, 클라우드 컴퓨팅, 머신러닝 등 데이터 중심 기술들이 네트워크의 핵심을 형성했던 반면, 2024년에는 보안 기술이 최대 중심성을 확보하면서 기존 데이터 기술들과 새로운 연결 구조를 형성하고 있다. 특히 보안 관련 클러스터(녹색)가 독립적인 영역을 형성하면서도 다른 모든 기술 클러스터들과 강한 연결성을 보이는 것은 보안이 모든 IT 영역에서 필수 요소로 자리 잡았음을 의미한다.

클러스터별 분석에서도 의미 있는 변화를 확인할 수 있다. 빨간색 클러스터는 두 시기 모두에서 핵심 기술군을 형성하며 전통적인 IT 인프라와 개발 기술들을 포괄하고 있어, 개발 역량이 IT 전문가의 지속적인 기본 소양임을 보여준다. 파란색 클러스터는 특정 전문 기술 영역을 나타내며 상대적으로 독립적인 위치를 유지하고 있는데, 이는 AI/머신러닝과 같은 첨단 기술이 높은 전문성을 요구하는 특화된 영역으로 자리 잡았음을 시사한다.

네트워크 구조의 밀도 변화도 중요한 시사점을 제공한다. 2024년 네트워크에서 관찰되는 높은 연결 밀도는 IT 직무에서 요구되는 기술의 복합성이 증가했으며, 단일 기술보다는 통합적 역량이 더욱 중요해졌음을 보여준다. 각 색상별 클러스터가 뚜렷하게 구분되면서도 상호 연결되는 패턴은 기술의 전문화와 동시에 영역 간 융합이 진행되고 있음을 나타낸다.

이러한 변화 양상은 IT 전문가들에게 새로운 숙련 개발이 요구된다는 시사점을 던진다. 보안 기술의 중심성 증대는 모든 IT 전문가가 보안에 대한 기본적 이해를 갖춰야 함을 의미하며, 네트워크 밀도의 증가는 자신의 전문 분야를 깊이 있게 발전시키면서도 관련 기술 영역에 대한 폭넓은 이해를 병행해야 한다는 시장의 요구를 시사한다. 결과적으로 2021년에서 2024년으

[그림 2-4] IT 분야 숙련 네트워크(2021년)



자료 : 온라인 구인공고 데이터를 바탕으로 저자 작성.

로의 변화는 한국 IT 시장이 보안을 중심축으로 하는 통합적 기술 생태계로 진화하고 있음을 보여준다.

3. 네트워크 특성 지표와 AI 노출도의 상관관계

가. 숙련 확산도 또는 숙련 보편성(Skill Ubiquity)

AI 시대의 도래와 함께 우리가 직면한 가장 핵심적인 질문은 “어떤 숙련이 진정으로 중요한가”라는 것이다. 이 질문은 전체 노동시장의 관점에서 제기될 수 있으며, 정보기술 분야나 제조업과 같은 특정 산업 영역에서도 동일하게 던져질 수 있다. 그러나 개별 숙련을 독립적인 단위로 바라보는 전통적 접근법으로는 이 복잡한 질문에 온전히 답하기 어렵다. 숙련은 고립된 채 존재하지 않으며, 다른 숙련들과의 긴밀한 상호작용 속에서 그 가치를

의미를 발현하기 때문이다. 따라서 본 절에서는 개별 숙련의 네트워크적 특성과 관계적 속성을 체계적으로 분석함으로써 AI 시대에 요구되는 핵심 숙련을 식별하고자 한다.

이 장에서는 숙련의 관계적 의미를 측정하는 두 가지 접근법을 제시하는데, 먼저 개별 숙련이 얼마나 다양한 직업 영역에서 보편적으로 요구되는가를 숙련의 확산도라고 명명하고 검토해보기로 한다.

숙련 확산도(Spread of Dispersion, SOD)는 숙련의 보편성(Skill Ubiquity)을 측정하는 지표로, 특정 숙련이 얼마나 다양한 직업에서 요구되고 있는지를 정량화한다. 이는 정보이론의 엔트로피 개념을 차용하여 숙련의 분포 패턴을 분석한 것으로, 숙련이 특정 직종에 집중되어 있는지 아니면 여러 직종에 고르게 분산되어 있는지를 나타낸다.

$$SOD = \frac{1}{N_j} \sum_{k=1}^M P_{j,k} \log P_{j,k}$$

$P_{j,k}$ = 직종 j 에서 숙련 k 가 포함된 비율

[수식 1] 숙련 확산도

단, 이 지표 자체는 숙련이 등장한 채용공고의 절대적 빈도를 고려하지 않는 한계가 있었다. 즉, 매우 적은 수의 채용공고에서만 언급된 숙련이라 하더라도, 그것이 서로 다른 직종에 분포되어 있다면 높은 확산도 값을 갖게 되는 문제가 발생했다. 이러한 측정상의 왜곡을 보정하기 위해 채용공고 빈도에 따른 가중치를 반영한 개선된 지표를 개발하여 사용하였다.

[그림 2-5]의 숙련 확산도(SOD) 그래프는 각 숙련군이 얼마나 다양한 직업 영역에서 범용적으로 요구되는지를 보여준다. 가장 높은 확산도를 보이는 상위 숙련군은 ‘고객및클라이언트지원’과 ‘관리및지원’ 숙련군이다.

‘고객및클라이언트지원’ 숙련군(69개)은 고객지원(59개)과 고객관리및모니터링(10개)으로 세분화되며, ‘글로벌 커뮤니케이션 수행하기’, ‘VIP 고객 응대하기’, ‘반도체 관련 고객 지원하기’, ‘해외 항공 고객 서비스 제공하기’ 등이 포함된다. 이들 숙련의 높은 중심성은 현대 비즈니스가 고객 중심으로 재편되면서 거의 모든 업종에서 고객 접점 관리가 핵심 경쟁력으로 부상했

음을 보여준다. 특히 ‘네트워크 상담실 운영하기’, ‘아웃바운드 콜 업무 수행하기’ 등은 디지털 전환과 함께 다양한 산업에서 필수적인 연결 고리 역할을 하고 있다.

‘관리및지원’ 숙련군(58개)은 사무자동화(25개)와 일반관리및사무업무(33개)로 구성되며, ‘데이터 분석 및 보고서 작성하기’, ‘Excel 및 PowerPoint 고급 활용하기’, ‘MS Word 문서 작성 및 편집하기’, ‘Notion 협업 도구 활용하기’ 등의 숙련들이 핵심을 이룬다. 이들의 높은 중심성은 디지털 업무 환경에서 이러한 기초적 관리 숙련들이 다른 모든 전문 숙련들의 효율성을 증폭시키는 승수 효과를 갖고 있음을 의미한다. ‘인터넷 정보 검색 및 활용하기’, ‘사무자동화 프로그램 운영 및 활용하기’ 등은 업종을 불문하고 거의 모든 직무에서 다른 전문 숙련들과 결합되어 시너지를 창출하는 기반 인프라 역할을 수행한다.

낮은 확산도를 보이는 숙련군은 ‘정책및사회과학’, ‘정부및공공행정’, ‘농업/원예및조경업’ 숙련군이다. 이들의 낮은 중심성은 해당 숙련들이 고도로 전문화되고 제한적인 영역에서만 활용되는 특성을 반영한다.

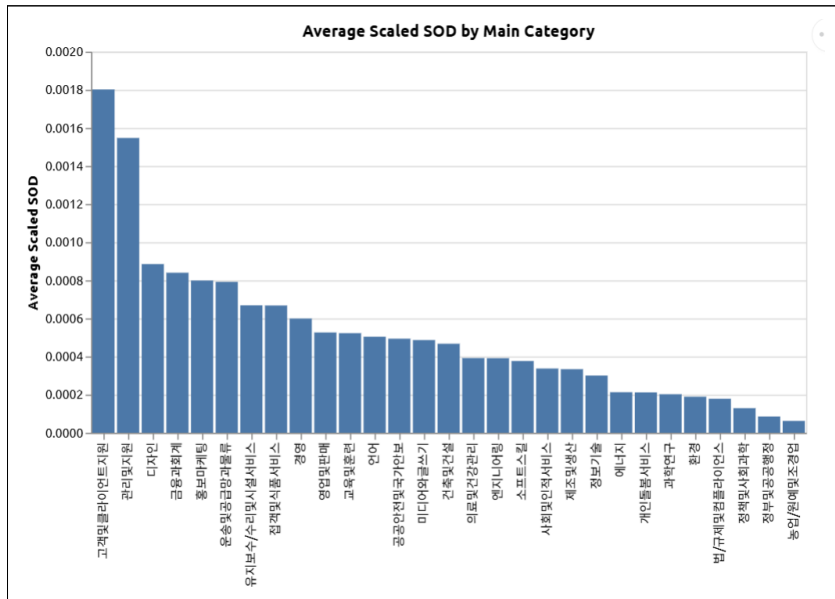
‘정책및사회과학’ 숙련군(15개)은 분과지식(8개)과 정책분석과연구및개발(7개)로 구성되며, ‘경제 이론 분석하기’, ‘정치 현상 분석하기’, ‘고고학 연구 수행하기’, ‘교육학 이론 활용하기’ 등의 고도로 학술적이고 이론적인 숙련들을 포함한다. 이들 숙련의 낮은 중심성은 주로 연구기관, 대학, 정책연구소 등 매우 제한된 영역에서만 요구되며, 다른 산업 분야의 실무적 숙련들과의 연결고리가 약하다는 것을 의미한다. ‘논문 작성하기’, ‘국책과제 수행하기’ 등은 학계와 정책연구 분야 내부에서는 핵심적이지만, 전체 노동시장의 숙련 생태계에서는 주변적 위치를 차지한다.

‘정부및공공행정’ 숙련군(14개)은 안보(5개)와 정부서비스(9개)로 세분화되며, ‘정부 및 공공기관 업무 수행하기’, ‘국방 업무 수행하기’, ‘군 전역 간부 경력 활용하기’, ‘정부기관 협력 네트워크 운영하기’ 등이 포함된다. 이러한 숙련들의 극도로 낮은 중심성은 공공부문의 특수성을 보여준다. ‘정부 지원사업 관리하기’, ‘정부 용역사업비 정산하기’ 등은 공공부문에서는 필수적이지만, 민간 영역과의 교차점이 제한적이어서 다른 숙련들과의 시너지 효과가 미미하다.

‘농업/원예및조경업’ 숙련군(25개)은 농업(17개)과 원예및조경업(8개)으로 구성되며, ‘한우 사육 및 관리하기’, ‘작물 생육 상태 조사하기’, ‘농축수산물 생산 관리하기’, ‘전지 작업 기술 활용하기’ 등의 전통적인 1차 산업 숙련들이 주를 이룬다. 흥미롭게도 ‘스마트팜 시스템 운영하기’와 같은 첨단 기술이 접목된 숙련도 포함되어 있지만, 여전히 농업이라는 특정 산업 내에 국한되어 있어 타 산업과의 연결성이 제한적이다. ‘농어촌 사회 현상 분석하기’, ‘화환 디자인 및 제작하기’ 등은 각각의 전문 영역에서는 중요하지만, 전체 숙련 네트워크에서는 고립된 섬과 같은 구조를 형성한다.

이러한 패턴은 노동시장의 숙련 양극화 현상을 보여주는 것이기도 하다. 상위 중심성을 갖는 숙련들은 산업 간 경계를 넘나들며 광범위한 연결성을 보이는 반면, 하위 중심성 숙련들은 특정 분야의 수직적 전문성은 높지만 수평적 확장성은 제한적이다. 특히 공공부문과 전통 산업의 숙련들이 디지털 경제의 네트워크 효과로부터 상대적으로 소외되어 있음을 시사하며, 이는 향후 인력 개발과 산업 전환 정책에서 중요한 고려사항이 될 것이다.

[그림 2-5] 숙련군별 숙련 확산도



자료 : 온라인 구인광고 데이터를 바탕으로 저자 작성.

나. 페이지랭크(PageRank) 기반 숙련 중심성

페이지랭크(PageRank) 기반 숙련 중심성은 숙련이 노동시장 내에서 실제로 얼마나 중요한 연결고리 역할을 하는지를 평가하는 지표이다. 이는 구글의 검색 엔진 핵심 알고리즘인 PageRank를 숙련 네트워크 분석에 적용한 것으로, 단순한 출현 빈도나 분포를 넘어서 숙련 간의 구조적 관계와 영향력을 측정한다.

PageRank 알고리즘은 1996년 래리 페이지와 세르게이 브린이 개발한 것으로, 웹페이지의 중요도를 해당 페이지로 향하는 링크의 수와 그 링크들의 품질을 종합적으로 고려하여 결정한다. 핵심 아이디어는 “중요한 페이지로부터 링크를 받는 페이지가 더 중요하다”는 것이다. 즉, 단순히 많은 링크를 받는 것보다도 신뢰할 만한 고품질 페이지로부터 링크를 받는 것이 해당 페이지의 중요도를 더 크게 높인다는 원리이다. 이 알고리즘은 각 페이지가 자신이 받은 PageRank 점수를 자신이 링크하는 다른 페이지들에게 균등하게 분배한다는 가정하에, 전체 웹 네트워크에서 안정적인 점수 분포에 도달할 때까지 반복 계산을 수행한다.

이러한 PageRank의 논리를 숙련 전이 분석에 적용하면, 특정 숙련이 노동시장 내에서 얼마나 핵심적인 허브 역할을 수행하는지를 정량적으로 평가할 수 있다. 여기서 ‘링크’는 숙련 간의 동시 출현 관계나 전이가능성으로 해석되며, 다양한 고중요도 숙련들과 연결된 숙련일수록 높은 중심성 점수를 받게 된다. 예를 들어, 많은 핵심 숙련들과 함께 요구되는 기초적 숙련이나, 여러 전문 영역을 연결하는 교량 역할을 하는 숙련들이 높은 PageRank 중심성을 보이게 된다. 이는 해당 숙련이 단순히 자주 언급되는 것을 넘어서, 실제로 다른 중요한 숙련들과의 시너지 효과를 통해 노동시장에서 구조적 중요성을 갖는다는 것을 의미한다.

[그림 2-6]에 나타난 바와 같이, 페이지랭크 중심성에서도 ‘관리및지원’ 숙련군이 압도적으로 높은 점수를 나타낸다. ‘고객및클라이언트지원’, ‘경영’, ‘언어’ 순으로 그 뒤를 잇는다.

‘관리및지원’ 숙련군(58개)의 압도적 중심성은 현대 업무 환경에서 이들이 수행하는 인프라적 역할의 중요성을 극명하게 보여준다. 사무자동화(25

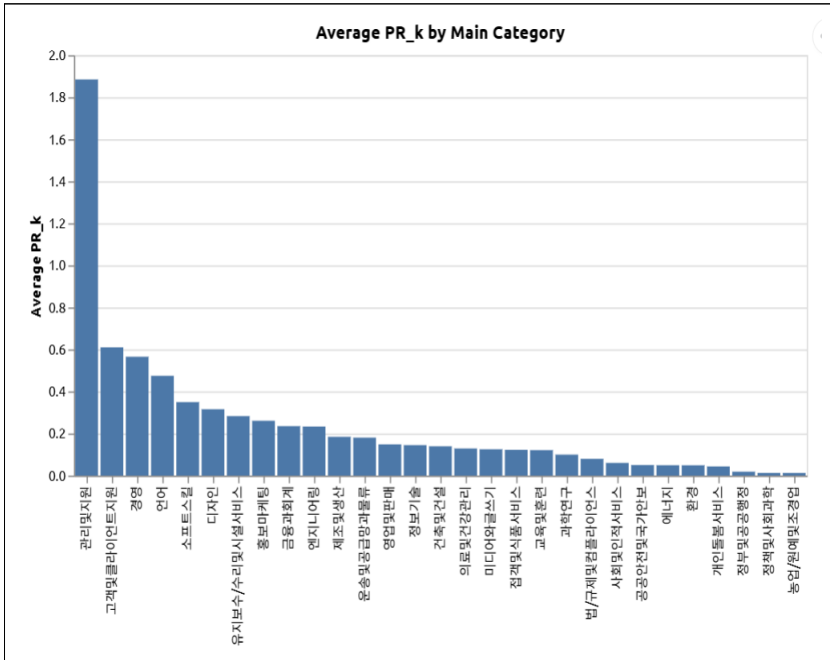
개)와 일반관리및사무업무(33개)로 구성된 이 숙련군은 ‘Excel 데이터 분석 및 시각화하기’, ‘MS Office 업무용 문서 작성하기’, ‘Microsoft Outlook 이메일 관리하기’, ‘사무자동화 프로그램 운영 및 활용하기’ 등 디지털 업무의 기본 토대를 형성하는 숙련들을 포함한다. 특히 주목할 점은 ‘회의록 작성 및 배포하기’, ‘문서 및 자료 발송 관리하기’, ‘주문서 작성 및 관리하기’ 등이 단순한 보조 업무를 넘어서 조직의 정보 흐름과 의사 결정 프로세스를 연결하는 핵심 허브 역할을 수행한다는 것이다. 이러한 관리 숙련들은 거의 모든 전문 분야에서 요구되면서 동시에 다른 고급 숙련들의 효과성을 증폭시키는 승수 효과를 갖고 있다.

‘고객및클라이언트지원’ 숙련군(69개)이 두 번째 높은 중심성을 보이는 것은 현대 비즈니스의 고객 중심적 전환을 반영한다. ‘민원 접수 및 응대하기’, ‘VIP 고객 응대하기’, ‘FAQ 콘텐츠 관리하기’, ‘서비스 품질 모니터링하기’ 등은 업종을 불문하고 고객 접점에서 핵심적인 가치 창출 역할을 담당한다. ‘반도체 관련 고객 지원하기’, ‘자동차 고장 및 사고 접수하기’ 등 산업별 특화된 고객 지원 숙련들도 포함되어 있지만, 이들 모두 기본적으로는 고객 관계 관리라는 공통 분모를 갖고 있어 다양한 전문 영역과의 연결고리 역할을 수행한다.

‘경영’ 숙련군(368개)은 전략및분석(80개), 운영및관리(179개), 인적자원관리(91개), 리더십및의사소통(18개)의 네 영역으로 세분화되며, ‘사업 성장 전략 개발하기’, ‘비즈니스 프로세스 재설계하기’, ‘밸류체인 분석 수행하기’, ‘PMI 합병 후 통합 관리하기’ 등 고도의 전략적 사고를 요구하는 숙련들을 포함한다. 이들의 높은 중심성은 경영 숙련이 단순히 관리직에만 국한되지 않고, 각 전문 분야에서 전략적 사고와 조직 운영 능력이 점점 더 중요해지고 있음을 시사한다. ‘직원 권한 부여하기’, ‘인적 자본 가치 향상 방안 수립하기’ 등은 조직의 모든 계층에서 요구되는 메타-스킬로 진화하고 있다.

‘언어’ 숙련군(35개)은 언어역량(25개), 언어교육(5개), 언어통역및번역(5개)으로 구성되며, ‘영어 의사소통하기’, ‘중국어 HSK 5급 수준 구사하기’, ‘베트남어 구사하기’, ‘영어 번역하기’ 등이 포함된다. 언어 숙련의 높은 중심성은 글로벌화와 다문화 사회로의 전환을 반영하며, 특히 ‘리스닝 능력 향상하기’, ‘발음 교정하기’ 등 기초적 의사소통 능력이 모든 전문 분야에서

[그림 2-6] 숙련군별 페이지랭크 중심성



자료 : 온라인 구인광고 데이터를 바탕으로 저자 작성.

필수적 기반 역할을 수행하고 있음을 보여준다.

이러한 결과는 범용성과 연결성의 조합이 숙련의 구조적 중요성을 결정하는 핵심 요인임을 확인해준다. 상위 중심성을 갖는 숙련들은 모두 특정 전문 분야에 국한되지 않으면서도 다른 고급 숙련들과의 시너지를 창출하는 기반 인프라 역할을 수행한다는 공통점을 갖고 있다.

다. 숙련 확산도와 페이지랭크 중심성의 AI 노출도와의 상관관계

다음으로 숙련의 확산도와 페이지랭크 중심성이 AI 노출도 점수와 갖는 상관관계를 살펴보았다. AI가 대신할 수 있는 숙련이라면 확산도나 중심성 측면에서 낮은 점수를 기록하지 않을까 예측해 볼 수 있기 때문이다. 그러나 예상과는 달리 이들 점수 간에 상관관계는 낮게 나타났다.

이러한 결과는 여러 각도에서 해석이 가능하다. 첫째, AI 대체의 시차적

특성으로 볼 수 있다. 2024년 시점에서 AI 기술이 급속도로 발전하고 있음에도 불구하고, 실제 노동시장에서의 본격적인 숙련 대체는 아직 초기 단계에 머물러 있을 가능성이 높다. 기술의 도입과 실제 업무 환경에서의 적용 사이에는 상당한 시간적 지연이 존재하며, 특히 조직의 관성과 제도적 저항이 이러한 변화를 더디게 만들고 있다.

둘째, AI 대체 가능성과 실제 중요성의 역설적 관계를 고려해야 한다. 오히려 AI가 쉽게 대체할 수 있는 숙련일수록 현재 노동시장에서 광범위하게 요구되고 있을 가능성이 있다. 예를 들어, 데이터 분석, 문서 작성, 번역 등의 숙련들은 AI의 대체 가능성이 높지만, 동시에 디지털 전환 과정에서 모든 산업에 걸쳐 필수적인 기반 숙련으로 자리 잡고 있다. 이는 보완성 효과(complementarity effect)로 설명될 수 있는데, AI 기술의 발전이 해당 숙련의 대체보다는 오히려 수요 증가를 가져오고 있는 것이다.

셋째, AI 노출도 측정의 한계도 고려되어야 한다. 현재의 AI 노출도 지표가 기술적 대체 가능성에만 초점을 맞추고 있다면, 실제 업무 환경에서의 맥락적 복잡성, 인간 간 상호작용의 중요성, 창의적 판단의 필요성 등을 충분히 반영하지 못할 수 있다. 높은 확산도와 중심성을 갖는 숙련들은 종종 이러한 맥락적 요소들이 핵심적인 역할을 하는 영역에 위치해 있을 가능성이 크다.

넷째, 네트워크 효과의 보호막 현상도 주목할 만하다. 높은 중심성을 갖는 숙련들은 다른 여러 숙련들과 복합적으로 연결되어 있어, 단독으로 대체되기보다는 전체 숙련 생태계 내에서 통합적으로 진화할 가능성이 높다. 이는 AI가 개별 태스크는 잘 수행할 수 있지만, 복합적이고 맥락적인 업무 환경에서는 여전히 인간의 조정 역할이 필수적임을 시사한다.

따라서 상관관계의 부재는 AI 대체가 아직 본격화되지 않았다는 단순한 해석을 넘어서, AI와 인간 숙련 간의 복잡한 상호보완적 관계와 노동시장의 구조적 특성을 반영하는 것으로 이해하는 편이 더 적절할 것이다.

제4절 숙련의 직업 간 전이가능성(Transferability)

1. 직종 간 숙련 전이 확률 측정

기술 발전과 산업 구조의 다각화로 인해, 특정 직종에서 요구되는 숙련이 다른 직종으로 얼마나 쉽게 전이될 수 있는지 평가하는 것이 중요해졌다. 본 연구의 첫 번째 목적은 직종 간 기술 유사성과 기술 전이가능성에 대한 체계적인 분석이다. 특히, 숙련의 확산 속도 및 산업 간 전이가능성을 분석하여, 숙련의 흐름을 보다 정교하게 예측하는 것이 연구의 핵심 목표 중 하나이다. 예를 들어, 데이터 분석 숙련이 IT 직종뿐만 아니라 헬스케어 관련 직종으로 확장되는 패턴을 분석할 수 있다.

직업 간 이동성을 분석하기 위해서는 먼저 직종별로 요구되는 숙련의 유사성을 정량적으로 파악할 필요가 있다. 직종 간 숙련 유사도(Occupation-Skill Similarity)는 두 직종이 요구하는 숙련의 정도가 얼마나 비슷한지 측정한다. 이는 직종을 단순한 산업 분류가 아니라 숙련의 조합으로 이해하는 접근으로, 최근 노동경제학에서 점차 확산되고 있는 분석틀이다. 예컨대 Muneeppeerakul et al.(2013)은 지역 고용 패턴의 공동 출현(co-occurrence)을 활용하여 직업 간 네트워크를 구축하고, 이 “직업 공간(occupation space)”에서 이동 가능성을 분석하였다. 이어서 Kok and ter Weel(2014), Alabdulkareem et al.(2018) 등은 직업을 구성하는 숙련 벡터 간 유사성을 활용하여 직종 간 근접성을 측정하는 방법을 발전시켜 왔다. 이는 본 장의 직종별 유사도 측정 접근법과 맥락을 같이하며, 직종을 숙련 단위로 세분화함으로써 기존의 산업 분류를 넘어 보다 미시적인 이동 가능성을 포착할 수 있다.

이러한 유사도 측정 방식이 실제 노동 이동성과 임금 결과를 설명한다는 실증적 근거도 축적되어 있다. Gathmann and Schönberg(2010)는 독일 노동시장을 대상으로 직종 간 숙련 유사도가 높을수록 전환 과정에서 임금 손실이 적음을 확인하였으며, Poletaev and Robinson(2008) 역시 미국 노동자

의 데이터를 통해 유사도가 낮은 직종 이동일수록 임금 손실이 크다는 사실을 보여주었다. Levy Yeyati and Montané(2020)는 구인·이직 데이터를 활용하여 직업 간 유사성 지표를 구축하고, 이 지표가 전환 후 임금과 양의 상관관계가 있음을 제시하였다. 즉, 숙련 유사도가 높은 직무 간 이동은 인적자본의 손실을 줄이고 전환을 원활히 한다는 점에서, 직종 간 숙련 유사도 ($OSS_{i,j}$)는 직종 전환 가능성을 평가하는 데 타당한 지표임이 확인된다.

$$OSS_{i,j} = \frac{\sum_k W_{i,k} \cdot W_{j,k}}{\sqrt{\sum_k W_{i,k}^2} \times \sqrt{\sum_k W_{j,k}^2}}$$

여기서,

OSS = 직종 i 와 직종 j 간의 숙련 유사도

$W_{i,k}$ = 직종 i 에서 숙련 k 가 요구되는 가중치($TF-IDF$ 기반)

$W_{j,k}$ = 직종 j 에서 숙련 k 가 요구되는 가중치

[수식 2] 직종 간 유사도 계산

그러나 단순 유사도 지표에는 한계도 있다. $OSS_{i,j}$ 는 직무 간 상대적 숙련 구조만 반영하기 때문에 업스킬링과 다운스킬링 방향성을 구분하지 못하며, 절대적 숙련 수준의 차이를 무시할 수 있다. 예컨대 판매 관리자와 판매원은 유사한 숙련 비중을 가질 수 있으나 관리자는 훨씬 높은 수준의 숙련을 요구한다. 그럼에도 불구하고 단순 유사도는 이를 동일하게 평가하는 한계가 있다. Nedelkoska et al.(2015)은 이러한 문제를 보완하기 위해 직종 전환에 필요한 추가 교육 연한을 추정하여 이동 방향성을 반영한 숙련 거리 지표를 제안한 바 있다. 본 연구는 이러한 선행 연구의 한계를 보완하기 위해, $OSS_{i,j}$ 에 더해 네트워크 기반 중심성(PageRank)과 전이 확률 모형을 결합하여 보다 정교한 분석을 시도한다.

특히 직종과 숙련의 관계를 직종-숙련 이중 모드 네트워크로 구축한 뒤, PageRank 알고리즘을 적용하여 노동시장 내에서 중심적 역할을 수행하는 핵심 숙련을 식별한다. PageRank는 본래 웹페이지 중요도를 평가하기 위해 개발되었으나, 네트워크 구조상 연결 정도와 연결의 질을 동시에 반영할 수

있다는 점에서 숙련 분석에도 적합하다. Vista et al.(2020)은 구인공고 데이터를 활용하여 네트워크 중심성이 높은 숙련일수록 더 다양한 직무 전환을 촉진한다는 점을 보였고, Stephany and Teutloff(2024)는 특정 숙련의 가치는 다른 숙련과의 보완적 관계성에서 파생된다는 사실을 실증하였다. 또한 Alabdulkareem et al.(2018)은 숙련 네트워크가 고임금·고가치 숙련과 저임금·저가치 숙련으로 양극화되는 현상을 발견했으며, 네트워크의 중심부에 위치한 숙련이 광범위한 직종 전환을 가능케 하는 교량 역할을 한다는 점을 밝혔다. 따라서 본 연구에서 제시하는 PageRank 기반 숙련 중요도(PR_k)는 단순 빈도 기반 지표보다 노동시장에서의 실제 활용 가능성을 더 잘 반영할 수 있다.

$$PR_k = (1 - d) + d \sum_j \frac{PR_j}{d_j}$$

여기서,

PR_k = 숙련 k 의 중요도

d = 감쇠계수

j = k 와 연결된 직종

d_j = 직종 j 가 연결된 직무의 개수

[수식 3] 숙련 중심성의 측정

마지막으로 직종 간 숙련 유사도와 숙련 중요도를 기반으로, 직무 전환을 확률적으로 설명하기 위해 마코프 전이모형(Markov Transition Model)을 적용한다. 이 모형에서 각 직종은 상태(state)로 정의되며, 특정 직종에서 다른 직종으로 이동할 확률은 $OSS_{i,j}$, PR_k , 그리고 목표 직종의 고용 규모를 반영하여 산출된다.

$$TP_{i,j,k} = OSS_{i,j} \times PR_k \times \frac{N_j}{N_{total}}$$

여기서,

$TP_{i,j,k}$ = 직종 i 에서 직종 j 로 숙련 k 가 전이될 가능성

PR_k = 특정 숙련 k 의 PageRank 값

N_j = 직종 j 의 채용공고 수

N_{total} = 전체 직종 채용공고 수

[수식 4] 숙련 전이 확률의 측정

여기서 $TP_{i,j,k}$ 는 직종 i 에서 j 로, 숙련 k 를 매개로 전이될 확률을 의미한다. 즉, 두 직종 간 요구 숙련이 유사하고($OSS_{i,j}$), 해당 숙련이 노동시장 전체에서 중심성을 가지며(PR_k), 동시에 목표 직종의 규모가 클수록(N_j), 실제 전이 확률은 높아진다.

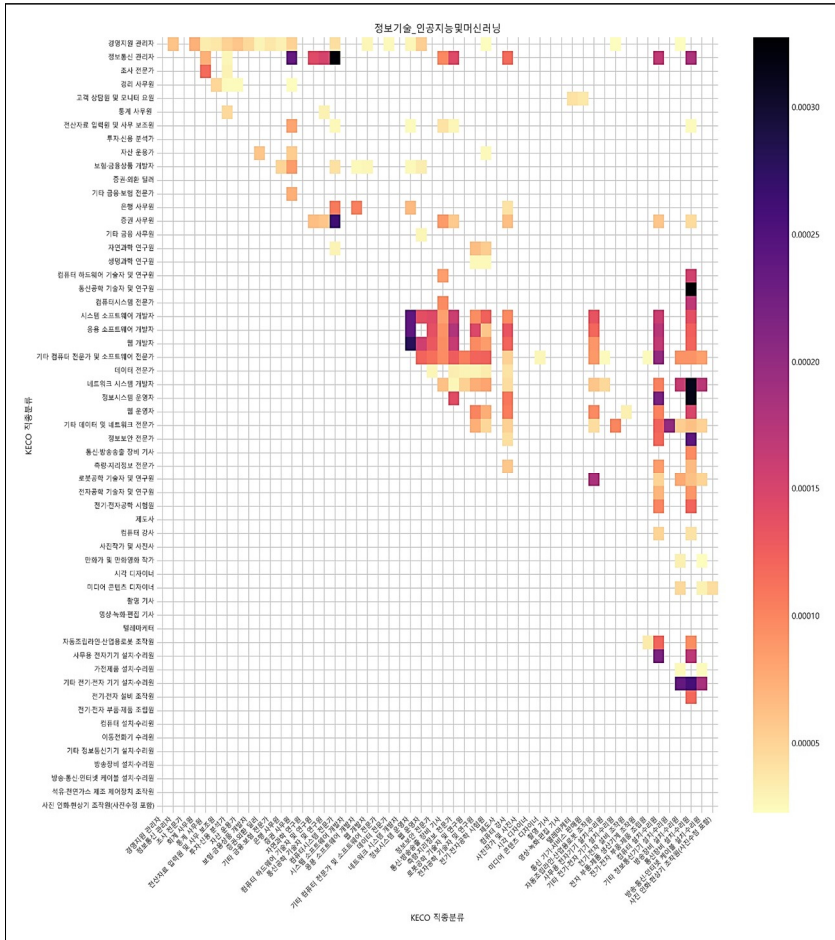
이러한 마코프 전이모형은 노동시장의 직업 이동을 네트워크상 확률적 과정으로 설명한다는 점에서 기존 연구와 맞닿아 있다. del Rio-Chanona et al.(2021)은 미국 노동시장을 대상으로 직업-전이 네트워크를 구축하고, 자동화 충격에 대응하는 직업 이동을 시뮬레이션한 결과, 실제 노동시장과 유사한 이동 패턴(예: 베버리지 곡선)을 재현하는 데 성공하였다. 이는 직업 이동을 1차 마코프 과정으로 모델링하는 접근이 노동시장 행태를 설명하는 데 타당함을 보여준다. 더 나아가 이 연구는 네트워크 내에서 밀접하게 연결된 직종은 다양한 이동 경로를 제공하는 반면, 고립된 직종은 전환 기회가 제한되어 구조적 취약성을 가진다는 점도 확인하였다. 즉, 본 장의 전이 확률 모형은 단순히 과거의 이직 데이터에만 의존하지 않고, 잠재적 이동 경로와 숙련의 매개적 역할까지 반영한다는 점에서 기존 접근을 보완한다.

2. 직업 간 숙련 전이가능성 측정 결과

가. IT 산업을 넘어선 AI 숙련의 확산

최근 연구에 따르면 인공지능(AI) 및 머신러닝(ML) 관련 숙련에 대한 수요가 대부분의 산업과 직종에서 급격히 증가하고 있다. 이는 AI/ML 기술이 전통적인 IT 업계의 경계를 넘어 금융, 제조, 전자, 미디어 등 다양한 분야에서 활용되고 있음을 시사한다. 실제로 Alekseeva et al.(2021)의 연구에 따르면 2010년대 미국 노동시장의 데이터를 보면, AI 숙련 수요는 IT 직군에서 가

[그림 2-7] 직업 간 속련 전이가능성 시각화



자료 : 저자 작성.

장 높지만 건축/엔지니어링, 과학, 경영 직군 등 그 외의 분야에서도 크게 증가한 것으로 나타났다. 이러한 확산 양상은 AI 기술이 특정 산업에 국한된 전문기술이 아니라 범용적으로 적용되는 기술임을 보여준다. 다시 말해, AI/ML은 범용 기술(General Purpose Technology)의 특성을 지니며, 개별 응용 분야에서의 활용뿐만 아니라 경제·사회 전반에 걸쳐 광범위하고 근본적인 영향을 미치고 있는 것으로 평가된다(Brynjolfsson, 2017). 특히 Babashahi et al.(2024)에 따르면 AI 기술의 도입은 거의 모든 직업군의 업무 내용과 속

런 요구에 영향을 미치는 변화를 가져오고 있다는 지적도 있다.

나. 금융 분야와 IT 인력의 교류 : 핀테크를 통한 융합

분석 결과, IT 전문가와 금융업 간의 활발한 인력 교류가 두드러진 현상으로 나타났다. 예컨대 응용 소프트웨어 개발자와 증권 사무원 간의 인력 전이와, 웹 개발자와 증권 사무원 간의 전이가 상대적으로 높은 전이 확률을 보였는데, 이는 금융권과 IT 분야 간 인력 이동이 빈번함을 의미한다. 이러한 현상은 핀테크(FinTech) 산업의 성장과 금융업의 디지털 전환이 실제 노동 이동을 통해 구현되고 있음을 방증한다. Jiang et al.(2025)의 연구에서도 핀테크가 전통적 산업 경계를 허물며, 금융 지식과 IT 기술을 겸비한 인재에 대한 수요를 창출하고 있음을 확인할 수 있다. 핀테크 도입이 금융과 기술의 융합 숙련을 요구함에 따라 기업들은 신규 채용과 내부 재교육을 통해 인력을 업스킬링하고, 조직 혁신을 추진하는 것으로 보고되었다. 이처럼 금융도 메인 지식을 보유한 인력이 IT 전문성을 습득하거나, 반대로 IT 배경의 인력이 금융 지식을 습득하여 융합형 인재로 거듭나는 패턴이 나타나고 있다. 실제 분석에서도 증권 사무원에서 컴퓨터시스템 전문가로의 전이, 은행 사무원에서 정보통신 관리자로 전이가 유의미하게 확인되었는데, 이는 금융 분야 인력이 AI/ML 등 정보기술 역량을 갖추어 새로운 역할로 이동하는 사례라 할 수 있다. 이러한 금융-IT 인력 교류는 핀테크 혁신의 인력 측면 변화를 보여주는 동시에, 산업 간 복합 숙련에 대한 수요 증가를 뒷받침한다.

다. 제조 및 전자 산업에서의 AI 융합

제4차 산업혁명의 영향 아래, 스마트 제조업 및 전자 산업에서도 AI/ML 기술과 전통적 하드웨어 중심 직종 간의 융합이 뚜렷하게 나타나고 있다. 본 분석에서 자동조립라인·산업용로봇 조작원에서 컴퓨터시스템 전문가로의 전이, 전자공학 기술자 및 연구원에서 컴퓨터시스템 전문가로의 전이가 유의미하게 나타난 것은 단순한 자동화가 아닌, 기술 수요의 재구성과 복합 역량의 확산을 보여준다. 이는 기존 기계 조작 중심의 직종들이 데이터 분석

이나 소프트웨어 응용 역량을 포함한 융합형 인재로 전환되는 경향을 반영한다.

이러한 변화는 해외 연구들에서도 일관되게 확인된다. Xie et al.(2021)은 중국 제조업을 분석하여 AI 도입이 노동 수요에 비대칭적 영향을 주며, 특히 중간 숙련과 복합 기술을 요구하는 직무에서 새로운 수요가 증가한다고 보고하였다. 이는 본 연구에서 관찰된 기계적 자동화 직종에서 디지털 기반 직종으로의 전이가 단순 현상이 아니라 구조적 변화임을 뒷받침한다. 또 다른 연구인 Cui(2024)는 중국 텐진의 전자·제조업체를 대상으로, 생성형 AI가 생산 자동화뿐 아니라 제품 설계, 브랜드 구축, 마케팅 등 다양한 영역에 영향을 주고 있다고 밝혔다. 이는 제조업과 전자 산업이 더 이상 단순 생산 중심으로 머무르지 않고, AI 기술을 활용해 시장 대응과 부가가치 창출까지 확대되고 있음을 보여준다.

독일 제조업을 대상으로 한 Heimberger et al.(2025)의 연구 역시 같은 맥락을 제시한다. 연구에 따르면 조직의 기술적·제도적 준비도가 높을수록 AI 채택이 빠르게 이루어지며, 이 과정에서 숙련 구조와 직무 배치가 변화하고 있었다. 이는 하드웨어 중심의 직무가 소프트웨어 및 시스템 중심 직무로 재편되는 현상과 직접적으로 연결된다. 더 나아가 Nelson et al.(2023)의 체계적 문헌 고찰은 제조업 내 AI 활용이 예측 유지보수, 품질 검사, 공정 모니터링, IoT 통합 등 다양한 영역에서 확산되고 있으며, 이로 인해 노동자들의 역할 변화와 기술 재구성이 광범위하게 일어나고 있음을 보여준다.

결국 제조 및 전자 산업에서 AI/ML의 융합은 생산 현장의 자동화를 넘어 기술 숙련의 재편, 직무의 확장, 조직 운영 방식의 변화를 동시에 이끌어내고 있다. 본 연구에서 확인된 직종 간 전이 확률은 이러한 국제적 추세와 맥을 같이 하며, 한국 노동시장에서 나타나는 지능형 기술로의 숙련 전환이 세계적 변화의 일환임을 실증적으로 보여준다.

라. AI/ML 기술의 범용성 : 융합 숙련과 혁신적 변화

위와 같은 결과는 AI/ML 기술이 범용 기술로서 작동하고 있음을 실증적으로 보여준다. AI/ML 관련 숙련은 특정 산업이나 직종에 국한되지 않고 금

용, 제조, 전자, 미디어 등 광범위한 분야로 확산되고 있으며, 기존 전문성과 결합하여 새로운 형태의 융합 숙련을 창출하고 있다. 이는 AI/ML 기술이 단순한 특화 기술이 아니라, 기존 산업과 직무를 근본적으로 변화시키는 혁신적 기술임을 의미한다. 실제 경제학계에서도 AI를 전기나 내연기관과 같은 역사적 범용 기술에 비견하며, 개별 응용에서의 효용보다도 경제 전반의 구조 변화를 이끄는 잠재력에 주목하고 있다(Brynjolfsson, 2017). Soni et al.(2020)의 연구에서는 이미 거의 모든 산업이 AI 기술을 활용하고 있고, 기업들이 스스로를 ‘AI 기업’으로 표방하는 추세까지 나타난다고 지적한다. 이러한 범용성으로 인해 AI 도입에 따른 파급 효과가 금융, 제조, 서비스, 교육 등 전 영역에서 관찰되고 있으며, 그 과정에서 새로운 직무와 숙련의 형태가 계속해서 나타나고 있다. 또한 AI 기술 확산의 영향으로 기존 직무의 역할과 요구 역량이 재편되면서, 기술 발전은 단순히 새로운 직종을 만들어내는 것에 그치지 않고 기존 직종의 성격을 근본적으로 변화시키고 있음을 알 수 있다.

마. 직종 간 전이 네트워크

본 연구에서 파악된 AI/ML 숙련의 직종 간 전이는 단순한 일대일 이동이 아닌 복잡한 네트워크 구조를 형성하고 있다. 특정 직종들은 다수의 다른 직종과 연결되며 숙련 전파의 허브(hub) 역할을 수행하는데, 이러한 허브 직종을 통해 다양한 기술 배경을 가진 인력들이 서로 이어지고 교류함으로써 전체 숙련 확산이 가속화되고 있다. 노동시장 네트워크에 대한 데이터 기반 모형 연구에 따르면, 직업 간 이동 가능 경로의 구조는 노동자가 새로운 일자리로 전환할 수 있는 기회에 중요한 영향을 미치며, 네트워크상에서 위치에 따라 전이 경로가 많은 직종과 적은 직종의 차이가 큰 것으로 나타난다(del Rio-Chanona et al., 2021). 이는 곧 네트워크 중심에 위치한 직종일수록 다양한 방향으로 인력 이동을 연결하여 숙련 확산의 관문 역할을 할 수 있음을 시사한다. 본 분석에서 관찰된 바와 같이, AI/ML 숙련과 관련하여 몇몇 직종들이 이러한 중심 노드로 기능하고 있으며, 이들은 산업 내외의 지식과 기술을 연결하여 융합형 인재의 배출을 촉진하고 있다. 이러한 네

트위크 효과는 AI/ML 관련 인력 양성 전략 수립에 있어 특정 허브 직종의 전략적 중요성을 부각시킨다.

궁극적으로, 본 연구 결과는 기술 진화가 직종 구조에 미치는 영향을 명확히 보여준다. 하드웨어 중심의 전통적 기술 직종에서 소프트웨어와 융합된 지능형 기술 직종으로의 전환이 활발히 일어나고 있으며, 이는 기술 발전이 단순히 새로운 직종을 창출하는 것을 넘어 기존 직종의 내용과 요구 역할을 근본적으로 변화시키고 있음을 의미한다. 또한 AI/ML 관련 숙련의 직종 간 전이는 산업 경계를 넘나드는 광범위한 융합 현상이 실증적으로 확인되었다. AI/ML 기술은 범용 기술로서 기존 전문성과 결합하여 새로운 형태의 융합 숙련을 만들어내고 있으며, 이 과정에서 특정 직종들이 허브 역할을 수행하여 전체적인 네트워크 효과를 창출하고 있다. 이러한 결과는 AI/ML 시대의 인력 개발 정책이 단순한 기술 교육을 넘어 산업 간 융합과 직종 간 전이를 고려한 통합적 접근이 필요함을 시사한다. 특히 허브 역할을 하는 직종에 대한 전략적 투자와 융합형 인재 양성을 위한 정책적 지원이 AI/ML 관련 숙련의 효과적 확산을 위해 필수적일 것으로 판단된다. 이는 정부와 교육기관, 기업이 협력하여 교차 분야(skill crossover) 교육 프로그램을 개발하고, 재교육 및 전환 지원 시스템을 강화해야 함을 의미한다. 이러한 통합적 인력 양성 전략을 통해서만 산업 전반에 걸친 AI/ML 기술의 포용적 확산과 그로 인한 경제·사회적 가치 창출을 극대화할 수 있을 것이다

제5절 요약과 결론

본 연구는 한국 숙련사전을 기반으로 6,558개 숙련에 대한 AI 노출도를 측정하고, 네트워크 분석을 통해 숙련 간 구조적 특성과 직업 간 전이가능성을 체계적으로 분석하였다. AI 노출도 분석 결과, ‘관리및지원’과 ‘정보기술’ 숙련군이 가장 높은 노출도를 보인 반면, ‘개인돌봄서비스’와 ‘소프트스킬’은 가장 낮은 노출도를 나타냈다. 특히 반복적이고 규칙성이 높은 업무일수록 AI 대체 가능성이 높았으며, 인간의 감정적 지능과 창의성, 신체적

접촉이 필요한 영역은 여전히 AI 대체가 어려운 것으로 확인되었다.

숙련 네트워크 분석에서는 2021년과 2024년 IT 분야 숙련 구조의 중요한 변화가 관찰되었다. 2021년에는 '빅데이터 컴퓨팅'이 네트워크의 중심이었으나, 2024년에는 '통합 보안 관리'가 최대 중심성을 갖게 되어 데이터 중심 시대에서 보안 중심 시대로의 패러다임 전환을 보여주었다. 이는 디지털 전환 가속화와 사이버보안 위협 증가라는 현실적 변화를 반영하는 것으로, IT 전문가들에게 보안 기술에 대한 기본적 이해가 필수가 되었음을 시사한다.

네트워크 특성 지표 분석에서는 숙련 확산도와 페이지랭크 중심성 모두에서 '고객맞클라이언트지원'과 '관리및지원' 숙련군이 최상위를 차지하였다. 이는 현대 비즈니스 환경에서 고객 중심적 접근과 디지털 업무 처리 능력이 산업 전반에 걸쳐 핵심 인프라 역할을 수행하고 있음을 보여준다. 흥미롭게도 이러한 네트워크 중심성과 AI 노출도 간에는 강한 상관관계가 나타나지 않았는데, 이는 AI 대체 가능성이 높은 숙련일수록 현재 노동시장에서 광범위하게 요구되는 역설적 상황을 반영한다.

직업 간 숙련 전이가능성 분석에서는 AI/ML 기술이 범용 기술로서 IT 영역을 넘어 금융, 제조, 전자 등 다양한 산업으로 확산되는 현상이 확인되었다. 특히 핀테크를 통한 금융-IT 인력 교류, 스마트 제조업에서의 하드웨어-소프트웨어 융합, 그리고 전자 산업에서의 AI 기술 접목이 두드러진 특징으로 나타났다. 이러한 융합 현상은 단순한 직종 이동을 넘어 기존 직종의 성격을 근본적으로 변화시키고 있으며, 특정 허브 직종을 통해 숙련 확산이 가속화되는 네트워크 효과를 창출하고 있다.

본 연구의 결과는 AI 시대 인력 개발 정책에 중요한 시사점을 제공한다. 첫째, AI 기술 발전이 모든 숙련을 획일적으로 대체하는 것이 아니라 숙련의 성격에 따라 차별적 영향을 미치므로, 숙련별 맞춤형 대응 전략이 필요하다. 둘째, 네트워크 중심성이 높은 기반 숙련에 대한 지속적 투자와 보안 기술의 범용적 교육이 요구된다. 셋째, AI/ML 기술의 범용성을 고려하여 산업 간 융합과 직종 간 전이를 지원하는 통합적 재교육 시스템이 구축되어야 한다. 특히 허브 역할을 하는 직종에 대한 전략적 지원과 교차 분야 교육 프로그램 개발을 통해 AI 기술의 포용적 확산과 경제·사회적 가치 창출을 극대화할 수 있을 것이다.

제3장

숙련-직업 간 네트워크의 '임계적' 양극화

제1절 서론

최근 숙련을 독립적인 개인 속성이 아니라 상호 연결된 체계로 이해하려는 네트워크 접근이 숙련 연구 분야에서 빠르게 확산되고 있다. 이러한 접근은 숙련 간 연관성과 보완성이 노동시장 내 구조적 변화, 직업 간 이동, 그리고 불평등의 동학을 설명하는 핵심 요인임을 전제한다. 특히 Alabdulkareem et al.(2018)과 Henning et al.(2025)은 숙련의 관계적 구조를 각각 다른 방식으로 분석함으로써 숙련 체계의 복합적 속성을 보여준다. 전자는 기능적 상호 의존성을 반영한 숙련 보완성(functional complementarity) 기반 네트워크를 구축하여 숙련의 구조적 위계를 분석한다. 후자는 비정형 텍스트 데이터로부터 의미적 유사성을 학습하여 의미론적(semantic) 직업 간 숙련 연관성(relatedness) 네트워크를 구성하고, 이러한 숙련 연관성과 경제적 성과 간 관계를 탐색한다.

본 연구는 이러한 두 접근법의 장점을 통합하여, 숙련과 직업을 상호 연결된 기능적 네트워크로 분석하고자 한다. 이를 위해 한국고용정보원이 발간한 '2017~2020년 한국직업정보(Korea Network for Occupations and Workers, KNOW) 재직자조사' 자료를 활용해 한국 노동시장의 숙련과 직업 구조를 네트워크 관점에서 재구성한다. KNOW 자료는 약 2만 명의 재직자

를 대상으로 직업별 숙련의 중요도와 수준을 조사한 국내 최대 규모의 데이터로, 숙련의 인지적·사회적·정서적·신체적 속성을 포괄적으로 반영한다는 점에서 기존 O*NET 기반 연구보다 한국적 현실에 더 적합하다.

본 연구의 목적은 첫째, 루뱅(Louvain) 알고리즘을 적용하여 숙련 간 보완성 네트워크 내에서 위상적으로 결속된 숙련 군집을 식별하고, 이를 통해 숙련 체계의 구조적 양극화와 기능적 분화를 규명하는 것이다. 둘째, 직업별 핵심 숙련 리스트를 추출하고, 이를 벡터화하여 직업 간 기능적 숙련 연관성 행렬을 구성함으로써, 직업 간 숙련 구조의 위계적 관계를 시각화한다. 셋째, 이러한 기능적 보완성·연관성 행렬을 기반으로 최소신장나무(Minimum Spanning Tree, MST)와 임계값(threshold) 기반 핵심 연결망(backbone) 네트워크를 구축하여, 전체 네트워크의 최소 연결 구조와 임계적 결속 구조 간 차이를 시각적으로 비교한다. 넷째, 임금 수준과 격차의 요인 분석에서 직업별 숙련 유형 구성과 숙련 연관성이 미치는 효과를 추정한다.

본 연구는 숙련의 구조적 상호 의존성을 통해 노동시장 내 불평등과 구조적 전환의 동학을 이해하려는 시도이다. 이를 통해 한국의 숙련과 직업 네트워크가 어떤 방식으로 결속되고 분화되는지를 규명하고, 인지적·사회적·신체적 요소가 결합된 숙련 구조가 노동시장의 양극화, 임금 격차, 직업 구조의 재편에 미치는 함의를 분석한다. 다시 말해, 본 연구는 숙련을 협의의 기술 능력으로 한정하지 않고, 노동시장에서 요구되는 인지적·사회적·정서적·신체적 역량의 통합체로 확장하여 이해하며, 이를 바탕으로 한국형 숙련 체계의 구조적 특성과 불균형의 근원을 네트워크 관점에서 조명하고자 한다.

이를 위해, 제2절에서는 네트워크 관점에서의 숙련 연구를 간략히 검토하고 주요 선행 연구의 흐름과 한계를 정리한다. 제3절에서는 본 연구에서 활용하는 자료와 분석 방법을 제시하며, 이를 토대로 숙련 및 직업 네트워크의 구조적 특징과 양극화 패턴을 실증적으로 분석하고, 수행된 분석 결과를 종합적으로 해석하고, 마지막으로 제4절의 소결에서는 주요 분석 결과를 요약하고 함의와 시사점을 제시한다.

제2절 선행 연구

숙련과 노동시장 구조를 이해하기 위한 최근 연구는 개별 숙련을 고립된 것으로 보기보다는 상호 연계된 관계망으로 파악하는 방향으로 나아가고 있다. 사회 네트워크 분석은 이러한 관계적 시각을 구현할 수 있는 대표적 방법으로, 숙련 간 상호 의존 구조와 정보 흐름 패턴, 그리고 노동시장에서의 숙련 재편 과정과 전이 구조를 이해하는 데 활용되고 있다(Alabdulkareem et al., 2018; Maghsoudi, 2024; Henning et al., 2025).

이러한 네트워크 접근은 숙련을 하나의 시스템으로 간주하여, 직업 내에서 두 숙련이 함께 요구되는 빈도를 기반으로 숙련 간 연계 강도와 보완성과 직업 간 숙련 연관성을 측정하고, 여러 중심성 지표를 통해 각 숙련이 전체 네트워크에서 차지하는 위치와 영향력을 평가한다(Alabdulkareem et al., 2018; Henning et al., 2025). 이러한 네트워크 기반 분석은 노동시장 양극화를 단순히 임금 격차가 아닌 숙련 구조의 분절화로 해석할 수 있게 하며, 숙련 간 거리를 활용해 직업 전환 경로와 이동성의 제약 조건을 설명하는 근거로 활용될 수 있다

Alabdulkareem et al.(2018)은 이러한 네트워크 관점을 숙련 연구에 선구적으로 도입하고 있다. 그들은 미국의 O*NET 데이터를 활용하여 각 직업의 숙련 중요도 집합을 추출하고, 이를 기반으로 숙련 간 보완성 관계를 연결망 형태로 구성한 '숙련 공간(skillscape)'을 구축한다. 이러한 네트워크 접근을 통해 데이터 기반 방식으로 숙련 간 상호 의존 구조를 규명하고, 숙련이 서로 밀접히 연결된 몇 개의 명확한 군집으로 구분된다는 사실을 확인한다. 특히 이러한 분석은 숙련 구조 분화가 직업 계층 양극화와 중간층 붕괴(hollowing of the middle class)와 밀접히 연관되어 있음을 보여주며, 노동시장 양극화 기제를 숙련 네트워크 관점에서 설명하는 새로운 통찰을 제시한다.

Maghsoudi(2024)는 컴퓨터 과학 분야를 대상으로 LinkedIn 구인 자료를 분석하여 숙련 네트워크를 구축한다. 약 7,700건 구인광고에서 도출된

1,300여 개 숙련 간 공출현 관계를 분석한 결과, 컴퓨터 과학의 업무는 제너럴리스트, 소프트웨어 개발, 인프라·보안, 임베디드 시스템 등 네 개 군집으로 구분된다. 특히 커뮤니케이션, 영어, SQL이 중심 노드로 나타나, 기술적 역량과 더불어 연성 숙련(soft skill)이 숙련 구조 내에서 핵심 역할을 하고 있음을 보여준다. 이 연구는 일반적 숙련 수요가 감소하고 전문화된 숙련이 점차 중심으로 이동하는 패턴을 확인함으로써, 네트워크 분석이 숙련 수요 변화의 방향성을 조기에 탐지하는 도구가 될 수 있음을 제시한다.

Henning et al.(2025)은 분석 단위를 개별 업무에서 지역경제로 확장하여, 숙련 연관성과 지역 성과 간 관계를 탐색한다. 스웨덴의 대규모 구인 데이터를 바탕으로 Doc2Vec 임베딩과 코사인 유사도를 이용해 직종 간 숙련 유사도 행렬을 구축하고, 이를 통해 지역의 숙련 일관성(skill coherence)을 측정한다. 분석 결과, 숙련 일관성이 낮을수록, 즉 숙련 구성이 다양할수록, 평균 임금과 생산성이 높게 나타나며, 이는 고학력 근로자 비중과 밀접하게 연관된다. 반면에, 숙련 연관성이 높은 업무로 이동하는 근로자 비율이 높은 지역 또한 임금 수준이 높게 나타나, 지역 내 다양성과 전환 가능성 공존이 해당 지역의 경제적 성과를 좌우하는 주요 요인 중 하나임을 보여준다. 이 연구는 숙련 네트워크 분석이 지역의 경제적 회복력과 혁신 역량을 진단하는 유용한 지표로 활용될 수 있음을 예증한다.

한편, Huang et al.(2021)은 산업 간 숙련 전이가능성을 분석하기 위해, 현대 산업 근로자의 숙련 구조를 중심으로 산업 간 이동 경로를 추정한다. 미국의 O*NET 데이터를 활용해 숙련 간 네트워크를 구성한 결과, 현대 산업에서 활용되는 연성 숙련(예: 능동적 경청, 말하기, 비판적 사고 등)은 다른 산업에서도 유효한 전이가능한 숙련(portable skills)으로 나타난다. 특히 현대 산업과 보건 과학 분야 간 연계가 가장 강하게 관찰되어, 해당 산업 근로자가 IT 부문보다 보건 부문으로 전환할 가능성이 높음을 시사한다. 이러한 결과는 네트워크 접근이 산업 간 전환성과 재교육 경로 설계에 실질적 근거를 제공할 수 있음을 보여준다.

국내 연구로 이현경(2025)은 국내 AI 숙련 네트워크 분석의 방법론적 제약을 비판적으로 검토함으로써, 이 분야 연구의 한계와 시사점을 제시한다. 기존 해외 연구가 LinkedIn이나 Lightcast 등 글로벌 플랫폼 데이터를 활용

해 AI 숙련을 측정하는 방식과 달리, 한국의 구인 광고 시장은 데이터 편향이 크고 포괄성이 낮다는 점을 실증적으로 제시한다. 예를 들면, 국내 AI 관련 구인공고의 약 3분의 1이 주요 구직 플랫폼에 게시되지 않아, 네트워크 분석의 표본 대표성 부족이 결과 해석의 일반화를 제한할 수 있음을 지적한다. 이 연구는 데이터 기반 네트워크 분석의 신뢰성을 높이기 위해 국가 단위의 숙련 데이터 인프라 구축과 다원적 자료 통합 전략이 필요함을 강조한다.

제3절 실증 분석

1. 분석 방법

가. 숙련 간 보완성과 직업 간 연관성 네트워크 구축

본 연구는 숙련 간 보완성 네트워크와 직업 간 숙련 연관성 네트워크를 각각 구축한 Alabdulkareem et al.(2018)과 Henning et al.(2025)의 방법론을 준용하여, 두 유형의 네트워크를 구축하고자 한다. 이하에서는 두 연구에서 제시한 분석 절차를 간략히 요약하고, 두 네트워크 접근법의 차이를 비교하여 제시한다.

Alabdulkareem et al.(2018)은 숙련 간 관계를 개별 목록이 아닌 상호 연계된 네트워크 구조로 파악해야 한다는 관점에서 미국 O*NET 데이터를 이용하여 숙련 네트워크를 선구적으로 구축한다. 이 연구는 직업별 숙련 구성을 데이터 기반으로 추출하고, 숙련 간 보완성 정도를 수치화하여 노동시장의 구조적 변화를 분석한다.

O*NET 데이터는 미국 표준직업분류(SOC)에 따른 672개 직종에 대해 161개 숙련(skill), 지식(knowledge), 능력(ability)의 상대적 중요도를 0~1 범위로 제시하며, 직업-숙련 간 관계를 연도별로 측정된 직업-숙련 행렬($onet(j, s)$)을 제공한다. 이들은 O*NET 자료를 활용하여 현시비교우위지

수(Revealed Comparative Advantage, RCA)를 통해 각 직업의 특화 숙련을 추출하고, 이러한 특화 숙련 간 공출현(co-occurrence) 관계를 기반으로 보완성 지수를 산정하고, 이를 토대로 숙련 간 네트워크를 구축한다.

각 직업이 특정 숙련에 얼마나 상대적으로 특화되어 있는가를 파악하기 위해 RCA 지수를 산출하는데, 그 식은 등식 (1)과 같다.

$$RCA_{j,s} = \frac{x_{j,s} / \sum_s x_{j,s}}{\sum_j x_{j,s} / \sum_{j,s} x_{j,s}} \quad (1)$$

여기서 $x_{j,s}$ 는 직업 j 에서 숙련 s 의 중요도이며, 0과 1 사이에 있다. 이 등식에서 $RCA_{j,s} > 1$ 이면 해당 숙련이 그 직업에서 평균 이상으로 강조된 특화 숙련으로 간주된다. 이 조건을 충족할 경우 $e_{j,s} = 1$, 그렇지 않으면 0으로 이진화하여 직업-숙련 행렬을 단순화한다. 그런데 이 연구는 기존 연구와 달리 $RCA_{j,s} > 1$ 인 특화 숙련의 값을 그대로 유지하여 직업-숙련 간 가중행렬을 구성한다.

동일 직업에서 자주 함께 사용되는 두 숙련 간 보완성은 조건부 공출현 확률을 기반으로 등식 (2)와 같이 정의된다(Alabdulkareem et al., 2018: 2). 이는 두 숙련이 동시에 등장하는 직업 비율을 측정하는 값으로, 값이 1에 가까울수록 두 숙련이 함께 요구되는 빈도가 높다.

$$q(s_1, s_2) = \frac{|j: e_{j,s_1} = 1, e_{j,s_2} = 1|}{\max(|j: e_{j,s_1} = 1|, |j: e_{j,s_2} = 1|)} \quad (2)$$

이러한 숙련 쌍 간 보완성 지수를 행렬로 구성하면, 절절은 숙련, 간선은 $q(s_1, s_2)$ 로 정의된 숙련 간 보완성 네트워크가 된다. 보완성 지수가 높은 숙련 쌍은 동일 직업 내에서 함께 발휘될 때 생산성을 높이거나 상호 습득이 쉬운 숙련으로 볼 수 있다. 예컨대, '데이터 분석'과 '통계 지식'은 서로 높은 보완성이 있지만, '수학적 추론'과 '신체 근력'은 거의 함께 나타나지 않는다.

반면에, Henning et al.(2025)은 숙련 간 공출현 행렬 대신, 숙련 유사

성을 기반으로 직업 간 연관성 네트워크를 구축한다. 이는 스웨덴 대규모 구인공고 데이터에 기반한 실제 노동시장 구조를 반영한다는 점에서 Alabdulkareem et al.(2018)과 차별적이다. Henning et al.(2025)의 분석은 직업별 숙련 정보를 벡터화한 후에, 의미적 연관성을 파악하기 위해 Doc2Vec 기반 코사인 유사도(cosine similarity)를 이용하여 직업 간 연관성 행렬을 구축하고, 마지막으로 MST를 통해 핵심 연계 구조를 시각화하는 절차로 이루어진다.

이들은 우선 각 구인공고에서 추출된 숙련 어휘를 정제된 뒤, 직업-숙련 발생 행렬을 구축한다. 이후 각 직업 내에서 상대적으로 많이 등장한 특화 숙련을 식별하기 위해 Alabdulkareem et al.(2018)과 같이 RCA 지수를 이용해 $RCA > 1$ 인 숙련을 해당 직업의 특화 숙련으로 간주하고 이를 추출한다.

각 직업의 숙련 리스트를 Doc2Vec(Le and Mikolov, 2014) 임베딩으로 벡터화하고, 등식 (3)과 같이, 직업 간 코사인 유사도를 산출하여 직업 간 연관성 행렬을 생성한다. 이는 숙련이 유사할수록 직업 간 이동 가능성이 높다는 가정에 근거한다.

$$R(i, j) = \frac{v_i \cdot v_j}{\|v_i\| \|v_j\|} \quad (3)$$

여기서 v_i 와 v_j 는 각각 직업 i 와 j 의 숙련 벡터다.

요약하면, Henning et al.(2025)은 구인공고와 같은 비정형 텍스트 데이터를 활용해 직업 간 숙련 연관성 네트워크를 구축한다. 이들은 구인공고에서 추출한 숙련 어휘를 Doc2Vec 임베딩으로 벡터화하고, 코사인 유사도를 통해 직업 간 관련성을 계산하여 직업-숙련 연관성 행렬을 구축한다. 이때 임베딩은 단어 간 문맥적 의미를 학습하여 잠재 의미공간을 형성하고, MST를 통해 네트워크의 핵심 구조를 시각화한다.

반면에, 여기서 사용하게 될 자료는 이미 숙련별 중요도가 수치화된 정형 데이터이므로, 텍스트 임베딩 과정을 거치지 않고 숙련 벡터 간 기능적 유사성을 직접 계산한다. 즉, 문맥 의미를 학습하여 유사성을 추정하는 대신, 각 직업이 요구하거나 중시하는 숙련의 구성 패턴을 기반으로 직업 간 구조적 유사성을 측정한다. 이러한 접근은 Doc2Vec 기반의 의미적 유사성과는 달

리, 숙련 지표의 구조적 관계를 반영하여 직업 간 거리와 군집을 드러낸다.

따라서 Henning et al.(2025)의 방법이 비정형 텍스트로부터 잠재 의미공간을 학습하는 의미론적 네트워크를 구성한다면, 본 연구는 정형화된 숙련 중요도 지표를 이용해 기능적 네트워크를 구축한다는 점에서 차이가 있다. 이 접근은 데이터의 구조적 제약을 반영하면서도 숙련과 직업 간 연계 구조를 탐색할 수 있게 한다.

전술한 두 접근법은 모두 숙련을 네트워크 형태로 구조화하지만, 관심사가 다르다. Alabdulkareem et al.(2018)은 숙련 자체의 내적 결합 구조를 분석하여 숙련 체계의 기능적 양극화를 드러내지만, Henning et al.(2025)은 이러한 숙련을 매개로 한 직업 간 구조적 연관성을 탐색함으로써 노동시장의 공간적·직업적 일관성(coherence)을 규명한다.

두 연구 모두 숙련을 독립적 속성이 아닌 상호 연결된 체계로 이해한다는 점에서 의미가 있으며, 본 연구는 이 두 관점을 통합해 숙련-직업 간 네트워크를 기능적 관점에서 분석하고자 한다.

나. 루뱅 방식에 의한 네트워크 내 클러스터 식별

Alabdulkareem et al.(2018)은 숙련 간 보완성 행렬을 구축한 후, 루뱅 커뮤니티 탐지 알고리즘을 적용해 숙련 네트워크 내에서 상호 보완적인 숙련 클러스터를 식별한다. 분석 결과, 숙련은 '인지·사회적' 군집과 '감각·신체적' 군집으로 구분되며, 이는 현재 미국 노동시장에서 나타나는 양극화를 가시적으로 드러낸다.

이와 같이 숙련 보완성 또는 직업 간 관련성 네트워크를 구축한 이후, 네트워크의 위상적 구조를 해부하고 이해하기 위해 내부적으로 결집된 집단, 즉 군집을 식별하는 과정이 필요하다. 이 연구 역시 이러한 목적에서 Blondel et al.(2008)과 Lambiotte et al.(2010)이 제안한 루뱅 알고리즘을 적용하여, 숙련과 직업 네트워크 내 군집을 탐지하고 그 위계적 패턴을 분석한다.

대부분의 네트워크는 여러 개의 모듈(modules)로 구성되어 있으며, 각 모듈 내부에서는 연결이 조밀하지만, 서로 다른 모듈 간에는 연결이 상대적으

로 많지 않다. 이런 특성을 정량화하는 지표가 모듈성(modularity, Q)이다. 이는 네트워크를 일정한 수의 커뮤니티로 분할했을 때, 각 커뮤니티 내부의 연결 밀도가 전체 네트워크의 평균 연결 수준을 얼마나 넘어섰는지를 측정하며, 주어진 네트워크 구획의 질(quality of a partition)을 평가한다. 일반적으로 Q 값이 1에 가까울수록 내부 결집이 강하고 외부 연결이 약한, 즉 뚜렷하게 구획된 네트워크 구조를 시사한다. 이는 등식 (4)와 같이 직관적으로는 모듈 내 결절의 수와 그러한 결절의 기댓값을 차감한 것이다.

$$Q = \frac{1}{2m} \sum_{i,j} [A_{ij} - \frac{k_i k_j}{2m}] \delta(c_i, c_j) \quad (4)$$

여기서 A_{ij} 는 결절 i 와 j 간 연결 가중치, k_i 는 결절 i 에 연결된 모든 변의 가중치 합, m 은 네트워크 전체 변의 가중치 합이다. c_i 와 c_j 는 각각 결절 i 와 j 가 속한 군집을 의미하며, $\delta(c_i, c_j)$ 는 두 결절이 동일 군집에 속하면 1, 그렇지 않으면 0을 갖는 크로네커 델타(Kronecker delta) 함수다.

루벵 알고리즘은 Q 를 극대화하는 방향으로 네트워크를 반복적으로 분할하는데, 두 단계로 구성된다. 첫 단계에서 각 결절은 고유한 군집에 속한다고 가정한다. 그다음 인접한 결절로 이동했을 때 모듈성이 증가하는지를 평가하며, 모듈성이 더 커지는 경우 해당 이동을 유지한다. 이 과정을 모든 결절에 대해 반복하여 더 이상 모듈성이 개선되지 않을 때, 국지적 최적화(local optimization)로 소규모 군집이 형성된다.

두 번째 단계에서는 형성된 각 군집을 새로운 단일 결절로 축약하여 요약 네트워크를 구성한다. 이때 동일 군집 내 연결은 자기 루프(self-loop)로, 다른 군집 간 연결은 군집 간 가중치로 표현된다. 이후 동일 절차가 이 축약 네트워크에 반복적으로 적용되며, 모듈성이 더 이상 증가하지 않을 때까지 반복함으로써 위계적 군집 구조가 나타난다(Blondel et al., 2008).

루벵 알고리즘은 계산 효율성이 높고, 대규모 네트워크에서도 위계적 구조를 안정적으로 식별할 수 있다. 하지만, Lambiotte et al.(2010)은 네트워크 규모가 커질수록 작은 커뮤니티가 인식되지 않는 문제를 지적하며, 분석 목적에 따라 해상도(resolution) 매개변수(γ)를 조정하는 방법을 제안한다.

일반적으로 $\gamma=1$ 을 기준으로 하되, 값을 높이면 더 큰 규모의 군집이, 낮추면 더 미세한 하위 구조가 식별된다. 이러한 조정 과정을 통해 루뱅 방식은 복잡한 네트워크 내에서 다층적이고 위계적인 구조를 효율적으로 탐지할 수 있다.

요약하면, 루뱅 알고리즘은 네트워크의 위상적 구조를 기반으로 결절 간 결집도를 최적화함으로써 숙련이나 직업과 같은 결절이 서로 밀접히 연관된 하위구조를 식별한다. 이 과정에서 도출된 군집은 숙련 체계 내부의 구조적 결집, 분절화, 그리고 기능적 분화를 파악하는 데 중요한 단서를 제공한다.

다. 숙련과 직업 간 네트워크 구조의 시각화

숙련-직업 간 네트워크 구조를 시각화하기 위해서는, 복잡한 연결 관계 속에서 핵심 연결망을 식별하는 과정이 필요하다. 네트워크의 핵심 연결망은 전체 구조를 대표하면서도 불필요한 연결을 제거하여서 관계의 본질적 패턴을 드러내는 최소 연결망을 의미한다. 여기서는 이를 구현하기 위해 두 가지 접근 방식을 병행한다.

첫째는 MST를 활용하는 방식이며, 둘째는 임계값(threshold)을 설정하여 일정 수준 이상의 연결만 남기는 방식이다. 두 방법 모두 네트워크의 복잡성을 줄이면서, 상호 관련성이 높은 숙련 또는 직업 간 핵심 구조를 추출하기 위한 것이다.

MST 방식은 네트워크를 구성하는 모든 결절을 최단 거리 또는 최대 유사도 기준으로 연결하되, 불필요한 순환(cycle)을 제거하여 가장 단순한 형태의 연결망을 생성한다(Onnela et al., 2003). 이때 각 결절은 분석 단위(예: 직업 또는 숙련)를 의미하고, 두 결절 간 변(edge)은 보완성 또는 연관성을 나타낸다. MST는 이러한 관계를 위상학적 구조 형태로 표현함으로써, 복잡한 네트워크 속에서도 핵심 연결 경로를 식별할 수 있도록 한다. 즉, 모든 결절이 최소 비용 또는 최대 유사도로 연결된 하나의 나무(tree) 형태로 구조화되며, 이는 네트워크의 핵심 골격 구조를 파악하는 데 자주 활용된다.

임계값을 적용하는 방식은 네트워크 밀도가 높거나 복잡도가 높은 경우에 사용된다. 보완성 또는 연관성 행렬에서 특정 임계값 이상(예: 0.60, 0.65

또는 0.70)을 만족하는 관계만 남기면, 전체 연결망 중에서 강한 연계 관계만 시각적으로 강조할 수 있다. 이러한 방식은 관계 분포가 연속적일 때 유용하며, 약한 연결을 제거함으로써 주요 군집 간 상호작용 구조를 명확히 드러낼 수 있다.

전술한 바와 같이, MST는 그래프상 모든 결절을 최단 거리 또는 최대 유사도 기준으로 연결한 최소 연결망으로 복잡한 네트워크를 단순화해 결절 간 연관 구조를 명확히 드러낸다(Onnela et al., 2003; 정준호, 2014). 따라서 MST는 방대한 정보망 속에서 개별 요소 간 관계를 불필요한 경로 없이 가장 효율적으로 연결함으로써, 네트워크 구조의 본질적 관계와 패턴을 직관적으로 파악할 수 있도록 하는 도구다. 이에 대해 간단히 기술하면 다음과 같다.

직업을 여러 속련으로 구성된 속성(attributes) 집합으로 보면, 각 직업 i 에 여러 속련 요소로 구성된 속성 벡터 $X_i = (x_{i1}, \dots, x_{in})$ 가 있으며, x_{is} 는 직업 i 에서 속련 s 의 중요도나 강도를 나타낸다. 이러한 속성 벡터 집합의 위상학적 구조는 일련의 꼭짓점(V)과 변(E)으로 구성된 연결 그래프 $G = (V, E)$ 를 결정한다. 두 속련 s_i 와 s_j 가 서로 연관되어 있다면, 대응하는 꼭짓점 v_i 와 v_j 는 하나의 변 e_{ij} 으로 연결된다. 두 속성 벡터 간 비유사성은 유클리드 거리로 정의되며, 이는 변 e_{ij} 의 가중치로 표현된다. 속련 s_i 와 s_j 간 유클리드 거리는 등식 (5)와 같다.

$$d(i, j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (5)$$

그래프에서 한 결절 v_i 에서 다른 결절 v_j 로의 경로는 v_i 와 그것을 잇는 변들의 연쇄로 구성된다. 그래프 G 의 모든 결절 쌍에 대해 적어도 하나의 최단 경로가 존재하며, 이러한 경로 구조를 기반으로 연결된 결절의 하위 집합이 정의된다.

이처럼 그래프상 모든 결절을 최단 경로로 연결하는 간결한 구조, 즉 MST를 구축할 수 있다면, 이를 통해 속련 간 또는 직업 간 핵심 연결망도 도출할 수 있다. 따라서 MST는 모든 결절을 포함하면서 회로(circuit)가 존재

하지 않는 연결 그래프 중, 총연결 비용의 합이 최소가 되는 그래프를 의미한다(Aho et al., 1983). 여기서 비용은 변의 가중치, 즉 숙련 또는 직업 간 거리의 합으로 정의된다.

요약하면, MST는 그래프의 복잡성을 줄이고, 즉 불필요한 변을 제거하여 간결화하고, 가장 효율적인 연결 구조만 남긴 단순화된 위상적 표현이다. 이는 네트워크 내에서 유사성이 높은 숙련 또는 직업 간 핵심적 연계 구조를 유지하면서, 의미 없는 약한 연결을 제거하여 숙련-직업 네트워크의 본질적·구조적 관계를 파악할 수 있게 한다.

Henning et al.(2025)은 직업 간 숙련 연관성 행렬을 기반으로 MST와 상위 0.5%의 강한 연결을 결합하여 네트워크의 핵심 구조를 시각화한다. 이는 네트워크의 구조적 뼈대를 도출하는 데 초점을 두며, 전체 연결망 중 필수적인 경로만을 남겨 직업 간 관계의 최소 연결망을 표현한다. 반면, Alabdulkareem et al.(2018)은 보완성 지수가 일정 임계값(예: 0.6)을 넘어서는 숙련 간 관계만을 남겨, 기능적으로 상호보완적인 숙련 집합의 결집 구조를 시각화한다. 전자의 MST 방식은 네트워크의 구조적 최소 연결성을, 반면에 후자의 임계값 기반 방식은 기능적 보완 관계의 집중성과 군집 구조를 각각 부각한다.

본 연구는 두 접근의 장점을 받아들여, MST를 통해 네트워크 구조적 골격을 파악하고, 동시에 임계값 기반 필터링과 루벡 알고리즘을 활용해 기능적으로 밀접한 숙련·직업군의 클러스터를 식별한다. 이를 통해 숙련과 직업 네트워크의 위상 구조를 입체적으로 분석하고, MST 기반 최소 연결성과 임계값 기반 결집 구조 간 차이를 시각적으로 보여준다.

2. 자료

본 연구는 한국고용정보원이 발간하는 'KNOW 재직자조사' 자료를 활용한다. 이 조사는 미국의 O*NET을 벤치마킹하여 2001년부터 매년 실시되고 있으며, 산업 현장에서 실제로 요구되는 지식, 업무 활동, 업무 수행 능력, 가치관, 성격, 업무 환경, 흥미 등 다양한 직업 특성을 체계적으로 측정한다.

2017년부터 2020년까지의 4개 연도에 걸쳐 약 2만 명 재직자를 대상으로

직업별 숙련의 중요도와 수준을 조사했으며, 연도별로 주요 항목이 상이하 게 구성되어 있다. 2017년에는 업무 활동, 2018년에는 업무 환경과 흥미, 2019년에는 성격과 지식, 2020년에는 업무 수행 능력과 가치관을 중심으로 조사가 이루어졌다. 분석 단위는 한국표준직업분류(KSCO-7) 세세분류 397 개 직종이며, 각 항목은 해당 직업을 구성하는 숙련의 중요도를 5점 리커트 척도로 측정되어 있다.

KNOW 재직자조사는 O*NET과 마찬가지로 직업의 숙련 구조를 체계적으로 파악하기 위한 기초자료로 설계되었는데, 일부 항목(예: 성격, 가치관, 업무 환경 등)은 전통적인 숙련 개념(업무·지식·능력)과 명확히 구분되기 어렵다. 예를 들면, 성격 영역에는 리더십, 사회성, 자기 통제와 같은 사회적·인지적 숙련 요소가 포함되며, 가치관 항목에는 자율성, 이타성, 신체활동 등 업무 수행과 밀접히 연관된 특성이, 업무 환경 항목에는 신체 조정, 균형 유지, 손과 도구 사용 등 물리적 숙련에 해당하는 내용이 다수 포함된다. 이러한 이유로 본 연구는 숙련을 단순히 업무·지식·능력으로 한정하지 않고, 직업 수행에 실질적으로 요구되는 ‘확장된’ 숙련 속성으로 정의한다.

이에 따라 위의 항목 전반을 포괄하여 각 직업의 숙련 벡터를 구성하고, 각 항목의 5점 척도 값을 평균해 직업별 숙련 중요도의 대푯값으로 활용한다. 또한 부가적인 회귀 분석에 이용되는 직업별 연령, 교육연수, 연봉 등의 설명변수 자료는 ‘2020년 KNOW 재직자조사’에서 동일한 직업 단위로 추출·연계한다.

이처럼 숙련을 광의적으로 접근하는 것은 숙련을 다차원적·통합적 속성으로 이해하고자 하는 최근의 연구 흐름(예: Alabdulkareem et al., 2018)과도 부합한다. 즉, 특정 항목을 인위적으로 배제하기보다 실제 직무 수행에서 작동하는 복합적 숙련 체계를 포착하려는 분석 목적과도 조응한다.

이처럼, KNOW 데이터를 활용하는 방식은 Alabdulkareem et al.(2018)이 O*NET 데이터를 이용해 구축한 숙련 보완성 네트워크와 방법론적으로 유사하고, 두 데이터는 모두 직업별로 세분화된 숙련 항목을 다차원 척도로 제시한다는 점에서 구조적으로 유사하다. 그런데, O*NET이 미국의 산업·직업 구조를 반영하고 있지만, KNOW는 한국 노동시장에 특화된 조사 체계를 통해 한국의 문화·산업 맥락에 적합한 숙련 구조를 반영한다는 점에서

차이가 있다. 특히 KNOW의 확장된 항목(예: 성격, 가치관, 업무 환경 등)은 한국의 직업 환경에서 실제로 요구되는 사회적·신체적·인지적 숙련의 복합 구조를 정밀하게 포착할 수 있을 것이다.

따라서 본 연구는 O*NET 기반 선행 연구의 분석 틀을 계승하면서도, KNOW 데이터의 포괄적 항목 구성을 활용하여 숙련의 실질적 기능성과 맥락적 다양성을 반영하는 네트워크 분석을 시도한다. 이 접근은 숙련을 협의의 숙련 개념에 한정하지 않고, 노동시장에서 요구되는 인지적·사회적·정서적·신체적 역량의 통합체로 확장해 이해하고자 한다.

3. 분석 결과

가. 숙련 간 보완성 네트워크

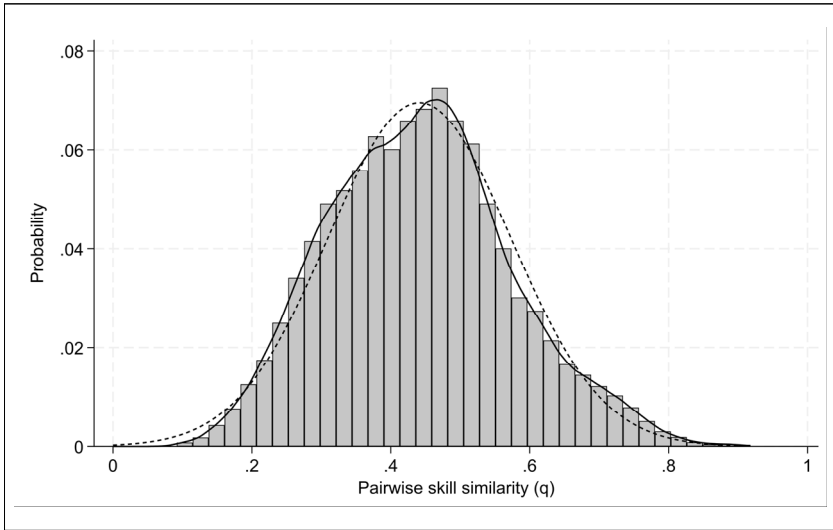
1) 숙련 간 보완성 네트워크 구축과 시각화

미국은 숙련 간 보완성 지수 분포가 쌍봉형(bimodal) 형태로 나타나, 보완성이 낮은 집단과 높은 집단이 명확히 구분되지만(Alabdulkareem et al., 2018), [그림 3-1]에서 보는 바와 같이, 한국은 단봉형(unimodal) 분포를 보이며, 전반적으로 숙련 관계 강도가 균등하게 분포하는 구조를 보여준다. 그러나 보완성 지수가 0.6을 초과하는 강한 관계 구간으로 한정하면, 분포의 비대칭성이 점차 나타난다. 이는 전체적으로 균형적인 구조 속에서도 강한 보완성 중심의 군집이 형성되는 '임계적(critical)' 양극화의 징후를 보여준다.

[그림 3-2]는 임계값 0.6을 기준으로 숙련 간 보완성 관계만을 남긴 네트워크를 시각화한 것이다. 이는 전체 숙련 관계 중에서도 가장 강한 상호보완적 연결만을 남긴 것으로, 숙련 체계의 양극화가 한층 뚜렷하게 드러난다. 각 결절은 개별 숙련을, 연결선은 해당 숙련 쌍의 보완성 관계를 의미하며, 결절의 크기는 중심성 지숫값에 비례한다.

아이겐벡터 중심성(eigenvector centrality)은 네트워크 내에서 다른 영향력 높은 숙련과의 연결 강도를 반영하는 지표로, 숙련의 핵심 영향력을 나타낸다. 이 값이 큰 숙련은 주로 업무 활동과 업무 수행 능력 영역에 집중되

[그림 3-1] 숙련 간 보완성 지수 분포



주: 실선은 정규분포 곡선이고, 점선은 커널 밀도 곡선을 나타냄.
자료: 저자 작성.

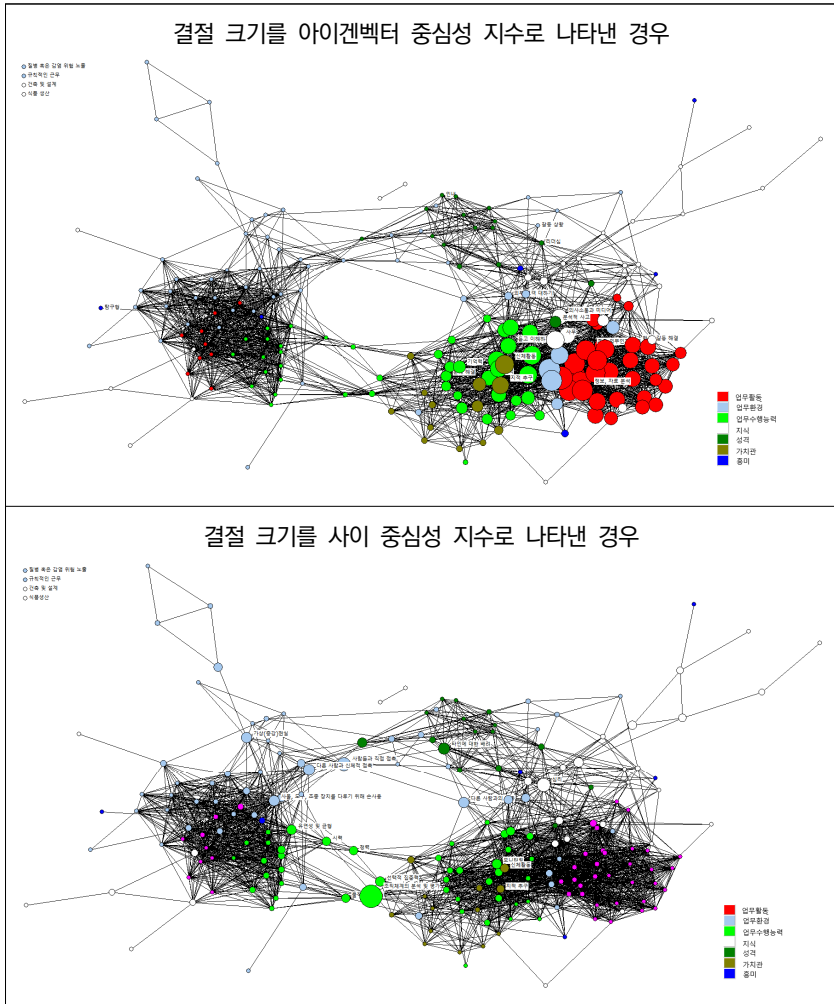
어 있으며, 정보 처리, 컴퓨터 활용, 문서 작성, 자료 분석, 의사소통 등과 같은 인지기적 숙련이 핵심을 형성한다.

반면, 사이 중심성(betweenness centrality)은 서로 다른 클러스터 간 최단 경로 위에 위치하는 정도를 측정함으로써, 네트워크 내에서 흐름과 전이를 증개하는 매개 숙련을 식별한다. 사이 중심성이 높은 숙련에는 ‘조직 체계의 분석 및 평가’, ‘사람과의 직접적 상호작용’, ‘배려 및 협동’, ‘선택적 집중력’, ‘모니터링’ 등이 포함된다. 이들은 사회·인지적 숙련과 감각·신체적 숙련 사이를 연결하며, 전체 숙련 체계의 교량 역할을 한다. 특히 ‘조직 체계의 분석 및 평가’ 항목은 가장 높은 사이 중심성을 보여, 분리된 숙련 클러스터 간 정보의 흐름을 유지하는 핵심 연결축으로 기능함을 시사한다.

이러한 패턴은 한국의 숙련 네트워크가 완전히 분리된 이중 구조라기보다는, 두 개의 주요 클러스터가 몇 개의 가교 숙련으로 연결된 ‘중심-교량-주변부’ 구조로 되어 있음을 보여준다. 즉, 정보나 기술이 확산될 때는 중심부의 업무 활동과 수행 능력이 영향을 주도하되, 실제 전이와 연결은 이들 매개 숙련이 담당하는 것이다. 특히 ‘조직 체계의 분석 및 평가’나 ‘심리’,

‘업무 환경’ 및 ‘성격’ 관련 숙련은 각 군집 간 전이의 효율성을 높이는 핵심 매개 역할을 하며, 이들이 부재하면 네트워크의 연결 경로가 단절될 위험이 커진다. 이처럼 한국의 숙련 네트워크는 두 개의 결집된 클러스터와 이를 잇는 소수의 가교 결절로 구성된 ‘부분적’ 양극화 구조를 띤다고 볼 수 있다.

[그림 3-2] 임계치(0.6) 기반 숙련 간 보완성 지수의 네트워크 구조



주: 각 결절은 숙련을 나타내며, 결절 간 연결은 보완성 지수에 기반하고, 원 크기는 각 결절의 아이겐벡터와 사이 중심성의 상대적 크기를 반영함.
 자료: 저자 작성.

또한, 각 속련 간 연결선을 하나씩 제거했을 때 네트워크가 분리되는지를 점진한 결과, 구조적 단절을 초래할 수 있는 결정적 연결선(edge-cut disconnectors)이 총 11개로 확인된다. 이는 네트워크의 통합성을 유지하는 핵심 연결선으로 서로 다른 속련 집단 간 연결을 유지하는 가교 또는 통로 역할을 한다.

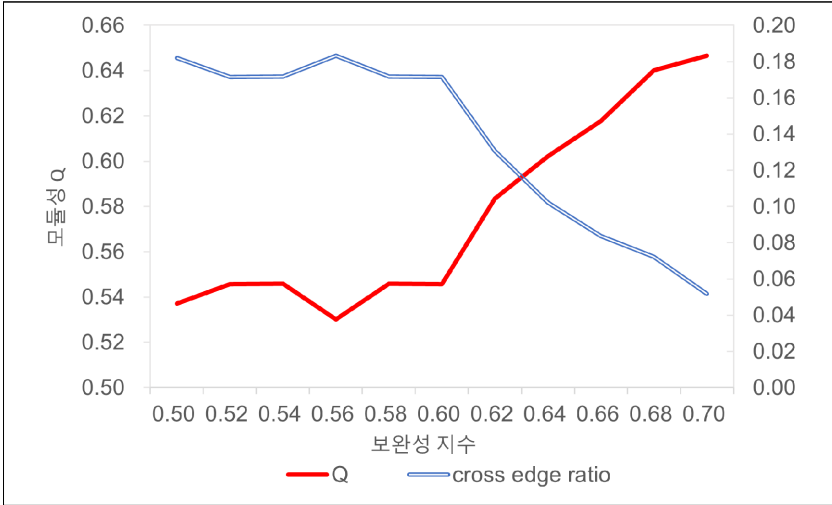
이러한 위상학적 특징은 속련 간 보완성의 전반적인 균형 분포 속에서도 강한 보완성 중심의 핵심 연결이 공존하고 있음을 시사한다. 따라서 한국의 속련 체계는 전체적으로 균등한 관계망 위에서 작동하지만, 보완성 지수가 일정 수준을 넘어서면 일부 속련 클러스터가 밀집되어 '부분적' 양극화로 전개되는 측면이 있다.

이러한 구조적 변화는 [그림 3-3]에서도 확인된다. 가중 모듈성 Q 값은 각 클러스터 내 연결의 실제 비중이 무작위 연결망에서 기대되는 비중보다 얼마나 높은지를 나타내는 지표로, 일반적으로 0.3~0.7 범위의 값은 명확한 클러스터 구조의 존재를 의미한다(Blondel et al., 2008). 임곗값이 0.6일 때 Q 값은 0.55로 나타나, 속련 네트워크가 무작위 연결이 아닌 내부 결집도가 높은 구조임을 보여준다. 또한, 임곗값 0.6 이상에서 네트워크가 급격히 재편되는 양상이 나타난다. 이 구간부터 Q 값이 상승하고, 동시에 교차 연결 비율(cross-edge ratio)이 감소한다. 후자는 상이한 클러스터들을 잇는 연결선 비중을 의미하므로, 이 값이 낮으면 내부 결속이 강화된다는 것을 의미한다. 이는 약한 보완성이 있는 속련 간 관계가 제거되고 강한 보완성 중심의 속련 집단이 더욱 결집되는 네트워크의 위계적 분화, 즉 양극화가 진행되고 있음을 시사한다.

또한, 클러스터 내부의 평균 연결 강도는 0.390, 클러스터 간 평균 연결 강도는 0.003으로 계산되어, 내부 연결 강도가 외부보다 약 125배 높다. 이는 속련이 같은 클러스터 내에서는 밀접하게 연결되어 있으나, 다른 클러스터와는 거의 연결되지 않음을 보여준다. 이에 따라 네트워크는 내부적으로는 조밀하고 외부적으로는 연결이 많지 않아, 속련 집단 간 경계가 뚜렷하다.

한편, 각 클러스터가 외부와 얼마나 많은 연결이 있는지를 나타내는 지표인 평균 전도도(mean conductance)를 계산해 보면, 0.134로 낮은 편이다.

[그림 3-3] 속련 간 보완성 지수 변화에 따른 모듈성 지수(Q)의 변화



자료 : 저자 작성.

이는 클러스터 간 경계가 뚜렷하고 상호 침투가 적음을 의미한다. 바꿔 말하면, 속련 간 네트워크가 강한 보완성 중심으로 재편되면서 클러스터 간 교류는 감소하고, 내부 결집도는 더욱 강화되고 있음을 보여준다.

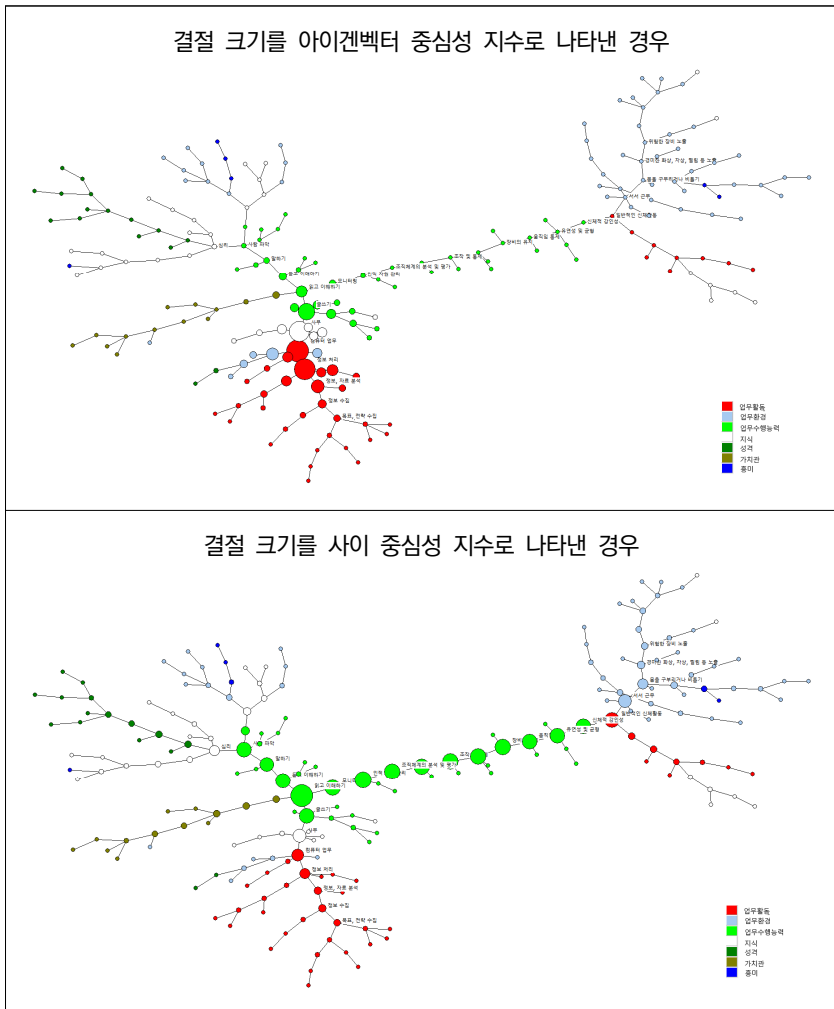
요컨대, 속련 네트워크는 겉으로는 단봉형 분포를 보이지만, 보완성 지수가 0.6을 초과하는 강한 관계만을 고려하면, 내부 집단이 외부와 단절된 형태로 재편되는 '임계적인' 전환이 일어난다. 이는 표면적으로 균등해 보이는 속련 간 구조 속에도 강한 보완성을 중심으로 한 양극화된 하위 네트워크가 존재함을 보여준다.

[그림 3-4]의 MST 기반 속련 네트워크는 [그림 3-2]의 임계치(0.6) 기반 네트워크와 구조적으로 유사한 형태를 보인다. 이는 속련 간 보완성 관계가 이미 강한 연결 중심으로 형성되어 있어, 약한 연결을 제거하거나 최소 연결망으로 축약하더라도, 네트워크의 핵심 구조가 크게 변하지 않음을 의미한다. 다시 말해, 속련 간 관계망이 내재적으로 높은 결집과 위계적 중심 구조로 되어 있음을 보여준다.

MST는 전체 연결 중 최소한 경로만 남겨 네트워크의 구조적 뼈대를 도출하는 방법임에도 불구하고, 주요 클러스터 분포와 중심 결절의 배치가 임계

치 기반 네트워크와 거의 비슷하게 나타나고 있다. 이는 두 접근법 모두 강한 보완성 중심의 핵심 연결망을 포착하고 있으며, 네트워크가 다수의 약한 연결보다는 소수의 핵심 연결로 유지되는 ‘중심-교량-주변부’ 구조를 띠고 있음을 뒷받침한다.

[그림 3-4] 속련 보완성 지수에 기반한 MST 네트워크 구조



주: 각 결절은 속련을 나타내며, 결절 간 연결은 보완성 지수에 기반하고, 원 크기는 각 결절의 아이겐벡터와 사이 중심성의 상대적 크기를 반영함.
 자료: 저자 작성.

특히 사이 중심성을 기준으로 한 MST 시각화에서, 업무 수행 능력과 업무 환경 범주 숙련이 클러스터 간을 연결하는 주요 가교 역할을 하고 있다. 이는 앞서 임계치(0.6) 기반 네트워크에서 관찰된 결과와 일관되며, 두 숙련 집단 간 전이가 주로 업무 수행 활동 및 작업 환경 관련 숙련을 매개로 이루어진다는 점을 보여준다. 따라서 한국의 숙련 체계에서 실질적 전이와 확산은 인지적·사회적 숙련이 아니라 작업 수행 및 환경 조정 능력을 중심으로 이루어진다는 점에서 특기할 만하다.

또한, 아이겐벡터 중심성은 네트워크 내에서 영향력이 큰 다른 숙련과 얼마나 밀접히 연결되어 있는가를 나타내는 지표로, 주로 인지적 숙련에 해당하는 업무 활동 항목(예: 정보 처리)에서 높게 나타난다. 이는 인지적 숙련이 네트워크 전반의 핵심 연결축으로 작동하며, 숙련 체계의 중심적 위치를 차지하고 있음을 시사한다.

MST 분석 결과는 임계치 0.6 초과에서 이미 관찰된 임계적 양극화 단면을 다시 한번 확인해 준다. 한국 숙련 네트워크는 강한 보완성 중심의 결속을 통해 안정적인 위계 구조를 형성하고 있으며, 네트워크 복잡도를 줄이더라도 핵심 군집 패턴과 중심 연결 구조가 유지된다. 또한, MST는 이러한 구조가 효율적이면서도 제한된 전이 경로를 갖고 있음을 보여준다. 불필요한 연결을 제거하고 핵심 경로만 남기는 MST의 특성상, 숙련 간 전이는 일부 주요 숙련, 특히 업무 수행 능력과 업무 환경 관련 숙련을 매개로 이루어진다. 따라서 이들 숙련이 변화하면 전체 네트워크의 확산과 재편에 큰 영향을 미칠 수 있음을 시사한다.

2) 루빙 방식에 따른 군집 판별 : 숙련 유형 구분

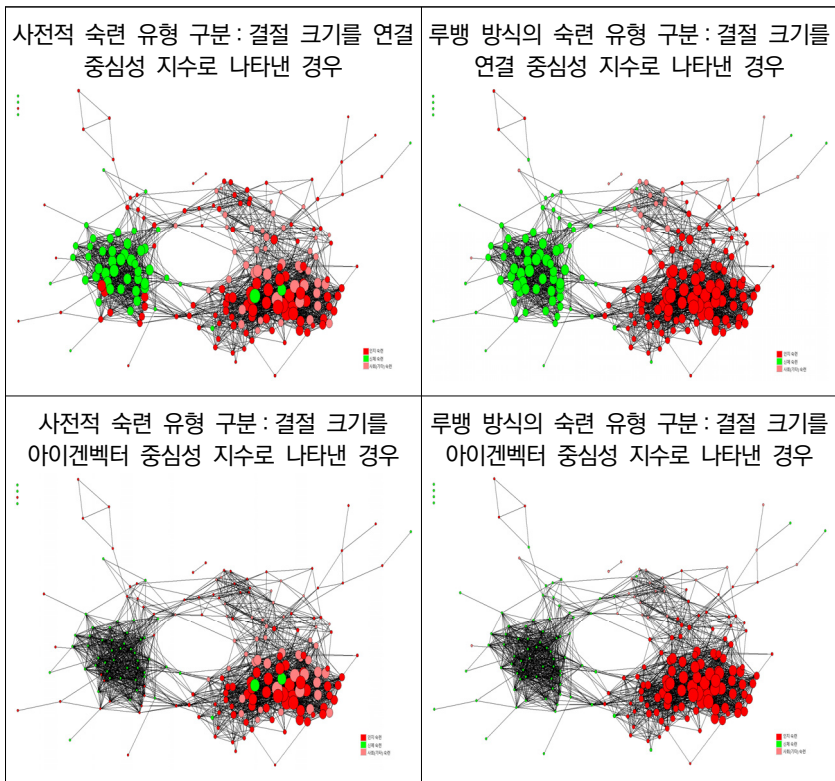
숙련의 양극화 분석은 일반적으로 숙련을 선형적으로 구분하여(예: 인지·비인지 또는 인지·신체·사회 숙련 등) 통계적으로 검정하는 방식에 기반한다(예: Acemoglu and Autor, 2011; Gathmann and Schönberg, 2010). 예컨대, 사고·판단·정보 처리 업무는 인지 숙련으로, 물리적 조정이나 감각-운동 능력은 신체 숙련으로, 상호작용·협업·감정노동은 사회 숙련으로 분류된다. 그러나 이러한 사전적 구분은 숙련 간 실제 결합 양상을 충분히 반영하지 못한다. 전술한 바와 같이, Alabdulkareem et al.(2018)은 숙련

간 보완성을 토대로 루뱅 알고리즘을 적용해, 미국의 숙련 양극화를 자료 기반 귀납적 네트워크 구조로 제시한 바 있다.

루뱅 방식은 의미적 속성 대신 숙련 간 보완성 행렬의 연결을 이용해, 직무 수행에서 함께 요구되는 숙련의 결속 패턴을 탐지한다. 따라서 인지 숙련이더라도 사회적 숙련과 자주 결합된다면, 동일 군집으로 분류되지만, 기능적으로 결합되지 않는 숙련은 다른 클러스터로 분리된다. 이는 숙련의 기능적 결합을 반영하는 구조적이고 위상학적 접근이다.

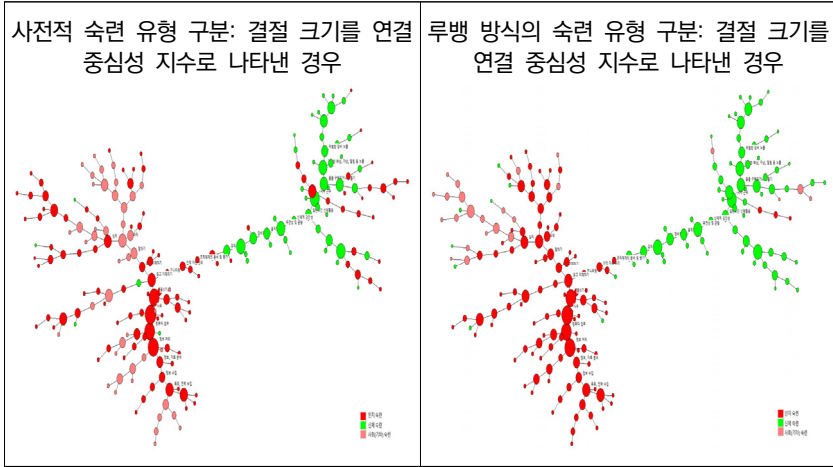
[그림 3-5]와 [그림 3-6]은 각각 사전적 숙련 구분과 루뱅 방식에 따른 군집화를 숙련 보완성 네트워크에 투영한 결과다. 루뱅 알고리즘을 적용한 결

[그림 3-5] 숙련 유형 구분을 투영한 임계치(0.6) 기반 숙련 간 보완성 네트워크 구조



주: 각 결절은 숙련을 나타내며, 결절 간 연결은 보완성 지수에 기반하고, 원 크기는 각 결절의 아이겐벡터와 사이 중심성의 상대적 크기를 반영함.
 자료: 저자 작성.

[그림 3-6] 숙련 유형 구분을 투영한 숙련 간 보완성 MST 네트워크 구조



주: 각 결절은 숙련을 나타내며, 결절 간 연결은 보완성 지수에 기반하고, 원 크기는 각 결절의 연결 중심성의 상대적 크기를 반영함.

자료: 저자 작성.

과, 전체 210개 숙련은 세 개의 클러스터, 즉 인지, 신체, 사회·기타 숙련으로 구분되며, 특히 인지·사회 숙련과 신체 숙련 간 경계가 명확하게 드러난다. 이는 Alabdulkareem et al.(2018)의 결과와 유사하게, 숙련 네트워크가 인지-신체 숙련 축을 중심으로 양극화된 이중 구조를 형성하고 있음을 보여준다.

[그림 3-5]의 임계치(0.6) 기반 보완성 네트워크에서는 이러한 구분이 더욱 명확하고, [그림 3-6]의 MST 분석에서도 거의 동일한 군집 패턴이 유지된다. 이는 약한 연결을 제거하더라도 네트워크의 핵심 구조가 변하지 않음을 의미한다. 즉, 숙련 간 보완성 관계가 이미 강한 연결 중심으로 조직되어 있으며, 인지와 신체 숙련 간 분리가 구조적으로 고착되어 있음을 시사한다.

[그림 3-5]와 [그림 3-6]의 양쪽 그래프를 비교해 보면, 루뱅 방식은 의미적 유사성이 아니라 실제 연결 패턴, 즉 숙련 간 보완성 관계에 기반해 군집을 형성한다는 점에서 사전 분류와는 다르다. 그 결과, 사전 분류에서는 동일 숙련 범주 내에서도 연결이 약해 '섬' 형태의 단절된 구조가 관찰되지만, 루뱅 방식에서는 이러한 단절이 거의 사라지고, 숙련 간 연계가 강화된 결속적인 네트워크 구조가 나타난다.

이러한 유형 구분의 불일치는 루뱅 알고리즘이 숙련의 개념적 유사성보다는 직무 수행에서 함께 요구되는 결합 패턴을 기반으로 군집을 식별하기 때문이다. 실제 업무 환경에서 상이한 숙련이 서로 협력적으로 작동하는 빈도와 강도를 반영함으로써, 의미상으로는 구분되는 숙련이라도 기능적으로 긴밀히 연결된 관계망으로 재배치된다. 따라서 루뱅 방식에서는 사전 구분에 따른 범주 간 경계가 완화되고, 숙련 간 기능적 통합이 강화된다. 이는 실제 노동시장과 직무 현장에서 숙련이 독립된 속성으로 존재하기보다 상호 의존적 체계로 작동한다는 점을 시사한다.

또한, [그림 3-5] 그래프의 절절 크기는 아이젠벡터 중심성을 반영하는데, 정보 처리와 같은 인지 숙련이 네트워크의 중심부를 차지한다. 반면, 신체 숙련은 주로 주변부에 분포하며, 일부 숙련만이 인지 숙련 집단과 느슨하게 연결되어 있다. 이는 한국의 숙련 네트워크가 전반적으로 인지 숙련 중심의 위계 구조를 형성하고 있음을 시사한다. 그리고 [그림 3-5]의 상단 그래프와 [그림 3-6]에서 나타난 연결 중심성 분석 결과를 보면, 인지·사회 숙련과 신체 숙련이 양쪽으로 대등하게 분포하며 이중적 양극화 구조를 보인다. 그러나 아이젠벡터 중심성을 기준으로 하면, 네트워크 내 영향력은 여전히 인지 숙련이 주도하고 있으며, 이는 숙련 체계가 표면적으로는 균형적이나 구조적으로는 인지 숙련 중심의 비대칭적 결속을 내포하고 있음을 함의한다.

한편, 의미 기반의 연역적 분류와 자료 기반 귀납적 분류 간에는 상당한 불일치가 있다(표 3-1 참조). 이러한 불일치는 숙련 간 연결 구조의 비대칭성, 즉 노동시장에서 숙련이 결합되는 기능적 상호 의존 패턴의 차이에 기인한다. 전체 210개 숙련 항목 중 약 42%에 해당하는 88건에서 두 방식 간 불일치가 나타났으며, 그중 절반 이상이 업무 환경(24건)과 업무 활동(15건) 범주에 집중되어 있다. 이는 루뱅 군집이 숙련 간 '기능적' 보완성을 기준으로 형성되므로, 실제 직무 수행 맥락에서 환경 요인과 활동 요소가 결합된 복합 숙련 구조를 반영한 결과로 볼 수 있다.

또한, 지식(16건)과 업무 수행 능력(13건) 범주에서도 불일치가 적지 않다. 특히 '공학·기술·디자인·건축·물리·화학' 등 STEM 기반 지식 항목이 루뱅 군집에서 신체 숙련으로 이동하는 양상이 관찰된다. 이는 이론적 지식이 실제 현장 맥락에서 장비 제어·품질 관리·기술 설계 등 실행 숙련

과 결합되어 활용된다는 점을 시사한다. 루뱅 방식은 이러한 항목을 단순한 '육체적 활동'이 아니라, 현장 기반 '기술-운영' 숙련으로 포착하고 있는 것으로 보인다.

불일치의 또 다른 축은 전체의 약 27%에 해당하는 인지와 사회 숙련 간 교차(57건)다. 이는 인지적 판단과 사회적 상호작용이 실제 직무 수행 과정에서 긴밀히 통합되어 있다는 점을 함의한다. 예를 들면, '설득', '협상', '가르치기', '사람 파악' 등은 인지적 분석과 대인 소통이 결합된 '사회적 인지 숙련'으로 작동한다. 루뱅 군집은 이러한 숙련을 개념적으로 구분하지 않고, 실제 상호작용의 빈도와 협업 구조를 반영하여 동일한 군집으로 재배치한 것으로 보인다.

〈표 3-1〉 사전 숙련 유형 구분과 루뱅 방식 숙련 유형 구분 간 불일치

숙련 범주	사전과 루뱅 방식 구분 간 불일치(건)	사전과 루뱅 방식 구분에서의 이동(건): 인지·사회(기타) ↔ 신체 숙련
업무 활동	15 (41)	2 (41)
업무 환경	24 (57)	14 (57)
업무 수행 능력	13 (44)	6 (44)
지식	16 (33)	7 (33)
성격	11 (16)	1 (16)
가치관	5 (13)	1 (13)
흥미	4 (6)	0 (6)
계	88 (210)	31 (210)

주: 괄호 안의 수치는 KNOW 재직자조사에 나타난 범주별 숙련 항목 수임.
자료: 저자 작성.

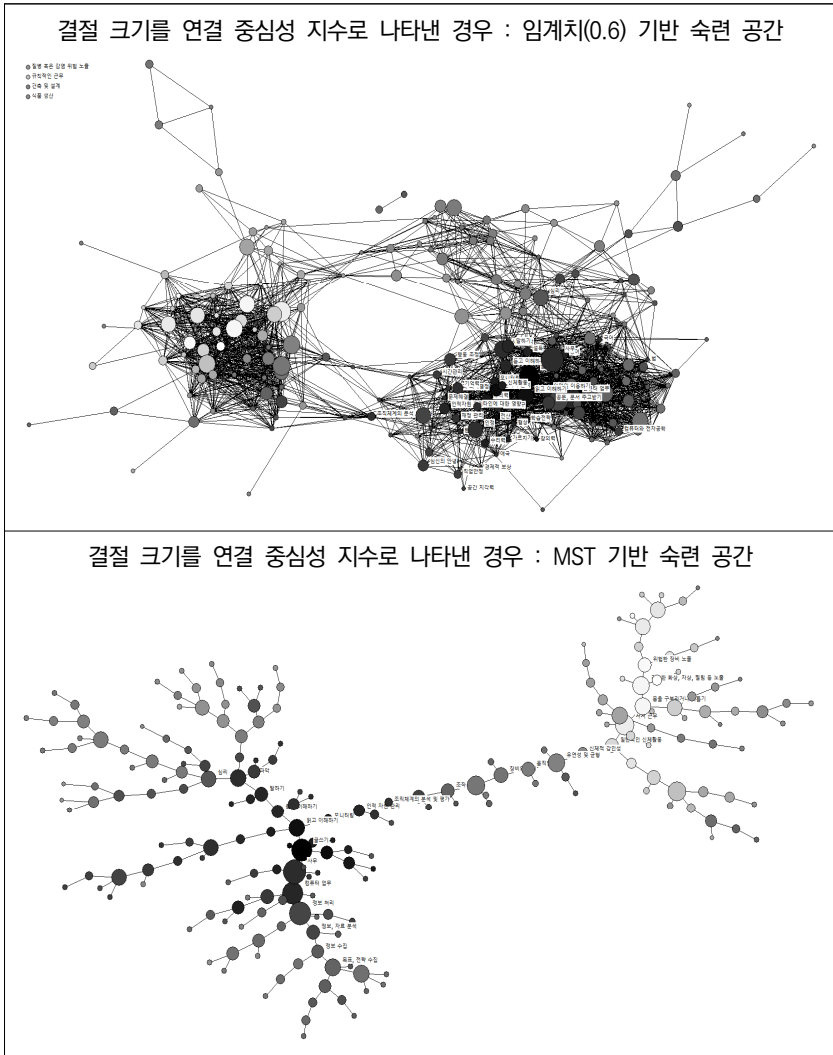
이처럼, 연역적 분류와 루벨 군집 간 불일치는 단순한 분류 오차가 아니라, 숙련 체계의 기능적 통합과 재조정을 보여주고 있다. 전체 불일치 중 약 15%는 인지·사회 숙련이 신체 숙련으로 이동한 경우로, 이는 추상적 인지 능력이 현장의 기술·운영의 실행 역량으로 전환되는 과정을 함의한다. 이들 항목의 다수는 ‘정밀도’, ‘품질 관리’, ‘장비제어’, ‘안전관리’ 등 운영·제어 중심의 실무형 숙련에 속한다.

따라서 전체 불일치의 42% 내외는 사회적 인지 또는 기술-운영으로 수렴하는 ‘기능적 통합형’ 재배치다. 이는 단순한 분류상의 차이라기보다, 한국 숙련 체계 내에서 인지·사회·기술적 요소가 교차하며 형성되는 ‘틈새적’ 숙련 영역을 포착한 것으로 볼 수 있다. 루벨 군집은 표면적으로는 ‘사회적 인지-기술 운영’으로 이어지는 새로운 숙련 연쇄를 드러내지만, 이는 기존 ‘인지·사회-신체’ 축 사이에서 작동하는 기능적 결합과 재조정에 가깝다. 다시 말해, 불일치는 한국 숙련 구조의 기능적 통합의 경계를 보여주지만, 그 내부 메커니즘은 여전히 인지적 판단과 신체적 실행이라는 두 축 위에서 작동하고 있음을 시사한다.

[그림 3-기는 앞서 제시한 임계치(0.6) 기반 숙련 보완성 네트워크와 MST 기반 네트워크에 평균 교육연수를 투영한 결과다. 결절의 색상 농도는 교육연수를 나타내며, 색이 짙을수록 교육훈련 강도가 높음을 의미한다. 두 네트워크 모두에서 교육연수가 인지·사회 숙련 집단에 집중되고, 신체 숙련 집단으로 갈수록 그 수준이 뚜렷이 낮아지는 양극화 패턴이 확인된다. 임계치(0.6) 기반 네트워크에서는 인지·사회 숙련 클러스터(왼쪽)가 상대적으로 짙은 회색을 띠지만, 신체 숙련 중심 클러스터(오른쪽)는 옅은 색으로 표현된다. MST 기반 네트워크에서도 동일한 경향이 반복되며, 이처럼 구조가 단순화된 이후에도 교육연수의 집중도가 인지·사회 숙련과 신체 숙련 간 구분과 높은 상관성을 유지하고 있다.

이는 교육·훈련 기회의 분포가 숙련 네트워크의 구조적 분화와 밀접하게 연동되어 있음을 시사한다. 다시 말해, 인지·사회 숙련은 고차적 판단·분석·대인 상호작용이 요구되는 영역으로 지속적 재교육과 역량 확충이 필수적인 핵심 숙련 클러스터이지만, 신체 숙련은 기술적 실행 중심의 반복 수행에 기반한 낮은 교육연수 수준의 주변부 숙련 클러스터로 나타난다. 이

[그림 3-7] 교육연수를 투영한 숙련 간 보완성 네트워크 구조



주: 각 결절은 숙련을 나타내며, 결절 간 연결은 보완성 지수에 기반하고, 원 크기는 각 결절의 연결 중심성의 상대적 크기를 반영함.

자료: 저자 작성.

처럼, 숙련 공간에서의 교육연수 분포는 루뱅 군집 기반 인지·사회·신체 숙련 간 이중 구조와 거의 일치하며, 이는 숙련의 양극화가 교육·훈련 투자 격차로 시각적으로 재현되고 있음을 명확히 보여준다.

3) 숙련 유형이 임금 수준에 미치는 효과

〈표 3-2〉와 〈표 3-3〉은 '2020년 KNOW 재직자조사' 자료를 이용해 직업 별 숙련 구성, 즉 인지·신체·사회 숙련 비중이 로그 연봉에 미치는 영향을 OLS로 추정한 결과다. 숙련의 의미적·기능적 구분을 모두 고려하기 위해, 〈표 3-2〉는 사전적 유형 구분을, 〈표 3-3〉은 루벨 군집 기반의 구분 방식을 적용한다. 종속변수는 로그 연봉이며, 통제 변수로는 총숙련량, 교육연수, 연령, 연령 제곱항을 포함하며, 분석 단위는 한국표준직업분류 세세분류상 직업이다.

두 표의 결과는 매우 일관된 패턴을 보인다. 인지 숙련 비중은 임금에 유

〈표 3-2〉 숙련 유형이 로그 임금에 미친 효과 : 사전적 유형 구분의 경우

종속변수 : 로그 연봉	모형 1	모형 2	모형 3	모형 4	모형 5
인지 숙련 비율	4.511*** (6.44)	-	-	6.892*** (4.12)	4.194*** (3.59)
신체 숙련 비율	-	-2.151*** (-6.37)	-	1.515* (1.84)	2.866*** (5.38)
사회(및 기타) 숙련 비율	-	-	0.951 (1.39)	-	-
총숙련량	-	-	-	-	0.002*** (5.54)
연령	-	-	-	0.086** (2.27)	0.076*** (5.44)
연령 제곱	-	-	-	-0.001** (-2.03)	-0.001** (-1.97)
교육연수	-	-	-	-	0.074** (2.38)
상수	5.847*** (15.80)	8.740*** (105.95)	8.027*** (50.67)	2.257* (1.75)	1.411* (1.82)
R ²	0.195	0.106	0.006	0.247	0.569
N	385	385	385	385	385

주 : 1) 클러스터 오차(KSCO_7 기준)를 적용하여 추정한 것임.

2) *, **, ***는 각각 10%, 5%, 1% 수준에서 통계적으로 유의함을 나타냄.

3) 괄호 안의 수치는 t값임.

자료 : 저자 작성.

의한 양(+)¹의 영향을 미치고, 신체 숙련 비중은 유의한 음(-)²의 영향을 보이며, 사회(또는 기타) 숙련 비중의 효과는 상대적으로 약하거나 통계적으로 유의하지 않다. 즉, 인지 숙련 집약도가 높은 직업일수록 임금 수준이 높고, 신체 숙련 집약도가 높은 직업일수록 낮은 임금을 받는 양극화 패턴이 존재한다.

사전적 구분(표 3-2)에서 인지 숙련의 계수는 4.5~6.9로, 모두 1% 수준에서 통계적으로 유의한 양(+)³의 효과를 보인다. 반면 신체 숙련은 모형 2에서 -2.2로 유의한 음(-)⁴의 효과를 나타내지만, 모형 4와 모형 5에서는 부호가 양(+)⁵으로 전환되고 통계적으로 유의하나, 그 크기는 인지 숙련 효과보다 낮다. 이러한 패턴은 Alabdulkareem et al.(2018)의 분석 결과와도 일관되며, 신체 숙련의 임금 효과가 양(+)⁶의 방향으로 나타나더라도 상대적으로 약하다는 점을 보여준다. 사회 숙련은 모형 3에서 약한 양(+)⁷의 계수를 보이지만, 통계적으로 유의하지 않다.

루뱅 방식(표 3-3)에서도 이러한 방향성과 유의성은 대체로 유지된다. 인지 숙련은 약 1.9~3.7의 유의한 양(+)⁸의 효과를 보이며, 신체 숙련은 모형 6에서 -1.5로 1% 수준에서 유의한 음(-)⁹의 효과를 보이나, 모형 10에서는 약 2.0으로 유의한 양(+)¹⁰의 효과로 전환된다. 그러나 그 절대적 크기는 여전히 인지 숙련 효과보다 낮다.

이와 같이, 숙련 구분 방식에 따른 세부 차이에도 불구하고 인지 숙련 프리미엄과 신체 숙련 디스카운트의 패턴은 견고하게 유지된다. 이는 숙련 네트워크 분석에서 확인된 '인지·사회 숙련 중심-신체 숙련 주변'의 이중 구조가 임금 결정 구조에서도 일관되게 반영되고 있음을 시사한다.

한편, 통제 변수의 결과 역시 전반적으로 일관된 방향성을 보인다. 연령은 임금에 양(+)¹¹의 영향을 미치다가 연령제곱항이 음(-)¹²으로 유의하게 나타나 전형적인 생애주기 효과가 확인된다. 교육연수 또한 1% 수준에서 유의한 양(+)¹³의 효과를 보여, 교육이 임금 프리미엄으로 작용함을 알 수 있다. 더 나아가 총숙련량이 많을수록 임금이 높아지는 관계도 통계적으로 유의하게 나타나, 숙련의 유형뿐 아니라 숙련의 종합적 축적 정도 자체가 임금 수준을 결정하는 중요한 요인임을 함의한다.

이러한 임금 결정 구조는 인지·사회 숙련이 정보 처리, 판단, 대인 상호

〈표 3-3〉 숙련 유형이 로그 임금에 미친 효과 : 루벨 유형 구분의 경우

종속변수 : 로그 연봉	모형 5	모형 6	모형 7	모형 9	모형 10
인지 숙련 비율	1.865*** (6.93)	-	-	3.690** (2.18)	3.367*** (3.26)
신체 숙련 비율	-	-1.486*** (-5.65)	-	1.559 (0.96)	1.957*** (5.57)
사회(및 기타) 숙련 비율	-	-	-2.054 (-1.49)	-	-
총숙련량	-	-	-	-	0.002*** (5.76)
연령	-	-	-	0.120*** (3.01)	0.072** (2.32)
연령 제곱	-	-	-	-0.001*** (-2.74)	-0.001* (-1.90)
교육연수	-	-	-	-	0.077*** (5.66)
상수항	7.258*** (52.14)	8.751*** (92.33)	8.526*** (47.47)	3.041* (1.94)	1.877** (2.51)
R ²	0.117	0.080	0.019	0.179	0.571
N	385	385	385	385	385

주 : 1) 클러스터 오차(KSC0_7 기준)를 적용하여 추정된 것임.

2) *, **, ***는 각각 10%, 5%, 1% 수준에서 통계적으로 유의함을 나타냄.

3) 괄호 안의 수치는 t값임.

자료 : 저자 작성.

작용 등 고차 기능 수행과 결합되어 교육·훈련 투자가 집중되고 높은 임금 수준으로 보상되는 반면, 신체 숙련은 반복적·절차적 수행 중심으로 자동화와 기계화의 대체 압력이 높아 상대적으로 낮은 임금 수준에 머무는 구조임을 보여준다.

또한 루벨 방식을 적용한 추정 결과가 사전적 구분과 거의 동일하다는 점은, 숙련의 의미적 구분이 달라지더라도 임금 구조의 방향성은 변하지 않는다는 점을 시사한다. 이는 숙련 네트워크의 위계적 이중 구조가 노동시장의 보상 체계에도 그대로 투영되고 있음을 보여준다.

요컨대 숙련의 양극화는 단순한 숙련 구조상의 분절을 넘어 경제적 보상 체계의 이중화로 나타나고 있다. 다시 말해, 한국의 숙련 체계는 '인지 숙련 프리미엄-신체 숙련 디스카운트'라는 임금 양극화를 내재하고 있으며, 이는 숙련의 의미적 범주나 기능적 결합 패턴을 초월해 일관되게 관측되고 있다.

나. 직업 간 연관성 네트워크

1) 직업 간 연관성 네트워크 구축과 시각화

앞서 언급한 바와 같이, 본 연구는 Henning et al.(2025)의 틀을 기능적 관점으로 조정해, 직업별 숙련 중요도 벡터의 코사인 유사도에 기반한 직업 간 숙련 연관성 네트워크를 구축한다. 이러한 직업 간 연관성 지수를 분포로 나타낸 것이 [그림 3-8]이다. [그림 3-8]의 분포는 전반적으로 단봉형이며 정규 분포에 거의 가깝다, 즉 다수 직업이 중간 수준의 연관성에 모여 있음을 의미한다. 한편, 연관성 지수가 대략 0.65를 넘는 연관성이 높은 구간에서는 빈도가 크게 줄어 서로 밀접하게 연결된 일부 핵심 직업군만이 남아 있음을

[그림 3-8] 직업 간 연관성 지수 분포



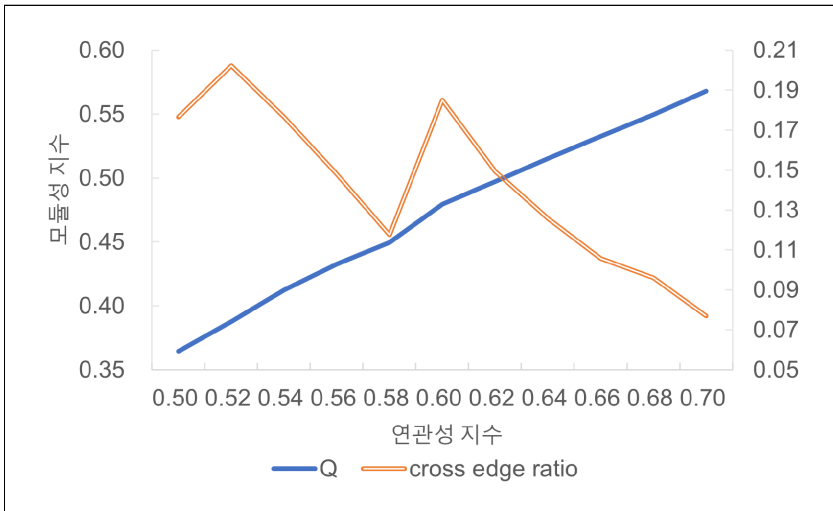
주: 실선은 정규분포 곡선이고, 점선은 커널 밀도 곡선을 나타냄.
자료: 저자 작성.

시사한다. 하지만, 이런 히스토그램만으로 직업 양극화를 단정할 수는 없다. 다만 이 상단 꼬리 구간을 분석 임계치로 삼아 핵심 연결망(예: MST)과 모듈성·교차 연결 비율 등 네트워크 지표를 함께 보면, 연관성이 높은 직업군의 결집·분리 여부를 판단할 수 있다.

[그림 3-9]는 직업 연관성 지수 변화에 따른 Q 와 교차 연결 비율 추이를 나타낸 것이다. 직업 연관성 지수가 높아질수록 Q 는 0.36에서 0.57까지 계속해서 높아지는데, 이는 네트워크가 점차 명확한 군집 구조를 형성하고 있음을 시사한다. 다시 말해, 결절이 무작위로 연결되어 있던 단계에서 벗어나, 내부적으로 밀접하게 연결된 클러스터가 뚜렷이 구분되는 방향으로 구조가 재편되고 있음을 보여준다. 반면, 교차 연결 비율은 직업 연관성 지수가 증가함에 따라, 변동이 약간 있기는 하나, 0.20에서 0.08로 점차 감소하고 있다. 이는 군집 간 연결선이 줄어들면서 내부 결속이 강화되고, 집단 간 경계가 명확해지고 있음을 의미한다. 이러한 양상은 네트워크의 분화가 진전되고 있음을 시사한다.

루빙 알고리즘을 적용한 가중 네트워크 분석 결과, 직업 연관성 지수 임계치가 0.65일 때, 그 네트워크의 모듈성 지수는 0.528로 나타난다. 이는 네

[그림 3-9] 직업 간 연관성 지수 변화에 따른 모듈성 지수(Q)의 변화



자료 : 저자 작성.

트위크가 비교적 뚜렷한 군집 구조를 형성하고 있음을 보여주는 값으로, 내부 결집과 집단 간 분리가 동시에 확보된 상태다. 또한, 평균 가중 내부 밀도와 집단 간 밀도를 계산해 보면, 평균 가중 내부 밀도는 0.238로, 집단 간 밀도 0.004에 비해 약 63배 높다. 이는 각 군집 내 연결 강도가 매우 높고, 군집 간 연결은 제한적이라는 점을 확인시켜 준다. 평균 가중 전도도를 계산해 보면, 이는 0.200으로 낮게 나타나, 군집 간 경계가 비교적 뚜렷이 분리되고 있음을 다시 한번 보여준다. 구조적 단절선(edge-cut disconnectors)은 총 3개로, 일부 핵심 연결선이 제거되면, 군집 간 분리가 심화될 가능성이 있다.

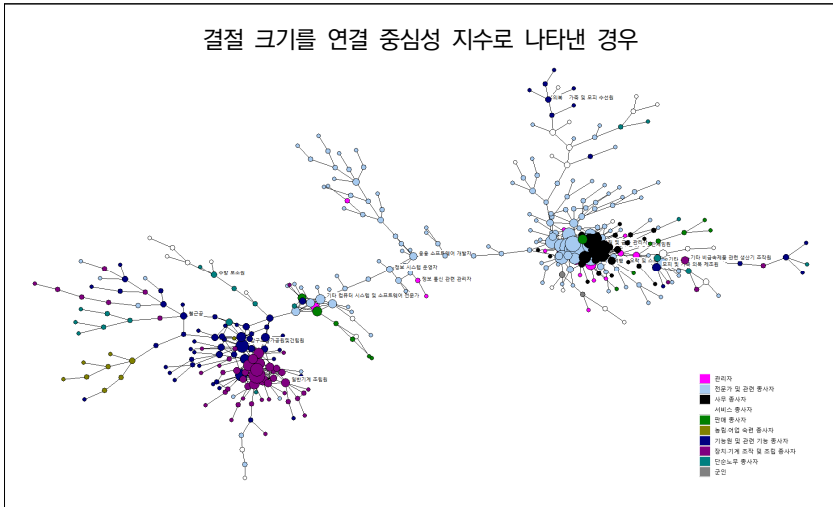
요약하면, 연결성 지수가 높아질수록 네트워크는 다수의 느슨한 연결 구조에서 벗어나 소수의 강하게 결속된 군집으로 전환되고 있다. 이는 직업 간 연관 구조가 점차 직업 간 관계망에서 중심성이 높은 군집으로 집중되는 현상과 일관된 패턴을 보여주며, 네트워크가 점차 위계적인 구조로 전이되고 있음을 시사한다.

[그림 3-10]은 한국표준직업분류 세세분류 397개 직업 간 숙련 연관성 지수를 이용해 MST를 구성하고, Henning et al.(2025)처럼, 여기에 상위 0.5% 연결선을 부가해 네트워크의 핵심 구조를 시각화한 것이다. 각 결절의 크기는 연결 중심성을 나타내며, Henning et al.(2025)이 제시한 직업 공간이 직업 간 '의미적' 유사성을 보여주는 것이라면, 이 분석의 MST는 '기능적' 차원에서 직업 간 연관성을 나타낸 직업 공간이다.

MST는 모든 직업 결절을 최소한의 연결선으로 묶어 네트워크의 기본 골격을 구성하므로, 이를 통해 직업 체계 전반의 핵심 연결 구조를 가장 단순한 형태로 확인할 수 있다. 따라서 MST는 네트워크 밀도나 가중치와 관계없이 직업 간 연관 구조의 뼈대를 드러내는 분석적 출발점으로 기능한다.

연결 중심성은 특정 직업이 다른 직업과 얼마나 폭넓게 연결되어 있는지를 나타내는 지표로, 직업 네트워크에서 해당 직업이 가지는 연결 허브로서 기능적 중요성을 의미한다. 연결 중심성이 높을수록 그 직업은 다른 직업과의 숙련, 기술, 기능적 연관성이 넓게 분포해 있으며, 이는 직업 간 이동성, 기술 확산, 노동시장 적응력의 잠재적 중심축으로 작용함을 시사한다. 따라서 중심성이 높은 직업군은 단순히 빈도로 많이 등장하는 직종이 아니라, 네

[그림 3-10] 직업 간 연관성 MST 네트워크 구조



주: 각 결절은 직업을 나타내며, 결절 간 연결은 연관성 지수에 기반하고, 원 크기는 각 결절의 연결 중심성의 상대적 크기를 반영함.

자료: 저자 작성.

트위크 전체의 상호 연결 구조를 유지하고 다른 직업 간 전이를 매개하는 핵심 결절로 이해될 수 있다.

그림에서 확인되었듯이, 직업 간 연관성 지수 네트워크는 앞서 숙련 보완성 지수 네트워크와 유사하게 ‘중심-교량-주변부’의 삼중 구조를 형성하고 있다. 다만, 직업 연관성 네트워크의 MST에서는 교량부(bridging zone)의 위상학적 구조가 보다 복잡하게 전개되어 있으며, 이는 직업 간 기능적 연관이 다층적으로 얽혀 있음을 함의한다.

각 결절의 색상은 한국표준직업 대분류를 반영한다. 한쪽에는 ‘관리자’, ‘전문가 및 관련 종사자’, ‘사무 종사자’ 등으로 구성된 화이트칼라 집단이 결집되어 있고, 다른 한편에는 ‘장치·기계 조작 및 조립 종사자’, ‘기능원 및 관련 기능 종사자’ 등 블루칼라 집단이 밀집되어 있다. 양 집단을 연결하는 교량부에는 주로 엔지니어, 관리자, 전문직 등 전문가 계층 직업이 위치하며, 이들이 화이트칼라와 블루칼라 직종 간 숙련 연관성을 매개하고 있다.

연결 중심성이 높은 주요 직업으로는 ‘유리 제조 및 가공기 조작원’, ‘강구조물가공원 및 건립원’, ‘기타 시각 및 공연 예술가’, ‘단조원’, ‘인사 및 노사

관련 전문가', '보험 심사원 및 사무원', '변호사', '회계사', '의회의원·고위 공무원 및 공공단체임원', '교육 관리자', '보험 및 금융 관리자', '인문과학 연구원' 등이 확인된다. 이들은 화이트칼라와 블루칼라 영역을 모두 포함하며, 전체 네트워크의 결속과 상호 연계를 유지하는 핵심적 매개 결절로 기능한다. 특히 이러한 중심성이 높은 직업은 기술과 숙련의 교차점에서 네트워크의 구조적 안정성을 유지하며, 직업 체계 내 정보·기능 흐름의 중심 경로를 형성한다는 점에서 중요하다.

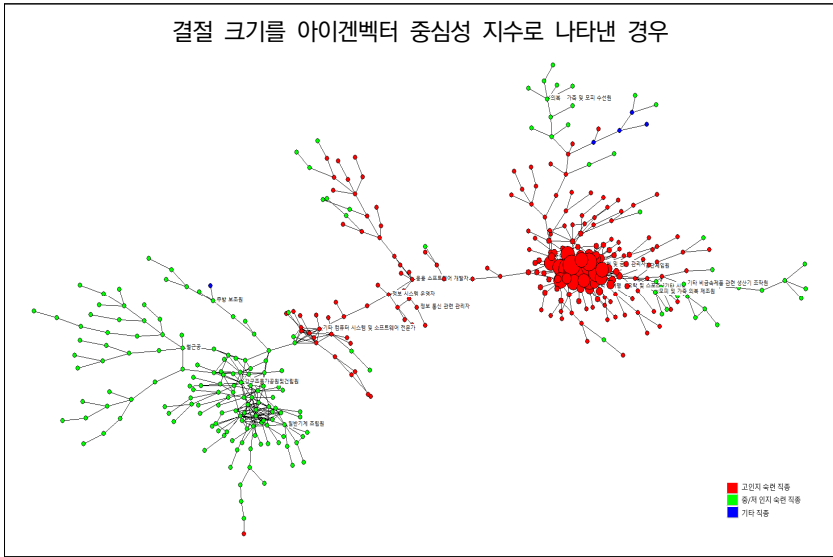
MST는 연결 강도의 임계치를 적용하기 전 전체 네트워크 구조를 가장 단순화된 형태로 제시하는데, 이 최소 연결망에서도 이미 양극화된 패턴이 분명히 나타난다. 즉, 화이트칼라와 블루칼라 직업군이 각각 독립된 클러스터로 나타나고, 양자를 연결하는 교량부의 존재가 제한적이라는 점에서, 네트워크가 점차 이원화된 구조로 전개되고 있음을 시사한다. 이는 앞서 제시한 분석에서 임계치가 0.65를 넘어서면 명확하게 관찰되는 양극화 구조의 기초적 형태가 MST에서 이미 내재되어 있음을 보여준다.

2) 루빙 방식에 따른 군집 판별 : 직업 유형 구분

397개 직업에 루빙 알고리즘을 적용해 직업 간 연관성을 군집화한 결과, 세 개의 클러스터가 식별된다. 첫 번째 클러스터는 총 205개 직업으로 구성되며, '의회의원·고위공무원 및 공공단체임원', '기업 고위 임원', '정부행정 관리자', '경영지원 관리자', '마케팅 및 광고홍보 관리자', '연구 관리자' 등이 이에 속한다. 두 번째 클러스터는 187개 직업으로, '건설 및 광업 관련 관리자', '통신 및 방송 송출 장비 기사', '건축가', '건설자재 시험원' 등 기술·기능 기반 직종이 중심을 이룬다. 마지막으로, 세 번째 클러스터는 5개 직업으로 구성된 소규모 집단으로, '안마사', '돌봄 서비스 종사원', '웨이터', '가사 도우미', '육아 도우미' 등이 포함된다.

이 세 클러스터를 각각 '고인지 숙련 직종', '중·저인지 숙련 직종', '기타 직종'으로 명명한다. 이러한 구분은 선형적으로 직업 유형을 설정한 것이 아니라, 자료 기반의 귀납적 접근을 통해 도출된 결과라는 점에서 의미가 있다. 이는 직업 간 기능적 연관성 지수를 통계적으로 추출하고, 이를 군집화하여 숙련 구조의 내재적 패턴을 경험적으로 식별한 것이다.

[그림 3-11] 직업 유형 구분을 투영한 직업 간 연관성 MST 네트워크 구조



주: 각 결절은 직업을 나타내며, 결절 간 연결은 연관성 지수에 기반하고, 원 크기는 각 결절의 아이겐벡터 중심성의 상대적 크기를 반영함.
 자료: 저자 작성.

이러한 세 가지 직업 유형을 반영하여 구성한 직업 간 연관성 MST는 [그림 3-11]에 제시되어 있다. 이 그림은 네트워크 구조를 아이겐벡터 중심성 지수를 기반으로 시각화한 것으로, 중심성이 높을수록 네트워크 내에서 구조적 영향력이 크다는 것을 의미한다. 분석 결과, 고인지 숙련 직종, 즉 화이트칼라 중심의 직업군이 전체 네트워크에서 가장 높은 아이겐벡터 중심성을 보여주고 있다. 특히 ‘회계사’, ‘변호사’, ‘보험 및 금융 관리자’, ‘세무사’, ‘인사 및 노사 관련 전문가’, ‘의회의원·고위공무원 및 공공단체임원’, ‘손해 사정사’ 등이 높은 중심성을 보이며, 이는 이들 직종이 다른 직업군과의 관계망 속에서 가장 큰 구조적 영향력을 행사하고 있음을 보여준다. 다시 말해, 이러한 고인지 숙련 직종은 네트워크 중심부에 위치하여 다른 직업군으로의 기술 확산, 숙련 전이, 정보 흐름의 핵심 경로를 형성하고 있으며, 직업 체계의 기능적 핵심 직업군으로 작동한다.

[그림 3-11]에 투영된 루벨 알고리즘에 의한 귀납적 군집화와 아이겐벡터 중심성 분석은 직업 간 연관성이 단순한 기술적 유사성을 넘어 숙련 수준에

따른 기능적 계층화를 내포하고 있음을 보여준다. 환언하면, 직업 체계는 고인지 숙련 중심의 화이트칼라 집단을 정점으로, 중·저인지 숙련 집단과 기타 집단으로 분리되는 위계적이고 양극화된 구조를 형성하고 있다.

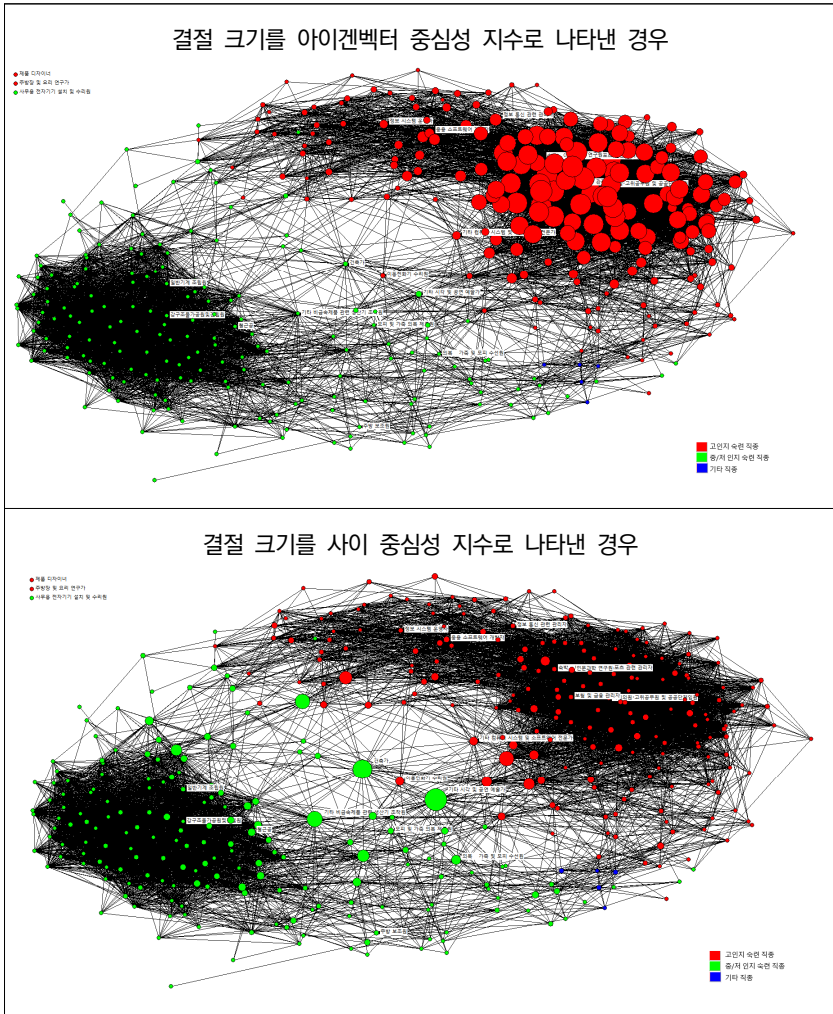
임계치 0.65를 초과하는 직업 간 연관성 네트워크를 시각화한 [그림 3-12]는 직업 유형별 군집이 명확히 분리된 양극화 구조를 보여준다. 상위 0.65를 초과하는 연결선만을 남긴다는 것은 직업 간 연계 중에서도 가장 강한 기능적 연관성을 유지하는 관계만을 추출한 것이며, 이로써 네트워크는 느슨한 연결망에서 벗어나 연관성이 높은 직업군 중심의 결집 구조로 전환된다. 이러한 변화는 앞서 제시된 MST 및 루뱅 군집 구조에서 이미 예비적으로 드러난 것이나, 임계치 적용을 통해 그 양극화가 한층 더 부각된 것이다.

[그림 3-12]는 고인지 숙련 직종(적색)과 중·저인지 숙련 직종(녹색)이 명확하게 구분되어 있으며, 두 집단을 연결하는 교량부(bridging zone)는 여전히 존재하지만, 그 폭이 현저히 좁아져 있다. 앞서 숙련 보완성 네트워크와 비교하면, 중간 영역의 구조가 다소 복잡하게 얽혀 있으나, 전체적으로는 화이트칼라(고인지 숙련)와 블루칼라(중·저인지 숙련) 집단이 각각 밀집된 형태를 유지하고 있다. 이는 직업 간 숙련의 상호 보완 구조가 약화되고, 기능별 전문화가 강화되고 있음을 시사한다.

아이젠벡터 중심성을 기준으로 보면, MST 분석에서와 마찬가지로 '보험 심사원 및 사무원', '회계사', '보험 및 금융 관리자', '교육 관리자', '변호사', '인사 및 노사 관련 전문가' 등이 높은 값을 보여준다. 이들은 네트워크 내에서 다수의 중요한 직업과 직접 연결될 뿐만 아니라, 중심성 높은 다른 직업과도 긴밀히 연계된 구조적 중심에 위치한다. 이러한 직업군은 네트워크 결속을 유지하며, 기능적 영향력을 다층적으로 확산시키는 핵심 축으로 작용한다.

반면, 사이 중심성이 높은 직종은 '기타 시각 및 공연 예술가', '건축가', '기타 비금속제품 관련 생산기 조작원', '기타 전문 서비스 관리자', '건설 및 광업 관련 관리자', '통신공학 기술자 및 연구원', '환경미화원 및 재활용품 수거원', '소규모 상점 경영 및 일선 관리 종사원' 등으로 나타나고 있다. 이들은 서로 다른 숙련 집단을 연결하는 교량적 위치를 차지하며, 직업군 간 정보·기술·숙련의 흐름을 매개하는 역할을 한다. 즉, 이들은 직업 체계 내

[그림 3-12] 숙련 유형 구분을 투영한 임계치(0.65) 기반 직업 간 연관성 네트워크 구조



주: 각 결절은 숙련을 나타내며, 결절 간 연결은 연관성 지수에 기반하고, 원 크기는 각 결절의 아이겐벡터와 사이 중심성의 상대적 크기를 반영함.
 자료: 저자 작성.

기능적 통로로 두 영역 간 전이를 가능케 하는 구조적 다리로 작용한다.

이러한 결과는 노동시장 내 직업 구조가 점점 이원화된 숙련 체계로 수렴하고 있음을 보여준다. 고인지 숙련 직종은 네트워크 중심에서 상호 긴밀히 연결되어 높은 영향력을 유지하지만, 중·저인지 숙련 직종은 주변부로 밀

려나면서 상대적으로 낮은 결속도를 보인다. 그 사이에서 교량적 직종이 존재하지만, 그 수와 영향력은 점차 제한적으로 축소된다. 결국 임계치 0.65를 넘어서는 네트워크에서 나타나는 구조적 패턴은 직업 간 기능적 연계가 특정 고속련 집단에 집중되고 다른 직군과의 상호작용이 단절되는 직업 구조의 양극화를 여실히 보여준다. 이는 숙련의 계층화가 노동시장의 구조적 불균형으로 이어질 가능성을 함의한다.

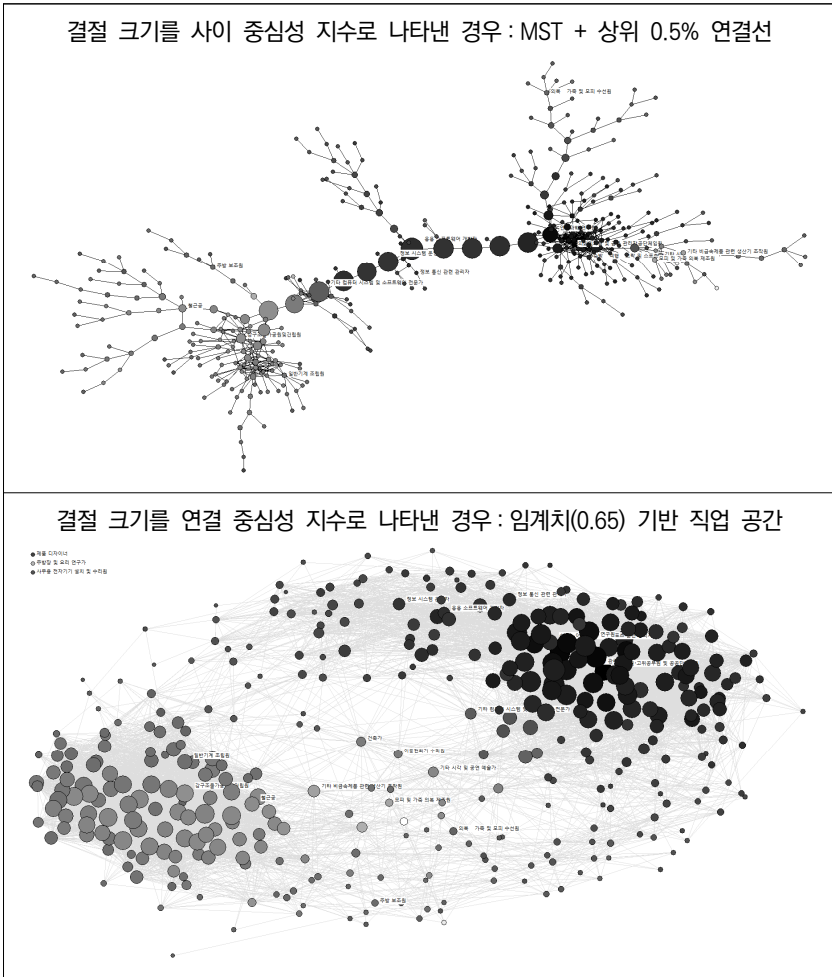
[그림 3-13]은 앞서 숙련의 양극화 분석에서 루빙 군집화로 도출된 인지 숙련 유형을 기준으로 직업별 숙련 구성에 가중치를 적용하여 산출한 인지 숙련 비중을 MST 구조와 임계치 0.65 이상 네트워크에 투영한 것이다. 두 그림 모두 색조의 진함은 인지 숙련 비중이 높음을 의미하며, 직업별 기능적 연관 구조 속에서 인지 숙련도가 어떤 방식으로 분포되어 있는지를 시각적으로 보여준다.

MST 구조에서 사이 중심성이 높은 직업군은 '응용 소프트웨어 개발자', '미디어 콘텐츠 디자이너', '작가', '만화가 및 만화영화 작가', '정보 시스템 운영자', '기자 및 언론 관련 전문가', '기타 컴퓨터 시스템 및 소프트웨어 전문가', '웹 개발자', '네트워크 시스템 개발자' 등으로 나타난다. 이들은 주로 지식정보·창의 산업 분야의 전문가 직종으로, 네트워크 내에서 지식과 숙련의 전이를 매개하는 교량적 역할을 한다. MST의 구조상 이들 직종은 두 대규모 집단을 연결하는 중심 경로를 따라 배치되어 있으며, 직업 체계 전반의 기능적 연결성을 유지하는 핵심적 중개 지점으로 작용한다.

이에 비해 임계치 0.65 이상 네트워크에서는 연결 중심성이 높은 직업이 주로 화이트칼라 직종에 집중되어 있다. 구체적으로는 '보험 심사원 및 사무원', '교육 관리자', '회계사', '마케팅 및 광고홍보 관리자', '보험 및 금융 관리자', '변호사', '기획 및 마케팅 사무원' 등이 중심에 위치하며, 일부 '유리 제조 및 가공기 조작용', '건설 및 광업기계 설치 및 정비원', '단조원', '강구조물가공원 및 건립원' 등 일부 제조·기능직도 높은 중심성을 보인다. 그러나 전체적으로 볼 때, 화이트칼라 직종의 중심성 지수가 상대적으로 훨씬 높아, 네트워크의 영향력 구조가 고인지 숙련 직업군을 중심으로 조직되어 있음을 확인할 수 있다.

두 네트워크를 비교하면, MST에서는 교량부를 중심으로 화이트칼라와

[그림 3-13] 루뱅 방식의 인지 숙련 비중을 투영한 직업 간 연관성 네트워크 구조



주: 각 결절은 숙련을 나타내며, 결절 간 연결은 연관성 지수에 기반하고, 원 크기는 각 결절의 사이와 연결 중심성의 상대적 크기를 반영함.
 자료: 저자 작성.

블루칼라 직종이 혼재된 연속적 구조를 보이지만, 임계치 0.65 이상에서는 이 경계가 한층 분명해진다. 즉, 고인지 숙련 직업군이 네트워크의 중심을 차지하고, 중·저인지 숙련 직업군은 주변부로 밀려나는 이중적 양극화 구조가 형성된 것이다.

이 결과는 루뱅 군집화로부터 산출된 인지 숙련 비중이 직업 간 기능적

연관성을 구분하는 핵심 기준 변수임을 보여준다. 직업별 인지 숙련 비중이 높을수록 네트워크 내 중심성 지수도 동반하여 올라가며, 이는 숙련의 인지적 차원이 직업 간 연계 구조의 위계를 결정하는 주요 요인으로 작용함을 시사한다. 결국, 한국의 직업 체계는 고인지 숙련 직업을 중심으로 기능적 연결망이 형성되고, 중·저인지 숙련 직업군은 점차 주변화되는 지식 집약적 이중 구조로 전환되고 있다.

요약하면, 루뱅 기반 인지 숙련 구조와 중심성 지표 분석은 직업 간 연관성이 단순한 기술적 유사성을 넘어 숙련 수준의 인지적 위계화를 반영함을 보여준다. 이는 한국의 직업 체계가 점차 고숙련 중심의 핵심 직업군을 중심으로 재조직되고, 숙련의 사회적 확산보다는 집중을 통해 직업 구조의 양극화를 심화시키고 있음을 보여준다.

3) 직업 간 연관성이 직업 간 임금 수준 차이에 미치는 효과

〈표 3-4〉는 '2020년 KNOW 재직자조사' 자료를 이용하여, 직업 간 연관성이 로그 임금 격차에 미치는 영향을 중력모형(gravity model) 형태로 추정된 결과다. 분석 단위는 한국표준직업분류 세세분류 기준 직업이며, 추정의 강건성 확보를 위해 직업 클러스터 수준에서 오차항을 조정한다. Henning et al.(2025)이 지역의 일관성(coherence)과 임금 사이에 음(-)의 관계를 실증적으로 제시한 것과 달리, 이 분석은 가용한 세밀한 지역 자료가 부재하여 직업 간 숙련 연관성 네트워크를 통해 유사한 구조적 관계를 검증한다.

모형 1은 기본 모형으로, 직업 간 연관성, 교육연수 차이, 연령 차이를 포함한다. 직업 간 연관성 계수는 0.111로 1% 수준에서 유의한 양(+)의 값을 보인다. 이는 숙련이 비슷할수록 같은 노동시장 내 경쟁이 심화되어 임금 격차가 벌어지는 평균적 경쟁 효과를 의미한다. 따라서 숙련 연관성이 높을수록 임금 수준이 상향 평준화되기보다는 차별화된다.

모형 2는 앞서 숙련의 양극화 분석에서 루뱅 알고리즘을 이용해 직업 간 기능적 연관성으로부터 도출된 인지 숙련 비중 변수를 활용한다. 이는 직업 간 숙련 구조를 관계적 관점에서 추정된 것으로, 실제 노동시장에서의 기술 및 지식 연계, 산업 내 기능 분화를 명확하게 반영한다. 이 모형에서 직업 간 연관성 계수는 1.411로 1% 수준에서 유의한 양(+)의 값을 보인다. 이는 단순

〈표 3-4〉 직업 간 연관성이 직업 간 로그 임금 차이에 미친 효과

종속변수 : 직업 간 로그 연봉 차이	모형 1	모형 2
직업 간 연관성	0.111*** (3.18)	1.411*** (2.72)
직업 간 교육연수 차이	0.062*** (9.04)	0.061*** (9.12)
직업 간 연령 차이	0.007*** (3.69)	0.007*** (4.12)
직업 간 평균 인지 숙련 비중	-	3.209*** (4.94)
직업 간 인지 숙련 비중 차이	-	-0.407 (-1.63)
상호작용항 [직업 간 연관성×평균 인지 숙련 비중]	-	-2.670*** (-2.89)
상수항	0.092*** (3.51)	-1.539*** (-4.13)
R ²	0.157	0.218
N	73,920	73,920

주 : 1) 클러스터 오차(KSCO_7)를 적용하여 추정함.

2) *, **, ***는 각각 10%, 5%, 1% 수준에서 통계적으로 유의함을 나타냄.

3) 괄호 안의 수치는 *t*값임.

자료 : 저자 작성.

한 경쟁 효과를 넘어, 노동시장에서 나타나는 양극화, 즉 고숙련 직종 내부의 집중과 경쟁을 포착한 것이다. 이는 앞서 MST와 임계치(0.65) 네트워크에서 확인된 화이트칼라-블루칼라 이중 구조와 부합되는 결과다. 두 분석 모두 고인지 숙련 직종이 네트워크 중심부를 차지하고, 중·저숙련 직종은 주변부로 밀려나는 경쟁적 이중 구조를 보여준다. 따라서 모형 2의 양(+)의 연관성 효과는 ‘직업 공간의 구조적 밀집 → 내부 경쟁 강화 → 임금 격차 확대’ 과정을 실증한다.

그러나 상호작용 항(직업 간 연관성×평균 인지 숙련 비중)은 음(-)으로 유의하게 나타나, 인지 집약도가 높을수록 숙련 연관성이 임금 격차를 완화하는 방향으로 작동함을 보여준다. 즉, 고숙련 직업군에서는 직업 간 숙련 연관성이 경쟁을 심화하기보다는 보완적 안정화 기제로 전환된다. 이러한 조

건부 효과는 노동시장 내 구조적 이중성을 반영하며, 고숙련 영역에서는 보완적 수렴이, 저숙련 영역에서는 경쟁적 차별화가 우세한 이중적 구조가 공존하고 있음을 시사한다.

통제 변수의 효과에 대해 살펴보면, 직업 간 교육연수 차이는 모든 모형에서 일관되게 양(+)으로 유의하며, 교육 수준 격차가 클수록 임금 격차가 확대된다. 연령 차이 또한 양(+)으로 유의하여, 세대 간 숙련·임금 구조 차이가 직업 간 불평등에 기여함을 보여준다. 평균 인지 숙련 비중은 유의한 양(+)의 부호를 보이며, 인지 집약도가 높은 직업 쌍일수록 평균 임금 수준이 높고 격차도 크다는 점을 시사한다. 반면에, 인지 숙련 비중 차이는 음(-)의 부호이지만, 직업 클러스터 수준의 표준 오차를 적용한 결과, 통계적으로는 10% 수준에서 비유의하게 나타난다. 계수의 유의성이 제한적임에도 불구하고, 인지 집약도 차이가 큰 직업 쌍에서는 임금 격차가 상대적으로 작게 나타난다. 환언하면, 고숙련과 저숙련 직종 간에는 이미 노동시장 구조가 분리되어 있어 직접적인 경쟁이 약하다는 것을 의미한다.

이러한 분석 결과는 관련 다양성(related variety) 논의와 밀접하게 연결된다(Frenken et al., 2007; Boschma, 2017). 모형 2의 양(+)의 연관성 계수는, 숙련 연관이 집중될수록 경쟁이 강화되고 임금 격차가 확대되는 경쟁적 이중 구조를 반영한다. 이는 직업 체계가 유사한 연결성을 기반으로 결집되지만, 그 결집이 내부 경쟁을 심화시키는 양면적 구조임을 함의한다. 즉, 유사한 네트워크 결집이 내부 경쟁을 내재화하는 것이다. 이는 직업 간 숙련 연관성이 단순히 협력적 구조가 아니라, 내부 분화와 임금 양극화를 동시에 심화시키는 '연관성 속의 경쟁' 구조임을 보여준다. 이는 노동시장 양극화, 특히 고숙련 화이트칼라 직종 집단의 임금 상위 집중과 부합한다.

또한, 분석 결과는 노동시장의 임금 격차가 단순한 숙련 수준의 차이가 아니라, 숙련 네트워크 내 위치와 연결 구조의 차이에서 비롯됨을 보여준다. 전체적으로는 고인지 숙련 직종이 임금 격차를 주도하며, 전술한 '고숙련 집중 → 내부 경쟁 심화 → 임금 양극화 확대'라는 경로가 유지된다. 그러나 상호작용항이 음(-)이라는 점은, 이러한 거시적 불평등 구조와는 다른 층위에서 고숙련 내부의 경쟁 완화와 보완적 조정이 부분적으로 작동함을 함의한다. 즉, 임금 격차 확대를 주도하는 것은 고숙련 집단이지만, 그 내부의

작동 메커니즘은 경쟁적이라기보다 조정적이라는 점에서 다층적 구조를 이룬다.

제4절 소 결

본 연구는 한국고용정보원이 발간한 '2017~2020년 KNOW 재직자조사' 자료를 활용하여, 숙련 보완성과 직업 간 숙련 연관성 지수를 네트워크 관점에서 분석함으로써 한국 노동시장의 숙련 및 직업 양극화 구조를 이론적으로 탐색했다. 이러한 분석은 MST와 루뱅 알고리즘을 통해 네트워크의 위상 구조와 클러스터 결집 양상을 규명하고, 나아가 임금 결정 과정에서 이러한 구조적 요인이 어떻게 작용하는지를 간단한 회귀 분석을 통해 실증적으로 검증했다.

이로부터 밝혀진 분석 결과를 요약하면 다음과 같다.

첫째, 숙련 간 보완성 및 연관성 지수의 분포는 전체적으로 정규형에 가깝지만, 네트워크 위상 구조에서는 '중심-교량-주변부'의 삼중 구조가 나타난다. 즉, 한국의 숙련 체계는 표면적으로는 연속적이지만, 내부적으로는 인지·사회 숙련과 신체 숙련으로 이원화된 구조를 형성하고 있다. 보완성 지수가 일정 임계치(약 0.6)를 넘어설 경우, 일부 숙련 클러스터가 결속되어 외부와 단절되는 '임계적 양극화' 상태로 전이된다. 이러한 구조는 숙련 체계가 안정성과 중심성을 유지하면서도, 강한 보완성이 축적되는 영역을 중심으로 구조적 분화가 심화되고 있음을 의미한다.

둘째, 루뱅 알고리즘을 적용한 결과, 전체 숙련은 인지·신체·사회(기타) 숙련의 세 유형으로 구분되며, 특히 인지-신체 숙련 축을 중심으로 명확한 이중 구조가 형성된다. 인지 숙련은 사회적 숙련과 결합하며 네트워크의 중심을 이루지만, 신체 숙련은 주변부로 밀려나 낮은 결속도를 보인다. 이는 한국의 숙련 체계가 인지 숙련 중심의 위계적 결속 구조를 내포하고 있음을 보여준다. 또한 교육연수 분포 역시 인지 숙련 클러스터에 집중되어 있으며, 신체 숙련 집단으로 갈수록 급격히 감소하는 교육·훈련 격차 기반의 구

조직 양극화가 확인된다.

셋째, 397개 직업의 연관성 네트워크를 분석한 결과, MST의 시각화에서 화이트칼라(고인지 숙련)와 블루칼라(중·저인지 숙련) 직업군이 분리된 구조가 나타난다. 루뱅 알고리즘을 적용하면 세 개의 직업군, 즉 고인지 숙련, 중·저인지 숙련, 기타 직업군이 도출되며, 이는 고숙련 중심의 위계적 직업 체계를 시사한다. 임계치(0.65) 이상의 네트워크에서는 고숙련 직종의 중심 집중과 주변부 직군의 분리가 더욱 뚜렷해지며, 한국의 직업 구조가 점차 '인지 숙련 중심-기타 숙련 주변'의 이중적 구조로 수렴하고 있음을 보여준다. 이때 직업별 인지 숙련 비중은 네트워크 중심성과 밀접하게 연동되며, 숙련의 인지적 차원이 직업 간 위계와 연결 구조를 결정하는 핵심 요인으로 작동한다.

넷째, 임금 수준과 그 격차의 결정요인의 분석 결과, 인지 숙련 비중은 유의한 양(+)의 효과를, 신체 숙련 비중은 음(-)의 효과를 보이며, 사회 숙련의 효과는 상대적으로 미약하다. 이는 '인지 숙련 프리미엄-신체 숙련 디스카운트' 구조가 임금 결정 과정 전반에 일관되게 작용하고 있음을 시사한다. 한편, 임금 격차의 요인 분석에서 직업 간 연관성 계수의 양(+)의 효과는 숙련 유사성이 집중될수록 경쟁이 강화되고 임금 격차가 확대되는 '연관성 속의 경쟁 구조'를 반영한다. 그러나 상호작용항(직업 간 연관성×평균 인지 숙련 비중)의 음(-)의 계수는, 고숙련 내부에서는 오히려 유사성이 경쟁 완화와 조정 기능을 수행함을 보여준다. 즉, 거시적으로는 고숙련 집중이 임금 양극화를 주도하지만, 미시적으로는 내부 조정이 병행되는 다층적 구조가 존재한다.

이러한 분석 결과로부터 '임계적' 양극화의 의미를 뒤집어 볼 필요가 있다. 한국의 숙련과 직업 체계는 전반적으로 보완성과 연관성이 높은 정규형 분포를 보이지만, 그 내부 구조에서는 이미 고숙련·고인지 직종의 결집과 저숙련 직종의 주변화가 병행되는 임계적 양극화가 진행되고 있다. 숙련 네트워크의 보완성 지수가 일정 수준을 넘어서면 일부 숙련 클러스터가 내부적으로 결속되며, 외부와의 연결성이 급격히 약화된다. 이때 네트워크의 분포는 연속적으로 보이지만, 위상 구조는 이미 이중적 안정상태, 즉 고숙련 중심의 핵심 영역과 저숙련 주변부로 분리되어 있다. 이는 한국의 숙련 체

계가 아직 미국식 쌍봉형 양극화에 이르지지는 않았지만, 겉으로의 연속성과 내적 분절이 공존하는 잠재적 균열 단계에 진입했음을 의미한다.

직업 네트워크에서도 동일한 현상이 반복된다. 직업 간 연관성 지수의 분포는 전체적으로 균형적이나, MST 및 임계치 0.65 이상의 네트워크에서는 이미 화이트칼라와 블루칼라 직종의 위계적 이원화가 뚜렷하게 나타난다. 루뱅 알고리즘으로 분류된 세 개의 직업군, 즉 고인지 숙련, 중·저인지 숙련, 기타 직업군은 서로 명확한 경계를 이루며, 특히 고인지 숙련 직종이 네트워크의 중심에서 높은 영향력을 행사하고 있다. 이러한 현상은 숙련 네트워크에서의 인지 숙련 중심 구조와 정확히 병행된다. 즉, 숙련의 인지적 위계화가 직업 구조의 기능적 분리를 매개하면서, 숙련과 직업 모두에서 중심-주변의 이중적 구조가 강화되고 있다. 인지 숙련과 직업의 집중은 숙련과 직업 네트워크의 결속도를 강화하지만, 동시에 외부 확산과 새로운 연결의 가능성을 제한한다. 그 결과, 유사한 숙련 간 긴밀한 연계를 의미하는 높은 연관성과 분업적 협업을 기반으로 한 강한 보완성이 결합되어도, 이는 새로운 숙련이나 산업으로의 확산으로 이어지지 못하고 오히려 고숙련 중심의 집중화와 저숙련 주변부의 고립화를 심화시킨다. 다시 말해, 보완성은 숙련 체계를 단단히 묶어주지만, 그 결속이 과도할 경우 새로운 연결을 차단하고, 연관성은 직업 간 지식과 숙련의 순환을 촉진하지만, 유사성이 높을수록 내부 경쟁을 강화하는 방향으로 작동하는 ‘보완성과 연관성의 역설’이 나타난다.

이러한 임계적 양극화는 숙련과 직업의 관계가 단순히 기술적 특성의 차이가 아니라, 인지적 속성과 네트워크 연결성의 결합에 따라 재편되고 있음을 시사한다. 숙련 차원에서 인지적 역량이 핵심 축으로 작동하면서 신체 숙련을 주변화시키고, 직업 차원에서는 이 인지적 위계가 노동시장 내 기능적 분화를 구조화한다. 그 결과, 임금 구조 또한 ‘인지 숙련 프리미엄-신체 숙련 디스카운트’의 형태로 고착되며, 노동시장의 불평등이 숙련 구조의 위계와 정합적으로 연결된다. 다만, 다른 층위에서 보면, 고숙련 영역 내 연관성은 단순히 경쟁을 심화시키기보다 체계 내부의 조정과 균형을 촉진하는 방향으로 작용한다. 즉, 거시적 차원에서는 고숙련 중심의 불균형과 집중이 심화되지만, 미시적 차원에서는 이러한 불균형을 일정 부분 흡수하고 완화

하는 보완적 구조가 함께 존재하는 다층적 균형 상태가 형성되고 있다.

정책적으로 이러한 결과는 단순한 고숙련화나 인지 중심 교육 강화만으로는 불평등을 완화하기 어렵다는 점을 시사한다. 숙련 체계의 균형적 발전은 보완성과 연관성의 임계점을 관리하는 정책적 조정 능력에 달려 있다. 즉, 이러한 숙련 집중이 체계 결속으로 기능할 때는 긍정적이지만, 과도할 경우 구조적 단절을 초래한다는 점에서, 숙련 간 연결성의 다양성과 교차성을 유지하는 것이 핵심 과제다. 따라서 한국의 숙련·직업 정책은 고숙련 집단의 내부 경쟁 완화와 중·저숙련 집단의 연결성 강화, 그리고 인지 숙련 중심 구조에 대응하는 다층적 네트워크 균형 전략으로 전환되어야 한다. 이는 숙련 체계의 안정성과 포용성을 동시에 확보하기 위한 접근으로 임계적 양극화 완화에 기여할 수 있을 것으로 생각된다.

제 4 장

숙련 네트워크에서 브로커리지의 역할 : 노동시장 매개 구조 분석

제1절 연구 배경과 목적

AI 기술 확산과 디지털 전환은 노동시장의 숙련 구조를 급격하게 재편하고 있다. AI 관련 기술들은 기업의 생산 및 물류, 마케팅, 인사 구조를 근본적으로 혁신하도록 추동하고 있으며, 기업들은 이러한 신기술을 도입하기 위하여 필요한 인력에 대한 채용공고를 내고 있다. 본 연구는 이러한 채용 공고에서 요구하는 숙련을 네트워크 방법론을 통해 분석하여 기술 숙련 지도를 만들고, 나아가 새롭게 재편되는 노동시장의 숙련수요의 구조를 파악하고자 한다.

기존 연구는 주로 특정 숙련의 전체 숙련 네트워크상에서의 중심성에 주목해 왔다. 하지만, 산업과 직종으로 갈라지고 구획되는 숙련 간 경계를 가로지르는 중개(brokerage) 역할에 대한 분석은 상대적으로 부족하다. 본 연구는 숙련 네트워크의 산업/직종으로 구성되는 블록 간 연결을 중개하는 브로커 숙련에 주목하며, 이들이 노동의 수요 및 임금에 미치는 영향을 탐색한다.

특히 Gould and Fernandez(1989)의 브로커리지 역할 유형과 Burt(1992)의 제약 지수(constraint index)를 적용하여, 구조적 중개 위치에 있는 숙련의 특성과 전략적 중요성을 밝히고, 이러한 구조적 특성이 기술 수요 및 그

보상(임금)에 어떤 영향을 끼치는지를 회귀분석을 통해 실증적으로 분석한다. 곧, 본 연구는 기존의 숙련 네트워크 분석 프레임에 새로운 구조적 위치 개념인 브로커리지를 도입하여 숙련 네트워크상의 블록 간 매개 구조에 대한 보다 정교한 분석 도구를 제시하고자 한다. 이를 통해, 노동시장에서 수요 및 보상이 어떤 메커니즘을 통해 결정되는지에 관해 새로운 관점을 제시하고자 한다.

제2절 이론적 배경

1. 브로커리지 이론

Burt(1990)는 두 거래 상대가 연결되지 않은 공간인 ‘구조적 공백(structural hole)’이 발생할 때, 이들을 연결함으로써 이득을 얻는 ‘중개’가 생겨난다고 지적한다. 이때 ‘구조적 공백’에 얼마나 근접하느냐, 얼마나 많은 구조적 공백에 위치해 있느냐에 따라 사회적 자본의 양이 결정된다(Burt, 2005). 사회적 자본은 행위자에게 특정한 사회구조적 유리함(advantage)을 제공한다(Bourdieu and Wacquant, 1992; Coleman, 1988, 1990). 네트워크상의 구조적 공백을 이용한 우월함/경쟁적 이점(competitive advantage)을 사회적 자본의 형태로 보유한 행위자가 브로커이다(황지영 외, 2025).

숙련 네트워크에서 브로커에 해당하는 숙련은 숙련들 혹은 그 숙련을 보유한 자들 간의 정보 비대칭(Akerlof, 1970)을 해소하고, 숙련 간 친화성(기능적 유사성 및 상보성)을 확보하여 두 숙련 집단을 잠재적으로 연결하여 지역적 혹은 산업마다 산개한 숙련(집단)을 중개하는 역할을 한다.

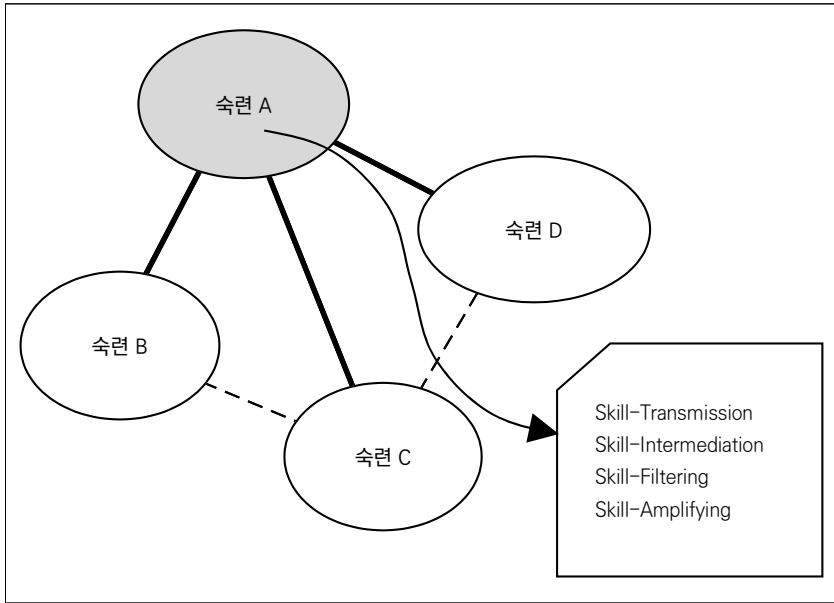
이러한 브로커리지가 가능한 것은 각각의 숙련 간에 잠재적으로 구축되어 있는 숙련 간 ‘유사성’에 기반한다. 즉, 중개는 유사성을 통해 네트워크의 구멍을 메움으로써 숙련 혹은 숙련 보유자들을 연결하는 행위이다. 브로커 혹은 중개자(intermediary)는 숙련 간 유사성을 그들 고유의 자본으로 하고, 구조적으로 고립된 숙련(보유자) 간 정보 비대칭(Akerlof, 1970)과 과도한 탐

색/전환 비용(searching cost, Mortensen, 1982)으로 인한 숙련 간 결합의 실패 가능성을 이용하여, 양측에 적절한 양의 '숙련 간 상보적 유사성'이라는 매개 가능성을 공급함으로써 숙련 간 결합을 가능케 한다. 곧, 이러한 브로커 위치의 숙련(보유자)은 특정 숙련들의 구매자와 공급자 사이의 거래 실패 가능성을 완화시키고 조절(coordinate)하는 역할을 한다(Spulber, 1996; 황지영 외, 2025). 마지막으로, 이들은 전략적으로 구매자와 공급자 간의 연결-분리 구조를 유지함으로써 구조적 공백 아래 그들의 이윤을 장기적으로 확보한다.

[그림 4-1]은 숙련 A(예: 솔루션·SI·CRM·ERP)가 수많은 기술들의 네트워크에서 어떻게 브로커리지 역할을 수행할 수 있는지를 보여준다. 솔루션은 업무 효율성을 높이기 위해 데이터 분석, 회계, 인사 관리 등의 맞춤형 프로그램을 제공하는 업무이다. SI(System Integration)는 고객사의 요구에 맞춰 공공기관과 같은 조직에 IT 인프라 서비스를 제공하는 업무이다. CRM(Customer Relationship Management)은 고객 정보를 데이터베이스로 만들어 영업 관련 정보를 분석하여 제공하는 업무라면, ERP(Enterprise Resource Planning)는 재무, 회계, 생산, 인사, 구매 등 기업 조직의 다양한 업무를 하나의 통합된 시스템으로 관리하는 솔루션을 제공하는 업무이다. 이 숙련들은 기업이나 산업 조직 내부에서 인사나 생산 조직들의 정보와 자원의 흐름, 의사 결정을 지원하는 통합적인 서비스를 제공함으로써 다양한 조직 간의 매개 역할을 한다. 또한, 산업 블록 간에서도 온라인 유통업체와 금융업체, 제조업(그림 4-1의 숙련 B, 숙련 C, 숙련 D) 간의 데이터와 물류의 흐름을 조정하고 통합-관리하는 솔루션을 제공한다. 이 브로커리지 숙련이 없었으면 연결되지 않은 채 각자의 업무를 수행했을 조직이나 숙련 보유자가 이 통합 인프라 서비스로 인해 연결되어 협업을 수행할 수 있게 되는 것이다.

이러한 브로커리지 구조를 네트워크상에서 어떻게 포착할 수 있을까? 기존 연구는 크게 두 갈래로 나뉜다. Gould and Fernandez(1989)의 브로커리지 유형 연구와 Burt(1992) 구조적 공백 연구들은 유사한 브로커리지 구조를 다소 다른 방식의 이론과 측정 도구를 이용해서 포착한다. 본 연구는 다음의 순서로 진행된다. 먼저, 기존의 네트워크 이론에서 많이 보고되는 중심성 메저들을 소개한다. 다음으로, Gould and Fernandez의 5가지 브로커리

[그림 4-1] 숙련 네트워크 공동체의 브로커리지 구조



자료 : Lee and Goh(2023), p.25, Figure 1을 기반으로 저자 작성.

지 유형을 설명한 후 Burt의 ‘제약’ 관련 변수를 3가지 정도 추가로 논의할 것이다. 마지막 장에서는 이 브로커리지 변수들을 degree centrality와 교육 변수, 대산업 고정효과 더미, 대인관계 기술 등을 통제한 회귀분석 결과를 보여준다.

2. 브로커리지를 이용한 기존의 연구들

조직 및 네트워크 연구들은 네트워크상의 구조적 위치가 어떻게 조직 내부 혹은 조직 내부와 외부 간의 권력, 영향력, 성과 등에 영향을 미치는지를 탐구해 왔다(Burt, 1992; Laumann, 1976; Coleman, 1990). 특히 최근의 네트워크 연구들은 제약, 의존 개념을 중심으로 네트워크상의 ‘매개’와 ‘중개’의 역할이 어떻게 서로 다른 구조적 위치에 놓인 주체들을 연결해 줌으로써 영향력을 발휘하는지를 설명해 왔다.

소셜 네트워크 관련 문헌들에서 브로커는 두 연결되어 있지 않은 주체들

을 연결해 주는 매개자 위치에 있는 자로 규정하고, 그러한 단절과 연결의 구조를 브로커리지로 규정한다. Granovetter는 이제 고전의 반열에 오른 그의 1973년 논문에서 약한 연결망이 더 효과적으로 새로운 정보와 자원을 이전시키며 강한 연결망은 그 내부의 견고하고 중복적인 관계들로 인해 그러한 기능을 하지 못함을 밝혔다(Granovetter, 1973). Gould and Fernandez (1989)는 Granovetter의 정보 이전 아이디어에 구조적 시각을 덧붙여 5가지 브로커리지 구조를 제시했다(아래 'Gould and Fernandez의 브로커리지 유형' 항 참조).

보다 최근의 연구들은 브로커리지 액션에 주목하여 네트워크의 컨택 포인트들에 결합하는 과정, 그 과정에서 컨택 포인트들 사이를 매개하기 시작하는 과정, 동시에 컨택 포인트들 사이의 분리를 유지하는 과정을 집중적으로 분석하고 있다(Obstfeld et al., 2014; Grosser et al., 2017).

브로커리지는 두 가지 양면성을 갖고 있다. 정보와 자원으로부터 소외된, 다양한 거래 상대나 협업의 상대를 갖지 못한, 구조적으로 열악한 상황에 있는 행위자들에게 컨택 포인트를 마련해 줌으로써 시장의 생산자나 소비자 네트워크 혹은 시민사회와 정치사회의 정치적 동원 네트워크 등에 점점 마련해 줌으로써 소외된 주체나 자원(이 논문에서는 숙련)을 동원해 내는 역할을 하는 것은 긍정적인 면이다. 하지만, 이러한 동원 메커니즘은 '구조적 착취'(structural exploitation)의 가능성을 동시에 내포하고 있다(Stovel and Shaw, 2012). 네트워크에 접속된 행위자가 브로커를 통하지 않고서는 정보와 자원을 가진 다른 행위자들에 접근할 수 없기 때문에 브로커가 중간에서 편취를 해도 불이익을 감수할 수밖에 없는 것이다. 브로커는 이러한 구조를 유지함으로써 편취를 장기화할 수 있기 때문에, 구조를 지속시키고자 하는 브로커와 이 구조를 뛰어넘어 정보와 자원 보유자와 직접 거래하려는 구조적 소외자 사이에 갈등이 발생할 수 있다. 이러한 구조를 지속시키고자 하는 브로커에 대한 의존성(dependency)이 높고, 이 브로커를 경유해야만 가치 있는 자원에 접근할 수 있는 자가 많아질수록, 브로커의 파워는 증대된다(Marsen, 1982, 1983; Coleman, 1973).

Burt는 이러한 기존의 브로커리지 연구를 한층 확장시켜 구조적 공백하에서의 브로커리지의 전략적 가치를 측정하는 모델을 완성한다(1992,

2004). Burt는 ‘제약(constraint)’이라는 지표를 통해 개인의 네트워크가 한 집단에 얼마나 집중되어 있는지(즉, 그로 인해 구조적 공백이 존재하는지)를 측정했다. Burt는 구조적 공백이 많은 네트워크가 관리자의 성과에 긍정적인 영향을 줄 수도 있음을 보여준다. 동시에 제약이 큰, 즉 구조적 공백이 부족한 네트워크(응집성이 높은 네트워크)도 집단의 공통 이해와 정체성을 강화하는 장점이 동시에 있을 수 있음을 인정한다. Uzzi의 연구(1997)에 따르면, 직무나 산업, 기업 조직 내부의 응집성이 너무 강해서 외부와의 연결이 부족하거나, 너무 약한 상태에서 외부와의 브릿지만 활성화될 경우 모두 기업 성과가 좋지 않음을 보여준다. 그는 응집성과 외부와의 연결성을 모두 중간 정도로 갖고 있는 조직이 가장 성과가 좋음을 실증함으로써 집단 내부 응집성과 외부와의 브로커리지의 균형이 중요함을 강조한다. Ahuja(2000)는 (Burt의) 예측과 달리 파트너들 간 상호 연결이 없는 개방적 네트워크가 협업 및 조정 비용을 상승시켜 오히려 혁신을 저해할 수 있음을 밝혔다. 맥락에 따라 브로커리지/구조적 공백 구조는 부정적 효과를 나타낼 수도 있는 것이다.

이러한 소셜 네트워크와 브로커리지 개념의 고전들에 기반하여 최근의 연구들은 대규모 숙련 네트워크에서 블록모델과 중개구조를 탐색하고 있다. Alabdulkareem et al.(2018)은 도시 및 산업의 스킬 네트워크에서 기술 군집들이 사회·인지 대 감각·물리 군집으로 나뉘는 양극화 현상을 보이고 있으며 고임금 직무가 전자에 집중되고 있음을 발견했다.

최근에는 O*NET 데이터를 이용하여 직무와 숙련의 복잡성과 임금과의 관계를 설명하는 연구들도 존재한다. 이들은 직업과 기술을 연결한 이중모드 네트워크를 구축해 각 직업의 기술 복잡성(complexity)과 기술 일관성(coherence)이 임금에 끼치는 영향을 밝혀내거나(Aufiero et al., 2024), 통계적으로 유효한 연결(statistically validated network)의 직업별 ‘fitness’와 ‘complexity’가 임금과 어떤 관련을 맺는지를 테스트한다(Aufiero et al., 2024).

또 다른 연구는 영국의 구인공고를 바탕으로 구축한 6,500만 건의 채용공고를 분석해서 기술 발전과 함께 등장하는 새로운 스킬 클러스터들의 지역·시간별 변화의 추이를 발견해 냈다(Liu et al., 2025). 이와 유사한 다른 연

구는 컴퓨터 공학 분야 채용공고에 대한 네트워크 분석을 통해 어떤 기술이 중심적인 위치에 있는지, 어떤 기술들이 특정 커뮤니티 블록을 구성하는지를 밝혀 냈다(Maghsoudi, 2024). 기술 구성에 대한 네트워크 연구들은 특정한 기술의 조합을 가진 이들이 다른 직업으로 더 쉽게 이동할 수 있음을 보여주기도 한다. 이러한 기술의 구성에 대한 연구들은 재훈련 프로그램의 설계를 위한 정책적 방향을 제시해 줄 수 있을 것이다(Waters and Shutters, 2022).

숙련 네트워크에서 더 자주 쓰이는 숙련은 그 소유자와 소유자가 속한 기업으로 하여금 다른 연관 산업으로 다각화를 할 수 있도록 해준다. Neffke and Henning(2013)은 이러한 skill relatedness를 노동 이동을 통해 추정함으로써 기업의 인력이동이 숙련의 전이가능성을 측정하는 대리지표로 기능할 수 있음을 밝혔다. 또한 이 연구는 숙련 또는 기술 공출현 네트워크에서 스킬 간 공출현 빈도(가중치)를 반영하여 중심성을 재정의할 필요가 있다는 점, 그리고 중심 노드들 사이의 강한 연결 집단화(rich club)가 네트워크의 통제 구조(control structure)를 형성할 수 있다는 관점을 제공한다. 만약 숙련 네트워크에서 특정 기술들이 서로 강한 상호작용을 많이 한다면, 그 기술들이 중심적 역할을 하거나 통제적인 영향력을 행사할 가능성이 있다.

3. 중심성 지표들의 한계와 브로커리지의 독자적 기여

중심성 지표들은 네트워크상에서 다양한 중심성 정의를 충족시키는 노드들의 구조적 중요성을 밝히는 데 유용한 도구이다. 예를 들어 연결 중심성(degree centrality)은 연결 수의 많고 적음에 따른 중심성을, 매개 중심성(betweenness centrality)은 경로상에 매개 위치에 따른 중심성을, 근접 중심성(closeness centrality)과 고유벡터 중심성(eigenvector centrality)은 전체 네트워크 안에서 이웃들에게 갖는 중요성에 따른 전반적 위치의 중심성을 측정한다. 하지만, 이들 지수는, 그 연결이 동질적 집단 안에만 집중되어 있는지, 아니면 이질적 집단 사이에 걸쳐 있는지를 구분하지 못한 채 특정 집단 간 경계에 의해 구획된 노드들 간의 유사성과 차이로 인해 발생하는 구조적 맥락을 포착하는 데에는 한계를 갖는다.

이에 반해, 브로커리지 지수들은 노드가 블록 내부에서는 동질적이지만, 블록 간에는 이질적인 집단(산업과 직종) 간의 관계를 어떻게 매개하는지를 정량화한다. 예를 들면, 동일한 수의 연결을 가진 두 기술이라도, 하나는 동일 산업 내 기술들과만 연결되어 있고, 다른 하나는 서로 다른 산업이나 직종을 연결하고 있다면, 후자가 기업에게 갖는 가치는 훨씬 클 수 있다. 후자가 타 산업이나 직종에서 그 기술을 통해 기술이 원래 속한 산업에 접촉하거나 그 산업으로부터 중요한 기술 셋을 뽑아내려는 수요가 높을 것이고, 그에 따른 산업과 직종 간 융합과 매개를 통한 혁신의 가능성 또한 높을 것이다. Gould and Fernandez(1989)의 다섯 유형(조정자, 이동자, 문지기, 대표자, 연결자)은 이러한 맥락을 세밀하게 구분할 수 있는 도구를 제공한다.

4. 브로커리지를 통한 경계 횡단의 전략적 가치와 보상 구조(임금) 분석

기존 연구는 “중심에 위치한 숙련일수록 (그 숙련의 높은 수요와 범용성으로 인해) 높은 보상을 받는다”는 경향을 지지해 왔다. 그러나 브로커 숙련은 이러한 중심성과 무관하게, 경계 매개의 희소성(scarcity) 때문에 독자적인 보상 프리미엄을 창출할 수 있다. 경계 매개의 희소성은 기술이 속한 내부 산업/직종에서 외부 산업/직종과 기술이 연결되고자 하는 수요가 발생할 경우, 경계 매개 위치로 인한 수요압박으로 그 희소한 가치가 더욱 올라가는 경우이다. 외부 산업/직종에서 그 기술이 속한 내부 산업/직종과 연결되고자 할 때도 경계 매개 위치로 인한 수요압박은 마찬가지로 현상이 나타난다. 특정 경계 매개에 위치한 숙련을 경유하지 않고는 내부의 다른 숙련으로 연결이 불가능할 경우, 경계 매개에 위치한 숙련은 더욱 희소성과 보상이 올라가게 된다. 결국, 브로커 숙련은 단순히 현재의 수요를 반영한다기보다는, 신기술의 확산과 산업 간 융합을 통한 새로운 산업의 출현을 예비하는 일종의 ‘촉매’이자 ‘거간’의 역할을 선제적으로 반영하는 지표라 볼 수 있다.

브로커리지 지표를 도입하면, 산업 간 혹은 직종 간 교차 영역에서 새롭게 등장하는 기술들의 전략적 중요성을 포착할 수 있다. 예를 들어, AI+물류 기술, 데이터 분석+인사 관리 기술은 기존 산업/직종 분류로는 서로 별 상

관없는 분야로 분리되어 있지만, 브로커리지를 통해 경계 연결 노드로 드러날 수 있다.

5. Gould and Fernandez의 브로커리지 유형

Gould and Fernandez(1989)는 네트워크상의 브로커 역할을 다음 다섯 가지 유형으로 구분한다: 조정자(coordinator), 이동자(itinerant), 문지기(gatekeeper), 대표자(representative), 연결자(liaison).

먼저 조정자(coordinator) 역할은 [그림 4-2]의 첫 번째 패널에서 확인할 수 있듯이, 동일한 산업이나 직무 카테고리 내부에서 다른 두 개인, 조직, 혹은 숙련(기술)을 연결하는 중개자 역할을 의미한다. 이러한 중개자는 같은 카테고리나 묶인 블록 집단 내에서 다른 두 주체나 기능, 기술을 연결하는 역할을 수행한다. 예를 들어, 특정 기업 조직이나 개발자 네트워크 그룹 안에서 특정 프로그래밍 언어를 잘 아는 사람이 특정 프로젝트 수행 시 조직 혹은 네트워크의 독립된 작업 단위나 개인을 묶어 주는 ‘조정자’ 역할을 수행할 수 있다. 장기적으로 이러한 기술 혹은 기술을 보유한 개인은 조직의 리더 혹은 관리자 역할을 수행할 가능성이 높으며, 더 높은 보상을 받을 가능성이 높다.

다음으로 이동자(itinerant)는, [그림 4-2]의 두 번째 패널에서 보여지듯이, 같은 산업 혹은 직무 블록에 속한 두 개인이나 기술을 연결하지만, 정작 이동자 본인은 그 집단의 구성원이 아닌 경우이다. 서로 연결이 안 되어 있는 두 조직에 외부 전문가가 필요한 기능이나 지식을 제공하며 두 조직 혹은 두 조직에 속한 개인을 연결하여 업무를 추진하는 예를 들 수 있다. 숙련 네트워크에 적용할 경우, 자동차 디자인 관련 숙련 혹은 그 숙련의 보유자가 생산조직 내부의 서로 연결되어 있지 않은 개인 혹은 하부부서를 연결하여 자동차 디자인을 관철시키는 경우가 이에 해당된다.

[그림 4-2]의 세 번째 패널에서 보이듯이, 문지기(gatekeeper)는 외부 집단에 속한 개인 혹은 기술과 자신이 속한 집단의 개인 혹은 기술을 연결하되, 그 개인과 기술/지식이 자신의 집단으로 들어올 때 적합한 내용들을 추려내는(sorting and selecting) 역할을 한다. 문자 그대로, 브로커가 외부 지

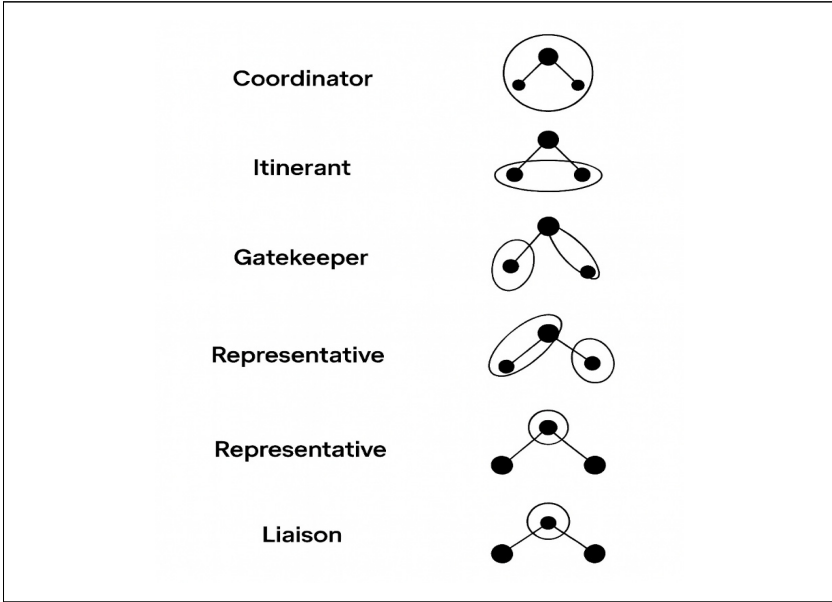
식과 자원의 집단 내부로의 유입을 관리하고 통제하는 ‘문지기’ 역할을 하는 경우이다. AI 관련 신기술이 외국계 기업을 통해 유입되었을 때, 특정 기업의 관련 전문가가 그 기술을 조직 내부로 번역하여 적용하고 소개하는 역할을 할 수 있다. 바로 그 전문가가 문지기 역할을 하는 셈이다. 이러한 문지기는 신기술의 한계와 가능성을 파악하고 있음과 동시에, 조직의 고유한 특성과 문화를 이해하고 있어, 신기술을 어떻게 조직에 녹여낼 수 있을지에 관한 구체적 로드맵을 짤 수 있는 자이다.

네 번째 유형인 대표자(representative)는 조직을 대표하는 정보와 아이디어, 이해관계 사항의 송출자이다. 이들은 집단의 대변자이자 대표자로 외부 조직과의 관계를 수립하고, 조율하며, 협상한다. 노동자들을 대표하여 자본가나 정부와 협상하는 노동조합 지도자들이 대표자 역할의 전형이다. 특정 기업의 내부 숙련을 외부 협력업체나 납품업체와 소통하며 이해시켜서 거래를 가능하도록 만드는 것 또한 대표자가 수행하는 역할이다.

마지막으로, 연결자(liaison)는 스스로는 어느 집단에도 속하지 않은 채로 서로 다른 두 집단에 속한 개인 혹은 기술을 연결하는 역할을 한다. 앞서 이동자(itinerant)와 다른 점은, 이동자는 같은 조직에 속한 두 개인을 밖에서 연결하지만, 연결자는 다른 조직에 속한 자들을 역시 밖에서 연결한다는 점이 다르다. 곧 이동자에 비해 훨씬 이질적인 두 개인 혹은 조직을 연결하는 역할을 할 가능성이 높다. (비합법) 이주민과 일용 혹은 단기 노동인력을 필요로 하는 업체를 연결해주는 이주민 취업 브로커들이 이러한 상호 이질적인 두 조직이나 개인들을 연결시켜주고 중간 수수료를 챙기는 연결자(liaison) 브로커들이다. 숙련 네트워크에 적용할 경우, 자동차 디자인 관련 숙련 혹은 그 숙련의 보유자가 완전히 서로 다른 두 부서, 예를 들면 마케팅 부서와 생산 공정 부서를 연결하여 자동차 디자인을 관철시키는 경우가 이에 해당된다.

이상 다섯 유형의 브로커리지를 통해 숙련 네트워크를 이해하게 될 때, 우리는 숙련이 각각의 독립된 자산이 아니라, 네트워크를 통해 서로 연결되어 있고, 각 숙련의 구조적 위치에 따라 서로 다른 방식으로 조직 간 정보와 자원을 매개하는 역할을 한다는 것을 알 수 있다. 본 연구는 이러한 브로커리지 개념을 숙련 네트워크에 적용함으로써, 특정 기술 역량이 조직 내부와 외부, 산업과 직종 내부와 외부로 가로지르는 ‘협력’의 양태와 구조를 드러

[그림 4-2] Gould and Fernandez' Brokerage Types(Gould and Fernandez, 1989)



자료 : 저자 작성.

내고, 보다 효율적인 조직 간 협력을 촉진할 수 있다고 본다.

이러한 역할은 다음 수식으로 정량화할 수 있다. 숙련 s 의 브로커리지 역할 수는 다음과 같다:

$$B_{GF}(s) = \sum_{i \in G} \sum_{j \in H} \delta_{i \neq j} \cdot e_s^{(i \rightarrow j)}$$

이 수식에서 $e_s^{(i \rightarrow j)}$ 는 숙련 s 가 그룹 i 의 노드와 그룹 j 의 노드를 연결하는 중개 연결 수를 의미하고, $\delta_{i \neq j}$ 는 두 그룹이 서로 다를 경우 1의 값을 가지며, 그렇지 않으면 0이 된다. 이 수식은 숙련이 얼마나 자주 다른 블록 간 연결을 매개하고 있는지를 정량적으로 평가할 수 있는 도구로 작용한다.

6. Burt의 구조적 홀 이론과 제약 지수

본 연구는 브로커리지 지수들과 유사한 이론적 속성을 갖는 또 다른 네트워크

크 지수로 Burt의 구조적 홀 개념과 그 관련 지표인 구조적 제약(constraint) 지수를 사용한다. Burt는 네트워크 내 구조적 공백(structural hole)을 메우는 위치에 있는 행위자가 정보에의 접근 속도에서 더 우위를 갖는 경향을 발견했다. 이러한 구조적 공백은 네트워크상에서 상호 연결이 적거나 존재하지 않아 다른 브로커를 통해야 다른 행위자(숙련)에 도달할 수 있는, 상대적으로 고립된 집단들 사이에 존재한다. 이러한 구조적 공백을 연결하는 행위자는 고립된 집단들을 연결해 줌으로써 이전에 존재하지 않았던 수요와 거래를 창출할 수 있는 위치에 서게 된다. 이들은 정보의 비대칭성(Akerlof, 1970)을 이용하여 새로운 자원과 정보를 비대칭적으로 확보할 수 있는 구조적 위치를 보유하게 된다. 이론적으로 이러한 위치는 높은 사회적 자본과 더 나은 성과로 이어진다.

제약 지수는 특정 노드가 연결된 이웃들과 얼마나 중복적으로 연결되어 있는지를 측정한다. Burt(1992)에 따르면, 제약 지수는 다음과 같이 정의된다:

$$Constraint\ Score = \sum \left(P_{ab} + \sum_c P_{ac} P_{cb} \right)^2, \quad c \neq a, b$$

위 수식에서 p_{ab} 는 노드 a 에서 b 로의 연결 강도의 비율이다. $p_{ac} * p_{cb}$ 는 간접 경로를 통해 연결된 관계까지 고려한 중복성을 의미한다. 전체 식은 직접 연결과 간접 연결이 모두 높을수록 제약이 크다는 점을 수치화한 것이다. 따라서 제약 지수가 낮을수록, 해당 노드는 서로 연결되지 않은 영역을 매개하는 구조적 홀 위에 위치해 있다고 해석할 수 있다.

숙련의 브로커리지 역량은 다른 숙련들의 조밀한 존재(그리고 구조적 홀의 부재)로 인해 제약될 수 있다(Burt, 1992). 이는 대단히 밀접하게 상호 연관된 숙련들의 장기간 발달 때문일 수 있다. 반대로, 전체 숙련 공동체의 미 발달이나 숙련이 다른 숙련들로부터 심각하게 주변화되는 상황에서도 숙련의 브로커리지 역량은 제한될 수 있다(Lee and Goh, 2023).

첫째, [그림 4-3]의 3-a에서는 숙련 A와 숙련 C 사이의 연결(P_{ac})이 숙련 B의 존재로 인해 제약을 받는다. 숙련 A는 숙련 C와 직접적인 관계를 맺고 있지만, 숙련 B가 숙련 C에게 자원 전달의 대체 경로를 제공한다. 또한 경로 P_{bc} 는 숙련 A가 숙련 B와 숙련 C 사이에서 중개할 기회를 빼앗는다. 따라서

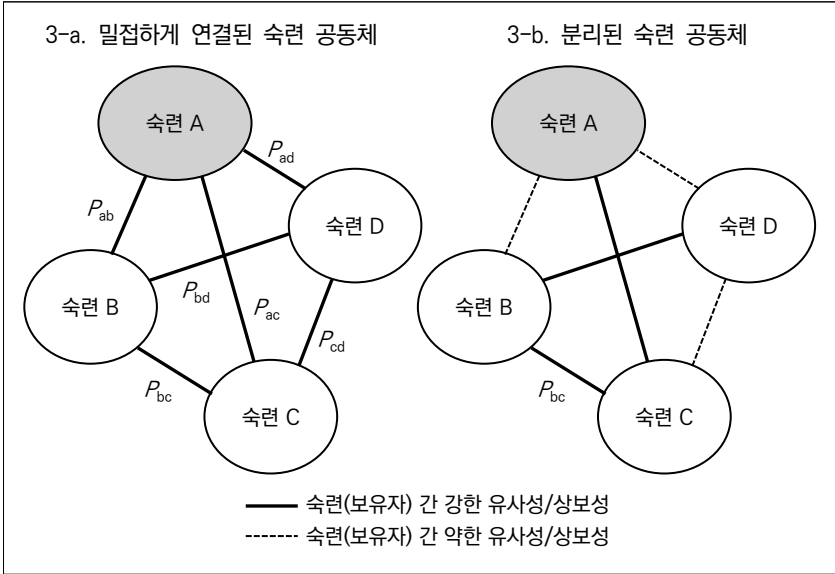
제3의 숙련이 숙련 A가 직접 연결을 유지하는 상대 숙련들에게 대체 경로를 제공할 경우, 숙련의 브로커리지 역량은 제약된다.

둘째, [그림 4-3]의 3-b에서는 숙련이 전체 숙련 공동체로부터 주변화됨으로써 브로커리지 역량이 제약된다. 숙련 A는 숙련 C와의 단일 연결만 가지며, 숙련 B와 D와는 단절되어 있다. 따라서 숙련 A는 숙련 공동체 내에서 숙련 C와 숙련 D 사이에 존재하는 구조적 홀(즉, 두 숙련 간의 관계 부재)을 활용할 수 없다. 다른 한편, 숙련 A는 숙련 B에 도달하기 위해 숙련 C에 의존해야 하며, 이는 숙련 A의 숙련 C에 대한 협상력을 크게 제한한다. 3-b 그림에서는 숙련 B만이 전체 숙련 공동체 내에서 숙련 C와 숙련 D 사이에서 효과적인 브로커 지위를 확보하고 있다(또 다른 브로커 지위는 숙련 A와 B 사이에서 숙련 C가 확보하고 있다).

이처럼 [그림 4-3]의 두 경우 모두에서 숙련 A는 전체 숙련 공동체의 자원 동원과 전달 과정에서 부분적으로 혹은 완전히 배제된다. 이러한 유형의 숙련 공동체나 블록에서는 숙련 혹은 숙련 보유자들이 다른 숙련과 연합하거나 정보 및 자원 교환 통로를 확보하는 데 어려움을 겪게 된다. 버트의 구조적 공백을 고려한 제한성(constraint) 지수는 Gould and Fernandez의 브로커리지 유형 지수와는 달리, 다른 제약 요소의 존재(대안적 브로커의 존재 혹은 부재)를 측정함으로써(제약 요소의 부재를 통해) 브로커 역량을 측정하는 간접적인 브로커 역량 지수인 셈이다.

이 외에 Burt가 제안한 3가지 다른 지수인 유효규모(effective size), 효율성(efficiency), 위계성(hierarchy) 지수를 함께 테스트한다. 유효규모는 한 노드가 가진 연결 중 중복되지 않는 고유한 이웃 수를 나타낸다. 총이웃의 수에서 중복된 이웃의 수를 빼서 계산한다. 이 지수가 클수록 중복이 적어지기 때문에 브로커리지가나 구조적 공백의 크기는 더 커질 수 있다. 효율성은 유효규모를 전체 연결 수(degree)로 정규한 값이다. 관계가 얼마나 낭비 없이 새로운 자원 혹은 숙련에 도달하는지 정도를 측정한다. 마지막으로 위계성은 한 노드(숙련)가 가진 관계가 특정 소수의 이웃에 의존하는지를 측정한다. 높은 위계성일수록 자원과 정보의 공급 경로가 소수에 의존하며 편향되는 구조임을 의미한다. 제약 지수가 중복의 정도를 측정한다면, 위계성은 의존의 집중도를 측정한다.

[그림 4-3] 숙련 공동체에서 브로커리지 역량과 제약



주 : P_{ab} 는 A와 B라는 두 결사체 사이의 관계의 비례적 강도(Burt, 1992)를 의미함 (Lee and Goh, 2023).
 자료 : 저자 작성.

제3절 자료 및 방법론

자료는 2024년 IT 일자리 온라인 구인공고를 토대로 구축한 숙련 네트워크에 기반한다. 해당 데이터의 사이즈가 100만 개가 넘는 대규모 데이터 셋이기에 컴퓨팅 파워를 고려하여 10,000개의 숙련만을 무작위로 추출하여 분석하였다. 사례별 결측값 삭제 및 컴퓨팅 파워를 고려하여, 브로커리지 지표를 이용한 회귀분석에서는 3,018개의 숙련이, 구조적 공백을 이용한 회귀분석에서는 1,062개의 숙련이 사용되었다. 종속 변수인 임금은 로그를 취했으며, 하위 브로커리지 개수들 또한 로그를 취해서 이분산성(heteroskedasticity)을 통제하고자 했다.

회귀모델의 통제변수로는 구인공고상에 제시된 학위 요구사항과 주류

(top industry) 산업군 더미, 연결 중심성(degree centrality)을 투입했다. 주류 산업군(top industry dummy=1)은 네트워크 내에서 가장 빈번하게 나타나는 산업군에 속한 숙련이라는 의미이며, 그렇지 않은 경우 상대적으로 덜 출현하는 숙련(top industry dummy =0)이라는 의미이다. 연결 중심성은 한 숙련직무가 다른 숙련들과 얼마나 많이 직접 연결되어 있는지를 측정하는 지표이다.

제4절 분석 결과

1. 기술통계 분석 결과

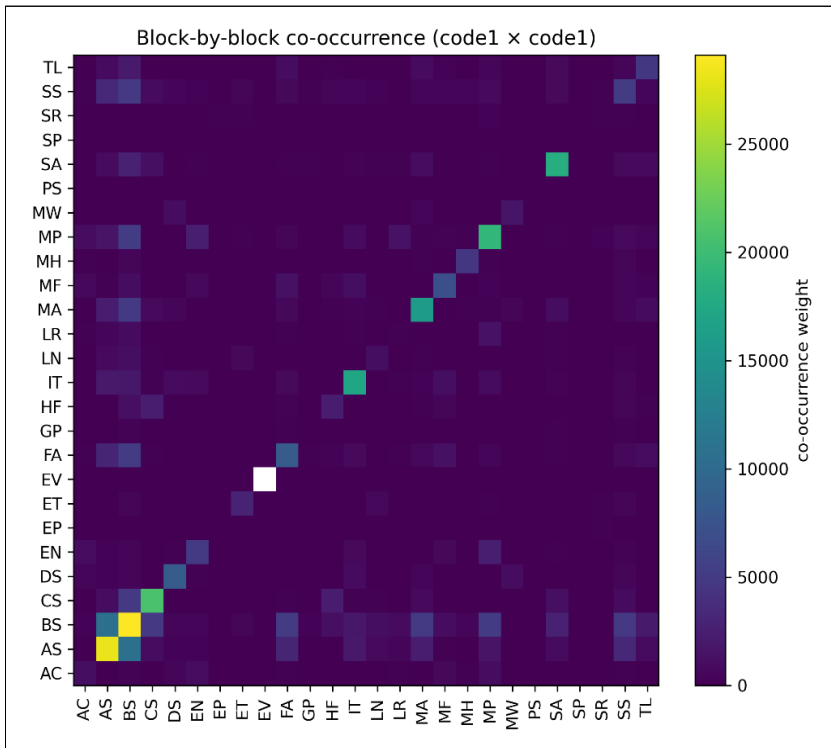
[그림 4-4]의 히트맵의 행과 열은 산업 및 직무 블록을 나타낸다. 색깔은 스킬의 공출현(같은 블록 내 스킬들끼리의 연관도)을 나타낸다. 대각선은 동일 블록 내부에서의 공출현을, 대각선 바깥은 다른 산업 및 직무 블록 간의 공출현 빈도를 나타낸다. 대각선의 밝은 부분들은 산업 및 직종 블록 내부에서 하위 스킬들 간 공출현이 많이 관찰됨을 의미하고, 대각선 외부의 어두운 부분들은 블록들 간에 공출현이 거의 관찰되지 않음을 의미한다. [그림 4-4]는 스킬의 공출현이 산업·직종 내부의 현상이고, 산업·직종을 가로질러 유용한 스킬은 그리 많지 않음을 시사한다. 따라서, 이들 산업·직종 내부에서는 조정자 역할(coordinator role)의 브로커리지가 출현할 가능성이 높다. 물류/운송(TL)이나 연구개발(SR)에서는 블록 내부에서도 공출현 빈도가 낮고, 다른 블록과의 공출현 빈도 또한 낮다. 이들 분야의 스킬은 각 산업·직종 내부 수요에만 응답하는 경향이 크다고 볼 수 있다.

좌측 하단에 회계·재무(AC) 블록이 비즈니스/경영지원(BS) 및 고객서비스/영업관리(CS)와 두텁게 연결되어 있다. 어느 기업이건 이 세 영역은 기업 내부에 필수적으로 존재하는 관리, 지원, 영업 관련 스킬들이기 때문에 각 영역에서 다른 영역의 스킬들이 높은 빈도로 공출현하는 현상은 놀랍지 않다.

주목할 현상은 마케팅/홍보(MA), 영업/판매(SA)와 함께 정보기술(IT) 관련 스킬들이 자체 산업·직무 블록 내에서 유의미하게 높은 스킬 공출현을 보여주고 있으며, 이 직무 블록들과 앞의 비즈니스/경영지원(BS) 및 행정지원 사무(AS)의 공출현 빈도가 밝은 청색으로 잡히는 것은 현장에서 디지털/IT, 마케팅, 세일즈가 경영 및 행정 지원 사무와 함께 요구됨을 의미한다.

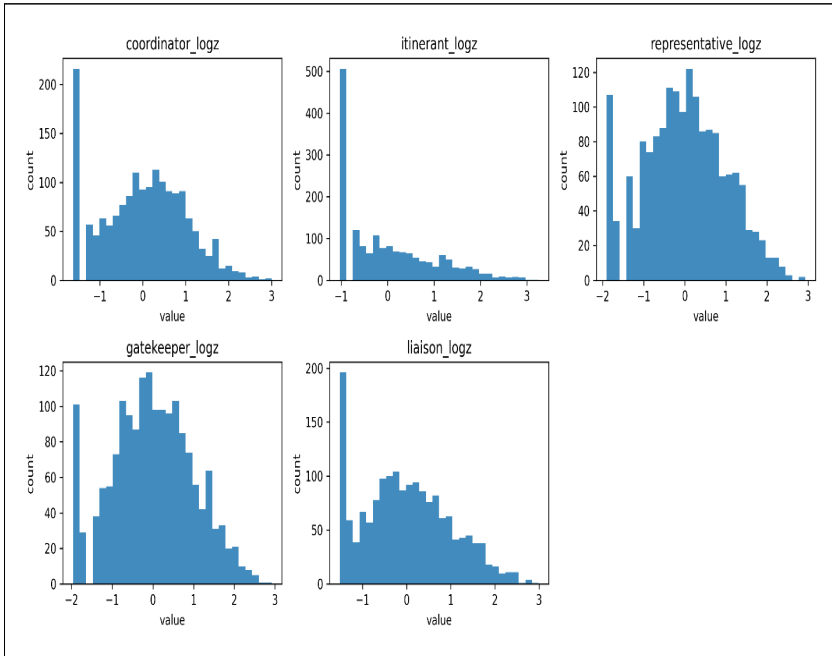
요약하면, 전체 산업 및 직종군들 중 핵심 코어 블록은 회계/재무(AC), 경영지원(BS), 고객센터서비스(CS)이며, 이들과 연계 블록인 정보기술(IT), 마케팅(MA), 영업(SA) 분야의 스킬들은 블록 간 브로커리지를 형성할 잠재력들을 갖고 있다. 특히 앞서 논의한 블록 내부의 스킬들을 블록 외부와 연결시켜 주는 스킬 커뮤니티의 대표 역할(representative role)이나 문지기 역할(gatekeeper role)이 이들 산업·직종 블록 간에 출현할 가능성이 높다.

[그림 4-4] 산업 및 직종 간 브로커리지 히트맵



자료 : 저자 작성.

[그림 4-5] 각 브로커리지 하위 유형의 출현 빈도

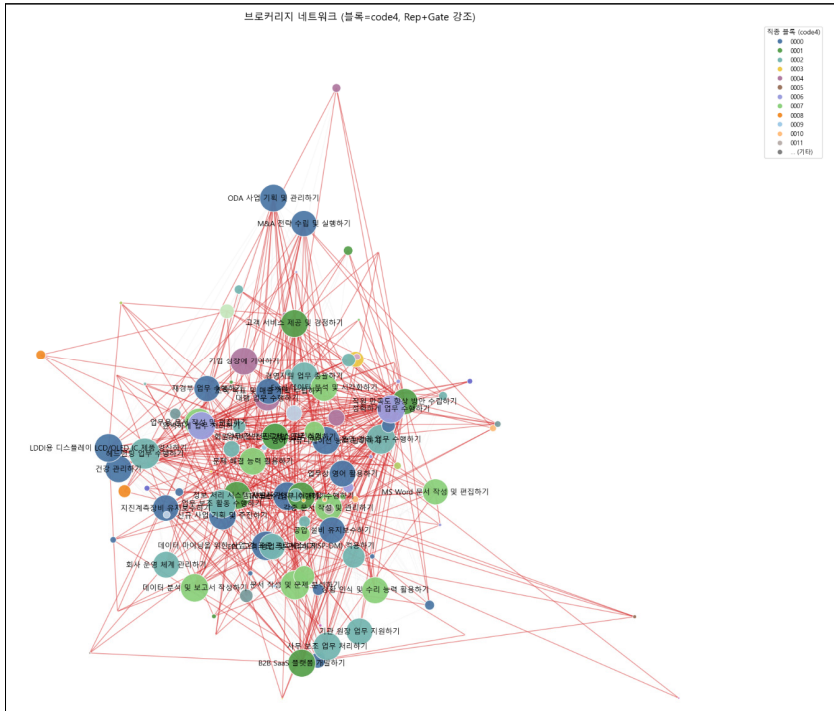


자료 : 저자 작성.

[그림 4-5]는 각 브로커리지의 하위 유형들의 전반적 출현 빈도를 보여준다. 먼저, 모든 지수의 왼쪽에 로그 변환 전 0 혹은 값이 없는 경우들이 몰려 있다. 그리고 각 브로커리지 유형마다(itinerant 유형을 제외한 나머지 네 유형의 경우) 0값을 정점으로 정규분포 모양을 띠지만, 동시에 오른쪽으로 긴 꼬리를 가진 약간 비대칭적인 분포를 보여준다. 조정자, 대표자, 문지기 역할 브로커리지 유형들의 경우, -1에서 +2 사이에서 좌우로 넓게 분포하며, liaison의 경우 좀 더 완만하게 퍼진 분포를 보여준다. Itinerant는 대부분 0 이하에 몰려 있고, 오른쪽으로 긴 꼬리를 가지며 아주 극소수 스킬만 높은 값을 갖는 분포를 보인다.

요약하면, 대부분의 숙련은 브로커리지 역할을 거의 하지 않는다. 다만 특정 숙련군들만이 각 브로커리지 역할 유형에서 중심적 위치를 차지하고 있음을 보여준다. 특히, itinerant나 liaison과 같이 복수의 산업과 직종을 오가며 브로커 역할을 하는 숙련은 매우 한정되어 있음을 알 수 있다.

[그림 4-6] 브로커리지 지수(대표자 + 문지기) 기준 상위 40개 숙련의 네트워크 시각화



자료 : 저자 작성.

[그림 4-6]은 직종 블록(code4) 기준으로 숙련의 공출현 네트워크에서 브로커 역할(문지기 + 대표자)을 하는 상위 40개 숙련과 여타 숙련들 간의 네트워크 구조를 시각화한 것이다. 구인광고별 스킬 묶음으로 무방향 네트워크를 구축한 후, 각 숙련을 직종 코드로 나뉜 블록에 배치하여 블록 내부와 블록 외부 간의 브로커리지 관계를 구별할 수 있도록 하였다. 다음으로 Gould and Fernandez(1989)의 브로커리지 다섯 역할을 계산한 뒤 대표와 문지기를 합한 값을 핵심 지표로 삼아 상위 40개 스킬을 선택하고, 40개를 제외한 나머지 숙련 중에 80개를 무작위로 뽑아 40개 주변에 뿌렸다. 또한, 블록 간 총연결 강도를 기준으로 중심 좌표를 잡았다. 노드 색은 직종 블록을 반영하며, 노드의 크기는 대표자 및 문지기 역할의 합계(의 제곱근)에 비례하도록 하였다. 회색 엷지는 블록 내부 연결, 빨간 엷지는 블록 간 연결이며, 이 중 가장 강한 블록 간 연결들만을 남겼다. 크기가 크고 라벨이 붙은

노드들은 직종의 경계를 넘어 숙련 정보를 이전시킬 수 있는 대표자 혹은 문지기 역할을 주로 하는 숙련들이다. 빨간 엣지(블록 간 연결)가 많을수록 다양한 직종을 잇는 매개숙련 역할을 하는 것이며, 회색 엣지(블록 내 연결)가 밀집된 경우는 블록 내부의 결속이 강함을 의미한다. 블록 간 연결을 많이 하는 상위 5개 숙련은 엑셀, 파워포인트/프레젠테이션 작성, 문서 작성 및 편집(MS word), 프로젝트 관리/업무 조정, 데이터 분석/보고서 작성이다. 어느 산업인지 여부와 상관없이 골고루 쓰일 수 있는 범용 기술이 가장 브로커 역할을 많이 하는 것이다.

이제 이렇게 각 산업과 직종 내부에 존재하는 브로커리지 구조가 실제로 그 구조로 인해 각 숙련에 대한 보상의 변이에 영향을 끼치는지를 테스트해 본다. 만일, 특정 유형의 브로커리지가 더 구조적으로 열등한 위치에 있는 숙련이 더 우월한 위치에 있는 숙련에게 의존하는 구조를 만들어 내고 유지하는 역할을 한다면, 특정 유형의 브로커리지 구조를 점유하고 있는 숙련과 그 숙련의 보유자에게 더 높은 임금이 주어질 것이다. 다음의 회귀분석 모델들은 이러한 가설을 다양한 브로커리지 지수 및 구조적 공백 지수들을 통해 검증한다.

2. 회귀분석 결과

가. 중심성 변수들의 임금에의 영향

〈표 4-1〉은 로그 임금을 네트워크 중앙성 지표들에 대해 OLS 회귀를 돌린 결과이다. 모든 모델에서 대인관계/사회적 숙련, 탐산업(가장 케이스 수가 많은 산업) 고정효과 더미변수, 교육 수준을 통제했다. 연결 중심성, 사이 중심성, 근접 중심성, 페이지랭크 중심성 등 모든 중심성 지수들은 아이겐벡터 중심성을 제외하고는 강력한 양의 상관관계를 보여주었다($p < 0.001$). 이는 숙련이 네트워크상에서 다른 숙련들과 연결이 많을수록(연결 중심성), 다른 숙련들의 사이에 놓일수록(사이 중심성), 다른 숙련들과 근접할수록(근접 중심성), 여타 높은 랭크를 가진 주요 숙련들과 연결될수록(페이지랭크 중심성), 더 높은 임금을 보장함을 의미한다.

〈표 4-1〉 중심성 지수들과 임금(로그) 간 관계에 대한 OLS 회귀분석

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal
Soft	-0.1294*** (0.0355)	-0.1103** (0.0357)	-0.1252*** (0.0356)	-0.1076** (0.0359)	-0.1154** (0.0367)	-0.0957** (0.0368)	-0.0978** (0.0367)	-0.0804* (0.0369)	-0.1100** (0.0362)	-0.0922* (0.0363)
edu_level	0.0473*** (0.0129)	0.0422** (0.0129)	0.0419** (0.0132)	0.0372** (0.0132)	0.0504*** (0.0128)	0.0444*** (0.0129)	0.0514*** (0.0128)	0.0448*** (0.0129)	0.0501*** (0.0128)	0.0442*** (0.0129)
fe_top_industry		0.1567*** (0.0320)		0.1571*** (0.0321)		0.1707*** (0.0320)		0.1824*** (0.0320)		0.1685*** (0.0320)
degree_logz	0.0525*** (0.0080)	0.0455*** (0.0079)								
closeness_logz			0.0594*** (0.0108)	0.0535*** (0.0106)						
betweenness_logz					0.0325*** (0.0084)	0.0264*** (0.0075)				
eigenvector_logz							0.0076 (0.0051)	0.0063 (0.0055)		
pagerank_logz									0.0388*** (0.0069)	0.0332*** (0.0067)
Adj. R ²	0.0146	0.0201	0.0169	0.0226	0.0092	0.0159	0.0059	0.0138	0.0106	0.0172
N	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000

주 : Standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001.
 자료 : 저자 작성.

나. 브로커리지 구조의 임금에의 영향

〈표 4-2〉는 임금(로그 취함)을 브로커리지 변수들에 대해 회귀를 돌린 결과를 보고한다. 모든 모델에서 교육 수준, 연결 중심성(degree centrality)과 대인관계 혹은 사회적 숙련(soft skill)을 통제했으며, 탑산업 더미 변수를 통제 한 모델과 통제하지 않은 모델을 나란히 배치하였다. 먼저 공통의 통제 변수 효과를 보면, 연결 중심성은 전체적으로 유의하지 않았다(단, 이 변수와 브로커리지 변수들의 상관계수값(correlation coefficient)이 0.7을 넘기 때문에 다중공선성이 발생했다고 볼 수도 있으므로 연결 중심성이 효과가 없다고 단정하기는 어렵다). 모델 (3)과 (9)에서만 양의 유의한 효과가 관찰되었고, 이는 네트워크상에서 더 많은 연결을 가진 숙련일수록 높은 임금을 받는 경향이 있음을 의미한다. 또한 숙련 공고에서 요구하는 교육 수준은 임금 수준과 강한 양의 상관관계를 보였다. 이상의 통제 변수들을 모두 모델에 넣은 상태에서 〈표 4-2〉는 브로커리지 변수들을 하나씩 넣고 테스트하였다.

사회적 숙련의 경우 모든 모델에서 일관되게 음의 계수 값이 추정되었다. 유의도 수준도 상당히 높음을 고려할 때($p < 0.01 \sim p < 0.001$), 사회적 숙련이 고임금/전문직 직무보다는 서비스업의 저숙련 직무에서 강조되고 있음을 암시한다.

모델 (2), (4), (6), (8), (10)에서 산업(대산업 분류) 고정효과를 포함시켰으며 모두 강력한 양의 효과($p < 0.001$)를 보여준다. 이들은 솔루션·SI·CRM·ERP 등에 속하는 직무들로, 이러한 직무들에서 임금 수준은 상대적으로 더 높았다.

조정자(coordinator) 역할의 경우, 모델 (1)과 (2)에서 강한 양의 효과를 보였다($p < 0.001$). 이는 동일 블록 내에서 정보와 자원을 연결하는 역할을 하는 숙련 혹은 그 숙련의 보유자는 더 높은 임금을 받는다는 첫 번째 가설을 지지하는 결과이다. 반면, 조직 외부에서 조직 내부의 결절자들을 이어주는 이동자(itinerant) 역할은 모델 (3)에서 경계선에 걸친 유의도를 보여주었다($p < 0.05$). 다만, 산업 고정효과 더미를 넣을 경우(모델 (4)), 연결 중심성(degree centrality)의 효과가 약해지고 이동자 역할만 살아남았다.

〈표 4-2〉 브로커리지(Brokerage) 지수들과 임금(로그) 간 관계에 대한 OLS 회귀분석

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal
Degree centrality_logz	-0.0049 (0.0096)	-0.0039 (0.0096)	0.0319* (0.0141)	0.0161 (0.0141)	0.0026 (0.0112)	0.0004 (0.0112)	-0.0100 (0.0110)	-0.0128 (0.0110)	0.0315** (0.0120)	0.0184 (0.0118)
Soft skills	-0.0624 (0.0367)	-0.0574 (0.0367)	-0.1304*** (0.0352)	-0.1103** (0.0353)	-0.1080** (0.0354)	-0.0926** (0.0356)	-0.1064** (0.0351)	-0.0907* (0.0354)	-0.1336*** (0.0353)	-0.1145** (0.0354)
edu_level	0.0325* (0.0132)	0.0309* (0.0132)	0.0450*** (0.0131)	0.0387** (0.0132)	0.0385** (0.0135)	0.0347* (0.0135)	0.0374** (0.0133)	0.0334* (0.0134)	0.0453*** (0.0131)	0.0394** (0.0132)
fe_top_industry		0.0874** (0.0333)		0.1683*** (0.0326)		0.1397*** (0.0320)		0.1389*** (0.0321)		0.1663*** (0.0323)
coordinator_logz	0.0964*** (0.0129)	0.0883*** (0.0136)								
itinerant_logz			0.0263* (0.0170)	0.0368* (0.0172)						
representative_logz					0.0706*** (0.0152)	0.0650*** (0.0153)				
gatekeeper_logz							0.0870*** (0.0147)	0.0823*** (0.0147)		
liaison_logz									0.0283 (0.0148)	0.0359* (0.0148)
Adj. R ²	0.0331	0.0345	0.0151	0.0214	0.0221	0.0264	0.0258	0.0301	0.0154	0.0216
N	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000	3018.0000

주 : Standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001.
 자료 : 저자 작성.

대표자(representative) 역할은 모델 (5)와 (6)에서 매우 강한 양의 효과를 보여주었다. 이는, 블록(산업) 내부를 외부와 연결해주는 대표자 역할을 하는 숙련이 임금에 가장 안정적인 영향을 끼침을 의미한다. 외부의 정보를 내부로 실어나르거나 번역하고 통제하는 역할을 하는 문지기(gatekeeper) 역할을 하는 숙련들 또한 (모델 (7)과 (8)) 강력한 양의 효과를 가지며 높은 임금과 연관되었다. 마지막으로, 블록 간 연결을 외부에서 담당하는 연결자(liaison) 역할 또한 모델 (10)에서 통계적으로 유의미($p < 0.05$)하였으나, 조정자, 대표자, 문지기 역할만큼 인상적인 수준은 아니었다.

요약하면, 이상의 회귀분석 결과들은 산업 내부와 외부를 연결하는 브로커 역할들과 산업 내부의 브로커 역할이 숙련 네트워크에서 임금을 높이는 주요 요인임을 보여준다. 곧, 브로커리지 구조는 숙련들 간의 위계를 반영하면서 그룹 내부와 외부로 가로지르며 국지적 범용 역할을 하는 기술들에게 더 유리한 구조적 위치를 보장함으로써 더 높은 보상을 유도한다. 또한 브로커리지 구조는 숙련 네트워크 구조의 파악에 있어 어느 기술이 '중심'에 위치하는지 못지않게, 어느 기술이 다른 블록 내에서 소외된, 혹은 블록에 갇혀 있는 기술들을 외부 산업들과 호환될 수 있도록 연결해주는지 여부가 중요함을 알려준다. 향후 숙련 네트워크 연구는 네트워크의 두 구조를 함께 연구하는 방향으로 진전되어야 할 것이다.

다. 구조적 공백의 임금에의 영향

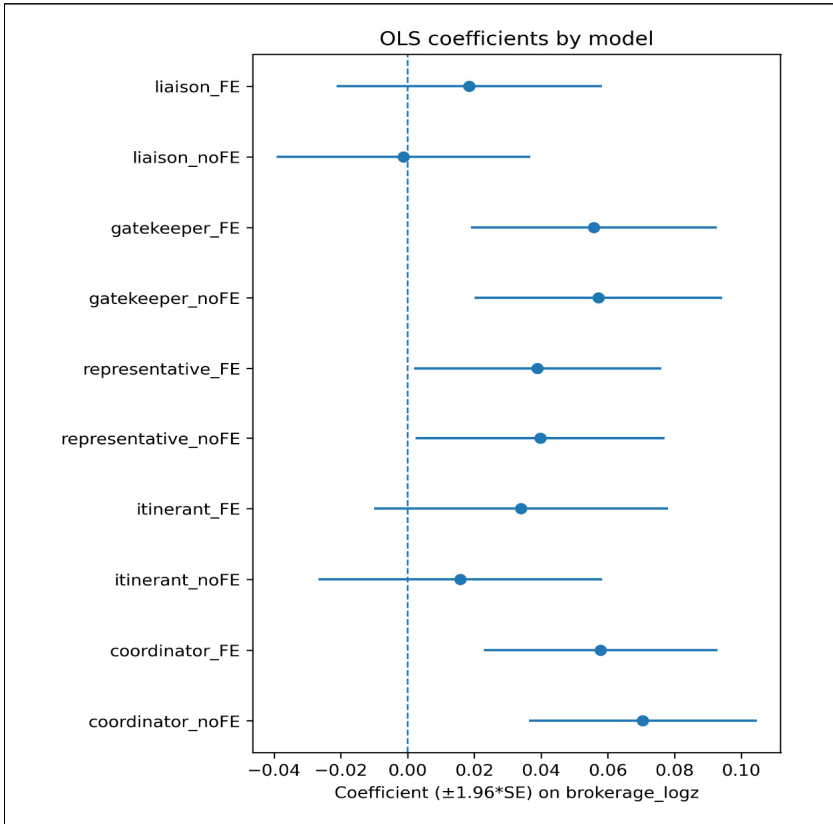
〈표 4-3〉은 임금(로그 취함)을 구조적 공백 변수들에 대해 OLS 회귀를 돌려서 얻은 결과를 제시한다. 앞의 브로커리지 분석과 같이 모든 모델에서 교육 수준, 연결 중심성(degree centrality)과 대인관계 혹은 사회적 숙련(soft skill), 탐산업 더미를 통제했다. 결과는 〈표 4-3〉의 모델 (1)에서 보듯이, Burt의 주요 변수들 중 제약(constraint)변수만이 통계적으로 유의미한 음의 결과를 보였다. 나머지 변수들(유효규모(effective size), 효율성(efficiency), 위계성(hierarchy) 지수들)은 통계적으로 무의미하였다. 이는, 다른 브로커에 의해 통제당하는, 구조적으로 열위에 있는 위치에 있는 숙련일수록 더 낮은 임금을 받게 됨을 의미한다.

〈표 4-3〉 구조적 공백(Structural Hole) 지수들과 임금(로그) 간 관계에 대한 OLS 회귀분석

	(1)	(2)	(3)	(4)
	ln_avg_sal	ln_avg_sal	ln_avg_sal	ln_avg_sal
degree_logz	-0.0305* (0.0127)	-0.0523 (0.1351)	-0.0194 (0.0140)	-0.0194 (0.0140)
Soft	-0.0540 (0.0658)	-0.0456 (0.0659)	-0.0507 (0.0657)	-0.0507 (0.0657)
edu_level	-0.0066 (0.0503)	0.0037 (0.0505)	-0.0032 (0.0513)	-0.0032 (0.0513)
fe_top_industry	0.1627*** (0.0477)	0.1743*** (0.0476)	0.1721*** (0.0477)	0.1721*** (0.0477)
Constraint	-1.0971*** (0.2613)			
effective_size		0.0006 (0.0016)		
Efficiency			0.3618 (0.2324)	
Hierarchy				-0.3618 (0.2324)
Adj. R ²	0.0168	0.0090	0.0105	0.0105
N	1062.0000	1062.0000	1062.0000	1062.0000

주 : Standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001.
자료 : 저자 작성.

[그림 4-7] 각 브로커리지 유형의 OLS 회귀 계수 유의성 및 크기 비교



자료 : 저자 작성.

제5절 결론 : 연구 기여 및 정책 함의

본 연구는 사람인에 실린 구인공고를 토대로 구축한 숙련 네트워크 데이터에 기반하여, 숙련 간 존재하는 브로커리지 및 구조적 공백 구조가 임금에 미치는 영향을 추정하였다.

예측대로, 브로커리지와 구조적 공백 지수들은 임금에 강한 양의 관계를 갖고 있었다. 중심성 지수들 중에는 연결 중심성, 사이 중심성, 근접 중심성,

페이지랭크 중심성이 통계적으로 유의미했다. 반면, 아이겐벡터 중심성은 통계적으로 무의미했다. 다른 위세를 가진 숙련과 연결되어 있다고 해서 그 숙련의 가치가 올라가지는 않는 것이다. 하지만 다른 숙련들과 더 많은 빈도(연결 중심성)로 연결되어 있거나, 숙련들 간의 연결망에 더 매개적인 위치를 점유하거나(사이 중심성), 다른 숙련들과의 연결망의 거리가 더 짧은 위치에 있을 경우(근접 중심성) 더 높은 임금을 제안 받는 경향이 있었다.

브로커리지 지수들 중에는 조정자 역할(coordinator role), 대표자 역할(representative role), 문지기 역할(gatekeeper role)을 수행하는 위치에 있는 숙련들이 더 높은 임금을 받는 경향을 보였다. 특히, 이러한 브로커리지 지수들은 중심성 지수들과는 달리, 산업과 산업 간의 매개자 역할을 통해 숙련의 지위를 측정하기 때문에, 산업 내 역할에 치중될 수밖에 없는, 혹은 산업 내 역할이 산업 간 역할을 지배하는 경향을 갖는 중심성 지표들과 달리 산업 간 융합을 통해 신산업의 출현을 예측할 수 있는 잠재력을 갖고 있다. 예를 들어, 아직 인공지능 기술이나 자동화, 데이터화 기술이 보편화되지 않은 특정 제조업과 특정 서비스업 사이를 특정 빅데이터 및 인공지능 적용 숙련이 매개할 경우, 향후 이 산업들의 인공지능 기반 자동화 경향이 출현하여 발달할 것임을 예측할 수 있다. 특히 대표자와 문지기 역할이 통계적으로 유의미한 결과를 보였다는 것은, 특정 산업과 다른 산업을 중개하는 역할을 하는 구조적 위치의 숙련이 향후 산업 발전에 있어 주요한 역할을 할 것임을 시사한다.

특정 직무 관련 훈련을 진행할 때, 개인 혹은 기업 수준에서 이러한 주요 숙련들의 교육에 집중적인 투자를 할 필요가 있다. 이러한 숙련들은 향후 지배적인 기술로 떠올라, 사업의 다각화(diversification)에 결정적인 역할을 함으로써 시장에서 수요가 폭증할 가능성이 크기 때문이다.

본 연구에서는 구인공고 데이터 전체를 대상으로 브로커리지 숙련을 계산하여 임금관계와의 상관관계를 테스트하였으나, 향후 연구에서는 IT 산업만을 따로 분석함으로써, 인공지능 기술이 어떻게 산업 내 숙련들 간의 브로커리지 관계를 매개하는지, 그러한 브로커리지 구조가 어떻게 임금 프리미엄을 창출하는지를 조사할 수 있을 것이다.

제 5 장

AI 숙련 네트워크와 AI 숙련-역량 정합성 평가

제1절 연구 배경과 목적

인공지능(AI)의 확산은 노동시장 전반에 걸쳐 기존의 숙련 수요 구조를 변화시키고 있으며, 이에 따라 교육훈련 시스템이 공급하는 역량과의 정합성 문제가 더욱 중요해지고 있다. 기존의 수요-공급 매칭 분석은 주로 텍스트 기반 유사도나 출현 빈도 중심으로 수행되어 숙련 간의 구조적 관계나 전이 가능성 등을 반영하기 어려웠다. 이 장에서는 숙련 간 네트워크 구조를 기반으로 숙련 수요와 역량 공급 간의 정합성을 분석하고, 이를 통해 교육훈련 체계의 개선 방향과 정책적 시사점을 도출하고자 한다.

오늘날 인공지능(AI)을 중심으로 한 기술 전환은 노동시장의 숙련 구조를 급속도로 재편하고 있다. 디지털 전환과 알고리즘 기반 자동화는 특정 직무에서 요구되는 숙련의 내용과 조합을 변화시키고 있으며, 이에 따라 어떤 숙련이 새롭게 요구되는지, 기존 숙련은 어떻게 전이될 수 있는지, 교육훈련은 이러한 수요에 얼마나 유연하게 대응하고 있는지에 대한 관심이 커지고 있다. 이러한 변화 속에서 가장 심각한 문제 중 하나는 바로 숙련 수급 불일치(skill mismatch) 문제이다. 이는 노동시장에서는 특정 숙련에 대한 수요가 증가하고 있음에도 불구하고, 교육 및 직업훈련 시스템이 적시에 대응하지 못함으로써 해당 숙련을 갖춘 인재가 부족한 현상을 의미한다. 반대로, 공급

되고 있는 교육이나 훈련이 노동시장에서 필요로 하지 않는 숙련에 집중될 경우 과잉 교육(over-education) 또는 훈련 낭비(training misalignment)로 이어질 수 있다. OECD¹⁾ 등 국제기구는 이러한 수급 불일치가 청년층의 노동시장 이행 지연, 중장년층의 이직 저항성, 지역 산업의 인재 확보난 등 여러 부작용을 낳고 있다고 경고하고 있다.

특히 AI 기술의 확산은 기존의 숙련 수요 예측 방식이 더 이상 유효하지 않다는 사실을 드러내고 있다. 예컨대, 데이터 분석, 머신러닝, 소프트웨어 개발 등의 AI 관련 숙련은 다양한 직무와 산업에서 동시다발적으로 요구되는 범용 숙련(general-purpose skills)의 성격을 가지거나, AI 숙련들끼리 서로 긴밀히 연결되어 요구되는 융합적-복합적 숙련의 성격을 가지는 경우가 많다. 따라서, 숙련 수요를 단일 키워드의 등장 빈도만으로 측정하거나, 교육훈련 과정을 단순 분류 체계로만 대응하는 방식은 네트워크 구조로 진화하는 숙련 환경을 충분히 설명하지 못한다.

더욱이 한국의 교육훈련 시스템은 여전히 공급자 중심의 설계가 지배적이며, 개별 커리큘럼이 노동시장 수요의 구조적 변화, 특히 숙련 간 연계성이나 전이가능성과 같은 관계적 속성을 충분히 고려하지 못하고 있다는 한계가 있다. 이로 인해 핵심 숙련의 공급 부족, 신기술 분야로의 전환 경로 부재, 기존 숙련군의 고립화 현상 등이 심화되고 있다. 예를 들어, 감각·물리적 숙련을 기반으로 한 직업군에 종사하는 근로자가 인지·사회적 숙련 중심의 고부가가치 직무로 전환하기 위해 필요한 중간 매개 숙련이 제공되지 않는 상황이 반복되고 있다.

이러한 맥락에서, 본 연구는 단순히 숙련 수요와 공급의 정합성을 분석하는 것을 넘어, 숙련 간 네트워크 구조, 숙련-직업 간 연계성, 그리고 교육훈련 내 숙련의 포함 구조를 동시에 고려함으로써, 보다 정교한 숙련 수급 구조 분석을 시도한다. 특히 온라인 구인공고, 교육과정 텍스트, 숙련 분류체계 등 이질적 자료를 통합하여 AI 숙련의 수급 구조를 규명하는 것을 목표로 한다. 이는 정책적으로도 수요 중심의 교육훈련 설계, AI 전환기 대응 전략 수립, 직업 전환 경로 설계 등에 실질적인 기여를 할 수 있을 것이다.

1) OECD, <https://www.oecd.org/en/topics/changing-skill-needs-in-the-labour-market.html>

제2절 이론적 배경 및 기존 연구

1. 숙련의 네트워크적 구조

전통적으로 숙련은 노동자의 개인 속성이나 특정 직업을 수행하기 위한 고립된 능력 단위로 간주되어 왔다. 기존의 숙련 분석은 주로 직업 수행을 위한 개별 능력 단위 또는 키워드 중심의 속성으로 숙련을 취급해왔다. 그러나 최근 연구들은 숙련이 단절된 기능의 집합이 아니라 직업 수행 과정에서 상호작용하는 속성을 지닌 관계적 구조(relational structure)로 존재함을 강조한다.

Alabdulkareem et al.(2018)은 이러한 접근을 체계화하기 위해 ‘Skillscape’라는 개념을 제안하였다. 이들은 O*NET 자료를 활용해 직업별로 요구되는 숙련 정보를 수집하고, 서로 다른 숙련들이 동일 직업 내에서 공존할 확률을 기반으로 숙련 간 보완성 지수 $\theta(s,s')$ 를 산출하였다. 이는 숙련들이 하나의 네트워크상에서 어떻게 연결되어 있는지를 파악할 수 있게 해준다. 이는 숙련이 단일 차원이 아니라, 네트워크상에서 상호작용하며 기능한다는 점을 계량화한 사례이다. 이러한 관계적 구조는 단순한 빈도 분석을 넘어, 숙련 중심성, 보완성, 클러스터화와 같은 네트워크 이론의 핵심 개념을 숙련 분석에 적용할 수 있는 기반을 제공한다. 특히 루뱅 알고리즘과 같은 커뮤니티 탐지 기법은 서로 밀접하게 연결된 숙련들의 하위집단을 식별하여, 숙련 클러스터 간의 특성과 위치성을 파악하는 데 유용하며, 고유벡터 중심성이나 페이지랭크 중심성은 노동시장 내에서 전략적 위치에 있는 숙련을 식별하는 데 유용하다. 이처럼 숙련 네트워크 분석은 직업 간 전이가능성, 숙련 양극화, 그리고 핵심 숙련 도출에 있어 기존보다 정교하고 구조적인 분석을 가능하게 한다.

또한 숙련 간 연결성은 직업 간 전이 경로 및 교육훈련 설계에 있어 핵심 정보로 활용될 수 있다(Dawson et al., 2021; Henning et al., 2023). Dawson

et al.(2021)은 구인공고 기반 스킬 집합 간의 유사도를 측정하는 ‘Skills Space’ 알고리즘을 제안하여, 숙련 네트워크 기반으로 직업 전환 경로 추천 시스템을 구축하였다. 특히 두 스킬 집합 간 유사도를 $\theta(s, s')$ 의 가중 평균으로 정의하고, 이를 바탕으로 실제 노동자 이직 데이터를 예측하는 데 활용함으로써, 숙련 간 연결성이 노동시장 이행성과와 강하게 관련되어 있음을 실증적으로 보여주었다. Henning et al.(2023)은 지역 내 직업 간 숙련 연관성과 지역 경제성과 간의 관계를 분석하며, 지역 수준의 숙련 정합성(coherence)이 낮을수록 평균 임금 및 생산성이 높지만, 직업 간 전이가 연관된(local job relatedness)지역에서의 이직은 더 나은 경제성과와 연결된다고 분석했다. 이는 단순한 교육 수준이 아니라 직업 간 연결성과 전이 경로가 중요하다는 점을 시사한다.

특히, Stephany and Teutloff(2024)는 기술 변화와 자동화의 시대에 숙련의 가치를 결정짓는 핵심 요소로 ‘보완성(complementarity)’이라는 점을 강조한다. 전통적으로 숙련은 고립된 개별 단위로 간주되었지만, 이들은 숙련이 네트워크상에서 상호작용하고 결합되는 방식에 따라 그 가치가 달라진다고 본다. 즉, 어떤 숙련이 다른 고가치 숙련과 얼마나 다양하게 조합될 수 있는지가 해당 숙련의 경제적 가치를 결정짓는 중요한 기준이라는 것이다. 이를 검증하기 위해 온라인 프리랜서 플랫폼의 5만 건에 달하는 프로젝트 데이터를 활용해 962개의 숙련 간 연결망을 구축했다. 숙련이 하나의 프로젝트 내에서 함께 사용된 빈도를 기반으로 연결망을 형성하고, 이를 바탕으로 보완성의 세 가지 측면(① 보완 숙련 수, ② 보완 숙련의 다양성, ③ 보완 숙련의 가치)을 계량화했다. 분석 결과, 특정 숙련의 ‘프리미엄(시장 가치)’은 단순한 수요-공급의 법칙을 넘어서 보완성에 크게 의존한다는 사실이 밝혀졌다. 예를 들어, 보완 숙련의 수가 많고, 이들이 다양한 유형이며 고가치일수록 해당 숙련의 프리미엄도 높았다. Python, Deep Learning, Business Coaching 등의 숙련은 다양한 고부가가치 숙련들과 결합되며 높은 시장 가치를 보였다. 특히 AI 관련 숙련은 평균적으로 21%의 임금 프리미엄을 기록하며, 일반 숙련(4%)보다 월등히 높은 경제적 가치를 가지는 것으로 나타났다.

또한 이 연구는 숙련의 ‘상대적 가치’에도 주목한다. 동일한 숙련이라도 어떤 숙련들과 함께 조합되는가에 따라 그 가치가 달라질 수 있다는 것이

다. 즉, 숙련은 그것이 기존에 어떤 숙련 집합 속에 포함되어 있는가에 따라 시너지 효과가 달라진다. 이처럼 숙련 간 조합 가능성과 전이 가능성은 재숙련 경로 설정에도 핵심적인 기준이 된다.

Stephany and Teutloff(2024)는 이러한 분석을 통해 노동시장 정보 기반의 맞춤형 재숙련 전략 수립이 가능하다고 제안한다. 개인의 기존 숙련 구조에 기반해 가장 높은 보완성과 경제적 가치를 가질 수 있는 신규 숙련을 추천하고, 재숙련 경로를 설계하는 데 이 데이터와 분석틀이 실용적으로 활용될 수 있다는 것이다.

2. AI 확산과 숙련 수요 구조의 재편

이러한 숙련의 네트워크 보완 구조는 AI 기술이 숙련 구조를 변화시키는 것을 파악하는 데 있어서 중요한 의미를 가진다. AI 기술의 보편화는 특정 고숙련 직업에 국한되지 않고 확산되면서, 산업 전반에 걸쳐 요구되는 범용 숙련(general-purpose skills)을 증가시키고 있다. 즉, AI 기술의 확산은 단순히 특정 기술 숙련의 수요 증가를 넘어, 노동시장에서 요구되는 숙련의 구조적 재편성을 야기하고 있다.

Taska et al.(2021)의 분석에 따르면, 머신러닝, 소프트웨어 개발, 스크립팅 언어 등 AI 관련 숙련은 다양한 산업군에서 동시에 수요가 증가하고 있으며, 이는 산업 구조와 무관하게 전이가능한 숙련이 점차 중심이 되고 있다는 점을 시사한다. 또한 AI 기술은 보완적 숙련과 함께 나타날 때 더 높은 임금 프리미엄이 관측되며, 숙련 간 구조적 연관성이 중요하다는 점을 강조했다.

Manca(2023)는 온라인 구인공고 데이터를 활용하여, AI 기술의 확산이 특정 숙련의 수요를 어떻게 변화시키는지 분석하였다. 특히 그는 고유벡터 중심성과 클러스터링 계수를 결합한 숙련 확산 지수(diffusion index)를 제안하였다. 이 지수는 AI 기술 수요에 따라 동시적으로 확산되는 핵심 숙련을 식별할 수 있는 유효한 지표로 기능한다. 즉, 특정 숙련이 노동시장 전반에 얼마나 빠르게 퍼지고 있는지를 정량화할 수 있다. 분석 결과, 머신러닝, 스크립팅 언어, 소프트웨어 개발 원칙 등이 AI 수요 확산과 밀접히 연관된 것

으로 나타났으며, 이러한 숙련들은 다양한 산업과 직업에서 공통적으로 요구되는 범용 숙련으로 부상하고 있었다. 이러한 분석은 숙련을 단순히 수요 여부의 문제로 보는 것이 아니라, 구조적으로 어떻게 연결되고 확산되며 기능군을 형성하는지를 종합적으로 이해할 필요성을 제기한다.

Hosseinioun et al.(2025)은 미국 O*NET 직업사전에 수록된 방대한 숙련 데이터를 분석하여, 숙련이 독립적으로 존재하는 것이 아니라 상호 의존성(dependencies)과 중첩성(nestedness) 속에서 조직되어 있음을 보여주었다. 그는 직업-숙련 매트릭스를 네트워크로 전환한 뒤, 특정 숙련이 다른 숙련과 함께 나타나는 조건부 관계를 통해 의존성을 추정하고, 생태학적 네트워크 분석에서 사용되는 중첩성 지표(NODF)를 적용하여 숙련 구조의 위계적-계층적 패턴을 계량화하였다. 의존성(dependencies)은 특정 숙련이 다른 숙련의 습득과 활용을 전제로 한다는 의미를 가진다. AI 숙련을 예로 들자면, 예컨대 기계학습 모델링 능력은 통계적 추론이나 프로그래밍 언어와 같은 기초 숙련을 전제로 해야만 형성될 수 있다. 따라서 숙련은 단일 기술의 수요만으로 설명되지 않으며, 다른 숙련과의 결합 구조 속에서 그 가치와 수요가 결정된다. 또한 그는 중첩성(nestedness)이라는 개념을 통해 숙련의 구조적 분포를 설명하였다. 네트워크 관점에서 일부 숙련은 광범위한 산업과 직업에서 공통적으로 요구되지만, 다른 숙련은 특정 영역에 한정적으로만 필요하다. 이러한 현상을 생태학적 네트워크 분석에서 차용한 nestedness 개념으로 설명하였다. 예를 들어 데이터 분석이나 수학적 모델링은 거의 모든 직무에서 요구되는 반면, 컴퓨터 비전은 일부 특정 직무에서만 필요하다는 점을 고려할 수 있다. 분석 방법으로는 미국 O*NET 직업사전 데이터를 활용하여 직무-숙련의 동시 출현 관계를 기반으로 네트워크를 구성하였다. 여기서 특정 숙련이 다른 숙련을 조건으로 하여 출현하는 조건부 확률을 계산해 의존성을 추정하였다. 이어서 생태학에서 종-서식지 관계를 파악할 때 사용하는 중첩성 지표인 NODF(Nestedness metric based on Overlap and Decreasing Fill)를 적용하여 숙련 네트워크의 중첩 구조를 계량화하여 분석하였다. 분석 결과, 일부 숙련은 광범위한 직업에서 반복적으로 요구되며 높은 중첩성을 보이는 반면, 다른 숙련은 특정 직업군에 제한적으로 분포하는 것으로 나타났다. 이러한 구조적 특징은 노동시장의 숙련

체계가 개별 기술 수요의 단순 집합이 아니라, 어떤 숙련이 다른 숙련을 전제로 하고 또 얼마나 다양한 직업에서 중첩적으로 활용되는가에 의해 형성된다는 점을 시사한다. Hosseinioun의 연구는 숙련 수요를 단일 차원에서 파악하기보다, 숙련들 간의 네트워크적 관계와 구조적 배치를 함께 고려해야 한다는 점을 부각시켰으며, 이는 노동시장 내 숙련의 보완성과 확산을 종합적으로 이해하는 데 중요한 기초를 제공한다.

3. 숙련 수요-공급 간 정합성 분석의 진화

숙련 수요와 교육훈련을 통한 역량 공급 간의 정합성(matching)은 고용정책 및 직업교육 분야에서 오래된 주제지만, 최근에는 단순히 키워드 유사성 수준을 넘어서 구조적 정합성(structural compliance)으로 분석 범주가 확장되고 있다.

전병유·장지연·심지환(2024)은 데이터사이언스 전공을 대상으로 대학 교육 커리큘럼 내 텍스트 자료를 분석하고, 이를 통해 온라인 구인공고 기반 숙련 수요와의 내용적 유사도를 측정하였다. 이들은 숙련 키워드 기반의 코사인 유사도를 활용해 수요-공급 간 매칭 정도를 정량화하고, 교육과정 내에서 수요 중심의 역량이 충분히 반영되고 있는지를 평가하였다. 이는 숙련 간 키워드 유사도를 기반으로 정합성을 정량화한 시도로서 의미 있지만, 숙련 간 구조적 관계(중심성, 숙련 위계 구조)를 반영하지 못한다는 한계를 가진다. 이와 유사하게 Pater et al.(2022)은 Mismatch Index와 Structural Compliance Index를 활용하여, 온라인 구인공고와 설문조사 자료를 결합한 정합성 평가를 제시하였다. 이 지수들은 단순한 유사도 정도가 아닌, 구조적 수준에서 수요와 공급이 얼마나 조응하는지를 판단하는 데 유용한 도구로 활용된다.

하지만 대부분의 기존 연구는 숙련 간 구조적 관계성과 위계 구조 등을 반영하지 못한 채, 개별 키워드의 등장 빈도나 정태적 유사도에 기반해 정합성을 측정해왔다. 따라서 실제로 핵심 숙련이 어떻게 연결되고, 교육과정은 이에 어떻게 대응하고 있는지에 대한 구조적 이해는 부족한 상황이다.

본 연구는 숙련 수요와 교육훈련 간 정합성을 평가함에 있어, 단일 숙련

수준의 유사성 비교를 넘어서 숙련 간 네트워크 구조를 반영하고자 한다. 네트워크 중심성이 높은 숙련이나 숙련 네트워크의 위계 구조에서 상위의 일반 숙련이 공급체계 내에 반영되어 있는지를 평가하는 것은 정책 설계에 있어 실질적인 시사점을 제공할 수 있다. 즉, 단순한 유사도 지표를 넘어, 교육 과정이 전략적 숙련을 반영하고 있는가에 대한 정교한 평가 체계를 제공할 수 있을 것이다.

제3절 자료와 방법론

1. AI 직업 숙련

AI 기술은 매우 빠르게 변화하고 있기 때문에 AI 숙련도 매년 새로운 숙련 키워드들이 등장하고 있다. OJP는 이러한 숙련을 파악하기에 적절한 자료라고 판단된다. 본 연구에서는 사람인과 잡코리아에서 스크래핑한 온라인 구인광고(이하, OJP, Online Job Posting) 자료를 활용한다. 기간은 2024년 1년간을 대상으로 한다. OJP에서 숙련 키워드는 OJP의 직무 설명 텍스트에서 자연어 처리, LLM 활용, 정규화, 수작업 등을 거쳐 추출하였다.

이렇게 추출한 숙련 키워드들 중에서 AI 숙련을 추출해야 한다. 본 연구에서는 AI 숙련 키워드를 좀 더 일반화하기 위해서 유사도 분석을 통해서 OJP 숙련 키워드들과 OECD 기준의 AI 숙련 키워드 리스트를 매핑하였다. 본 연구에서는 아래 [박스 1]에서 설명된 바와 같이 Baruffaldi et al.(2020)의 학술 기반의 포괄적이고 전통적인 AI 키워드 세트를 기본으로 하고, Squicciarini and Nachtigall(2021)의 산업 실무적 AI 키워드, Borgonovi et al.(2023)과 OECD.AI의 후속 보고서들의 생성형 AI 키워드, GitHub 토픽에서 도출된 최근 기술 진화를 추가로 포함시키는 방식을 채택하였다. 이렇게 구축된 OECD AI 키워드는 232개이다(이 키워드 자료는 부표 5-1 참고).

다음으로 이 OECD AI 키워드를 활용하여 AI 숙련 분류체계를 만들었다. OECD에서 이를 6개 군집으로 분류하고 있지만, 추가된 키워드들이 있기 때

[박스 1] OECD 숙련 키워드 리스트

OECD의 Baruffaldi et al.(2020)은 “Identifying and measuring developments in artificial intelligence: Making the impossible possible”에서 학술 논문과 특허의 텍스트 분석을 통해 인공지능 관련 개념을 식별하고, 약 200여 개에 달하는 포괄적인 AI 키워드 리스트를 제시하였다. 예를 들어 “machine learning”, “computer vision”, “speech recognition”, “reinforcement learning”, “swarm intelligence” 등 고전적인 AI 연구영역부터 “autonomous vehicle”, “cognitive computing”, “semantic web”과 같은 응용영역까지 포함되어 있어 당시로서는 가장 체계적이고 포괄적인 AI 숙련 키워드 세트를 제공하였다.

OECD의 Squicciarini and Nachtigall(2021)의 “Demand for AI skills in jobs: Evidence from online job postings”는 Baruffaldi et al.(2020)의 키워드 리스트를 기초로, 영국 BEIS(Department for Business, Energy & Industrial Strategy) 전문가들이 검증한 소프트웨어 및 프로그래밍 도구 목록을 결합하였다. 이 추가 리스트에는 Python, TensorFlow, PyTorch, Keras, scikit-learn, Caffe, Theano, MXNet 등 실제 산업 현장에서 자주 사용되는 AI 관련 소프트웨어·리포지토리가 포함되어 있어, 학술 연구에서 도출된 추상적 키워드를 보완하고 실무적 맥락을 강화한 특징이 있다.

최근 OECD.AI Policy Observatory와 연계된 Borgonovi et al.(2023)의 “Emerging trends in AI skill demand across 14 OECD countries” 및 「AI and the Future of Skills」 시리즈(Vol. 1, 2021; Vol. 2, 2023)는 이러한 기존 키워드 프레임워크를 토대로 최신 노동시장 데이터를 분석하면서, 생성형 AI의 부상에 따른 새로운 숙련 키워드 필요성을 강조하였다. 특히 2023년 이후 확산된 대규모 언어모델(LLMs), 파운데이션 모델, 트랜스포머(transformers), 확산 모델(diffusion models), 프롬프트 엔지니어링(prompt engineering), RAG(retrieval-augmented generation), 미세조정(fine-tuning, LoRA), RLHF/DPO, 안전성·정렬(alignment) 등은 기존 리스트에는 없던 새로운 범주로, 최신 기술 트렌드를 반영하는 생성형 AI 키워드 집합이 추가되었다는 점이 특징이다.

또한 OECD.AI Observatory는 GitHub 데이터에 기반한 지표를 통해 최신 AI 기술 토픽을 정기적으로 추출한다. 여기에는 “transformers”, “vision-

transformer”, “diffusion”, “multimodal-AI”, “explainable-AI”, “federated-learning”, “graph neural networks”, “fairness”와 “bias detection” 같은 주제가 포함된다. 이러한 GitHub 토픽 기반 키워드는 산업과 연구 커뮤니티에서 실제로 사용되는 코드와 저장소를 반영하기 때문에, 빠르게 진화하는 AI 생태계의 최신 동향을 포착할 수 있다는 강점을 지닌다.

문에, 본 연구에서는 AI 키워드들을 OpenAI 임베딩 모델(text-embedding-3-large)을 활용하여 영어로 정의하고, 이 정의문(definitions)에 대하여 텍스트 코사인 유사도 분석과 GPT-4와 Claude 등을 활용하여 다음과 같은 9개 대분류 군집(AI 속련 분류)으로 재분류하였다. 9개 대분류는 [그림 5-1]에서 보는 바와 같이 AI ML 모델과 알고리즘, AI 프레임워크 라이브러리, AI 시스템 공학 MLOps, 데이터 과학 분석, 자연어 처리, 컴퓨터 비전 이미지 인식, 자율 시스템, 로봇공학 자동화 제어, 책임 있는 AI 등이다.

이후 AI 키워드들을 이 AI 속련 대분류 체계에 매핑하기 위해 하이브리드 텍스트 매칭 기법을 적용하였다(자세한 매칭 기법은 [박스 2]를 참조할 수 있다). 먼저, 각 분류 정의문(description)과 키워드 설명을 OpenAI 임베딩 모델(text-embedding-3-large)을 이용해 벡터화하고, 이를 기반으로 코사인 유사도를 계산하여 상위 후보군을 도출하였다. 동시에 BM25 기반의 전통 정보검색 기법을 병행하여 단어 수준의 일치 가능성을 보완하였다. 이후 코사인 유사도와 BM25 점수에서 추출된 후보 집합에 대해 LLM 기반 재랭킹을 수행하였다. 구체적으로, 각 키워드 설명과 분류 정의문을 문장쌍으로 구성하여 GPT-4o 모델로부터 0~1 범위의 적합도 점수를 산출하고, 이를 코사인 유사도 및 BM25 점수와 결합하였다. 최종 판정 점수는 $0.6 * \text{LLM} + 0.3 * \text{코사인} + 0.1 * \text{BM25}$ 의 가중합으로 산출되었으며, 상위 1순위와 2순위 후보 간 점수 차이는 final margin으로 정의하여 분류 확신도의 보조 지표로 활용하였다. 최종 분류 기준은 $\text{final_score} \geq 0.65$ 이면서 $\text{final_margin} \geq 0.08$ 인 경우 자동 확정(auto-accept), $\text{final_score} < 0.55$ 또는 $\text{final_margin} < 0.03$ 인 경우 검수(review), 그 외의 경우 조건부 확정(spot-check)으로 분류하였다. 이러한 절차는 임베딩 기반 의미 유사도, 전통적 키워드 매칭, 그리고 LLM 재랭킹을 통합함으로써 단일 접근법 대비 정밀성과 신뢰성을 동시에

제고할 수 있는 매칭 결과를 제공할 수 있었다.

[박스 2] 텍스트 유사도 매핑 방법

본 연구에서 사용된 키워드 간 매핑이나 텍스트 유사도 분석에는 다음과 같은 방법을 적용하였다.

첫째, 입력 데이터는 OECD AI 키워드나 대학 AI 강좌 커리큘럼 키워드와 이에 대한 영어 정의문을 도출하고, 이를 입력 데이터로 넣어, BM25 인덱스를 구축하여 문헌 검색 기반의 어휘적(lexical) 유사도를 산출한다.

$$BM25(q,d) = \sum_{t \in q} IDF(t) \cdot \frac{f(t,d) \cdot (k_1 + 1)}{f(t,d) + k_1 \cdot (1 - b + b \cdot \frac{|d|}{avgdl}}$$

여기서 $f(t,d)$ 는 문서 d 내에서 토큰 t 의 빈도, $|d|$ 는 문서 길이, $avgdl$ 은 평균 문서 길이, k_1, b 는 하이퍼파라미터이다.

둘째, 의미적(semantic) 유사도를 반영하기 위해 멀티링구얼 문장 임베딩 모델인 distiluse-base-multilingual-cased-v2를 사용하여 모든 키워드 정의문을 벡터화하였으며, 각 강좌 설명에 대해서도 동일한 방식으로 임베딩을 산출한 후 코사인 유사도를 계산하였다. 임베딩 로딩이 불가능한 경우에는 TF-IDF 기반 대체 경로를 마련하여 안정성을 확보하였다.

$$EmbedSim(q,d) = \frac{v_q \cdot v_d}{\|v_q\| \|v_d\|}$$

셋째, 두 방식(BM25, 임베딩)에서 각각 상위 K 개의 후보를 추출한 뒤, Reciprocal Rank Fusion(RRF) 기법을 적용하여 상호보완적인 순위를 융합하였다. 이 과정에서 서로 다른 검색 방식에서 반복적으로 등장하는 후보는 가산 처리되므로, 단일 기법 대비 더 높은 재현율(recall)을 달성할 수 있다.

넷째, RRF 융합 결과에 대해 선택적으로 크로스인코더(Cross-Encoder) 재랭킹을 수행하였다. 이를 위해 ms-marco-MiniLM-L-12-v2 기반의 문장 쌍분류 모델을 활용하여 강좌 설명-키워드 정의문 쌍에 대한 문맥적 정합성을 예측하였다. 이 과정을 통해 상위 후보군에서 의미적 정밀도(precision)를 향상시켰다.

$$\text{RRF}(d) = \sum_{m \in M} \frac{1}{k + \text{rank}_m(d)}$$

다섯째, 선택적으로 자연어 추론(NLI) 모델을 이용하여 강좌 설명이 키워드 정의문을 포함(entail)하는 확률을 추가 점수로 반영하였다. NLI 모델은 xlm-roberta-large-xnli를 사용하였으며, entailment 확률을 보정 지표로 도입하였다. 또한, 필요시 대규모 언어모델(LLM)을 활용한 판정 혹은 제공하여, 인간 수준의 해석 기반 보정이 가능하도록 설계하였다.

여섯째, 최종 점수는 다중 지표의 가중합으로 산출하였다. 즉, 임베딩 유사도, BM25 점수 정규화값, 크로스인코더 점수, NLI 포함 확률, LLM 판정 점수 등을 사전에 정의된 가중치로 통합하여 최종 스코어(final_score)를 계산하였다.

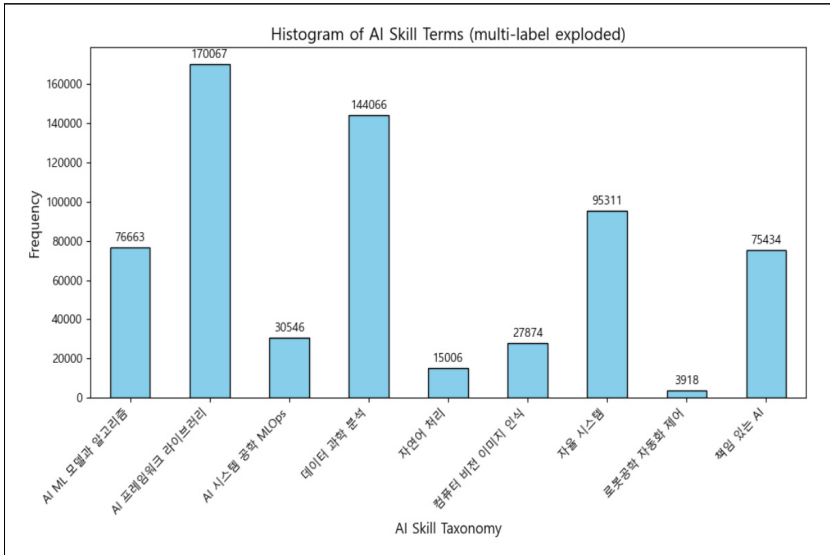
$$S_{\text{final}} = w_e \cdot S_{\text{embed}} + w_b \cdot S_{\text{bm25}} + w_c \cdot S_{\text{ce}} + w_n \cdot S_{\text{nli}} + w_l \cdot S_{\text{llm}}$$

이후, 유사도에 근거한 판정 기준을 적용하여 바로 채택, 검토 필요, 매칭 불가세 범주로 결과를 구분하였다. 임베딩 유사도가 일정 임계치를 초과하는 경우 자동 채택으로 분류하고, 그렇지 않더라도 최종 가중합 점수가 중간 임계치를 넘는 경우 검토 대상으로 분류하였다. 나머지는 매칭 불가로 처리하였다. 이러한 결과를 가지고 최종 매핑과 유사도를 판단하였다.

이러한 기준으로 매칭된 한국의 OJP 숙련 키워드는 613개였다. 이 중 IT 숙련 539개, 비IT 숙련 74개이다. 이 키워드를 AI 키워드라고 하고, OJP에서 AI 키워드가 2개 이상인 OJP를 AI-OJP로 정의하였다. 앞에서 검토한 OECD 연구에서도 AI 키워드가 2개 이상인 경우를 AI-OJP로 정의하였다. AI-OJP의 수는 17만 개(179,492)이다. 이 데이터를 대상으로 AI 숙련 대분류별 빈도와 AI 숙련 키워드별 빈도를 검토한 것이 [그림 5-1]과 <표 5-1>이다.

2024년 OJP 분석 결과, AI 숙련 키워드가 2개 이상 포함된 공고에서는 AI 프레임워크·라이브러리, AI/ML 모델과 알고리즘, 데이터 과학 분석, 자율 시스템과 같은 범용(generic) 숙련이 가장 높은 빈도를 보였다. 이는 산업 전반에서 공통으로 요구되는 기초적이고 범용적인 역량이 핵심적임을 보여준다. 반면, 자연어 처리, 컴퓨터 비전, 로봇공학 제어와 같은 특수(specific) 숙

[그림 5-1] AI 온라인 구인공고에서의 AI 속련 키워드의 분포(OJP당 복수 키워드 반영)



주: 2024년 AI-OJPs(AI 속련 키워드를 2개 이상 포함한 OJPs)에서의 AI 속련 분포를 나타냄.
 자료: 저자 작성.

련은 비교적 낮은 빈도를 보였는데, 이는 특정 산업과 직무에서 선택적으로 강조되는 특화 속련임을 의미한다. ‘책임 있는 AI’가 상위 키워드로 부상한 것은 AI 활용에서 윤리·안전성·사회적 책임의 중요성이 강화되고 있음을 시사한다. 종합하면, AI 채용 시장은 범용적 기술 속련 중심으로 형성되고 있고, 특수 응용 속련은 상대적으로 수요가 강하지 않은 것으로, 그리고 윤리적 활용 능력 등에 대한 수요도 적지 않게 나타나고 있는 것으로 보인다.

〈표 5-1〉에서 좀 더 세부적으로 보면, ‘AI ML 모델과 알고리즘’ 영역에서는 인지 인사이트 시스템(46,884회), 생성형 AI(33,042회), 자연 언어 생성(12,537회) 등이 상위에 위치하였다. 이는 최근 AI 모델링이 단순한 알고리즘 설계를 넘어, 생성형·다모달 AI 및 지능적 의사 결정 지원 시스템으로 확장되고 있음을 보여준다. ‘AI 프레임워크와 라이브러리’에서는 DeepLearning4j(240,036회), 지능형 인프라(205,668회), Python(57,735회) 등이 압도적으로 높은 빈도를 차지하였다. 이는 실무에서 프로그래밍 언어와 오픈소스 프레임워크에 대한 활용도가 매우 높음을 의미하며, 채용 시장에서 실질적으로

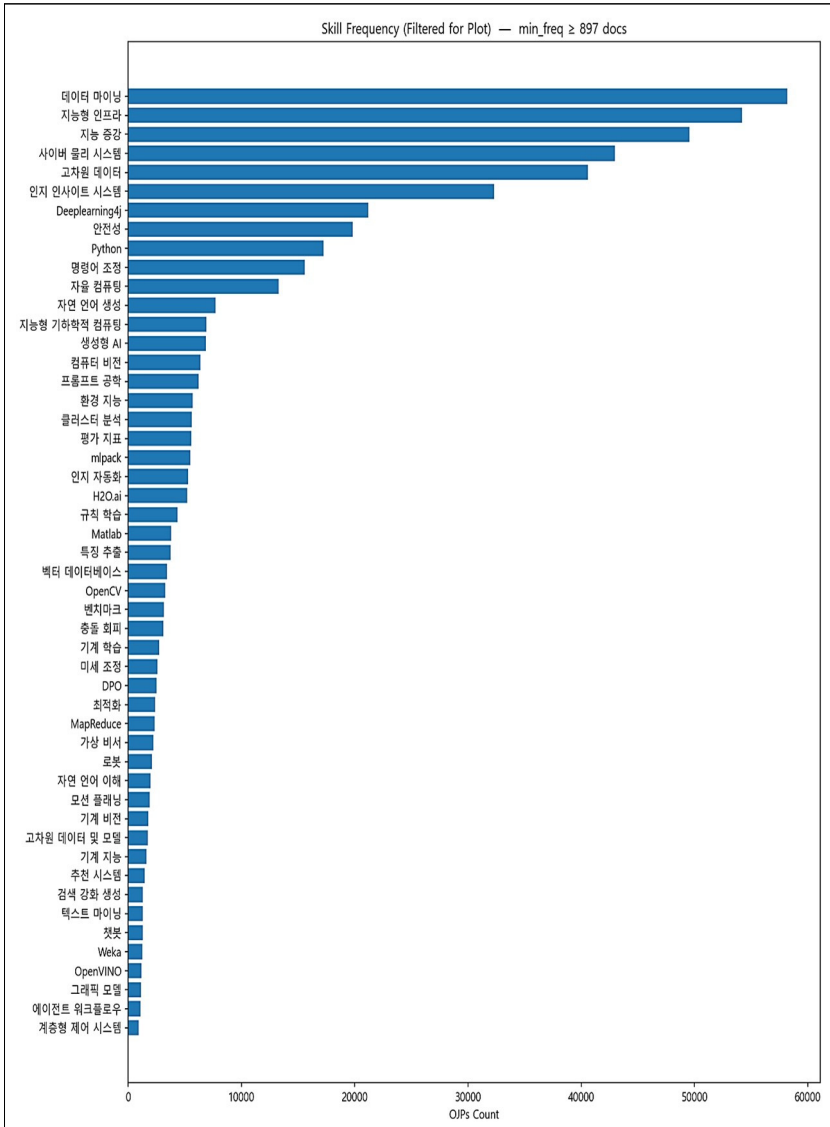
〈표 5-1〉 AI 숙련 분류별 빈도수 상위 10개 AI 키워드

	AI 숙련 분류	빈도수 상위 10개 숙련 키워드
1	AI ML 모델과 알고리즘	인지 인사이트 시스템 46884; 생성형 AI 33042; 자연 언어 생성 12537; 규칙 학습 7701; 평가 지표 7513; 기계 학습 4154; 그래픽 모델 2655; 추천 시스템 2478; 기계 지능 1856; 다모달 AI 848
2	AI 프레임워크 라이브러리	Deeplearning4j 240036; 지능형 인프라 205668; Python 57735; 명령어 조정 23226; H2O.ai 9463; mlpack 7796; 벡터 데이터베이스 7087; 벤치마크 6056; Theano 2395; 검색 강화 생성 1677
3	AI 시스템 공학 MLOps	최적화 2872; 미세 조정 2703; 모델 평가 131; 다목적 최적화 35; 메타 학습 7
4	데이터 과학 분석	데이터 마이닝 219699; 고차원 데이터 92264; 클러스터 분석 9408; MapReduce 8343; 고차원 데이터 및 모델 4629; Matlab 4338; 특징 추출 4231; 텍스트 마이닝 2502; Weka 1824; R 1108
5	자연어 처리	자연 언어 이해 2369; 챗봇 1326; NLTK 877; spaCy 719; 자연 언어 처리 339; 기계 번역 31; OpenNLP 24
6	컴퓨터 비전 이미지 인식	컴퓨터 비전 7722; 지능형 기하학적 컴퓨팅 7452; 기계 비전 7200; OpenCV 4493; OpenVINO 1547; 이미지 인식 1417; 감정 인식 607; 물체 탐지 498; 맹신호 분리 463; 이미지 처리 456
7	자율 시스템	사이버 물리 시스템 118366; 자율 컴퓨팅 37541; 환경 지능 17270; 인지 자동화 9915; 프롬프트 공학 7511; 충돌 회피 5832; 운동 계획 3150; 가상 비서 2918; 계층형 제어 시스템 2388; 장애물 회피 2101
8	로봇공학 자동화 제어	로봇 5037; 인간-로봇 상호작용 722
9	책임 있는 AI	지능 증강 56520; 안전성 31318; DPO 5689; 안전성 가이드라인 20

자료 : 저자 작성.

요구되는 핵심 역량임을 시사한다. 'AI 시스템 공학(MLOps)'은 최적화(2,872회), 미세 조정(2,703회)이 주를 이루어, 모델 개발 이후 운영·관리·최적화 단계의 중요성이 강조되고 있음을 확인할 수 있다. '데이터 과학 분석' 분야는 데이터 마이닝(219,699회), 고차원 데이터 분석(92,264회), 클러스터

[그림 5-2] AI 구인광고에서 나타나는 속련 키워드 수



자료 : 저자 작성.

분석(9,408회) 등이 핵심으로 나타났으며, 이는 방대한 데이터를 효율적으로 처리·분석하는 데이터 중심 역량이 AI 직무 전반의 기초임을 보여준다. 한편, ‘자연어 처리(NLP)’에서는 자연 언어 이해(2,369회), 챗봇(1,326회)이,

‘컴퓨터 비전 이미지 인식’에서는 지능형 기하학적 컴퓨팅(7,452회), 기계 비전(7,200회), OpenCV(4,493회) 등이 주요 키워드로 나타났다. 이들은 특정 도메인 응용에 해당하는 특수한(specific) 숙련으로, 산업별·서비스별 요구에 따라 선택적으로 강조되고 있음을 시사한다. ‘자율 시스템’ 영역은 사이버 물리 시스템(118,366회), 자율 컴퓨팅(37,541회) 등이 높은 비중을 차지하며, 실제 산업 현장에서 자율화·자동화 시스템 구현 능력에 대한 수요가 크다는 점을 보여준다. ‘로봇공학 자동화 제어’는 상대적으로 낮은 빈도를 보였지만, 로봇(5,037회), 인간-로봇 상호작용(722회)이 등장하여, 특정 제조·서비스 산업에서 요구되는 특수 역량을 드러냈다. 마지막으로, ‘책임 있는 AI’ 영역은 지능 증강(56,520회), 안전성(31,318회)이 상위에 위치하였다. 이는 AI 채용 수요 속에서 윤리·책임·안전성을 고려한 AI 활용 능력이 점차 중요한 역량으로 자리 잡고 있음을 보여준다.

아래 <표 5-2>는 AI-OJP 빈도수로 상위 30개 직업에서 직업별로 상위 3개 AI 숙련 키워드를 보여주고 있다. AI 숙련 수요는 응용 소프트웨어 개발자, 웹 개발자, 기획·마케팅 사무원 등 IT 개발과 데이터 기반 기획 직무에 집중되어 있다. 이들 직업군에서는 Deeplearning4j, 지능형 인프라, 데이터 마이닝 같은 기술 키워드가 핵심적으로 요구되었다. 한편, 정보시스템 운영자, 시스템 소프트웨어 개발자, 전자·전기공학 기술자 및 연구원 직무에서는 사이버 물리 시스템, Python, 프롭프트 공학 등이 나타나 하드웨어-소프트웨어 융합 역량과 운영·최적화 기술의 중요성을 보여준다. 또한, 회계, 총무, 물류, 고객 상담, 온라인 판매와 같은 일반 사무·서비스 직종에서도 데이터 마이닝, 지능 증강, 고차원 데이터 등 데이터 활용 능력이 꾸준히 요구되고 있어, AI 수요가 전문 IT 직종을 넘어 전 산업·직무로 확산되고 있음을 시사한다. 더불어 정보보안 전문가와 미디어 콘텐츠 디자이너 직군에서는 각각 AI-보안 융합 역량, AI-콘텐츠 제작 역량이 중요한 키워드로 등장하여, 특수 영역에서도 AI 숙련이 핵심 경쟁력으로 자리잡고 있음을 확인할 수 있다. 종합하면, OJP로 본 2024년 AI 채용시장은 소프트웨어·데이터 기반 직무 중심으로 형성되면서도, 사무·서비스·보안·콘텐츠 분야까지 AI 활용 역량을 요구하는 것으로 보인다.

〈표 5-2〉 직업별 AI 속련 키워드(상위 30개 직업)

KECO	직업명	빈도	비율	상위 3개의 속련명
1332	응용 소프트웨어 개발자	24,398	13.6	Deeplearning4j 23949; 지능형 인프라 14637; 사이버 물리 시스템 9675
1333	웹 개발자	18,686	10.4	Deeplearning4j 18562; 지능형 인프라 10310; 데이터 마이닝 7945
261	기획·마케팅 사무원	17,676	9.8	데이터 마이닝 14109; 고차원 데이터 10983; 생성형 AI 8502
1343	정보시스템 운영자	11,189	6.2	지능형 인프라 11899; 자율 컴퓨팅 6027; Deeplearning4j 4536
1532	전자공학 기술자 및 연구원	7,617	4.2	사이버 물리 시스템 6966; 지능형 인프라 2718; 프롭프트 공학 2143
1331	시스템 소프트웨어 개발자	6,881	3.8	사이버 물리 시스템 7444; 지능형 인프라 4127; Python 2753
283	자재·구매·물류 사무원	4,402	2.5	데이터 마이닝 2837; 지능 증강 2002; 고차원 데이터 1268
8419	기타 정보통신기기 설치·수리원	4,243	2.4	지능형 인프라 5836; 안전성 1783; 사이버 물리 시스템 1540
8329	기타 전기·전자 기기 설치·수리원	3,948	2.2	지능형 인프라 2259; 안전성 1655; 사이버 물리 시스템 1443
271	회계 사무원	3,637	2	데이터 마이닝 1963; 지능 증강 1835; 고차원 데이터 1255
4155	미디어 콘텐츠 디자이너	3,540	2	데이터 마이닝 1836; 인지 인사이트 시스템 1122; 지능형 인프라 1057
263	총무 사무원 및 대학 행정조교	3,373	1.9	지능 증강 2602; 데이터 마이닝 2386; 고차원 데이터 2296
1591	제도사	3,335	1.9	지능형 기하학적 컴퓨팅 2253; 사이버 물리 시스템 1744; 지능형 인프라 535
1531	전기공학 기술자 및 연구원	3,214	1.8	사이버 물리 시스템 2464; 인지 자동화 1438; 운동 계획 1032
1320	컴퓨터시스템 전문가	2,891	1.6	Deeplearning4j 2231; 지능형 인프라 1912; 데이터 마이닝 1407
292	고객 상담원 및 모니터 요원	2,794	1.6	데이터 마이닝 1283; 지능 증강 997; 지능형 인프라 778
1350	정보보안 전문가	2,739	1.5	지능형 인프라 3773; 안전성 2962; DPO 944

〈표 5-2〉의 계속

KECO	직업명	빈도	비율	상위 3개의 숙련명
272	경리 사무원	2,658	1.5	지능 증강 1686; 데이터 마이닝 1195; 고차원 데이터 1041
8411	컴퓨터 설치·수리원	2,624	1.5	지능형 인프라 1740; 안전성 1727; 자율 컴퓨팅 805
295	전산자료 입력원 및 사무 보조원	2,580	1.4	데이터 마이닝 1805; 지능 증강 1408; 고차원 데이터 885
243	상품 기획자	2,411	1.3	데이터 마이닝 3478; 고차원 데이터 1871; 지능 증강 1083
1341	데이터 전문가	2,358	1.3	데이터 마이닝 4625; 고차원 데이터 1103; 지능형 인프라 762
1339	기타 컴퓨터 전문가 및 소프트웨어 전문가	2,299	1.3	벤치마크 1294; 사이버 물리 시스템 815; 지능형 인프라 661
6153	온라인 판매원	2,177	1.2	데이터 마이닝 1809; 지능 증강 1131; 고차원 데이터 964
284	생산·품질 사무원	2,095	1.2	데이터 마이닝 2485; 고차원 데이터 1217; 지능 증강 1038
241	광고·홍보 전문가	2,015	1.1	데이터 마이닝 1453; 고차원 데이터 1113; 생성형 AI 1063
6121	기술 영업원	1,905	1.1	데이터 마이닝 1090; 지능형 인프라 849; 고차원 데이터 837
1511	기계공학 기술자 및 연구원	1,560	0.9	지능형 기하학적 컴퓨팅 944; 사이버 물리 시스템 694; 데이터 마이닝 446
262	인사·교육·훈련 사무원	1,558	0.9	데이터 마이닝 1114; 고차원 데이터 1060; 지능 증강 884
8340	전기·전자 설비 조작용	1,411	0.8	사이버 물리 시스템 749; 지능형 인프라 510; 안전성 507

자료: 저자 작성.

2. AI 교과 역량

한편 동일한 방법으로 교과 역량에 대한 키워드 추출과 역량 분류 체계를 구축하였다. 이를 위한 데이터는 전국 대학들의 AI 관련 강좌 커리큘럼을 수집하였다. 데이터는 두 가지 출처에서 수집되었는데, 하나는 각 대학의 자체

개설 강좌를 홈페이지 등을 통해서 수집된 강좌(비-KMOOC)이고, 다른 하나는 국가 차원의 온라인 교육 플랫폼인 K-MOOC에서 제공하는 강좌다. 홈페이지와 수작업을 통해 수집한 대학 AI 관련 강좌는 217개이고, K-MOOC 강좌는 260개로 전체 417개 강좌이다.

〈표 5-3〉은 대학별 강좌 분포를 보여준 것이다. 전체적으로 고려대학교, 서울대학교, 건국대학교, 경희대학교, 한양대학교, 성신여자대학교 등이 비-KMOOC 부문에서 높은 강좌 개설 수를 기록하고 있으며, K-MOOC 플랫폼에서는 성신여자대학교, 인하대학교, 제주대학교, 한동대학교, 포항공과대학교(POSTECH) 등이 두드러진 비중을 차지한다. 연구중심 대학(예: 고려대, 서울대, 포스텍, UNIST 등)은 비-KMOOC 영역에서 강세를 보인다. 이는 자체 연구 역량과 인프라를 바탕으로 학내외 학습자에게 직접 교육 기회를 제공하고 있는 것으로 보이고, 중소규모 대학 및 지역 대학(예: 제주대, 한동대, 포항공과대 등)은 K-MOOC를 통한 강좌 개설이 활발하다. 이는 대중적 접근성과 플랫폼 효과를 활용해 교육 기회를 넓히는 전략을 반영하는 것으로 보인다. 한편, 전공과 교양으로 구분해보면, 전공이 46.6%, 교양이 27.5% 정도로 되어 있다. 미분류를 고려하면, 전공이 61.3%, 교양이 36.2% 정도이다.

〈표 5-3〉 대학별 AI 강좌 분포

전체	강좌 수	비-KMOOC	강좌 수	KMOOC	강좌 수
고려대학교	27	고려대학교	16	성신여자대학교	8
서울대학교	21	서울대학교	13	인하공업전문대학	8
건국대학교	14	건국대학교	14	제주대학교	6
경희대학교	16	경희대학교	11	한동대학교	1
한양대학교	22	한양대학교	20	한국보건복지인재원	2
성신여자대학교	19	성신여자대학교	13	포항공과대학교	6
국민대학교	16	국민대학교	15	전남대학교	6
한국의국어대학교	4	한국의국어대학교	4	대구가톨릭대학교	3
한성대학교	6	한성대학교	3	서강대학교	3
서울과학기술대학교	11	서울과학기술대학교	11	(주)웅진씽크빅	2
덕성여자대학교	8	덕성여자대학교	7	EBS	18

〈표 5-3〉의 계속

전 체	강좌 수	비-KMOOC	강좌 수	KMOOC	강좌 수
GIST	6	GIST	6	JTBC	2
POSTECH	27	POSTECH	27	서울대학교	8
UNIST	11	UNIST	11	가톨릭대학교	1
단국대학교	15	단국대학교	15	부산대학교	8
충남대학교	13	충남대학교	13	국민대학교	1
연세대학교(미래)	8	연세대학교(미래)	8	고려대학교	11
연세대학교(신촌)	12	연세대학교(신촌)	12	선문대학교	6
인하공업전문대학	8			여수광양항만공사	1
제주대학교	6			한국과학기술원	2
한동대학교	1			세종대학교	10
한국보건복지인재원	2			한국연구재단 산업교육센터	9
포항공과대학교	6			중앙대학교	6
전남대학교	6			수원대학교	2
대구가톨릭대학교	3			충남대학교	3
서강대학교	3			한경국립대학교	1
(주)웅진씽크빅	2			전주대학교	3
EBS	18			한성대학교	3
JTBC	2			명지대학교	3
가톨릭대학교	1			인천가톨릭대학교	1
부산대학교	8			덕성여자대학교	1
선문대학교	6			IGM세계경영연구원	3
여수광양항만공사	1			신한대학교	1
한국과학기술원	2			서울시립대학교	1
세종대학교	10			인하대학교	1
한국연구재단 산업교육센터	9			상명대학교	7
중앙대학교	6			대구대학교	1
수원대학교	2			광운대학교(매치업)	5
충남대학교	3			광운대학교	4

〈표 5-3〉의 계속

전체	강좌 수	비-KMOOC	강좌 수	KMOOC	강좌 수
한경국립대학교	1			중부대학교	1
전주대학교	3			평택대학교	1
명지대학교	3			건국대학교(글로벌)	2
인천가톨릭대학교	1			이화여자대학교	10
IGM세계경영연구원	3			광주여자대학교	1
신한대학교	1			세명대학교	1
서울시립대학교	1			울산대학교	3
인하대학교	1			인천대학교	4
상명대학교	7			동아대학교	1
대구대학교	1			경희대학교	5
광운대학교(매치업)	5			호남대학교	6
광운대학교	4			서울여자대학교	1
중부대학교	1			동의대학교	2
평택대학교	1			단국대학교	3
건국대학교(글로벌)	2			대구한의대학교	1
이화여자대학교	10			동국대학교(WISE)	1
광주여자대학교	1			한양대학교	2
세명대학교	1			한국항공대학교	2
울산대학교	3			조선대학교	1
인천대학교	4			남서울대학교(매치업)	2
동아대학교	1			건설루션(주)	1
호남대학교	6			고려사이버대학교	3
서울여자대학교	1			서울사이버대학교	1
동의대학교	2			한국방송통신대학교	1
단국대학교	3			숙명여자대학교	1
대구한의대학교	1			한국원격대학협의회 AI융합교육원	4
동국대학교(WISE)	1			호서대학교	1
한국항공대학교	2			울산과학대학교	1
조선대학교	1				

〈표 5-3〉의 계속

전체	강좌 수	비-KMOOC	강좌 수	KMOOC	강좌 수
남서울대학교(매치업)	2			고려사이버대학교(매치업)	1
건설루션(주)	1			한남대학교	2
고려사이버대학교	3			한양사이버대학교	1
서울사이버대학교	1			지능정보산업협회	1
한국방송통신대학교	1			현대엔지비	1
숙명여자대학교	1			한국지능정보사회진흥원	1
한국원격대학협의회				성균관대학교	15
AI융합교육원	4			송실대학교	1
호서대학교	1			유원대학교	1
울산과학대학교	1			연세대학교(미래)	2
고려사이버대학교(매치업)	1			국립안동대학교	1
한남대학교	2			가톨릭관동대학교	1
한양사이버대학교	1			한국생산기술연구원	1
지능정보산업협회	1				
현대엔지비	1				
한국지능정보사회진흥원	1				
성균관대학교	15				
송실대학교	1				
유원대학교	1				
연세대학교(미래)	2				
국립안동대학교	1				
가톨릭관동대학교	1				
한국생산기술연구원	1				
합계	477		217		260

자료 : 저자 작성.

이 데이터를 가지고 앞에서 숙련 키워드 도출 및 분류 체계 구축과 동일한 방식으로 교과 커리큘럼에서 역량 키워드를 도출하고 9개의 분류 체계를 구축하였다. 9개 분류명은 수학 모델링과 알고리즘, 머신러닝과 딥러닝, 데이터 분석과 통계적 추론, 컴퓨터 비전과 시각 기술, 자율 시스템과 로보틱

〈표 5-4〉 대학별 AI 강좌 전공/교양 분포

	강좌 수	비율(%)	비율(미분류 제외)
전공 선택	93	19.5	25.7
전공 필수	78	16.4	21.5
전공 심화	51	10.7	14.1
교양	131	27.5	36.2
학문융합	2	0.4	0.6
자기 설계	2	0.4	0.6
연구과목	2	0.4	0.6
POSE-AI	2	0.4	0.6
자원·정보	1	0.2	0.3
미분류	115	24.1	
	477	100.0	100.0

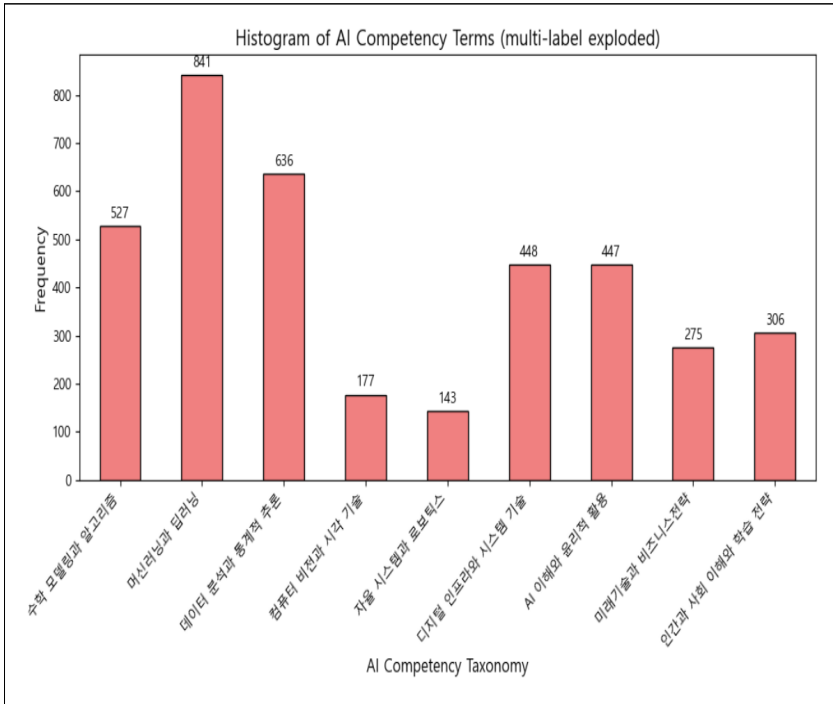
자료: 저자 작성.

스, 디지털 인프라와 시스템 기술, AI 이해와 윤리적 활용, 미래기술과 비즈니스 전략, 인간과 사회 이해와 학습 전략 등이다.

〈그림 5-3〉은 477개 AI 관련 강좌에서 다루는 AI 교과 역량의 빈도를 분석한 결과를 보여준다. 분석 결과, 강좌들이 가장 많이 다루고 있는 역량은 ‘머신러닝과 딥러닝’(841회)과 ‘데이터 분석과 통계적 추론’(636회)으로 나타났다. 이는 대학 교육에서 인공지능의 핵심 기술을 이해하고 실습하는 능력이 가장 중요한 학습 목표로 설정되고 있음을 보여준다.

그다음으로는 ‘수학 모델링과 알고리즘’(527회)이 높은 빈도를 차지하여, AI 학습의 기초가 되는 수학적 기반과 문제 해결 능력이 중시되고 있음을 알 수 있다. ‘디지털 인프라와 시스템 기술’(448회)과 ‘AI 이해와 윤리적 활용’(447회)도 비교적 높은 비중을 차지하는데, 이는 실제 AI 기술을 구현할 수 있는 시스템적 기반과 사회적 책임 및 윤리적 고려가 함께 강조되고 있음을 시사한다. 반면, ‘컴퓨터 비전과 시각 기술’(177회)과 ‘자율 시스템과 로보틱스’(143회)는 상대적으로 적게 다루어지고 있다. 이는 특정 응용 분야 중심의 역량이 아직은 제한적으로 포함되어 있음을 보여준다. ‘미래기술과 비즈니스 전략’(275회), ‘인간과 사회 이해와 학습 전략’(306회)은 기술적 역량과 더불어 사회적·전략적 이해를 강화하려는 시도로 해석할 수 있다.

[그림 5-3] 교과 역량 대분류별 강좌 수(강좌당 복수 역량 반영)



자료 : 저자 작성.

이는 앞에서 검토한 OJP에서 AI 숙련 키워드 분포와 큰 틀에서 유사한 것으로 볼 수 있다. 대학의 AI 교육이 기초 수학적 기반, 핵심 알고리즘, 데이터 분석 등 일반 역량 중심으로 이루어지고 있고, 응용 영역은 상대적으로 비중이 낮고, AI 윤리 부분도 적지 않게 다루어지고 있는 것으로 보인다. 큰 틀에서 직업 AI 숙련과 교육 AI 역량은 정합성이 있는 것으로 판단된다.

〈표 5-5〉는 9개 AI 교과 역량 대분류별로 강좌에서 가장 자주 등장한 상위 10개 키워드를 정리한 것이다. 전반적으로 ‘머신러닝과 딥러닝(딥러닝·신경망·강화학습 등)’이 가장 높은 빈도를 보이며, ‘수학 모델링과 알고리즘(알고리즘·최적화·확률·선형대수)’이 그다음으로 AI 학습이 수리적 기초와 핵심 알고리즘 이해에 강하게 집중되어 있음을 보여준다. ‘데이터 분석과 통계적 추론’에서는 데이터 분석·시각화·마이닝·회귀분석·전처리 같은 실무형 분석 역량이 고르게 상위에 배치되어, 데이터 기반 문제 해결

〈표 5-5〉 AI 교과 역량 분류와 상위 10개 역량 키워드(강좌당 복수 역량)

	AI 교과 역량 분류	빈도수 상위 10개 역량 키워드
1	수학 모델링과 알고리즘	알고리즘 97; 최적화 77; 확률 52; 선형대수 43; 고유벡터 38; 통계 37; 선형대수학 34; 행렬 33; 선형 대수 33; 확률 분포 31
2	머신러닝과 딥러닝	딥러닝 374; 머신러닝 192; 신경망 182; 기계 학습 164; 합성곱 신경망 148; 기계학습 141; 강화 학습 119; 순환 신경망 97; CNN 96; 자연어 처리 91
3	데이터 분석과 통계적 추론	데이터 분석 218; 빅데이터 118; 데이터 시각화 81; 데이터 마이닝 65; 회귀 분석 52; 데이터 전처리 44; 빅 데이터 35; 데이터 가공 30; 데이터 과학 30; R 프로그래밍 28
4	컴퓨터 비전과 시각 기술	증강현실 13; 가상현실 12; 의료 영상 12; 3D 렌더링 11; 객체 인식 11; 위치 추정 10; 포토닉 크리스탈 9; 광 공진기 9; 메타물질 9; 플라즈모닉스 9
5	자율 시스템과 로보틱스	로봇 기술 18; 이동 로봇 14; 스마트 팩토리 12; 소프트 로봇 10; 로봇 역사 10; 인공 근육 액추에이터 10; 공중 조종 로봇 10; 로봇 응용 10; 생체 모사 로봇 10; 자율주행 10
6	디지털 인프라와 시스템 기술	사물인터넷 62; 블록체인 37; 메타버스 32; 클라우드 컴퓨팅 28; 개인정보 보호 23; 네트워크 보안 21; 디지털 경제 20; 정보 보안 19; 사용자 경험 19; 디지털 회로 17
7	AI 이해와 윤리적 활용	인공지능 윤리 69; 데이터 윤리 23; 자동화 18; 사이버 범죄 16; 초지능 15; AI 창작 15; AI 응용 14; AI 윤리 14; 사회적 쟁점 14; 인공지능 교육 13
8	미래기술과 비즈니스 전략	특허 9; 지식재산권 8; 콘텐츠 생성 8; 사이보그 8; UX 평가 7; 인공지능 디자인 7; UX 설계 7; 사용자 중심 디자인 7; 인간과 AI 상호작용 7; 사용자 여정 맵 7
9	인간과 사회 이해와 학습 전략	비판적 사고 26; 스피치 교육 18; 의사소통 역량 15; 창의적 문제 해결 12; 기술 이해 9; 미디어 리터러시 9; 디지털 리터러시 9; 상황 인식 7; 수사학 7; 바이오피드백 7

자료 : 저자 작성.

역량이 필수 학습 목표로 자리 잡았음을 시사한다.

응용 영역 측면에서는 ‘컴퓨터 비전과 시각 기술’이 AR/VR·의료영상·객체 인식 등으로 폭을 넓히고, ‘자율 시스템과 로보틱스’는 이동 로봇·스마트 팩토리·자율주행 등 “시스템 통합형” 주제가 많이 다뤄진다. ‘디지털 인프라와 시스템 기술’에서는 사물인터넷·클라우드·블록체인과 함께 보

안·개인정보 보호가 상위에 올라, AI 구현을 뒷받침하는 인프라·보안 역량의 교육 비중이 적지 않음을 보여준다. ‘AI 이해와 윤리적 활용’에서는 인공지능 윤리·데이터 윤리가 중심을 이루며, ‘미래기술과 비즈니스 전략’은 특허·지식재산권·UX/사용자 경험 등 기술-비즈니스 연계를 강조한다. 마지막으로 ‘인간과 사회 이해와 학습 전략’은 비판적 사고·의사소통·디지털/미디어 리터러시 등 학습·소통 역량을 포함해, 기술 교육과 더불어 사회적·메타인지적 역량을 보완하고 있다.

종합하면, 현재 한국의 AI 대학 강좌 커리큘럼은 기초 수학-핵심 ML/DL-데이터 분석의 기반 위에 시스템·인프라·보안을 얹고, 응용 영역(비전·로보틱스)과 윤리·비즈니스·학습 전략을 모두 반영하고 있는 것으로 보인다.

3. 숙련 네트워크 지표

다음으로, 숙련 간 네트워크 구조를 파악할 수 있는 숙련 네트워크 지표를 구축한다.

본 연구에서의 네트워크는 OJP에 등장한 숙련들을 노드로, OJP에 숙련들이 공동으로 등장하는 것을 엣지로 하는 네트워크이다. 네트워크 연결 중심성(degree centrality)은 한 숙련이 얼마나 많은 다른 숙련과 직접 연결되어 있는지를 나타내는 가장 기본적인 지표다. 이 값이 높다는 것은 해당 숙련이 다양한 다른 숙련과 함께 빈번히 공존하며, 여러 기술 묶음의 출발점 혹은 교차점으로 기능함을 시사한다. 예컨대 “데이터 분석”이나 “Python 프로그래밍”은 많은 직무에서 폭넓게 요구되므로 다양한 숙련들과 나란히 제시되어 연결 중심성이 높게 관측된다. 반대로 “TensorRT 최적화”처럼 특정 맥락에 집중되는 숙련은 연결 상대가 적어 연결 중심성이 낮게 나타난다. 다만 네트워크 연결 중심성은 연결의 ‘수’를 강조할 뿐, 연결 상대의 중요도나 네트워크의 위계는 반영하지 못한다는 한계를 갖는다.

이 한계를 보완하는 것이 아이겐벡터 중심성(고유벡터 중심성, eigenvector centrality)이다. 이는 단순히 연결 수가 아니라 “중요한 노드와 연결되어 있는가”를 동시에 고려한다. 따라서 중간 정도의 연결 수를 갖더라도, 네트워크

크에서 영향력이 큰 허브 속련들과 긴밀히 연결된 경우 높은 고유벡터 중심성을 보일 수 있다. 예컨대 “API 설계”가 “시스템 설계”, “클라우드 아키텍처”, “데이터 모델링” 등 네트워크 상위의 핵심 속련들과 강하게 묶여 있다면, 연결 중심성은 보통 수준이더라도 아이겐벡터 중심성은 높게 나온다. 반대로 ‘엑셀 매크로’처럼 광범위하게 쓰이지만 상대적으로 주변부 기술들과의 동시 출현이 많은 경우에는 연결 중심성은 높아도 아이겐벡터 중심성은 제한적일 수 있다. 요컨대 연결 중심성은 ‘폭’을, 아이겐벡터 중심성은 ‘질’을 더 반영하는 지표라고 할 수 있다.

한편 본 연구에서는 특정 속련이 등장하는 공고들의 속련 맥락의 폭을 나타내는 맥락 중심성(context generality) 지표도 활용한다. 이는 특정 속련이 등장하는 공고들의 속련 맥락의 폭을 평균적으로 요약한다. 즉, 각 OJP 내에서 AI 속련들의 보유 비율(하나의 OJP에 속련이 10개 요구되는데 AI 속련이 3개 요구되면 0.3)을 먼저 구한 뒤, 해당 속련을 포함하는 공고들만 모아 그 평균을 다시 계산하고(예를 들어, Python이라는 속련이 포함된 OJP들만 모아 평균을 계산한다), 그 상대 비율을 구한 것이다. 이 값이 높다는 것은 그 속련이 보통 다양한 다른 속련들과 함께 요구되는 포지션에 등장한다는 것을 뜻한다. 예컨대 “MLOps”는 파이프라인, 배포 자동화, 모니터링, 데이터 엔지니어링 등과 묶여 제시되는 경향이 강해 맥락 중심성이 높다. 반대로 “CAN 버스 스트레스 테스트”처럼 특정 장치·도메인에 강하게 묶인 기술은 상대적으로 협소한 기술 묶음에서 등장하여 맥락 중심성이 낮게 관측된다. 이 지표는 빈도나 위계보다는 공고 수준의 동시 요구 폭(맥락의 두께)을 보여준다는 점에서 다른 중심성 지표들과 해석 축이 다르다. 이 값이 높으면, 그 속련이 등장하는 공고들은 대체로 다양한 다른 AI 속련들과 함께 기술 맥락을 이룬다는 뜻이고, 반대로 값이 낮다면 “그 속련은 좁은 기술 맥락에서 특정 상황에 국한되어 요구된다”는 뜻이다.

다음으로 속련 네트워크 구조의 계층성, 위계 구조에 집중하는 지표를 도출한다. 속련의 위계 구조에 관한 지표는 Hosseinioun et al.(2025)의 방법에 따라, LRC와 CS 지표를 도출한다.

LRC(Local Reaching Centrality)는 네트워크 방향 그래프의 위계를 직접 수치화한다. 어떤 속련에서 시작해 도달 가능한 후행 속련의 수(혹은 비율)

를 측정함으로써, 그 속련이 다른 속련들의 ‘전제’ 또는 ‘상위 개념’으로 얼마나 넓게 작동하는지를 포착한다. 이 값이 클수록 더 많은 특수 속련을 ‘아우르는’ 일반 속련이며, 값이 작을수록 특정 영역에만 종속되는 특수 속련으로 해석된다. 예컨대 “프로그래밍 기초”나 “데이터 구조/알고리즘”은 다양한 응용 기술로 뻗어 나가 높은 LRC를 보이지만, “CUDA 커널 튜닝” 같은 속련은 제한된 하위 집합에만 연결되어 LRC가 낮다. LRC는 단순한 연결 수보다 도달 범위(reach)와 위계(hierarchy)에 초점을 맞춘다는 점에서 연결 중심성이나 아이겐벡터 중심성 등과 다르다.

CS(Complexity Score)는 네트워크의 깊이와 중첩 구조까지 고려해 속련의 ‘복잡성’을 측정하려는 지표다. 구현 방식에 따라 차이가 있으나, 본 연구에서는 의존 그래프상에서 한 속련으로부터 다른 속련까지의 경로 길이를 고려하여 가까운 후행 속련에 더 큰 가중을 두고(예: 1/거리) 누적하는 접근을 취한다(해당 노드에서 다른 모든 노드로의 최단 거리 합의 역수). 예를 들어, Skill A가 “기초 통계”이고 “데이터 시각화(Skill D)” 속련으로 가려면 중간에 “판다스 활용(B)”, “matplotlib 사용(C)”이 필요하다면, 기초 통계만 갖고는 바로 시각화로 전이가 어렵고 3단계 경로가 필요하므로 경로가 길어지고, CS 값은 작게 나온다. LRC가 “얼마나 많은 후행 속련에 닿는가”를 단순 집계한다면, CS는 “얼마나 깊고 조밀한 계층 구조 속에 놓여 있는가”를 반영한다. 예컨대 “프로그래밍 기초”처럼 넓게 뻗되 얇은 층위로 연결된 속련과, “데이터 아키텍처 설계”처럼 다단의 전제 조건을 통해 깊은 경로로 이어지는 속련은 둘 다 LRC가 높을 수 있지만, 후자의 CS가 더 높게 평가될 수 있다. CS는 경로 구조에 민감하므로, 그래프 구축(공출현 임계값, 조건부 확률 비교, 스무딩 등) 단계의 설계가 결과에 실질적 영향을 미치게 된다.

속련 네트워크의 위계적 구조에서 상위에 있으면 “일반”적 속련이고, 하위에 있으면 “특수”한 속련이다. 앞에서 도출한 LRC 값으로, 일반, 중간, 특수 속련을 나눌 수 있다.

또한, 속련 네트워크의 ‘빈도’와 ‘방향’까지 고려하여 일반, 중간, 특수 속련으로 나눌 수 있다. 빈도는 해당 속련이 등장한 공고 비중(job_ratio)으로 파악할 수 있다. 방향의 경우, out-degree를 전체 차수(in+out)로 나눈 값인 RER(Relative Edge Ratio)이 해당 속련의 방향성 프로파일을 나타낸다. 1에

가까울수록 그 숙련이 다른 숙련을 ‘내보내는’ 일반적 출발점(원인, 전제)으로 기능함을, 0에 가까울수록 다른 숙련에 ‘의존하는’ 특수 도착점(결과, 응용)으로 기능함을 의미한다. 예컨대 “데이터 처리 프레임워크 설계”는 다양한 세부 구현으로 이어지므로 RER이 높고, “TensorRT 최적화”는 다른 기반 기술을 필요로 하는 말단에 놓여 RER이 낮다. LRC는 전체(전역) 도달 범위를 본다면, RER은 국소(local) 방향성을 나타낸다. LRC가 높으면서도 RER이 중간(≈ 0.5)인 경우는, 위계상 상위에 있지만 들어오고 나가는 의존이 균형 잡힌 연결 허브형 일반 숙련으로 해석할 수 있다.

G_score는 위계적 일반성(LRC), 방향성 일반성(RER), 그리고 시장 노출 빈도(job_ratio)를 하나의 척도로 종합하기 위해 설계한 복합 지표다. 구체적으로 job_ratio, LRC, RER을 각각 표준화(z-score)한 뒤 합산한다. 값이 높은 숙련은 (i) 공고에 자주 등장하고, (ii) 많은 후행 숙련을 아우르며, (iii) 방향성 측면에서도 ‘내보내는’ 성격이 강한, 소위 핵심·범용 기술의 특성을 보인다. 예컨대 “Python”, “데이터 모델링”, “API 설계”처럼 직무 스펙 전반에 빈번히 등장하고 다른 기술 요구의 출발점이 되는 숙련이 이에 해당한다. 반대로 낮은 G_score는 특정 하위 도메인에서만 쓰이는 말단 기술의 가능성이 높다. 다만 G_score에는 LRC와 RER이 포함되어 있어 이들과 긍정 상관을 띠는 것이 구조적으로 자연스럽다는 점, 그리고 빈도 성분(job_ratio)이 과도할 경우 시장 유행에 따른 단기 변동성을 함께 반영할 수 있다는 점을 유념해야 한다.

이러한 G_score 분포를 K-means(3군)로 비지도 학습하여 일반-중간-특수(“general-intermediate-specific”) 숙련으로 구분하였다. 이는 LRC 한 축만으로 자르지 않고, 빈도·위계·방향성을 동시에 고려해 구분한다는 점에서 실무적 직관과 부합하는 장점이 있다. 다만 군집은 데이터 분포에 따라 경계가 달라질 수 있으므로, 해석 시에는 샘플 분포와 대표 사례를 함께 제시하는 것이 바람직하다.

종합하면, 연결 중심성과 아이젠벡터 중심성은 각각 직접 연결의 폭과 연결 상대의 중요도를 보여주고, 맥락 중심성은 연결 상대의 폭을 나타낸다. LRC는 전역 위계의 도달 범위를, CS는 위계 경로의 깊이와 중첩성을 추가로 반영한다. LRC는 순수 위계 구조의 기준선을 제공하여 숙련의 위계 구조

를 ‘일반-중간-특수’로 분류할 수도 있고, G_score는 LRC, 빈도, 그리고 RER과 같은 국소 방향성(일반↔특수)까지를 반영하여 위계 구조를 보여줄 수 있다.

이러한 지표들을 활용하여, 일반(`general_level`), 중첩(`nested_level`), 비중첩(`un-nested_level`) 지표를 도출할 수 있다. 본 연구에서 제시하는 `general_level`, `nested_level`, `un-nested_level` 지표는 숙련 네트워크의 구조적 특성과 위계적 종속성을 반영하여 개별 숙련의 위상을 다차원적으로 파악하기 위해 도출된 것이다.

첫째, `general_level`은 특정 숙련이 다양한 직무와 직종에서 기본적으로 활용되는 범용성을 의미한다. 이 지표는 조건부 의존 네트워크에서의 도달 가능성이나 연결 정도와 같은 구조적 중심성, 특히 국소 도달 중심성(Local Reaching Centrality, LRC)과 연결중심성(degree centrality)을 토대로 산출된다. 예를 들어 Python과 같은 숙련은 머신러닝, 데이터 처리, 시각화, 인공지능 프레임워크 등 다수의 다른 숙련과 광범위하게 결합되어 나타나므로 높은 `general_level` 값을 가지게 된다.

둘째, `nested_level`은 특정 숙련이 다른 포괄적 숙련의 하위 요소나 종속적 기술로 반복적으로 등장하는 정도를 나타낸다. 이는 조건부 확률 기반의 조건부 의존 그래프(conditional dependency graph)를 구축한 뒤, 두 숙련 간의 확률적 종속 관계와 방향성을 추정함으로써 산출된다. 가령 “머신러닝 모델링”은 “데이터 분석”이나 “통계적 추론”과 같은 상위 숙련과 함께 나타날 때 종속적으로 위치하는 경우가 많아 `nested_level` 값이 높게 나타난다.

셋째, `un-nested_level`은 반대로 특정 숙련이 다른 숙련에 포함되거나 종속되지 않고 독립적으로 활용되는 정도를 보여준다. 이는 조건부 종속 네트워크에서 상위-하위 관계 간선이 거의 형성되지 않거나, 특정 산업에 특화되어 독립적으로 출현하는 숙련일수록 높은 값을 가진다. 예컨대, “CUDA 최적화”나 “ROS 네비게이션 모듈”과 같이 특정 분야에 한정된 특수 기술들은 다른 범용 숙련에 종속되지 않고 독립적으로 나타나기 때문에 `un-nested_level` 지표가 높게 측정된다.

이들 지표는 숙련 네트워크의 구조적 특성을 반영하는 동시에, 노동시장에서의 숙련 활용 패턴을 이해하는 데 기여한다. `general_level`은 광범위한

직업군에서 활용되는 범용 숙련을, nested_level은 특정 숙련체계 내에서 종속적·계층적으로 작동하는 숙련을, un-nested_level은 고유하고 독립적인 특수 숙련을 규정한다. 이를 통해 우리는 개별 숙련이 노동시장 내에서 어떠한 위치를 차지하는지를 다차원적으로 파악할 수 있으며, 이는 직업훈련 및 정책 설계에 있어 범용성과 특수성 간 균형을 조정하는 데 이론적 근거를 제공한다.

4. 숙련 네트워크 지표의 정책적 의미

본 연구에서 제시한 다양한 네트워크 지표들은 노동시장 변화와 직업훈련 정책 설계에서 상호보완적 의미를 가진다. 따라서 이들 지표는 노동시장 수요 변화에 대응하는 기초·핵심 숙련의 체계적 파악, 직업훈련 과정의 구조적 설계, 훈련 투자와 자원 배분의 우선순위 설정에 정책적으로 활용될 수 있다.

우선, 연결 중심성은 특정 숙련이 얼마나 많은 다른 숙련과 연결되는지를 보여줌으로써 노동시장 전반에서 기초적이고 범용적인 숙련을 식별할 수 있게 한다. 예컨대, 데이터 처리나 기본적인 소프트웨어 활용과 같은 숙련은 다양한 직무에서 공통적으로 요구되기 때문에 높은 연결 중심성을 보이며, 이는 기초 직업훈련이나 보편적 역량 강화 프로그램에 우선 편성해야 할 필요성을 시사한다.

한편, 아이겐벡터 중심성은 단순히 연결의 수가 아니라 중요도가 높은 숙련과 얼마나 연결되어 있는지를 반영한다. 이는 특정 숙련이 노동시장의 핵심 숙련군 내에서 가지는 영향력을 보여주며, 빈도는 낮더라도 전략적으로 중요한 숙련을 파악할 수 있게 한다. 따라서 아이겐벡터 중심성이 높은 숙련은 전략적으로 중요한 고급 직무훈련 과정이나 특화 과정에 포함시켜야 할 우선순위 숙련으로 평가될 수 있다.

맥락 중심성은 개별 숙련이 다른 숙련들과 함께 얼마나 자주 결합되는지를 보여주는 것으로, 직무 수행에 필요한 숙련 조합의 폭을 반영한다. 이 값이 높게 나타나는 숙련은 다양한 다른 숙련과 융합되어야 직무 수행이 가능하므로, 융합적이고 다학제적인 교육과정이 필요함을 의미한다. 반대로 맥

락 중심성이 낮은 숙련은 비교적 독립적으로 활용 가능하므로 단기 집중형 훈련 프로그램으로도 충분히 습득할 수 있다.

LRC는 숙련이 얼마나 많은 후행 숙련으로 확장되는지를 보여주어 해당 숙련의 위계적 일반성을 측정한다. LRC가 높은 숙련은 다양한 특수 숙련의 기반이 되기 때문에 직업 전환과 승급 경로를 확장할 수 있는 핵심 숙련으로 평가된다. 이러한 숙련은 노동시장 유연성을 높이는 데 기여하므로 기초 직업훈련이나 재교육 과정의 중심축으로 삼아야 한다. 반대로, LRC가 낮은 숙련은 특정 영역에 종속된 특수 숙련으로서 산업 변화에 따른 대체 위험이 크며, 전직 지원이나 재교육 정책에서 취약군으로 우선 고려해야 한다.

CS는 숙련이 네트워크에서 얼마나 중첩되고 깊은 계층 구조 속에 위치하는지를 보여주어 숙련의 학습 난이도와 경로 길이를 반영한다. CS가 높은 숙련은 습득 과정에서 여러 선행 숙련이 필요하며 학습 경로가 길기 때문에 장기 심화 훈련 과정에 적합하다. 반면 CS가 낮은 숙련은 비교적 단독으로 빠르게 습득 가능하므로 단기 훈련이나 전환 교육에 효율적으로 활용될 수 있다.

RER은 숙련 간 관계의 방향성을 반영하여 어떤 숙련이 다른 숙련의 전제 조건으로 작동하는지를 보여준다. RER이 높은 숙련은 훈련과정에서 선행 모듈로 배치되어야 하며, RER이 낮은 숙련은 후행 모듈이나 특수 직무군을 위한 전문 트랙에 적합하다. 따라서 RER은 직업훈련 과정의 모듈 간 순서를 합리적으로 설계할 수 있는 근거를 제공한다.

G_score는 빈도(job ratio), 위계(LRC), 방향성(RER)을 통합적으로 고려하여 숙련의 전반적 중요성을 측정하는 지표이다. G_score가 높은 숙련은 노동시장에서 반복적으로 요구되며 동시에 다양한 숙련으로 확장되는 핵심적인 위치에 있기 때문에 직업훈련과 인적자원개발 정책에서 최우선적으로 다루어야 할 대상이다.

다시 한번 요약하자면, 연결 중심성은 범용적 기초 숙련을 식별하여 보편적 훈련의 필요성을 보여주며, 아이겐벡터 중심성은 전략적 중요 숙련을 드러내어 특화 과정 설계에 활용된다. 맥락 중심성은 직무 수행에 필요한 숙련 조합의 폭을 반영해 융합형 교육과정의 근거가 된다. LRC와 RER은 숙련의 위계와 선후 관계를 규명함으로써 재교육 경로 및 훈련 모듈 순서 설계

〈표 5-6〉 숙련 네트워크 지표의 지표 특성과 정책적 함의

네트워크 지표		지표 특성	정책적 함의 (점수 값이 높은 경우)
연결 중심성	DC	얼마나 많은 숙련과 연결	기초 교육훈련이나 보편적 역량 강화 프로그램
아이겐벡터 중심성	EVC	중요한 숙련과 연결	전략적으로 중요한 고급 교육 훈련 과정
맥락 중심성	CC	직무 수행에 필요한 숙련과 함께 자주 결합	융합적이고 다학제적인 교육 과정
파급 중심성	LRC	얼마나 많은 후행 숙련으로 확장되는지	필수 교육훈련
복잡성 지수	CS	네트워크 위계에서 숙련 경로 길이	장기 심화 훈련 과정
숙련 방향성	RER	다른 숙련의 전제 조건으로 작동하는지	선형 교육훈련 과정 판별
G_score		빈도, 방향, 위계 동시 고려	최우선 교육훈련 과정

자료 : 저자 작성.

에 활용될 수 있다. G_score는 빈도와 위계, 방향성을 종합하는 종합 지표로 핵심 숙련을 우선적으로 제시한다. CS 지표는 숙련의 학습 난이도와 경로 길이를 반영해 훈련 기간과 심화 수준을 결정하는 데 기여한다.

결국, 이러한 지표들을 통합적으로 고려하는 것은 노동시장의 유연성과 안정성, 그리고 산업 맞춤형 혁신 역량을 동시에 확보하는 정책 전략을 마련하는 데 핵심적인 기여를 할 수 있다. 이러한 다양한 지표들을 결합하여 활용하면, 정책 입안자는 노동시장 변화에 따른 숙련 수요를 체계적으로 진단하고, 직업훈련의 내용과 구조, 자원 배분의 우선순위를 체계적으로 설정할 수 있다.

〈표 5-7〉은 AI 숙련 네트워크에서 중심성 상위-하위 10개 AI 숙련들을 나타낸 것이다. 중심성 지표 전반에서 상위 숙련으로 반복적으로 등장하는 핵심 기술들이 존재한다. 예컨대 Python, 딥러닝(DeepLearning4j 포함), 데이터 마이닝, 지능형 인프라, 인지 인사이트 시스템, 그리고 사이버 물리 시스템 등은 여러 지표에서 동시에 상위권을 차지한다. 이는 이들 숙련이 단순히

〈표 5-7〉 AI 숙련의 지표별 상위 및 하위 10개

	구인공고 수	연결 중심성	아이겐벡터 중심성	맥락 중심성	피급 중심성	복잡성 지수	일반 중심성
	job_count	degree Centrality	eigenvector	gen_mean_avg	lrc	cs	G_score
상위 10개	데이터 마이닝	지능형 인프라	지능형 인프라	인간-로봇 상호작용	지능형 인프라	지능형 인프라	데이터 마이닝
	지능형 인프라	데이터 마이닝	지능 증강	프롬프트 공학	데이터 마이닝	데이터 마이닝	지능형 인프라
	지능 증강	Python	사이버 물리 시스템	시각적 서보링	고차원 데이터	사이버 물리 시스템	지능 증강
	사이버 물리 시스템	사이버 물리 시스템	데이터 마이닝	OpenVINO	사이버 물리 시스템	인지 인사이트 시스템	사이버 물리 시스템
	고차원 데이터	지능 증강	안전성	패턴 인식	지능 증강	지능 증강	고차원 데이터
	인지 인사이트 시스템	명령어 조정	Python	딥 러닝	인지 인사이트 시스템	Python	인지 인사이트 시스템
	Deeplearning4j	인지 인사이트 시스템	명령어 조정	진화 알고리즘	Python	고차원 데이터	Deeplearning4j
	안전성	mlpack	인지 인사이트 시스템	RapidMiner	안전성	명령어 조정	안전성
	Python	클러스터 분석	mlpack	LLM(대규모 언어 모델)	Deeplearning4j	안전성	Python
	명령어 조정	고차원 데이터	Deeplearning4j	자연 언어 이해	자율 컴퓨팅	Deeplearning4j	명령어 조정

〈표 5-7〉의 계속

	구인공고 수	연결 중심성	아이겐벡터 중심성	맥락 중심성	피급 중심성	복잡성 지수	일반 중심성
	job_count	degree Centrality	eigenvector	gen_mean_avg	lrc	cs	G_score
하위 10개	메타 학습	메타 학습	학습 자동기	이미지 인식	다중 레이블 분류	Keras	메타 학습
	유전 알고리즘	인간 행동 인식	인간 행동 인식	장애물 회피	인과 추론	베이시안 모델 및 학습	유전 알고리즘
	학습 자동기	다층 퍼셉트론	유전 알고리즘	트랜스포머	지원 벡터 머신	인스턴스 기반 학습	학습 자동기
	인간 행동 인식	귀납적 모니터링	자동 인코더	감정 인식	모델 평가	인간 행동 인식	인간 행동 인식
	Keras	Kaldi	Keras	인과 추론	자율 주행 차량	귀납적 모니터링	Keras
	베이시안 모델 및 학습	진화 알고리즘	귀납적 모니터링	그래픽 모델	패턴 인식	Keras	베이시안 모델 및 학습
	인스턴스 기반 학습	Keras	Kaldi	인지 모델링	결정 트리	베이시안 모델 및 학습	인스턴스 기반 학습
	Kaldi	베이시안 모델 및 학습	베이시안 모델 및 학습	이미지 처리	지능형 에이전트	인스턴스 기반 학습	Kaldi
	진화 알고리즘	인스턴스 기반 학습	진화 알고리즘	모델 평가	이미지 처리	인간 행동 인식	진화 알고리즘
	귀납적 모니터링	자동 인코더	인스턴스 기반 학습	경로 계획	Jupyter	유전 알고리즘	귀납적 모니터링

자료: 저자 작성.

특정 산업이나 직무에 국한되지 않고, 네트워크 전반에서 다른 기술들과 강하게 연결되며, 노동시장에서 범용적이면서도 영향력이 큰 디지털 기반 역량임을 보여준다. 다시 말해, 중심성 지표가 달라지더라도 공통적으로 높은 순위를 보이는 숙련들은 인공지능 생태계에서 사실상 ‘허브 기술’로 기능하고 있음을 시사한다.

종합하면, 공통적으로 상위에 나타나는 숙련들은 AI·데이터 분석 생태계 전반을 지탱하는 기반 기술군이며, 특이하게 나타나는 숙련들은 지표별 특성에 따라 독립적·종속적 위치 혹은 맥락적 중요성을 드러낸다. 이러한 결과는 노동시장 내 숙련 수요가 단순히 개별 기술의 빈도에 의해 설명되지 않고, 숙련 간 연결망 구조와 맥락적 활용 패턴에 의해 달라진다는 점을 잘 보여준다.

제4절 AI 숙련 네트워크 구조

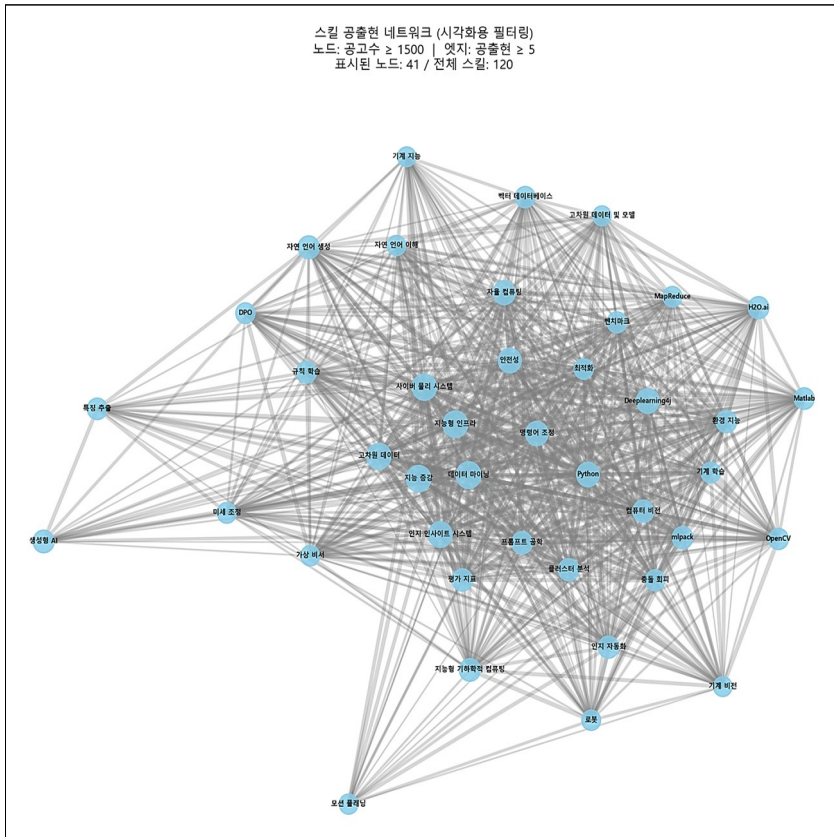
이 절에서는 앞에서 검토한 숙련 네트워크 지표들을 활용하여, 2024년 온라인 구인공고에 나타난 AI 숙련의 네트워크 구조를 검토한다. [그림 5-4]는 OJP에 등장한 AI 관련 숙련들의 공출현(co-occurrence) 네트워크를 시각화한 결과로, 각 숙련이 채용시장 속에서 어떤 방식으로 함께 요구되는지를 보여준다. 그림 속의 원형 노드는 각각 하나의 AI 숙련을 의미하며, 노드의 크기는 해당 숙련이 구인공고에 등장한 횟수, 즉 등장 빈도를 반영한다. 노드는 전체 구인공고 중 최소 897건 이상에서 등장한 숙련만 포함하여 빈도가 낮은 희귀 숙련을 제외하였고, 또한 하나의 공고(또는 강좌) 안에 두 숙련이 동시에 나타나면 두 노드(숙련) 사이에 간선, 엣지가 생기고, 함께 등장한 횟수가 많을수록 선이 더 두꺼워진다. 엣지는 두 숙련이 동일한 공고에서 동시에 나타나는 경우가 세 구인공고 이상에서 나타날 경우에만 연결하여 우연적 공출현을 배제하고 의미 있는 관계만을 반영하였다. 그 결과 상위 120개 숙련 중에서 출현 빈도가 높은 41개만 남겼고, 간선은 공존 횟수 5회 이상인 경우만 시각화에 반영했다. 채용시장에서 상대적으로 중요성이 높고

상호 결합이 뚜렷한 숙련들을 중심으로 시각화한 것이다.

이 그림은 채용시장에서 핵심적으로 요구되는 AI 숙련이 무엇인지, 어떤 숙련들이 함께 묶여 등장하는지, 그리고 어떤 숙련이 범용적이고 어떤 숙련이 특화되어 있는지를 직관적으로 보여준다. 다시 말해, 중심부의 큰 노드와 굵은 엣지는 현재 노동시장에서 AI 직무 수행을 위해 가장 중요한 역량을 반영하며, 주변부의 노드들은 특정 직무나 산업 분야에서 필요로 하는 특수한 응용 숙련을 나타낸다고 해석할 수 있다.

중앙에는 Python, 지능 증강, 지능형 인프라, 데이터 마이닝, 명령어 조정 등과 같은 범용 기반 숙련들이 허브를 이루고 있다. 이들은 학습 · 재직 · 전

[그림 5-4] AI 숙련 네트워크



자료 : 저자 작성.

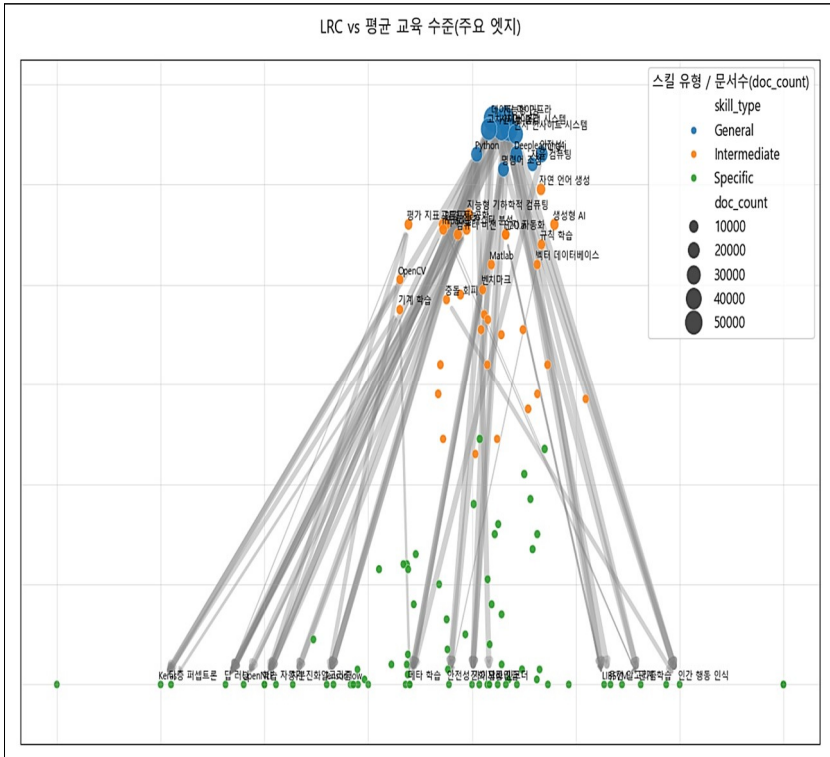
직 어디서든 가장 먼저 강화해야 하는 기초 플랫폼 숙련으로 볼 수 있다. 이들은 서로 긴밀히 연결되면서 동시에 주변의 여러 특화 영역과 교차하는 역할을 수행한다. 네트워크의 오른쪽 영역에는 컴퓨터 비전 및 이미지 인식과 관련된 숙련들이 모여 있다. 구체적으로 기계 비전, OpenCV, 컴퓨터 비전과 같은 숙련들이 이 구역을 차지하며, 중앙의 숙련들과 연결되면서도 자체적으로 응집된 클러스터를 형성한다. 이는 영상 처리 분야가 독립적이면서도 기반 기술에 강하게 의존하는 특성을 보여준다. 위쪽에는 자연 언어 생성, 자연 언어 이해, 벡터 데이터베이스 등 데이터 과학과 관련된 숙련들이 분포한다. 이들은 머신러닝과 Python 같은 허브 기술들과 연결되지만, 동시에 분석적·통계적 접근을 특징으로 하는 별도의 영역을 형성하는 것으로 보인다. 왼쪽 영역에는 생성형 AI와 관련된 숙련들이 뚜렷하게 자리한다. 생성형 AI가 특징 추출, 미세 조정, 가상 비서 숙련들과 연결되어 있는 것으로 보인다. 이는 생성형 AI가 기존 AI 숙련과 결합하며 새로운 클러스터를 형성하고 있음을 보여준다. 전체적으로 AI 숙련이 몇 개 영역으로 독립된 영역을 형성하면서도 중앙 핵심 숙련들과 통합되어 있는 구조를 보여주고 있다.

[그림 5-5]는 파급 중심성(LRC)을 세로축에, 평균 교육 수준을 가로축에 배치하여 AI 관련 숙련의 구조적 위계와 학력 요구 간의 관계를 보여준다. 또한, LRC를 기초로 숙련을 General-Intermediate-Specific으로 구분한 것도 표시하고 있다.

파급 중심성이 높은 숙련은 네트워크 구조에서 다수의 하위 기술로 이어지는 상위 개념적 허브로 작동하는데, 예컨대 Python, 데이터 마이닝, 사이버 물리 시스템, 인지 인사이트 시스템 등이 이에 해당한다. 이러한 숙련들은 공고 내에서 반드시 수십 개의 다른 기술과 함께 나열되지 않더라도, 다양한 특수 기술을 포괄하는 개념으로 기능하기 때문에 구조적으로 일반성이 높다. 동시에 이들 숙련은 가로축에서 상대적으로 오른쪽에 위치하는 경우가 많아, 대졸 이상의 고학력 직무와 밀접하게 연결된다는 점을 확인할 수 있다. 이는 곧 파급력이 큰 핵심 숙련일수록 채용시장에서 고학력 집단을 대상으로 한 직무의 기본 요구 조건으로 자리 잡고 있음을 시사한다.

중간 영역에 위치한 숙련들은 특정 응용 분야나 세부 도메인에 특화되어 있으며, 대표적으로 OpenCV, 기계 학습, Matlab, 자연 언어 생성 등이 포함

[그림 5-5] AI 숙련의 위계 구조(파급 중심성, LRC)와 교육 수준



자료 : 저자 작성.

된다. 이들은 고학력 직무에서 자주 등장하지만 동시에 전문 엔지니어링 포지션이나 프로젝트 기반 실무에서도 활용 가능하기 때문에, 평균 교육 수준이 중간 정도로 나타난다. 따라서 이러한 숙련은 응용·실무형 기술로서, 고학력보다는 실제 현장에서의 활용 경험이 더 큰 비중을 차지한다고 해석할 수 있다.

한편, 파급 중심성이 낮은 숙련들은 네트워크 구조상 말단에 해당하는 특수·세부 기술들로, Keras, TensorFlow, Perceptron, 특정 딥러닝 프레임워크, 강화학습 알고리즘, 안전 검증 기술 등이 여기에 포함된다. 이들은 구조적으로 다른 기술을 포괄하지 않으므로 파급력이 낮지만, 교육 수준에 따라 폭넓게 위치하고 있어서 저학력에 특화된 숙련도 있고, 대학원 이상의 학문적 배경이나 첨단 연구 직무와 밀접히 연결된 숙련도 있다.

이 그림에서 옛지는 단순한 연결선이 아니라 숙련 간 위계적 관계(nested skill hierarchy)를 의미한다. Hosseinioun et al.(2025)이 제시한 접근에 따르면, 특정 숙련이 다른 숙련의 기반이 되거나 더 상위·포괄적 숙련의 일부로 포함될 때 두 숙련 간에 종속 관계가 형성된다. 즉, 옛지는 두 숙련이 독립적으로 병렬적으로 존재하는 것이 아니라, 한 숙련이 다른 숙련의 전제 조건(precondition) 혹은 하위 구성 요소(component skill) 역할을 한다는 점을 보여준다. 예를 들어, “Python”이나 “데이터 분석”과 같은 범용 숙련은 여러 특수 숙련(예: “자연어 처리”, “컴퓨터 비전”)으로 연결될 수 있는데, 이는 해당 특수 숙련을 습득하고 활용하기 위해 먼저 기반이 되는 범용 숙련이 요구됨을 나타낸다. 그림에서 옛지가 위쪽(범용적이고 높은 LRC 값을 가지는 숙련)에서 아래쪽(보다 특화되고 낮은 LRC 값을 가지는 숙련)으로 뻗어 있는 구조는, 상위 숙련이 하위 숙련들을 포괄하고 있다는 위계성을 시각적으로 표현한다. 따라서 옛지는 숙련 네트워크에서 지식의 종속성과 전이가능성을 보여주는 중요한 구조적 요소로 해석할 수 있다. 요약하면, 옛지는 단순한 유사성이나 공출현(co-occurrence)이 아니라, “이 숙련을 가지면 다른 숙련으로 확장할 수 있다”라는 포괄적·계층적 관계를 나타낸다.

이러한 숙련의 네트워크 구조를 나타내는 그림은 몇 가지 시사점을 제공한다. 첫째, 파급 중심성이 높은 숙련들이 주로 고학력 직무에서 요구된다는 점은, 대학 및 대학원 교육과정에서 Python, 데이터 마이닝, 사이버 물리 시스템과 같은 핵심적이고 범용적인 숙련을 기초 커리큘럼의 중심으로 배치할 필요성을 보여준다. 이러한 숙련들은 다양한 하위 기술로 확산되는 허브 역할을 하므로, 교육과정 초기에 습득할 경우 이후의 전문화된 학습과 실무 적응력을 크게 높일 수 있다. 따라서 고등교육 단계에서는 구조적 허브 숙련을 기반으로 한 기초-응용-심화의 체계적 학습 경로를 설계하는 것이 중요하다.

둘째, 중간 수준의 파급 중심성을 가진 응용·실무형 숙련들(OpenCV, 기계 학습, Matlab 등)은 산업 현장에서 프로젝트 경험과 함께 실질적인 성과로 이어지는 경우가 많다. 이는 고학력 교육기관뿐만 아니라 직업훈련기관, 기업 내 재교육 프로그램에서도 충분히 다룰 수 있는 영역이며, 실습 중심의 단기 집중 과정과 연계될 필요가 있다. 따라서 이러한 숙련은 산학 협력

교육 프로그램, 전문대학 및 실무형 훈련 과정에서 비중 있게 다루어야 하며, 고학력 중심의 교육체계와 별도로 산업 현장 맞춤형 훈련 모델로 발전시킬 필요가 있다.

셋째, 파급 중심성이 낮지만 높은 교육 수준과 결합되는 특수 숙련들(Keras, TensorFlow, 강화학습, 특수 알고리즘 등)은 첨단 연구·개발 직무와 직결되며, 이는 고급 인력 양성 정책과 연구자 중심의 대학원 교육체계가 맡아야 할 영역이다. 이러한 숙련은 현장의 실무보다는 연구적 탐구와 신기술 개발과 연계되므로, 국가 차원의 AI 첨단 연구센터, 대학원 특화 과정, 박사급 인재 양성 프로그램을 통해 지원되어야 한다.

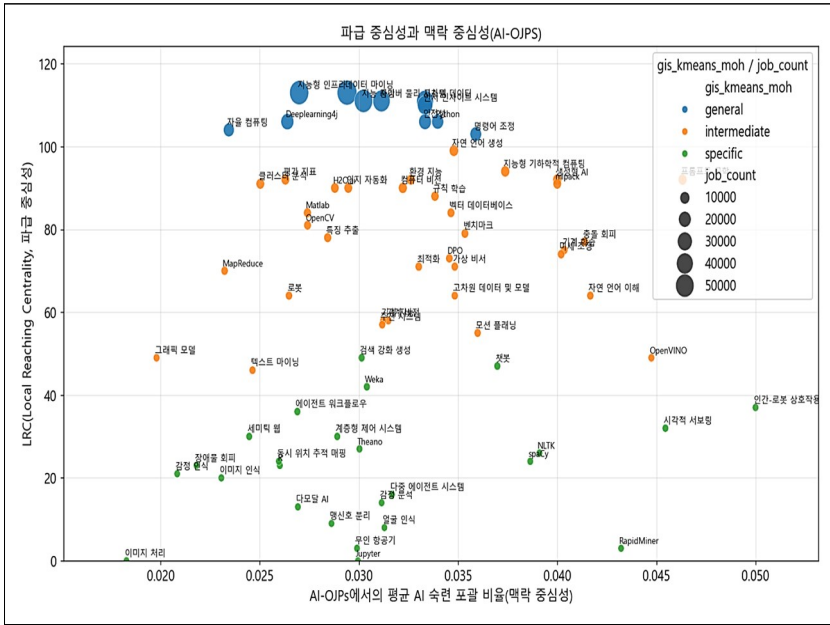
종합적으로 볼 때, 파급 중심성과 교육 수준 지표는 단순히 숙련의 수요를 파악하는 데 그치지 않고, 교육·훈련의 차별적 배치 전략을 설계하는 데 활용될 수 있다. 즉, 허브 성격을 가진 일반적 숙련은 기초 교육과정의 중심축으로, 응용·실무형 숙련은 산업 현장 맞춤형 훈련 과정으로, 특수 숙련은 첨단 연구개발 프로그램으로 각각 차별화하여 배치함으로써, 인적자원 양성과 산업 수요 간의 불일치를 줄일 수 있다. 이는 궁극적으로 AI 노동시장의 구조적 수요에 대응하면서 동시에 교육자원과 정책적 노력을 보다 효율적으로 배분할 수 있는 근거를 제공한다.

[그림 5-6]과 <표 5-8>은 파급 중심성과 맥락 중심성의 산점도를 보여주는 것이다. 먼저 위상단에는 명령어 조정, 자연 언어 생성, 지능형 기하학적 컴퓨팅, 생성형 AI 등이 분포한다. 이 구역은 두 지표가 모두 높은 숙련들로, 시장에서 핵심 허브이자 범용적 기반을 형성한다. 즉, 이들은 다양한 기술 맥락에서 반복적으로 요구되면서 동시에 다른 기술들과의 연결을 주도하는 중심적 역할을 한다.

좌상단에는 자율 컴퓨팅이나 DeepLearning4j와 같은 숙련이 위치한다. 이들은 네트워크 확산성(LRC)이 높아 다른 기술과의 연결성이 크지만, 채용 공고 내에서 동시에 언급되는 기술의 폭은 좁다. 이는 곧 허브적 성격을 가지면서도 특정한 기능에 집중된 기술임을 보여준다.

우하단에는 OpenVINO, spaCy, NLTK, RapidMiner, 시각적 서보링, 인간-로봇 상호작용과 같은 숙련들이 위치한다. 이들은 네트워크 확산력은 낮지만, 공고 내에서 다양한 AI 기술과 함께 등장하는 특성을 가진다. 다시 말해,

[그림 5-6] 파급 중심성과 맥락 중심성에 따른 속련 키워드의 분류



자료 : 저자 작성.

<표 5-8> 파급 중심성과 맥락 중심성에 따른 속련 분포와 그 의미 해석

구역	주요 속련 예시	특징	해석
좌상단 (LRC ↑, 맥락 중심성 ↓)	자율 컴퓨팅, Deeplearning4j	네트워크 확산 영향력은 크지만, 채용공고 속 동시 등장 속련 폭은 좁음	허브적 위치를 가지면서도 비교적 단일 기능 중심
우상단 (LRC ↑, 맥락 중심성 ↑)	명령어 조정, 자연 언어 생성, 지능형 기하학적 컴퓨팅, 생성형 AI	확산력과 맥락적 동반 등장 모두 높음	시장에서 가장 핵심적이며 범용성과 영향력을 동시에 지님
좌하단 (LRC ↓, 맥락 중심성 ↓)	감정 인식, 이미지 인식, 세미틱 웹, 다모달 AI, 무인 항공기	두 지표 모두 낮음	수요가 제한적이거나 아직 표준화되지 않은 니치 기술
우하단 (LRC ↓, 맥락 중심성 ↑)	OpenVINO, spaCy, NLTK, RapidMiner, 시각적 서보링, 인간-로봇 상호작용	확산 영향력은 낮지만 다양한 기술 묶음과 함께 언급됨	특화형 기술로, 포괄적 과업 속 보조적·전문적 위치

자료 : 저자 작성.

특정 프레임워크나 도메인에 귀속된 특화 기술로서, 포괄적 과업 묶음 속에서 보조적·전문적 역할을 수행한다고 볼 수 있다.

좌하단에는 감정 인식, 이미지 인식, 세미틱 웹, 다모달 AI, 무인 항공기 등이 자리한다. 이들은 파급성과 맥락적 동반 등장 모두 낮아, 아직 시장에서 수요가 제한적이거나 채용 표현이 표준화되지 않은 니치(niche) 기술 군으로 이해할 수 있다.

종합하면, 이 네 구역은 중심 숙련-허브 숙련-특화 보조 숙련-니치 숙련이라는 네 가지 유형으로 숙련을 구분할 수 있으며, 이는 노동시장 내에서 AI 관련 숙련의 역할과 전략적 가치가 어떻게 분화되는지를 보여줄 수 있다.

제5절 AI 숙련과 AI 역량의 연계와 정합성

대학 강좌 커리큘럼에서 제공되는 역량이 노동시장에서 요구되는 숙련과 어떠한 정합성을 가지는지를 분석하기 위해, 대학 강좌의 역량 키워드에 숙련 키워드를 매핑하였다. 이를 위해 먼저 숙련과 역량의 키워드를 GPT를 활용하여 영어 문장 텍스트로 변환한 후, OECD AI 키워드 정의문과 강좌 커리큘럼 설명 간의 의미적 유사성을 정량적으로 평가하였다. 매핑 방법은 앞에서 사용한 방법을 동일하게 적용하였다. BM25 기반 어휘적 유사도와 멀티링구얼 문장 임베딩을 활용한 의미적 유사도를 결합하고, 상위 후보군에 대해서는 Reciprocal Rank Fusion 기법을 적용하여 재현율을 제고하고, 크로스 인코더(cross-encoder) 기반 재랭킹과 자연어 추론(NLI) 모델을 통한 정합성 검증을 수행하여 정밀도를 강화하는 방법이다.

이렇게 AI 대학 강좌 역량 키워드에 숙련 키워드를 매핑한 결과를 활용하여 <표 5-9>와 같이, AI 숙련 대분류별로 온라인 구인광고(OJPs)에서 나타나는 빈도와, 해당 숙련이 실제 교육 강좌 내용 속에 포함된 정도를 비교할 수 있었다. 이를 통해 노동시장의 수요와 교육 공급 간의 정합성을 평가할 수 있다. AI OJP 중에서 학력 요구 조건이 대졸 이상인 경우만 따로 추출하여 분석한 결과도 포함하였다.

〈표 5-9〉 OJP와 대학 강좌에서의 숙련 빈도와 비율 비교

	빈도			비율		
	AI OJPs	AI OJPs (대졸 이상)	AI 강좌	AI OJPs	AI OJPs (대졸 이상)	AI 강좌
AI ML 모델과 알고리즘	76,663	19,051	544	(12.0)	(12.3)	(38.4)
AI 프레임워크 라이브러리	170,067	34,972	124	(26.6)	(22.6)	(8.7)
AI 시스템 공학 MLOps	30,546	7,651	85	(4.8)	(4.9)	(6.0)
데이터 과학 분석	144,066	40,949	253	(22.5)	(26.4)	(17.8)
자연어 처리	15,006	3,780	88	(2.3)	(2.4)	(6.2)
컴퓨터 비전 이미지 인식	27,874	8,958	108	(4.4)	(5.8)	(7.6)
자율 시스템	95,311	22,341	97	(14.9)	(14.4)	(6.8)
로봇공학 자동화 제어	3,918	1,232	38	(0.6)	(0.8)	(2.7)
책임 있는 AI	75,434	15,919	81	(11.8)	(10.3)	(5.7)
전 체	638,885	154,853	1,418	(100.0)	(100.0)	(100.0)

자료 : 저자 작성.

전체적으로 보았을 때, 두 영역에서 특정 숙련군이 차지하는 비중이 일치하지 않는 경우가 다수 확인되었다. 예컨대, OJPs에서는 AI 프레임워크 라이브러리(26.6%)와 데이터 과학 분석(22.5%)의 비중이 매우 높게 나타난 반면, 대학 강좌에서는 각각 8.7%, 17.8%에 그쳐 상대적으로 낮은 수준을 보였다. 반대로, AI ML 모델과 알고리즘은 OJPs에서 12.0%에 불과하지만 대학 강좌에서는 38.4%로 가장 높은 비중을 차지하여, 교육과 산업 수요 간 불일치가 뚜렷하게 나타난다. 또한, 책임 있는 AI의 경우 노동시장에서의 요구 비중은 11.8%로 비교적 높은 편이지만, 대학 강좌에서는 5.7%에 머물러 교육 현장에서의 반영이 부족함을 시사한다. 이와 대조적으로 자연어 처리나 컴퓨터 비전·이미지 인식은 두 영역에서 모두 낮은 수준으로 나타나, 해당 분야의 교육·수요 간 균형은 유지되고 있으나 절대적 비중은 크지 않음을

보여준다.

이러한 결과는 대학 강좌가 여전히 전통적·기초적 AI 모델 및 알고리즘 교육에 집중하는 반면, 산업 현장은 프레임워크 활용 능력이나 데이터 과학적 분석 능력에 더 큰 비중을 두고 있음을 시사한다. 따라서 교육과 노동시장의 적합성을 강화하기 위해서는, 대학 교육 과정에서 산업 수요가 높은 프레임워크·데이터 분석 역량을 보다 적극적으로 반영할 필요가 있음을 알 수 있다.

이러한 불일치는 정책적 함의를 지닌다. 첫째, 대학 교육과정은 여전히 기초적·전통적 AI 모델 교육에 치중하고 있으나, 산업 현장은 프레임워크 활용, 데이터 과학 분석, 책임 있는 AI와 같은 실질적이고 응용적인 역량을 더 중시한다는 점에서, 교육과정 개편이 필요하다. 둘째, 교육 공급이 산업 수요와 괴리되는 상황은 인력 미스매치를 심화시킬 수 있으며, 이는 AI 인재의 고용 가능성과 기업의 숙련 확보를 동시에 제약한다. 따라서 대학은 산업계와의 협력 체계를 강화하여 커리큘럼 설계 단계에서부터 노동시장 수요를 반영할 필요가 있다. 셋째, 책임 있는 AI와 같은 신흥 영역에 대한 교육적 투입을 확대하여 사회적·윤리적 요구에 대응하는 전문 인력을 양성해야 한다.

〈표 5-10〉은 각 네트워크 지표별로 상위 10개 숙련 키워드가 실제 AI 관련 교과 강좌에 어느 정도 반영되어 있는지를 보여준다. 전체 강좌 수는 477개이며, 이 가운데 상위 숙련이 포함된 강좌의 수와 비율을 지표별로 계산한 것이다.

분석 결과, job_count, degree centrality, eigenvector, G_score, cs지표에서 공통적으로 약 230개 강좌(전체의 약 48.2%)가 상위 숙련을 포함하고 있었다. 즉, 이들 지표에서는 두드러지는 숙련들이 실제 교육 과정 속에서도 비교적 높은 빈도로 다뤄지고 있음을 알 수 있다. 대표적으로 데이터 마이닝, 지능형 인프라, Python, 딥러닝 등 핵심 AI 기술들은 다양한 지표에서 반복적으로 등장하며, 실제 강좌에 일정 수준 반영되고 있다.

반면, gen_mean_avg에서는 상위 숙련이 반영된 강좌가 160개(33.5%)에 불과했다. gen_mean_avg는 숙련이 등장하는 맥락의 평균적 중요도를 반영하는 지표로, 특정 핵심 기술이 아니라 다양한 맥락에서 꾸준히 반복적으로

〈표 5-10〉 AI 네트워크 지표별 상위 10개 숙련 키워드의 AI 관련 강좌 반영 비율

지표	전체 강좌 수	포함 강좌 수	비율 (%)	상위 10개 숙련(역량에 매칭된 숙련)
job_count	477	230	48.2	데이터 마이닝, 지능형 인프라, 지능 증강, 사이버 물리 시스템, 고차원 데이터, 인지 인사이트 시스템, Deeplearning4j, 안전성, Python, 명령어 조정
degree centrality	477	230	48.2	지능형 인프라, 데이터 마이닝, Python, 사이버 물리 시스템, 지능 증강, 명령어 조정, 인지 인사이트 시스템, mlpack, 클러스터 분석, 고차원 데이터
eigenvector	477	230	48.2	지능형 인프라, 지능 증강, 사이버 물리 시스템, 데이터 마이닝, 안전성, Python, 명령어 조정, 인지 인사이트 시스템, mlpack, Deeplearning4j
gen_mean_avg	477	160	33.5	인간-로봇 상호작용, 프롭프트 공학, 시각적 서보링, OpenVINO, 패턴 인식, 딥 러닝, 진화 알고리즘, RapidMiner, LLM(대규모 언어 모델), 자연 언어 이해
lrc	477	221	46.3	지능형 인프라, 데이터 마이닝, 고차원 데이터, 사이버 물리 시스템, 지능 증강, 인지 인사이트 시스템, Python, 안전성, Deeplearning4j, 자율 컴퓨팅
cs	477	230	48.2	지능형 인프라, 데이터 마이닝, 사이버 물리 시스템, 인지 인사이트 시스템, 지능 증강, Python, 고차원 데이터, 명령어 조정, 안전성, Deeplearning4j
g_score	477	230	48.2	데이터 마이닝, 지능형 인프라, 지능 증강, 사이버 물리 시스템, 고차원 데이터, 인지 인사이트 시스템, Deeplearning4j, 안전성, Python, 명령어 조정

자료 : 저자 작성.

활용되는 숙련을 포착한다. 따라서 교육 과정 커리큘럼에는 상대적으로 덜 포함되거나, 강좌 기술 설명에서는 다소 간접적으로 다루어져 낮은 수치를 보이는 것으로 해석할 수 있다.

주목할 점은, 어떤 지표에서도 50%를 초과하여 반영된 경우가 없다는 사실이다. 즉, 상위 10개 숙련조차 절반 이상의 강좌에 포함되지 않았으며, 이는 대학 교육 과정과 실제 노동시장에서 요구되는 숙련 간 정합성이 제한적임을 의미한다. 산업 수요 측면에서 핵심적인 위치를 차지하는 숙련들이 교육 현장에서는 여전히 부분적으로만 반영되고 있는 셈이다.

따라서 이 결과는 단순히 특정 지표의 차이를 보여주는 데 그치지 않고, AI 관련 교육과정 전반이 노동시장의 핵심 수요를 충분히 수용하지 못하고 있음을 시사한다. 특히 맥락적 중요성이나 숙련 간 묶음 구조를 반영하는 지표에서 낮은 수치를 기록했다는 점은, 교육과정이 실제 산업 환경에서 요구되는 복합적·통합적 숙련 구조를 충분히 반영하지 못하고 있음을 보여주며, 향후 커리큘럼 개선과 산학 연계 강화의 필요성을 뒷받침한다.

〈표 5-11〉은 총 477개의 강좌에서 인공지능(AI) 관련 숙련 키워드가 얼마나 빈번하게 포함되는지를 분석한 결과로, 상위 20개 키워드가 일반 강좌와 K-MOOC 강좌로 구분되어 제시되어 있다.

분석 결과를 보면, 전체 강좌 기준으로는 ‘데이터 마이닝(210회)’, ‘기계 학습(173회)’, ‘딥 러닝(128회)’이 가장 많이 등장하여 핵심적인 AI 분석 및 모델링 기법이 교육 현장에서 광범위하게 다뤄지고 있음을 보여준다. 이어서 ‘인공 신경망(86회)’, ‘딥 벨리프 네트워크(56회)’, ‘강화 학습(45회)’ 등도 상위권에 포함되어, 기계 학습과 딥 러닝의 세부적 알고리즘과 응용 기술들이 교육 과정 속에서 중요한 위치를 차지하고 있음을 확인할 수 있다. 또한 ‘인간 인식 인공지능(45회)’, ‘설명 가능한 AI(44회)’, ‘인지 자동화(44회)’, ‘그래프 신경 네트워크(43회)’ 등은 AI 활용의 응용적·설명 가능성 측면을 강조하는 흐름을 반영한다.

이를 일반 강좌와 K-MOOC 강좌로 나누어 살펴보면, 일반 강좌에서는 ‘기계 학습(96회)’과 ‘딥 러닝(86회)’이 가장 많이 등장하여 학문적·연구적 관점에서의 AI 모델링 기법 학습이 주를 이루고 있다. 반면, K-MOOC 강좌에서는 ‘데이터 마이닝(130회)’이 압도적으로 많으며, 이어서 ‘기계 학습(77회)’, ‘딥 러닝(42회)’, ‘지능형 인프라(32회)’ 등이 뒤를 잇고 있다. 이는 K-MOOC 강좌가 상대적으로 실용적 데이터 활용 능력과 산업 현장 적용에 더 큰 비중을 두고 설계되어 있음을 시사한다.

〈표 5-11〉 AI 교육 강좌에서 상위 20개 숙련(역량에 매칭된 숙련)

	전체 강좌		일반 강좌		KMOOC 강좌	
1	데이터 마이닝	210	기계 학습	96	데이터 마이닝	130
2	기계 학습	173	딥 러닝	86	기계 학습	77
3	딥 러닝	128	데이터 마이닝	80	딥 러닝	42
4	인공 신경망	86	인공 신경망	62	지능형 인프라	32
5	딥 벨리프 네트워크	56	딥 벨리프 네트워크	45	인간 인식 인공지능	28
6	강화 학습	45	그래프 신경 네트워크	34	그래픽 모델	27
7	인간 인식 인공지능	45	Python	26	인공 신경망	24
8	설명 가능한 AI	44	설명 가능한 AI	26	강화 학습	21
9	인지 자동화	44	인지 자동화	26	가우시안 프로세스	18
10	그래프 신경 네트워크	43	최적화	25	설명 가능한 AI	18
11	Python	42	강화 학습	24	인지 자동화	18
12	그래픽 모델	42	기계 지능	23	로봇	17
13	자연 언어 처리	38	자연 언어 처리	23	Python	16
14	지능형 인프라	38	안전성 가이드라인	21	자연 언어 처리	15
15	최적화	37	인간 인식 인공지능	17	텍스트 마이닝	15
16	안전성 가이드라인	35	그래픽 모델	15	안전성 가이드라인	14
17	기계 지능	33	텍스트 마이닝	15	자율 주행 차량	14
18	가우시안 프로세스	32	가우시안 프로세스	14	명령어 조정	13
19	텍스트 마이닝	30	비지도 학습	14	생성적 적대적 네트워크	12
20	생성형 AI	23	안전성	14	생성형 AI	12

자료 : 저자 작성.

또한, K-MOOC에서 ‘로봇(17회)’, ‘자율 주행 차량(14회)’, ‘명령어 조정(13회)’ 등 물리적·공학적 응용 분야가 상위 20위 안에 포함된 반면, 일반 강좌에서는 ‘Python(26회)’, ‘최적화(25회)’, ‘텍스트 마이닝(15회)’ 등 도구적·분석적 숙련이 보다 강조된 차이를 보인다.

종합적으로 보면, 전체 강좌에서는 데이터 기반 분석과 기계 학습·딥 러닝 중심의 기술이 가장 두드러지게 교육되고 있으며, 일반 강좌는 상대적으로

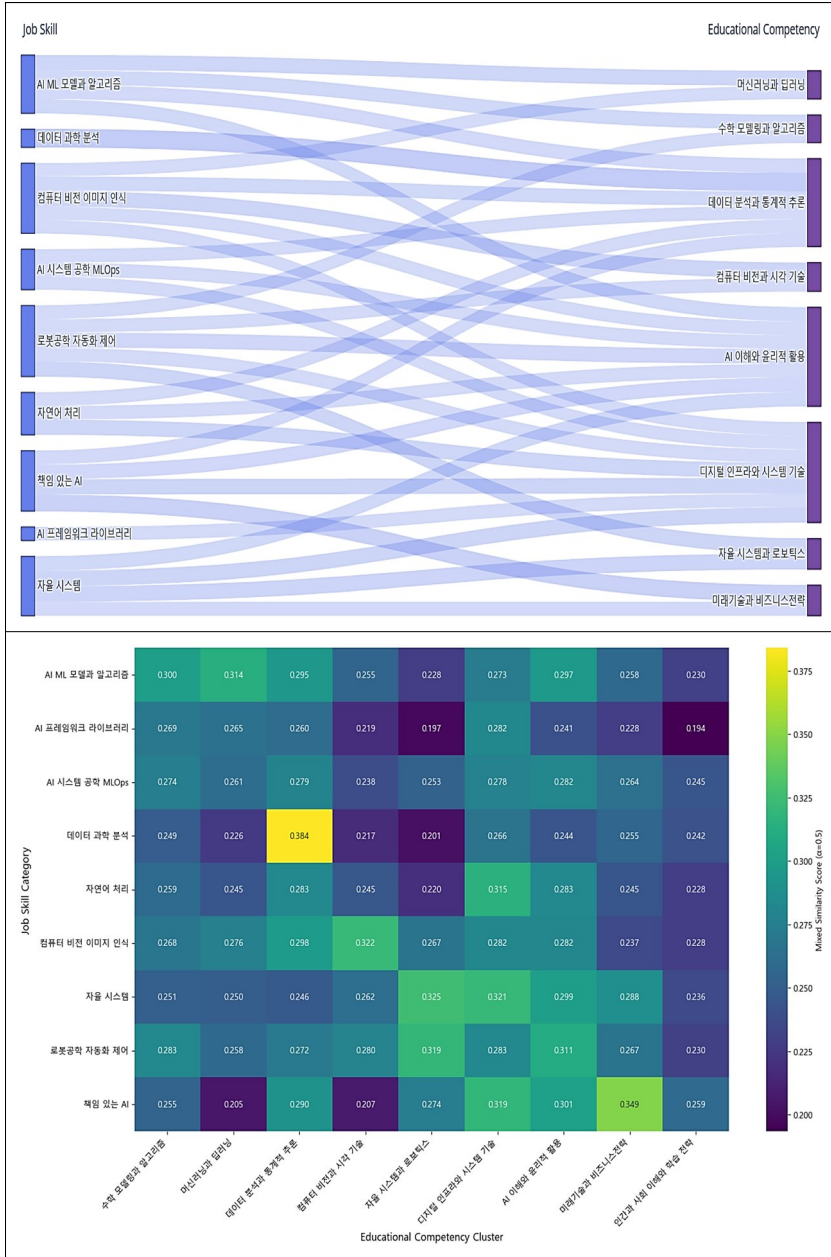
로 이론적·분석적 기초에 집중하는 반면, K-MOOC는 산업적 응용과 실용성을 강조하는 특징을 보인다. 이는 국내 AI 교육이 한편으로는 학문적 심화를 추구하면서도 다른 한편으로는 현장 수요에 대응하는 실용적 과정을 강화하고 있다는 점에서 상호보완적 구조를 형성하고 있음을 보여준다.

다음으로, 한국의 AI 대학 교육 역량이 AI 직업의 AI 숙련에 어떻게 대응하고 있는지를 보기 위해, 역량과 숙련의 유사도 분석을 하였다.

본 연구에서는 AI 직업 숙련과 AI 교육 강좌 역량 간의 연계성을 파악하기 위하여 숙련과 역량에 대한 영어 설명 텍스트를 기반으로 한 임베딩 유사도 분석을 수행하였다. 각 직업 숙련 및 교육 역량 키워드에 대해 부여된 영어 설명 문장을 임베딩 모델(Sentence-BERT(SBERT) 계열의 sentence-transformers/all-MiniLM-L6-v2)을 통해 의미론적 벡터 공간에 매핑하였다. 이를 통해 키워드 단어 수준의 단순 비교가 아닌, 해당 개념이 지닌 맥락적 의미와 개념적 함의를 반영한 유사도 측정이 가능하도록 하였다. 이후 직업별 스킬 설명과 교육과정 설명 간의 임베딩 유사도를 계산하고, 숙련과 역량 대분류 수준으로 집계하여 평균 유사도 지표를 도출하였다. 최종 산출된 지표를 활용하여 직업 숙련과 교육과정 역량 간의 구조적 연계성을 시각화하는 층적 다이어그램(Alluvial Diagram)과 히트맵(Heatmap)으로 시각화하였다. 히트맵에는 유사도 점수가 표시되어 있다.

[그림 5-7]에서 제시된 Job Skill과 Educational Competency 간의 유사도 분석 결과는 교육과 산업 현장 간의 정합성을 평가하는 중요한 근거를 제공한다. 전반적으로, [그림 5-7]에 나타난 직업 숙련과 교육 역량 간의 유사도 값은 대체로 0.20대 후반에 분포하였으며, 일부 매칭에서만 0.30을 상회하는 경향을 보였다. 이러한 결과는 교육체계가 산업 수요를 포괄적으로 반영하고 있음에도 불구하고, 세부 수준에서의 구체적인 대응은 아직 미흡하다는 점을 시사한다. 이는 산업 현장에서 요구되는 숙련과 교육과정에서 제공되는 역량 간의 직접적인 연계가 충분히 이루어지지 못하고 있음을 의미한다. 그러나 특정 영역에서는 상대적으로 높은 수준의 정합성이 관찰되었다. 이는 해당 교육과정이 산업적 요구를 비교적 잘 반영하고 있음을 나타낸다. ‘데이터 과학 분석’ 직업 숙련과 ‘데이터 분석과 통계적 추론’ 교육 역량 간의 유사도는 0.384로 가장 높은 수준의 정합성을 나타냈다. ‘책임 있는 AI’

[그림 5-7] AI 직업 숙련과 AI 강좌 역량 간의 총적 다이어그램(Alluvial Diagram)과 히트맵(Heatmap)



자료 : 저자 작성.

직업 숙련은 '미래기술과 비즈니스 전략' 교육 역량과 0.349의 유사도를 보였으며 이는 AI 윤리 및 전략 분야에 대한 산업적 중요성이 교육 과정에 어느 정도 반영되고 있음을 시사한다. '자율 시스템' 직업 숙련은 '자율 시스템과 로보틱스' 교육 역량과 0.325의 유사도를, '컴퓨터 비전 이미지 인식' 직업 숙련은 '컴퓨터 비전과 시각 기술' 교육 역량과 0.322의 유사도를 보여 높은 연관성을 확인시켰다. 이 외에도 '자율 시스템' 직업 숙련과 '디지털 인프라와 시스템 기술' 교육 역량(0.321), '로봇공학 자동화 제어' 직업 숙련과 '자율 시스템과 로보틱스' 교육 역량(0.319), 'AI ML 모델과 알고리즘' 직업 숙련과 '머신러닝과 딥러닝' 교육 역량(0.314) 등의 연결에서도 비교적 높은 유사도가 확인되었다.

반면, 다수의 다른 영역에서는 유사도가 0.20대 초반에서 중반에 머무르며 낮은 정합성을 보였다. 특히, 'AI 프레임워크 라이브러리' 직업 숙련의 경우, '자율 시스템과 로보틱스' 교육 역량(0.197) 및 '인간과 사회 이해와 학습 전략' 교육 역량(0.194)과의 유사도가 0.20 미만으로 나타나, 현장에서 중요하게 다뤄지는 숙련이 교육과정 내에서 충분히 대응되지 못하고 있음을 보여준다. 또한, '자연어 처리' 직업 숙련은 '자율 시스템과 로보틱스' 교육 역량과 0.220의 유사도를 보이는 등 전반적으로 낮은 값을 나타냈으며, '책임 있는 AI' 직업 숙련이 '머신러닝과 딥러닝'(0.205) 및 '컴퓨터 비전과 시각 기술'(0.207) 교육 역량과 낮은 유사도를 보인 점은 해당 교육 역량들이 AI 윤리 분야의 실질적인 요구를 포괄하지 못하고 있음을 시사한다.

대학 교과 역량 중심으로 볼 때, '데이터 분석과 통계적 추론', 'AI 이해와 윤리적 활용', '디지털 인프라와 시스템 기술' 등의 역량은 여러 AI 시장 숙련들과 상대적으로 높은 연결 강도를 보인다. 이는 해당 교육 역량들이 폭 넓은 산업 분야에서 활용 가능성이 높음을 의미한다. 반대로 '머신러닝과 딥러닝', '수학 모델링과 알고리즘', '컴퓨터 비전과 시각 기술', '자율 시스템과 로보틱스' 등의 역량은 특정 직업 숙련과 높은 유사도를 보임에도 불구하고, 다른 시장 숙련들과의 전반적인 연결 강도는 상대적으로 약하게 나타났다.

다음으로 앞에서 문제 제기했듯이, 직업 숙련-교육 역량의 정합성을 판단할 때, 단순히 유사도 수준의 정합성이 아니라 숙련의 네트워크 구조(네트워크

크 중심성과 숙련의 위계 구조)를 반영하는 네트워크 중심성 가중 정합성을 평가해보고자 한다. 앞에서 도출한 각 숙련의 중심성과 위계 점수를 정규화·결합하여 노드 가중치를 정의하고, 스킬-교육 설명문 임베딩 유사도를 직업×교과 대분류 수준으로 집계할 때 가중 평균을 적용하였다. 이렇게 하면, 네트워크상 전략적으로 핵심적인 숙련을 포함하는 교육과정 연결이 더 높은 정합성으로 평가되며, 단순 유사도 평균 대비 구조적 중요도를 반영한 차별적 평가가 가능하다.

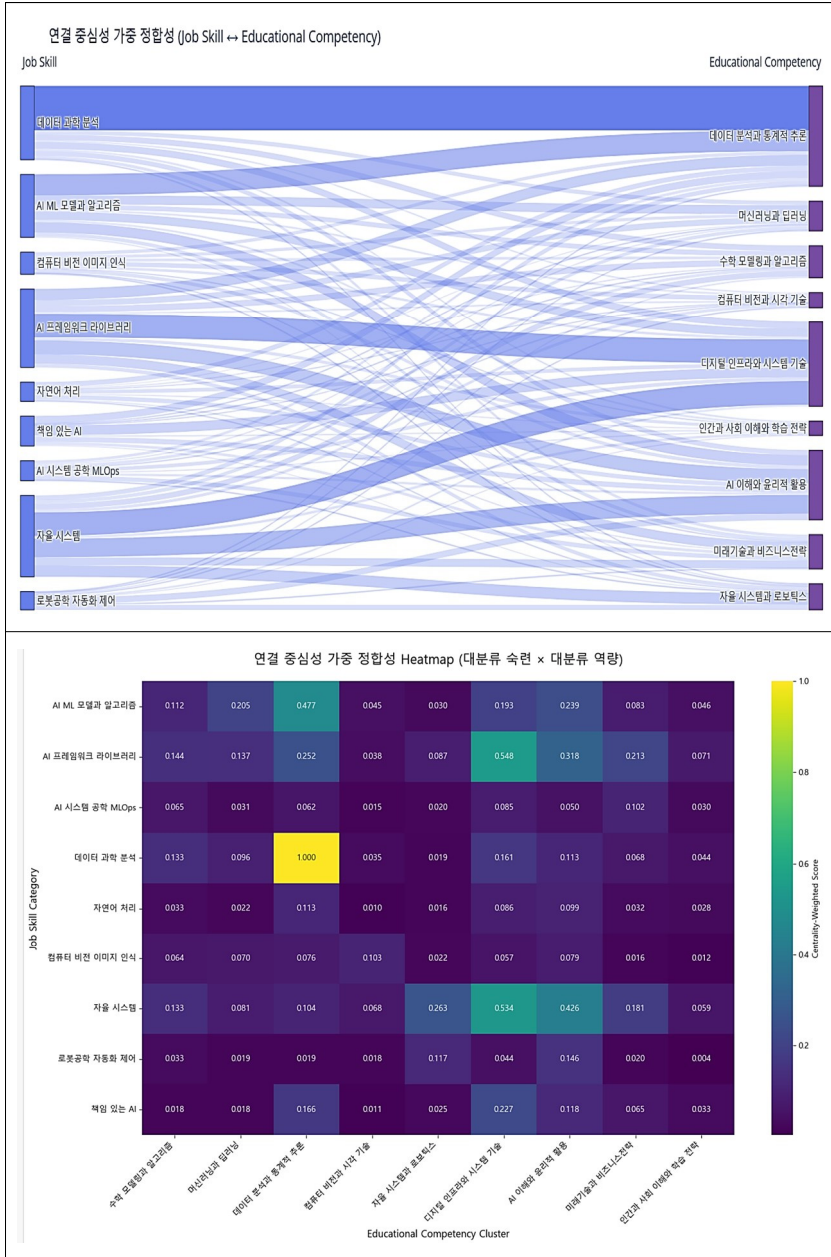
[그림 5-8]은 연결 중심성 가중 정합성을 보여주는 층적 다이어그램과 히트맵이다. 먼저, '데이터 과학 분석'은 가장 두터운 연결을 형성하고 있다. 특히 '데이터 분석과 통계적 추론'과 매우 강하게 연결되어 있으며, 동시에 '디지털 인프라와 시스템 기술', '머신러닝과 딥러닝', '수학 모델링과 알고리즘' 등과 같은 교과 역량과도 폭넓게 연결되어 있다. 이는 '데이터 과학 분석'이 산업 전반에서 중심성이 높은 숙련일 뿐 아니라, 교육과정에서도 기초 통계·수학적 기반부터 AI 응용까지 연속적으로 다뤄지고 있음을 보여준다. 또한, 대학의 '데이터 분석과 통계적 추론' 역량도 '데이터 과학 분석' 뿐만 아니라 'AI ML 모델과 알고리즘'이나 'AI 프레임워크 라이브러리' 숙련 등과도 강한 연결을 보여주고 있다.

'AI ML 모델과 알고리즘' 역시 '데이터 분석과 통계적 추론', '머신러닝과 딥러닝', '수학 모델링과 알고리즘' 역량과 굵직한 연결을 보인다. 이는 기계 학습 모델링이 교육과정에서 이론적 토대와 응용 기법에 걸쳐 반영되어 있다는 점을 시사한다.

'컴퓨터 비전 이미지 인식'이나 '자연어 처리' 등 응용 영역의 숙련은 대학의 다양한 AI 역량과 넓게 연결되어 있는 것으로 보인다. 산업 현장에서 특수 응용 영역은 대학 AI 강좌 역량과 강하게 연결되지 못하고 있음을 보여준다.

연결 중심성 가중 정합성은 연결 중심성이 높은 숙련일수록 굵고 다층적인 연결을 형성한다는 특징이 있다. 예컨대 '데이터 과학 분석'은 '데이터 분석과 통계적 추론'과 가장 강력한 연결을 형성하는 동시에, '디지털 인프라와 시스템 기술', '머신러닝과 딥러닝', '수학 모델링과 알고리즘' 등과도 폭넓게 이어졌다. 이는 단순히 과목 설명과의 유사성 때문이 아니라, 해당 숙련이 산업 전체에서 허브 역할을 하고 있다는 사실이 반영된 결과다. 또한,

[그림 5-8] AI 직업 숙련과 AI 강좌 역량 간의 총적 다이어그램과 히트맵 : 연결 중심성 가중 적합성



자료 : 저자 작성.

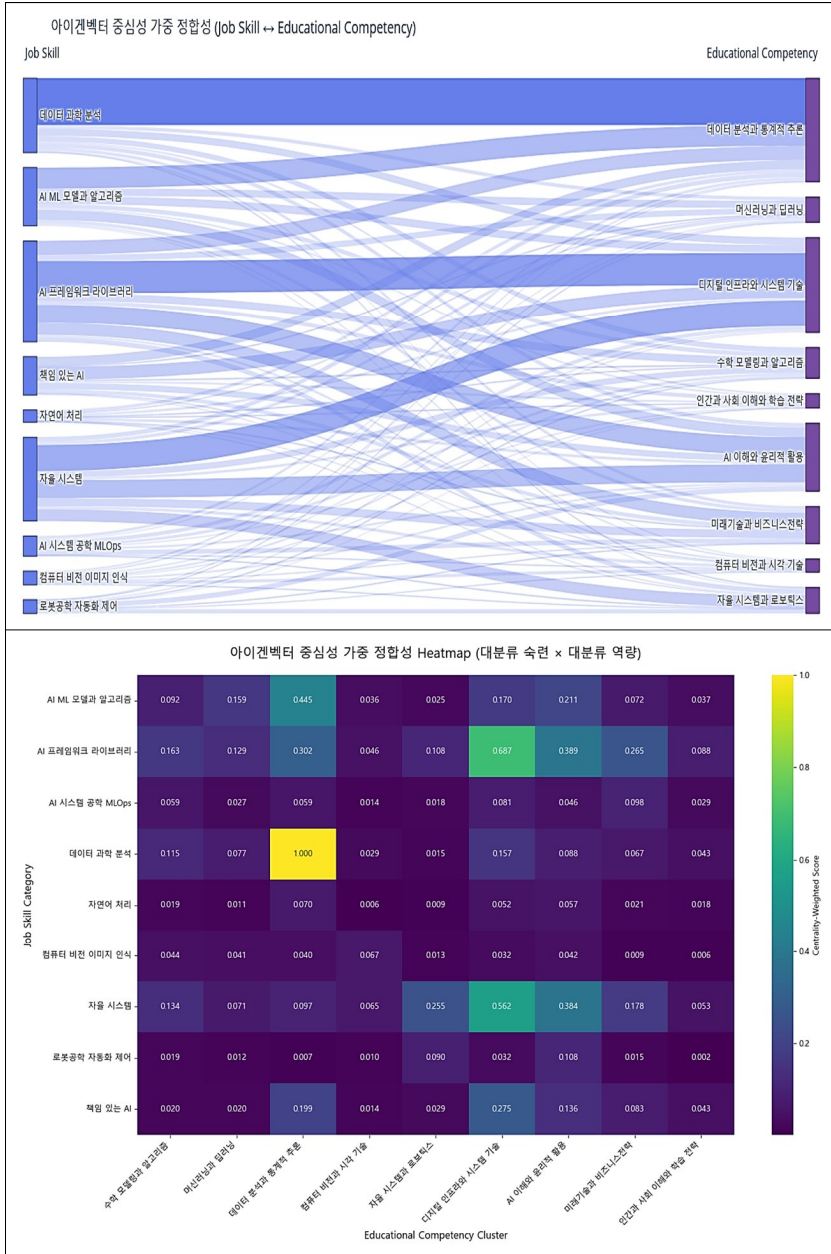
대학의 '데이터 분석과 통계적 추론' 역량이 '데이터 과학 분석'뿐 아니라 'AI ML 모델과 알고리즘', 'AI 프레임워크 라이브러리' 등 다른 핵심 숙련과도 강한 연결을 형성한 점은, 중심성 가중 정합성이 교육과 산업 간 전략적 결합 정도를 포착하는 데 유리함을 보여준다.

두 방식의 차이는 시사점에서도 뚜렷하게 드러난다. 코사인 유사도 기반 정합성에서 높은 점수를 기록했지만 중심성 가중 정합성에서는 낮게 평가되는 경우는, 특정 틈새 기술이나 전문화된 교과가 직무 설명과는 언어적으로 잘 맞으나 산업 생태계에서의 전략적 영향력은 제한적임을 의미한다. 예컨대, '컴퓨터 비전 이미지 인식'은 교과의 '컴퓨터 비전과 시각 기술'과 직접적으로 대응되어 유사도 정합성에서는 높은 점수를 보였으나, 산업 네트워크에서 다른 기술과의 연결성이 상대적으로 적어 중심성 가중 정합성에서는 낮은 수준으로 평가되었다. 유사한 사례로 '자연어 처리'도 '데이터 분석과 통계적 추론' 교과와는 어휘적으로 강하게 매칭되었으나, 전체 산업 네트워크에서의 연결 폭은 제한적이어서 가중 정합성에서는 두터운 연결로 나타나지 않았다. 이는 언어적 설명의 대응은 분명하지만 산업적 파급력은 상대적으로 좁음을 보여준다.

반대로, 중심성 가중 정합성에서 높은 연결을 보이는 경우는 산업 전반에서 파급력이 크고 응용 분야가 넓은 숙련이지만, 직무 설명과 교과의 텍스트 표현에서는 직접적으로 드러나지 않을 수 있다. 대표적으로 '자율 시스템'은 자율주행, 로봇틱스, 제어 기술 등과 긴밀히 연계되며 산업 네트워크 내에서 전략적 허브로 기능하기 때문에, 가중 정합성에서는 여러 교과 역량(예: '자율 시스템과 로봇틱스', '디지털 인프라와 시스템 기술')과 두터운 연결을 형성하였다. 그러나 코사인 유사도 기반에서는 언어적 설명의 제약으로 인해 일부 과목과만 밀접하게 연결되었다. 이와 유사하게 '데이터 과학 분석'은 '데이터 분석과 통계적 추론'과는 물론 '디지털 인프라와 시스템 기술', '머신러닝과 딥러닝' 등 다수의 교과와 굵은 연결을 형성하여 중심성 가중 정합성에서 가장 강한 영향력을 나타냈지만, 유사도 기반에서는 주로 통계·분석 교과와만 국한된 강한 연결이 포착되었다.

종합하면, 유사도 기반 정합성은 직무 설명과 교과 간의 텍스트적 대응을 중심으로, 중심성 가중 정합성은 산업 네트워크 내에서의 전략적 중요성을

[그림 5-9] AI 직업 속련과 AI 강좌 역량 간의 총적 다이어그램과 히트맵 : 아이젠벡터 중심성 가중 정합성



자료 : 저자 작성.

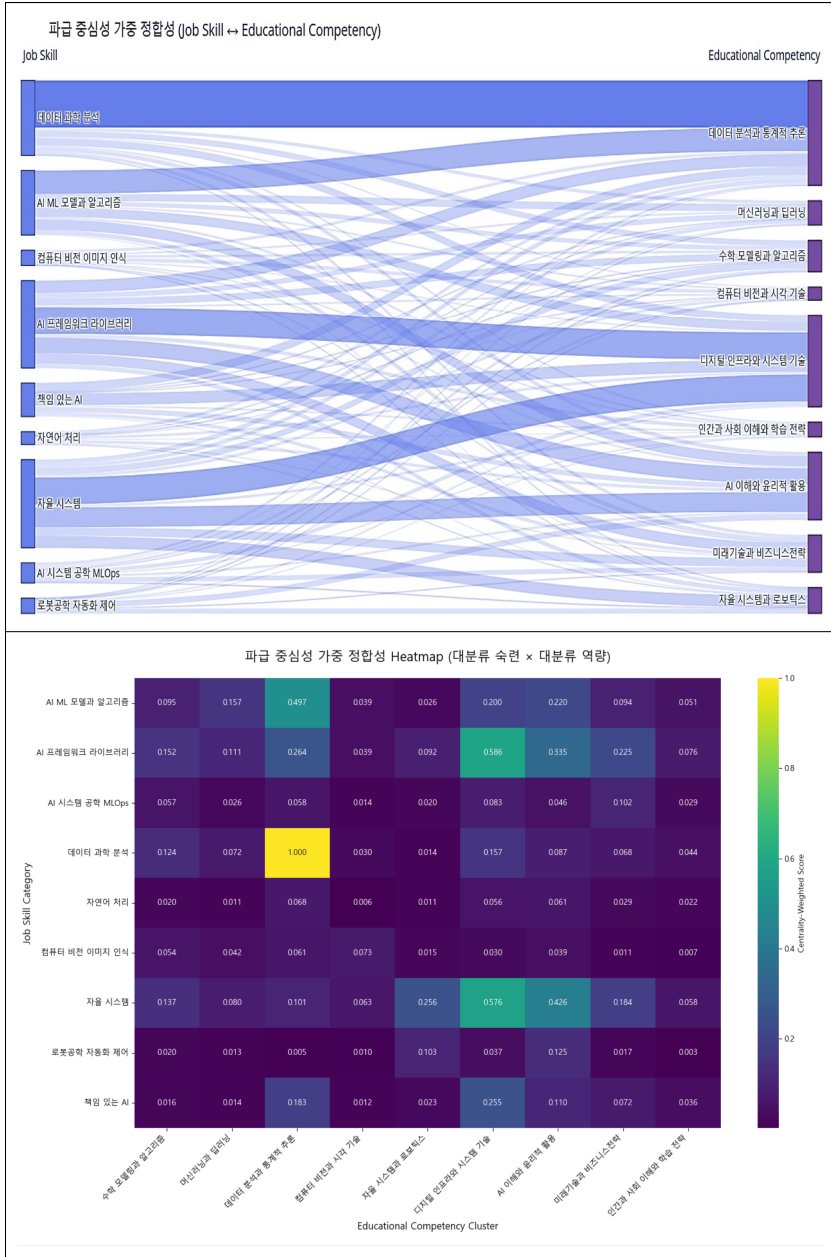
중심으로 측정된다는 점에서 차이가 있다. 따라서 교육과정-산업 간 정합성을 평가할 때는 두 지표를 병행하여, 표면적 유사성과 구조적 중요성을 함께 고려하는 것이 필요하다.

아이겐벡터 중심성(Eigenvector Centrality)은 단순 연결 수나 지역적 영향력을 넘어, “중요한 노드와 연결된 정도”를 반영하는 지표이다. 즉, 단순히 많이 연결된 숙련이 아니라, 이미 네트워크 내에서 중요성이 높은 숙련과 얼마나 강하게 연결되어 있는지에 따라 중심성이 결정된다.

아이겐벡터 중심성(EVC) 가중 정합성 결과는 연결 중심성 결과와 유사해 보이지만, 핵심 역량과의 연결 구조를 더 강조하는 것으로 보인다. ‘데이터 과학 분석’은 여러 교과 역량과 광범위하게 연결되어 있을 뿐 아니라, 특히 ‘데이터 분석·통계적 추론’, ‘머신러닝과 딥러닝’, ‘디지털 인프라와 시스템 기술’ 등 네트워크 내에서 중심성이 높은 역량들과 동시에 강하게 결합되어 있다. 이는 해당 숙련이 단순한 기술 요소를 넘어, AI 교육·훈련 체계의 중심 축을 구성하는 핵심 허브 역량들과 직접적으로 연결된 것으로 볼 수 있다. 다음으로 ‘AI 프레임워크·라이브러리’ 역시 ‘디지털 인프라와 시스템 기술’, ‘AI 이해와 윤리적 활용’, ‘미래기술과 비즈니스 전략’ 등과 밀접하게 연결되어 있다. 이는 프레임워크 역량이 단지 도구 활용 수준이 아니라, 인프라·윤리·전략 등 영향력 있는 교육 역량들과 함께 AI 역량 체계의 구조적 중심을 매개하는 역할을 하고 있음을 시사한다. 또한 ‘자율시스템’은 ‘디지털 인프라와 시스템 기술’ 및 ‘AI 이해와 윤리적 활용’과 굵은 흐름으로 연결되어 나타나며, 이는 자율시스템 관련 숙련이 특정 응용 영역에만 국한된 것이 아니라 AI 인프라와 책임 있는 활용이라는 핵심 축과 결합된 형태로 네트워크 중심부와 연동되어 있음을 보여준다. 마지막으로 ‘책임 있는 AI’는 연결 규모 자체가 가장 크지는 않지만, ‘AI 이해와 윤리적 활용’ 및 ‘디지털 인프라와 시스템 기술’과 같은 핵심 역량과 직접 연결되어 있다. 이는 EVC 관점에서 ‘책임 있는 AI’가 주변부 노드라기보다 중심 역량과의 결합을 통해 구조적 중요성이 강화되는 유형에 해당함을 의미한다. 다시 말해 연결 수가 많지 않더라도, 영향력 있는 역량과 함께 나타날 때 네트워크 내 전략적 위치가 높아질 수 있음을 보여주는 사례다.

반면, 자연어 처리, 컴퓨터 비전, 로봇공학 제어와 같은 특수(specific) 숙

[그림 5-10] AI 직업 숙련과 AI 강좌 역량 간의 총적 다이어그램과 히트맵 : 파급 중심성 가장 정합성



자료 : 저자 작성.

련은 비교적 낮은 빈도를 보였는데, 이는 특정 산업과 직무에서 선택적으로 강조되는 특화 숙련임을 의미한다. ‘책임 있는 AI’가 상위 키워드로 부상한 것은 AI 활용에서 윤리·안전성·사회적 책임의 중요성이 강화되고 있음을 시사한다. 종합하면, AI 채용 시장은 범용적 기술 숙련 중심으로 형성되고 있고 특수 응용 숙련은 상대적으로 수요가 강하지 않은 것으로, 그리고 윤리적 활용 능력 등에 대한 수요도 적지 않게 나타나고 있는 것으로 보인다.

종합하면, 연결 중심성은 네트워크 내에서 “얼마나 많은 연결을 맺고 있는가”를 보여주어 직업 현장에서 보편적으로 요구되는 숙련을 드러내는 데 유용하다. 반면 아이겐벡터 중심성은 “누구와 연결되어 있는가”라는 네트워크의 구조적 위치를 반영하여, ‘AI 프레임워크 라이브러리’, ‘자율 시스템’, ‘책임 있는 AI’처럼 허브와 긴밀히 연결된 숙련의 전략적 중요성을 포착한다. 따라서 교육과정 설계에서는 연결 중심성이 높은 숙련을 기초 교양 및 필수 과목으로 배치하고, 아이겐벡터 중심성(EVC)에서 높은 숙련을 핵심 트랙(core track)의 중심축으로 삼는 것이 효과적일 것이다.

파급 중심성(LRC) 기반 정합성 분석은 특정 숙련이 단순히 많은 역량과 연결되어 있는지(연결 중심성) 혹은 영향력 있는 역량과 직접 연결되어 있는지(EVC)를 넘어, 해당 숙련이 교과 역량 전반으로 얼마나 빠르고 넓게 파급될 수 있는지를 반영한다는 점에서 의미가 있다. 다시 말해, LRC는 숙련이 네트워크 내에서 확산의 출발점 혹은 매개 경로로 기능하는 정도, 즉 교육 역량 체계 전체와의 도달성과 전파 가능성을 강조한다.

분석 결과, ‘데이터 과학 분석’은 가장 높은 파급력을 보이며 네트워크의 핵심 숙련으로 나타났다. 특히 ‘데이터 분석·통계적 추론’과의 강한 결합을 중심으로, ‘AI ML 모델·알고리즘’ 및 ‘디지털 인프라와 시스템 기술’ 등 주요 교과 역량으로 연결이 확장되면서, 기초 통계·수학적 기반에서 AI 응용 영역으로 이어지는 학습 경로 전반을 효과적으로 매개하는 역할을 수행하고 있음을 시사한다.

이와 함께 ‘AI 프레임워크·라이브러리’와 ‘자율시스템’ 또한 높은 파급력을 보였다. ‘AI 프레임워크·라이브러리’는 ‘디지털 인프라와 시스템 기술’ 및 ‘AI 이해와 윤리적 활용’ 등과 밀접하게 연결되어 있어, 교육 및 산업 현장에서 관련 기술과 역량이 다양한 교과 영역으로 연쇄적으로 확산될 수 있

는 기반을 제공한다. '자율시스템' 역시 인프라 역량 및 윤리적 활용 역량과 연결된 흐름이 두드러져, 단일 응용 숙련을 넘어 다른 교과 역량으로 파급되는 경로를 생성하는 중간 허브로 작동한다.

흥미로운 점은 '책임 있는 AI'의 상대적 중요성이 LRC 관점에서 부각된다는 것이다. 연결 중심성이나 EVC 기준에서는 핵심 허브로 드러나지 않을 수 있으나, LRC에서는 핵심 교과 역량과의 연결을 매개로 확산 가능성이 상대적으로 강조되는 패턴이 나타난다. 즉, 직접 연결 규모가 크지 않더라도 '디지털 인프라와 시스템 기술' 및 'AI 이해와 윤리적 활용'과 결합함으로써, 다른 역량으로 영향이 확장될 수 있는 구조적 위치를 확보하고 있음을 보여준다.

반면 '자연어 처리', '컴퓨터 비전(이미지 인식)', '로봇공학 자동화 제어'와 같은 특수 응용 숙련은 LRC 기준에서도 상대적으로 낮은 정합성을 보인다. 이는 해당 숙련들이 특정 응용 분야에 집중되어 있으며, 교과 역량 전반으로 확산되는 경로가 제한적이기 때문에 교육 역량 체계 전체와의 연결성이 구조적으로 낮게 나타나는 특징을 반영한다.

정합성을 종합적으로 평가해보면, 먼저, 유사도 기반 정합성과 가중치를 반영한 정합성과의 차이가 존재한다. 단순히 숙련과 역량의 의미적 유사성을 반영하는 유사도 기반 정합성에서는 데이터 분석·통계적 추론, 머신러닝·딥러닝, 수학 모델링·알고리즘처럼 내용적으로 직접 관련된 기초·핵심 역량과의 정합성이 높은 것으로 나타났다. 반면, 네트워크 중심성이나 위계 중심성을 가중치로 반영하는 경우, AI 프레임워크·라이브러리, 자율 시스템, 책임 있는 AI와 같이 단순 유사도에서는 드러나지 않았던 역량이 AI 숙련과의 정합성이 높은 것으로 나타났다.

한편, 어떤 가중치를 사용하느냐에 따라 서로 약간의 차이가 존재한다. '데이터 과학 분석'과 '데이터 분석과 통계적 추론'의 강력한 결합은 세 지표에서 일관되게 나타나며, 전체 네트워크의 중심축을 형성한다. 반면 '자연어 처리', '컴퓨터 비전 이미지 인식', '로봇공학 자동화 제어'와 같은 특수 응용 영역은 세 지표 모두에서 낮은 정합성을 보여, 중심적인 교과 역량과의 연계성이 약하다는 점이 일관되게 드러난다. 이 때문에 겉으로 보기에 세 그림은 크게 다르지 않아 보인다. 그러나 각 지표가 반영하는 가중치의 성격

에 따라 교과 역량과의 정합성이 달라지는 사례도 분명히 존재한다.

우선 ‘AI 프레임워크 라이브러리’는 연결 중심성에서 단순히 다수의 연결을 가진 숙련에 불과했지만, EVC에서는 ‘디지털 인프라와 시스템 기술’, ‘AI 이해와 윤리적 활용’, ‘미래기술과 비즈니스 전략’과 같은 핵심 교과 역량과 밀접히 연결되면서 두 번째로 높은 중심성으로 부각되었다. LRC에서도 확산 가능성을 근거로 높은 비중을 차지하여, 이 숙련이 산업·교육 네트워크에서 전략적 매개 역할을 하고 있음을 보여준다. ‘자율 시스템’ 역시 유사한 변화가 나타난다. 연결 중심성에서는 ‘디지털 인프라와 시스템 기술’과 연결이 강조되었지만 전체적으로는 중간 정도의 위치에 머물렀다. 그러나 EVC에서는 중요 교과 역량과의 관계성이 반영되면서 중심성이 상위권으로 상승하였고, LRC에서는 다른 역량으로의 확산 경로를 형성하는 중간 허브로서의 성격이 드러났다. 또한 ‘책임 있는 AI’는 연결 수가 많지 않아 연결 중심성에서는 주변적인 위치에 그쳤다. 그러나 EVC에서는 ‘디지털 인프라와 시스템 기술’, ‘AI 이해와 윤리적 활용’ 등 네트워크 내 영향력이 큰 교과 역량과 결합하면서 전략적 중요성이 강화되었다. 나아가 LRC에서는 다른 영역으로의 파급 가능성이 평가되어, 단순 연결만으로는 드러나지 않았던 확산적 영향력이 상대적으로 강조되었다.

종합하면, ‘데이터 과학 분석’처럼 모든 지표에서 일관되게 핵심으로 평가되는 경우도 있지만, ‘AI 프레임워크 라이브러리’, ‘자율 시스템’, ‘책임 있는 AI’와 같이 지표에 따라 평가가 달라지는 교과 역량도 존재한다. 이들은 연결 중심성에서는 크게 드러나지 않지만, 아이겐벡터 중심성에서는 중요 숙련과의 전략적 결합이, 파급 중심성에서는 영향력 확산과 파급 경로가 강조되면서 차별적인 정합성을 보여준다. 이는 정합성 평가에서 단일 지표에 의존하기보다, 다양한 중심성 지표를 함께 고려해야 산업-교육 연계의 다층적인 성격을 포착할 수 있음을 시사한다.

제 6 장

요약과 결론

생성형 AI의 급속한 발전과 확산으로 인해 전통적인 업무 방식이 근본적으로 변화하고 있는 상황에서, AI 시대에 일터에서 요구되는 숙련은 무엇인가? 노동시장에서 경쟁력을 유지하기 위해 필요한 핵심 역량이 무엇인지에 대한 체계적인 분석과 전망이 절실히 요구되는 시점이다.

본 보고서는 숙련 변화를 연구하는 방법으로 네트워크 분석을 선택하였다. 숙련을 개별적으로 고립된 능력 단위로 간주하는 것보다, 관계적 개념이나 위계적 개념으로 이해할 때 보다 풍부한 인사이트를 얻을 수 있을 것으로 판단했기 때문이다. 네트워크 분석(Network Analysis)은 개체 간의 관계적 특성을 분석하는 데 적합한 기법으로, 복잡한 관계와 상호작용을 가진 데이터를 효과적으로 분석하고 시각화할 수 있다는 장점을 가진다. 이러한 분석을 통해 개별 숙련의 중요성을 파악함과 동시에, 숙련 간 상호작용과 숙련-직업 간 관계를 분석하였다.

제2장에서는 본 연구팀이 2024년에 만든 한국숙련사전을 기반으로 6,558개 숙련에 대한 AI 노출도를 측정하고, 네트워크 분석을 통해 숙련 간 구조적 특성과 직업 간 전이가능성을 분석하였다. AI 노출도 분석 결과, '관리및 지원'과 '정보기술' 숙련군이 가장 높은 노출도를 보인 반면, '개인돌봄서비스'와 '소프트스킬'은 가장 낮은 노출도를 나타냈다. 특히 반복적이고 규칙성이 높은 업무일수록 AI 대체 가능성이 높았으며, 인간의 감정적 지능과 창의성, 신체적 접촉이 필요한 영역은 여전히 AI 대체가 어려운 것으로 확인되

었다.

숙련 네트워크 분석에서는 2021년과 2024년 IT 분야 숙련 구조의 중요한 변화가 관찰되었다. 2021년에는 ‘빅데이터 컴퓨팅’이 네트워크의 중심이었으나, 2024년에는 ‘통합 보안 관리’가 최대 중심성을 갖게 되어 데이터 중심 시대에서 보안 중심 시대로의 패러다임 전환을 보여주었다. 이는 디지털 전환 가속화와 사이버보안 위협 증가라는 현실적 변화를 반영하는 것으로, IT 전문가들에게 보안 기술에 대한 기본적 이해가 필수가 되었음을 시사한다.

네트워크 특성 지표 분석에서 숙련 확산도와 페이지랭크 중심성 모두에서 ‘고객및클라이언트지원’과 ‘관리및지원’ 숙련군이 최상위를 차지하였다. 이는 현대 비즈니스 환경에서 고객 중심적 접근과 디지털 업무 처리 능력이 산업 전반에 걸쳐 핵심 인프라 역할을 수행하고 있음을 보여준다. 흥미롭게도 이러한 네트워크 중심성과 AI 노출도 간에는 강한 상관관계가 나타나지 않았는데, 이는 AI 대체 가능성이 높은 숙련일수록 현재 노동시장에서 광범위하게 요구되는 역설적 상황을 반영한다.

직업 간 숙련 전이가능성 분석에서는 AI/ML 기술이 범용 기술로서 IT 영역을 넘어 금융, 제조, 전자 등 다양한 산업으로 확산되는 현상이 확인되었다. 특히 핀테크를 통한 금융-IT 인력 교류, 스마트 제조업에서의 하드웨어-소프트웨어 융합, 그리고 전자 산업에서의 AI 기술 접목이 두드러진 특징으로 나타났다. 이러한 융합 현상은 단순한 직종 이동을 넘어 기존 직종의 성격을 근본적으로 변화시키고 있으며, 특정 허브 직종을 통해 숙련 확산이 가속화되는 네트워크 효과를 창출하고 있다.

이러한 연구결과는 AI 시대 인력 개발 정책에 중요한 시사점을 제공한다. 첫째, AI 기술 발전이 모든 숙련을 획일적으로 대체하는 것이 아니라 숙련의 성격에 따라 차별적 영향을 미치므로, 숙련별 맞춤형 대응 전략이 필요하다. 둘째, 네트워크 중심성이 높은 기반 숙련에 대한 지속적 투자와 보안 기술의 범용적 교육이 요구된다. 셋째, AI/ML 기술의 범용성을 고려하여 산업 간 융합과 직종 간 전이를 지원하는 통합적 재교육 시스템이 구축되어야 한다. 특히 허브 역할을 하는 직종에 대한 전략적 지원과 교차 분야 교육 프로그램 개발을 통해 AI 기술의 포용적 확산과 경제·사회적 가치 창출을 극대화할 수 있을 것이다.

제3장에서는 한국고용정보원이 발간한 '2017~2020년 KNOW 재직자조사' 자료를 활용하여, 숙련 보완성과 직업 간 숙련 연관성 지수를 네트워크 관점에서 분석함으로써 한국 노동시장의 숙련 및 직업 양극화 구조를 시론적으로 탐색했다. MST와 루뱅 알고리즘을 통해 네트워크의 위상 구조와 클러스터 결집 양상을 규명하고, 나아가 임금 결정 과정에서 이러한 구조적 요인이 어떻게 작용하는지를 간단한 회귀 분석을 통해 검증했다. 이로부터 밝혀진 분석 결과를 요약하면 다음과 같다.

첫째, 숙련 간 보완성 및 연관성 지수의 분포는 전체적으로 정규형에 가깝지만, 네트워크 위상 구조에서는 '중심-교량-주변부'의 삼중 구조가 나타난다. 즉, 한국의 숙련 체계는 표면적으로는 연속적이지만, 내부적으로는 인지·사회 숙련과 신체 숙련으로 이원화된 구조를 형성하고 있다. 보완성 지수가 일정 임계치(약 0.6)를 넘어설 경우, 일부 숙련 클러스터가 결속되어 외부와 단절되는 '임계적 양극화' 상태로 전이된다.

둘째, 루뱅 알고리즘을 적용한 결과, 전체 숙련은 인지·신체·사회(기타) 숙련의 세 유형으로 구분되며, 특히 인지-신체 숙련 축을 중심으로 명확한 이중 구조가 형성된다. 인지 숙련은 사회적 숙련과 결합하며 네트워크의 중심을 이루지만, 신체 숙련은 주변부로 밀려나 낮은 결속도를 보인다. 이는 한국의 숙련 체계가 인지 숙련 중심의 위계적 결속 구조를 내포하고 있음을 보여준다.

셋째, 397개 직업의 연관성 네트워크를 분석한 결과, MST의 시각화에서 화이트칼라(고인지 숙련)와 블루칼라(중·저인지 숙련) 직업군이 분리된 구조가 나타난다. 루뱅 알고리즘을 적용하면 세 개의 직업군, 즉 고인지 숙련, 중·저인지 숙련, 기타 직업군이 도출되며, 이는 고숙련 중심의 위계적 직업 체계를 시사한다. 임계치(0.65) 이상의 네트워크에서는 고숙련 직종의 중심 집중과 주변부 직군의 분리가 더욱 뚜렷해지며, 한국의 직업 구조가 점차 '인지 숙련 중심-기타 숙련 주변'의 이중적 구조로 수렴하고 있음을 보여준다. 이때 직업별 인지 숙련 비중은 네트워크 중심성과 밀접하게 연동되며, 숙련의 인지적 차원이 직업 간 위계와 연결 구조를 결정하는 핵심 요인으로 작동한다.

넷째, 임금 수준과 그 격차의 결정요인을 분석한 결과, 인지 숙련 비중은

유의한 양(+)의 효과를, 신체 숙련 비중은 음(-)의 효과를 보이며, 사회 숙련의 효과는 상대적으로 미약하다. 이는 '인지 숙련 프리미엄-신체 숙련 디스카운트' 구조가 임금 결정 과정 전반에 일관되게 작용하고 있음을 시사한다. 한편, 임금 격차의 요인 분석에서 직업 간 연관성 계수의 양(+)의 효과는 숙련 유사성이 집중될수록 경쟁이 강화되고 임금 격차가 확대되는 '연관성 속의 경쟁 구조'를 반영한다. 그러나 상호작용항(직업 간 연관성×평균 인지 숙련 비중)의 음(-)의 계수는, 고숙련 내부에서는 오히려 유사성이 경쟁 완화와 조정 기능을 수행함을 보여준다. 즉, 거시적으로는 고숙련 집중이 임금 양극화를 주도하지만, 미시적으로는 내부 조정이 병행되는 다층적 구조가 존재한다.

이러한 분석 결과로부터 '임계적' 양극화의 의미를 뒤집어 볼 필요가 있다. 한국의 숙련과 직업 체계는 전반적으로 보완성과 연관성이 높은 정규형 분포를 보이지만, 그 내부 구조에서는 이미 고숙련·고인지 직종의 결집과 저숙련 직종의 주변화가 병행되는 임계적 양극화가 진행되고 있다. 숙련 네트워크의 보완성 지수가 일정 수준을 넘어서면 일부 숙련 클러스터가 내부적으로 결속되며, 외부와의 연결성이 급격히 약화된다. 이때 네트워크의 분포는 연속적으로 보이지만, 위상 구조는 이미 이중적 안정상태, 즉 고숙련 중심의 핵심 영역과 저숙련 주변부로 분리되어 있다. 이는 한국의 숙련 체계가 아직 미국식 쌍봉형 양극화에 이르지 않는 것만, 겉으로의 연속성과 내적 분절이 공존하는 잠재적 균열 단계에 진입했음을 의미한다.

직업 네트워크에서도 동일한 현상이 반복된다. 직업 간 연관성 지수의 분포는 전체적으로 균형적이거나, MST 및 임계치 0.65 이상의 네트워크에서는 이미 화이트칼라와 블루칼라 직종의 위계적 이원화가 뚜렷하게 나타난다. 루뱅 알고리즘으로 분류된 세 개의 직업군, 즉 고인지 숙련, 중·저인지 숙련, 기타 직업군은 서로 명확한 경계를 이루며, 특히 고인지 숙련 직종이 네트워크의 중심에서 높은 영향력을 행사하고 있다. 이러한 현상은 숙련 네트워크에서의 인지 숙련 중심 구조와 정확히 병행된다. 즉, 숙련의 인지적 위계화가 직업 구조의 기능적 분리를 매개하면서, 숙련과 직업 모두에서 중심-주변의 이중적 구조가 강화되고 있다. 보완성은 숙련 체계를 단단히 묶어 주지만, 그 결속이 과도할 경우 새로운 연결을 차단하고, 연관성은 직업 간

지식과 숙련의 순환을 촉진하지만, 유사성이 높을수록 내부 경쟁을 강화하는 방향으로 작동하는 ‘보완성과 연관성의 역설’이 나타난다.

이러한 임계적 양극화는 숙련과 직업의 관계가 단순히 기술적 특성의 차이가 아니라, 인지적 속성과 네트워크 연결성의 결합에 따라 재편되고 있음을 시사한다. 숙련 차원에서 인지적 역량이 핵심 축으로 작동하면서 신체 숙련을 주변화시키고, 직업 차원에서는 이 인지적 위계가 노동시장 내 기능적 분화를 구조화한다. 그 결과, 임금 구조 또한 ‘인지 숙련 프리미엄-신체 숙련 디스카운트’의 형태로 고착되며, 노동시장의 불평등이 숙련 구조의 위계와 정합적으로 연결된다. 다만, 다른 층위에서 보면, 고숙련 영역 내 연관성은 단순히 경쟁을 심화시키기보다 체계 내부의 조정과 균형을 촉진하는 방향으로 작용한다. 즉, 거시적 차원에서는 고숙련 중심의 불균형과 집중이 심화되지만, 미시적 차원에서는 이러한 불균형을 일정 부분 흡수하고 완화하는 보완적 구조가 함께 존재하는 다층적 균형 상태가 형성되고 있다.

정책적으로 이러한 결과는 단순한 고숙련화나 인지 중심 교육 강화만으로는 불평등을 완화하기 어렵다는 점을 시사한다. 숙련 체계의 균형적 발전은 보완성과 연관성의 임계점을 관리하는 정책적 조정 능력에 달려 있다. 즉, 이러한 숙련 집중이 체계 결속으로 기능할 때는 긍정적이지만, 과도할 경우 구조적 단절을 초래한다는 점에서, 숙련 간 연결성의 다양성과 교차성을 유지하는 것이 핵심 과제다. 따라서 한국의 숙련·직업 정책은 고숙련 집단의 내부 경쟁 완화와 중·저숙련 집단의 연결성 강화, 그리고 인지 숙련 중심 구조에 대응하는 다층적 네트워크 균형 전략으로 전환되어야 한다.

제4장에서는 AI 기술 확산과 디지털 전환으로 급격히 재편되는 노동시장의 숙련 구조를 분석했다. 구인공고에서 요구되는 숙련들을 네트워크 방법론을 통해 시각화하고 브로커리지(brokerage)/구조적 공백(structural hole) 개념을 도입하여 임금과의 관계를 검증했다. 기존 숙련 연구가 주로 중심성 지표에 집중한 것과 달리, 본 연구는 산업·직종 블록을 가로지르는 브로커리지 역할에 집중한 것이 특징이다. 특히 Gould and Fernandez의 다섯 가지 브로커리지 유형과 Burt의 제약(constraint) 지수를 적용하여 숙련 네트워크상의 구조적 위치가 가지는 전략적 중요성을 측정하고, 대규모 IT 일자리 온라인 구인공고 자료(약 100만 건 중 10,000개 추출)를 사용해 회귀분석

을 실시하였다.

분석 결과, 엑셀, 파워포인트/프레젠테이션, 문서작성(MS Word), 프로젝트 관리/업무 조정, 데이터 분석/보고서 작성 등 범용성이 높은 기술들이 가장 강한 브로커리지 역할을 수행하며 직종 간 연결을 주도하는 것으로 드러났다. 브로커리지 하위 유형 중 조정자(coordinator), 대표자(representative), 문지기(gatekeeper) 역할의 숙련이 더 높은 임금을 받는 경향이 확인되었고, 구조적 공백(structural hole) 분석에서는 제약 지수만이 음의 효과를 보여 브로커에 의해 통제되는 숙련이 낮은 보상을 받음을 확인했다. 본 연구는 기존 중심성 지표가 포착하지 못한 산업 간 융합과 신기술 확산의 촉매로서 브로커 숙련의 중요성을 밝히며, 향후 재훈련과 교육 투자, 산업 다각화(diversification) 계획 수립에 전략적 함의를 제공한다.

브로커리지 숙련은 단순히 현재 수요를 반영하는 것이 아니라, 신기술 확산과 산업 간 융합을 통한 새로운 산업 출현을 예비하는 촉매 역할을 한다. 특히 대표자와 문지기 역할을 하는 숙련은 향후 산업 발전에 주요한 역할을 할 것으로 예상되므로, 개인 및 기업 차원에서 이러한 숙련에 대한 집중적 교육 투자가 필요하다. 이러한 숙련들은 사업의 다각화(diversification)에 결정적 역할을 하며 시장 수요가 증가할 가능성이 크다. 본 연구는 재훈련과 교육 투자, 산업 다각화 계획 수립에 있어 중심성뿐 아니라 경제 매개 역할의 중요성을 고려해야 함을 제시한다.

제5장에서는 숙련 간 네트워크 구조와 위계성, 그리고 교육과정 내 숙련 포괄성을 고려해 AI 관련 숙련의 수급 정합성을 분석하였다. 온라인 구인공고(사람인, 잡코리아)와 국내 대학 및 K-MOOC에서 개설된 AI 교육과정 텍스트를 수집하고, 텍스트 임베딩(text-embedding-3-large), BM25, GPT-4 기반 재랭킹 기법을 결합한 하이브리드 매칭 방식을 활용해 수요-공급 간 구조적 정합성을 평가하였다.

분석 결과, 산업 현장에서는 AI 프레임워크·라이브러리, AI/ML 모델, 데이터 분석 등 범용(generic) 기술 역량 수요가 높았고, 이는 기초 기술 기반의 인력이 다양한 산업에서 폭넓게 요구됨을 보여준다. 반면 자연어 처리, 컴퓨터 비전, 로봇틱스 등 특화(specific) 기술 역량은 상대적으로 수요가 낮았으며, 이는 해당 기술들이 특정 직무에 국한되기 때문으로 보인다. IT 및

데이터 직군에 AI 숙련 수요가 집중된 가운데, 일반 사무·서비스 분야에서도 AI 활용이 점진적으로 확산되는 추세다.

한편, 국내 대학과 K-MOOC의 AI 교육과정 477개를 분석한 결과, 머신러닝·딥러닝, 데이터 분석, 통계 추론, 수학 모델링 중심의 기초 기술 교육 비중이 높았고, AI 윤리, 디지털 인프라 등 다층적 주제도 일부 포함되어 있었다. 교육 내용은 일반 기술 역량에서는 산업 수요와 일정 수준 정합성을 보였지만, 생성형 AI, 강화학습, MLOps 등 최신 특화 기술은 교육과정에 거의 반영되지 않았다.

정합성 분석에 따르면, 120개의 숙련 키워드 중 교육과정에 어느 정도 반영된 키워드는 95개(47.5%)였으며, 범용 기술은 높은 일치치를 보인 반면, 생성형 AI, 빅데이터 플랫폼, 책임 있는 AI 등은 산업 수요에 비해 반영이 미흡했다. 직무 유형별로는 기술직군에서 정합성이 높았고, 기획·관리·서비스 직군에서는 낮았다.

본 연구는 단순 유사도 분석을 넘어, 연결 중심성, 아이젠벡터 중심성, 맥락 중심성, LRC, Complexity Score, RER 등 네트워크 기반 지표를 도입해 숙련의 위계성과 범용성을 다차원적으로 평가하였다. G_score를 활용해 숙련을 일반-중간-특수로 분류한 결과, 범용 숙련은 전이가능성과 연결성이 높아 재직자 훈련과 전직 설계에 적합했으며, 특수 숙련은 맞춤형 단기 교육에, 위계적 핵심 숙련은 전략적 재교육에 유용한 것으로 나타났다.

AI 숙련 공출현 네트워크 분석에서는 Python, 데이터 마이닝, 생성형 AI, 지능형 인프라 등이 중심 허브 역할을 하며 다양한 특수 숙련과 연결되고 있었고, 이는 통합적 교육 설계의 필요성을 보여준다. 특히 LRC 기반 분석에서 고학력 직무일수록 범용 숙련과의 연결성이 높았으며, 특화 숙련은 산업 맞춤형 또는 고급 연구직 훈련에 적합했다.

정합성 분석에서 의미 기반 유사도 점수는 전반적으로 낮았으며, 대부분의 직업-역량 조합에서 0.2대 후반의 유사도 값이 나타났다. 데이터 분석과 통계 추론은 높은 정합성을 보였지만, 자연어 처리, 컴퓨터 비전 등은 낮은 일치치를 보였다. 반면 중심성을 고려한 가중 정합성 분석에서는 AI 프레임워크, 자율 시스템, 책임 있는 AI 등이 전략적 중요 숙련으로 확인되었다.

이러한 결과는 정합성 평가에서 단순 유사도뿐 아니라 숙련 간 구조적 연

결성, 중심성, 파급력 등을 함께 고려해야 함을 보여준다. 예컨대 AI 프레임워크와 자율 시스템은 의미 유사도는 낮지만 구조적 연결성 지표에서는 높은 전략적 가치를 지닌 것으로 나타났다.

본 연구는 AI 채용 시장과 교육과정 간 숙련 정합성을 구조적으로 분석하고 시각화하여, 수요 기반 교육 설계, 전략적 숙련 공급, 직업 전환 경로 설정 등에 실질적인 시사점을 제공한다. 다만 몇 가지 한계도 존재한다. 첫째, 임베딩 기반 유사도는 언어적 정량화에는 유용하나, 산업 현장에서 요구되는 실제 역량이나 교육의 질적 요소를 완전히 반영하는 것은 아니다. 둘째, 직업 숙련과 교육 역량은 서로 다른 체계에 기반하므로 유사도가 낮더라도 반드시 불일치로 해석하긴 어렵다. 셋째, 중심성 기반 지표의 정책 적용에는 구체적 기준이 부족하며, 중심성과 노동시장 성과 간의 연계성에 대한 실증 검증이 추가로 필요하다. 또한 구인공고 데이터는 표현 표준화 부족과 특정 산업 중심 경향이 있어 대표성에 한계가 있고, 교육 분석 역시 키워드 중심이라 실제 교육 내용의 깊이나 교수법, 학습성과 등은 반영되지 않았다. 아울러 본 연구는 단면 분석이므로, 향후에는 시계열 분석을 통해 AI 기술 및 숙련 수요의 변화 동향을 지속적으로 추적할 필요가 있다.

본 보고서는 AI 시대 노동시장의 숙련 구조 변화를 네트워크 분석이라는 새로운 방법론적 렌즈를 통해 다층적으로 조명함으로써, 숙련을 고립된 개별 능력이 아닌 관계적·구조적 개념으로 재정립하고자 했다. 한국숙련사전과 구인공고 데이터를 기반으로 한 AI 노출도 분석과 숙련 네트워크의 중심성 변화 추적(제2장), 숙련 보완성과 직업 연관성을 통한 임계적 양극화 구조의 규명(제3장), 브로커리지 개념을 활용한 산업 간 숙련 매개 역할의 발견(제4장), 그리고 AI 교육과정과 채용 시장 간 구조적 정합성 분석(제5장)을 종합하여, 본 연구는 숙련의 단순한 보유 여부를 넘어 숙련 간 연결성, 위계성, 매개성이 노동시장 성과와 정책 설계에 미치는 영향을 실증적으로 제시하고자 했다. 특히 인지 숙련 중심의 위계적 결속과 신체 숙련의 주변화라는 이중 구조, 범용 기술의 브로커리지 역할과 그것이 임금에 미치는 차별적 효과, 그리고 교육-노동시장 간 구조적 불일치를 가시화함으로써, AI 기술 확산이 초래하는 숙련 재편이 단순한 대체나 보완의 문제가 아니라 노동시장 전체의 네트워크 위상과 연결 구조를 근본적으로 변화시키는 과정임

을 밝혔다. 이러한 발견은 향후 인력 개발 정책이 개별 숙련의 육성을 넘어 숙련 간 연결성의 다양성 유지, 브로커 숙련에 대한 전략적 투자, 그리고 교육과정의 구조적 재설계를 통해 포용적이고 균형 잡힌 숙련 생태계를 조성해야 함을 시사하며, 네트워크 관점이 노동시장 분석과 정책 수립에 있어 필수적인 분석틀로 자리매김할 수 있음을 보여준다.

참고문헌

- 이현경(2025), 「인공지능 노동력과 숙련(Skill) 정의 및 측정」, 『KISDI Perspectives』 1, 정보통신정책연구원.
- 장지연 · 전병유 · 정준호 · 이철승 · 심지환 · 안성준(2024), 『인공지능(AI) 발전의 고용효과』, 한국노동연구원.
- 전병유 · 장지연 · 심지환(2024), 「텍스트 분석을 통한 숙련 수요와 교과 역량 간의 매칭 분석: 데이터사이언스 전공을 중심으로」, 『한국진로창업경영학회지』 8(1), pp.195~217.
- 정준호(2014), 「주택시장의 네트워크 구조 분석: 수도권 아파트 매매시장의 사례」, 『한국경제지리학회지』 17(2), pp.280~295.
- 황지영 · 한나연 · 차필립 · 이철승(2025), 「비합법 이주노동의 중개와 '신뢰': 일용직 노동시장의 중개인들을 중심으로」, 미출간 논문.
- Acemoglu, D., and D. Autor(2011), "Skills, tasks and technologies: Implications for employment and earnings", *Handbook of Labor Economics* 4, pp.1043~1171.
- Aho, A. V., J. E. Hopcroft, and J. D. Ullman(1983), "Data Structures and Algorithms, Reading", MA: Addison-Wesley.
- Ahuja, Gautam(2000), "Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study", *Administrative Science Quarterly* 45(3), pp.425~455.
- Akerlof, George A.(1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics* 84(3), pp.488~500.
- Alabdulkareem, A., M. R. Frank, L. Sun, B. AlShebli, C. Hidalgo, and I. Rahwan(2018), "Unpacking the polarization of workplace skills",

- Science Advances* 4(7), eaao6030. <https://doi.org/10.1126/sciadv.aao6030>
- Alekseeva, L., J. Azar, M. Giné, S. Samila, and B. Taska(2021), “The demand for AI skills in the labor market”, *Labour economics* 71, 102002.
- Aufiero, Sabrina, Giordano De Marzo, Angelica Sbardella, and Andrea Zaccaria(2024), “Mapping job fitness and skill coherence into wages: an economic complexity analysis”, *Scientific Reports* 14(1), 11752.
- Babashahi, L., C. E. Barbosa, Y. Lima, A. Lyra, H. Salazar, M. Argôlo, ... and J. M. D. Souza(2024), “AI in the workplace: A systematic review of skill transformation in the industry”, *Administrative Sciences* 14(6), 127.
- Baruffaldi, S. H., M. Mebirouk, and F. Martínez-Plumed(2020), “Identifying and measuring developments in artificial intelligence: Making the impossible possible”, OECD Science, Technology and Industry Working Papers, 2020/05. OECD Publishing, <https://doi.org/10.1787/d62f5f47-en>
- Blondel, V. D., J.-D. Guillaume, R. Lambiotte, and E. Lefebvre(2008), “Fast unfolding of communities in large networks”, *Journal of Statistical Mechanics: Theory and Experiment*, P10008(arXiv:0803.0476 [physics.soc-ph]).
- Borgonovi, F., A. Ferrara, and S. Kampelmann(2023), “Emerging trends in AI skill demand across 14 OECD countries”, OECD.AI Policy Observatory Reports, <https://oecd.ai>
- Boschma, R.(2017), “Relatedness as driver of regional diversification: A research agenda”, *Regional Studies* 51(3), pp.351~364.
- Bourdieu, Pierre, and Loïc J. D. Wacquant(1992), *An Invitation to Reflexive Sociology*, Chicago: University of Chicago Press.
- Brynjolfsson, E., D. Rock, and C. Syverson(2017), “Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics”, (No. w24001) National Bureau of Economic Research.

- Burt, R. S.(1992), "Structural Holes: The Social Structure of Competition", Harvard University Press.
- Burt, Ronald S.(1990), *Structural Holes: The Social Structure of Competition*, Cambridge, MA: Harvard University Press.
- _____(2004), "Structural Holes and Good Ideas," *American Journal of Sociology* 110, no.2, pp.349~399.
- _____(2005), *Brokerage and Closure: An Introduction to Social Capital*, Oxford: Oxford University Press.
- Coleman, James S.(1973), "The Mathematics of Collective Action," *Heinemann Educational Books*.
- _____(1988), "Social Capital in the Creation of Human Capital," *American Journal of Sociology* 94, (Supplement): S95-S120.
- _____(1990), *Foundations of Social Theory*, Cambridge, MA: Harvard University Press.
- Cui, J.(2024), "Exploring the Impact of Generative AI on Cross-Border E-Commerce Brand Building in Chinese Tianjin's Manufacturing Sector", arXiv preprint arXiv:2411.17700.
- Dawson, N., M. A. Williams, and M. A. Rizoïu(2021), "Skill-driven recommendations for job transition pathways", *PLOS ONE* 16(8), e0254722, <https://doi.org/10.1371/journal.pone.0254722>
- del Rio-Chanona, R. M., P. Mealy, M. Beguerisse-Díaz, F. Lafond, and J. D. Farmer(2021), "Automation and occupational mobility: A data-driven network model", *Journal of The Royal Society Interface* 18(174), 20200898.
- Eloundou, T., S. Manning, P. Mishkin, and D. Rock(2023), "Gpts are gpts: An early look at the labor market impact potential of large language models", arXiv preprint arXiv:2303.10130.
- Frenken, K., F. Van Oort, and T. Verburg(2007), "Related variety, unrelated variety and regional economic growth", *Regional Studies* 41(5), pp.685~697.

- Gathmann, C. and U. Schönberg(2010), “How general is human capital? A task-based approach”, *Journal of Labor Economics* 28(1), pp.1~49.
- Gmyrek, P., J. Berg, and D. Bescond(2023), “Generative AI and jobs: A global analysis of potential effects on job quantity and quality”, ILO Working Paper 96.
- Gould, R. V. and R. M. Fernandez(1989), “Structures of Mediation: A Formal Approach to Brokerage in Transaction Networks”, *Sociological Methodology* 19, pp.89~126.
- Granovetter, Mark S.(1973), “The Strength of Weak Ties,” *American Journal of Sociology* 78, no.6, pp.1360~1380.
- Grosser, Travis J., Giuseppe Labianca, and Daniel J. Brass.(2017), “A Social Network Analysis of Positive and Negative Gossip in Organizational Life,” *Group & Organization Management* 42, no.3, pp.1~29.
- Heimberger, H., D. Horvat, A. Jäger, and F. Schultmann(2025), “Exploring AI Adoption in Manufacturing: An Empirical Study on Effects of AI Readiness”, *International Journal of Production Economics* 109733.
- Henning, M., R. Eriksson, P. Garefelt, H. Martin, and Z. Elekes(2023), “Job Relatedness, Local Skill Coherence and Economic Performance: A Job Postings Approach”, Utrecht University, Human Geography and Planning.
- _____ (2025), “Job relatedness, local skill coherence and economic performance: a job postings approach”, *Regional Studies, Regional Science* 12(1), pp.95~122.
- Hosseinioun, M., F. Neffke, L. Zhang, and H. Youn(2025), Skill dependencies uncover nested human capital, *Nature Human Behaviour*, pp.1~15.
- Huang, A. Y., T. Fisher, H. Ding, and Z. Guo(2021), “A network analysis of cross-occupational skill transferability for the hospitality industry”, *International Journal of Contemporary Hospitality Management* 33(12), pp.4215~4236. <https://doi.org/10.1108/IJCHM-06-2021-0749>

- Jiang, W., Y. Tang, R. J. Xiao, and V. Yao(2025), “Surviving the FinTech disruption”, *Journal of Financial Economics* 171, 104071.
- Kok, S., and B. ter Weel(2014), “Cities, Tasks, and Skills”, *Journal of Regional Science* 54(5), pp.856~892.
- Lambiotte, R., J.-C. Delvenne, and M.Barahona(2010), Multi-scale Modularity in Complex Networks, *arXiv:1004.4268* [physics. soc-ph].
- Le, Q. and T. Mikolov(2014), “Distributed representations of sentences and documents”, *In International conference on machine learning*, pp.1188~1196, PMLR.
- Lee, Cheol-Sung, and Taekyeong Goh(2023), “Union Brokerage and the Gender Gap in the Labor Market: A Cross-National Comparative Study of Associational Networks and Gendered Labor Force Participation in OECD Countries”, *International Journal of Comparative Sociology* 64(1), pp.22~56.
- Levy Yeyati, E. and M. Montané(2020), “Specificity of human capital: An occupation space based on job-to-job transitions”, *CID Working Paper No.379*, Harvard University.
- Liu, Zhaolu, Jonathan M. Clarke, Bertha Rohenkohl, and Mauricio Barahona(2025), “Patterns of co-occurrent skills in UK job adverts”, *PLOS Complex Systems* 2(2), e0000028.
- Maghsoudi, M.(2024), “Uncovering the skill sets required in computer science jobs using social network analysis”, *Education and Information Technologies* 29 (10), pp.12759~12780. <https://doi.org/10.1007/s10639-023-12241-2>
- Manca, F.(2023), “Six questions about the demand for artificial intelligence skills in labour markets”, OECD Social, Employment and Migration Working Papers No.287, OECD Publishing. <https://doi.org/10.1787/5d396948-en>
- Marsden, Peter V.(1982), “Brokerage Behavior in Restricted Exchange

- Networks,” In *Social Structure and Network Analysis*, edited by Peter V. Marsden and Nan Lin, pp.201~218, Beverly Hills, CA: Sage.
- _____(1983), “Restricted Access in Networks and Models of Power,” *American Journal of Sociology* 88, no.4, pp.686~717.
- Mortensen, Dale T.(1982), “The Matching Process as a Noncooperative Bargaining Game,” In *The Economics of Information and Uncertainty*, edited by John J. McCall, pp.233~258, Chicago: University of Chicago Press.
- Muneepeerakul, R., J. Lobo, S. T. Shutter, A. Gómez-Liévano, and M. R. Qubbaj(2013), “Urban economies and occupation space: Can they get “there” from “here”?”, *PLoS ONE* 8(9), e73676.
- Nedelkoska, L., F. Neffke, and S. Wiederhold(2015), “Skill mismatch and the costs of job displacement”, Harvard University, CID Working Paper.
- Neffke, Frank, and Martin Henning(2013), “Skill Relatedness and Firm Diversification”, *Strategic Management Journal* 34(3) pp.297~316.
- Nelson, J. P., J. B. Biddle, and P. Shapira(2023), “Applications and societal implications of artificial intelligence in manufacturing: A systematic review”, arXiv preprint arXiv:2308.02025.
- Obstfeld, David, Stephen P. Borgatti, and Jason Davis(2014), “Brokerage as a Process: Decoupling Third Party Action from Social Network Structure”, In *Contemporary Perspectives on Organizational Social Networks*, pp.135~159, Emerald Group Publishing Limited.
- OECD(2023), *OECD Skills Outlook 2023: Skills for a Resilient Green and Digital Transition*, OECD Publishing, Paris, <https://doi.org/10.1787/27452f29-en>.
- Onnela, J.-P., A. K. Chakraborti, Kaski, and K. A. Kertész(2003), “Dynamics of market correlations: taxonomy and portfolio analysis”, *Physical Review E* 68(056110).
- Pater, R., H. Cherniaiev, and M. Kozak(2022), “A dream job? Skill demand

- and skill mismatch in ICT”, *Journal of Education and Work* 35 (6-7), pp.641~665. <https://doi.org/10.1080/13639080.2022.2136562>
- Poletaev, M. and C. Robinson(2008), “Human capital specificity: Evidence from the Dictionary of Occupational Titles and displaced worker surveys, 1984-2000”, *Journal of Labor Economics* 26(3), pp.387~420.
- Soni, N., E. K. Sharma, N. Singh, and A. Kapoor(2020), “Artificial intelligence in business: from research and innovation to market deployment”, *Procedia Computer Science* 167, pp.2200~2210.
- Spulber, Daniel F.(1996), “Market Microstructure and Intermediation” *Journal of Economic Perspectives* 10, no.3 (Summer), pp.135~152.
- Squicciarini, M., and D. Nachtigall(2021), “Demand for AI skills in jobs: Evidence from online job postings”, OECD Science, Technology and Industry Working Papers, 2021/10. OECD Publishing. <https://doi.org/10.1787/7c63bc3e-en>
- Stephany, F. and O. Teutloff(2024), “What is the price of a skill? The value of complementarity”, *Research Policy* 53 (1), 104898, <https://doi.org/10.1016/j.respol.2023.104898>
- Stovel, Katherine, and Lynette Shaw(2012), “Brokerage,” *Annual Review of Sociology* 38, pp.139~158.
- Taska, B., S. Samila, J. Azar, M. Gine, and L. Alekseeva(2020), “The demand for AI skills in the labour market”, LinkedIn Economic Graph Research (2020, May 3).
- Uzzi, Brian.(1997), “Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness,” *Administrative Science Quarterly* 42, no.1, pp.35~67.
- Vista, A., E. Care, and H. Kim(2020), “Data-driven identification of skills for the future: 21st-century skills for the 21st-century workforce”, *SAGE Open* 10(2).
- Waters, K. and S. T. Shuttters(2022), “Skills-approximate occupations:

Using networks to guide jobs retraining”, *Applied Network Science* 7, 43. <https://doi.org/10.1007/s41109-022-00485-3>

Xie, M., L. Ding, Y. Xia, J. Guo, J. Pan, and H. Wang(2021), “Does artificial intelligence affect the pattern of skill demand? Evidence from Chinese manufacturing firms,” *Economic Modelling* 96, pp.295~309.

[부록]

〈부표 2-1〉 숙련분류체계 ver.2

대분류	중분류
건축및건설	건설관리
건축및건설	일반건설및건설노동
건축및건설	토목및건축공학
농업/원예및조경업	농업
농업/원예및조경업	원예및조경업
관리및지원	사무자동화
관리및지원	일반관리및사무업무
경영	리더십및의사소통
경영	운영및관리
경영	인적자원관리
경영	전략및분석
고객및클라이언트지원	고객관리및모니터링
고객및클라이언트지원	고객지원
디자인	공간디자인
디자인	그래픽및시각디자인
디자인	기타디자인
디자인	산업디자인
디자인	서비스디자인
디자인	영상및3D그래픽
에너지	발전
에너지	에너지과학및기술
에너지	지속가능에너지
엔지니어링	설계
엔지니어링	엔지니어링일반
엔지니어링	측정및시험
정책및사회과학	분과지식
정책및사회과학	정책분석과연구및개발
교육및훈련	교육서비스

〈부표 2-1〉의 계속

대분류	중분류
교육및훈련	교육지원
교육및훈련	교육콘텐츠
교육및훈련	기타교육
교육및훈련	직업교육
교육및훈련	학생지원및서비스
환경	컴플라이언스
환경	환경과학및기술
환경	환경및자원관리
금융과회계	금융서비스
금융과회계	금융시장및거래
금융과회계	기업금융
금융과회계	대출및신용
금융과회계	리스크관리및규제
금융과회계	보험및연금
금융과회계	세무및법규
금융과회계	자산및부동산
금융과회계	재무보고및분석
금융과회계	투자및자산관리
금융과회계	회계및감사
금융과회계	회계시스템
금융과회계	회계실무
정부및공공행정	안보
정부및공공행정	정부서비스
접객및식품서비스	식품서비스
접객및식품서비스	접객
접객및식품서비스	행사및시설
정보기술	IT전략및거버넌스
정보기술	네트워크및클라우드인프라
정보기술	데이터관리및분석
정보기술	소프트웨어개발및엔지니어링

〈부표 2-1〉의 계속

대분류	중분류
정보기술	신흥기술
정보기술	엔터프라이즈솔루션
정보기술	인공지능및머신러닝
정보기술	하드웨어및IoT
언어	언어교육
언어	언어역량
언어	언어통역및번역
법/규제및컴플라이언스	법률지원
법/규제및컴플라이언스	법집행및형사사법
법/규제및컴플라이언스	컴플라이언스
홍보마케팅	기획
홍보마케팅	디지털마케팅
홍보마케팅	마케팅
홍보마케팅	채널관리
홍보마케팅	콘텐츠마케팅
홍보마케팅	홍보
유지보수/수리및시설서비스	기기및차량
유지보수/수리및시설서비스	시설관리
유지보수/수리및시설서비스	장비관리
유지보수/수리및시설서비스	청소및관리
의료및건강관리	공중보건및질병예방
의료및건강관리	관리및운영
의료및건강관리	보조의료검사및기술
의료및건강관리	의료기기
의료및건강관리	일반의료검사및기술
의료및건강관리	제약
제조및생산	생산관리
제조및생산	제조및생산
제조및생산	제품관리
제조및생산	조립및포장등수작업

<부표 2-1>의 계속

대분류	중분류
미디어와글쓰기	엔터테인먼트및음악
미디어와글쓰기	연극및공연예술
미디어와글쓰기	영상
미디어와글쓰기	출판
미디어와글쓰기	콘텐츠개발및관리
개인돌봄서비스	돌봄
개인돌봄서비스	미용서비스
개인돌봄서비스	피트니스
공공안전및국가안보	보안
공공안전및국가안보	환경안전보건관리
영업및판매	B2B(기업간)영업
영업및판매	B2C영업
영업및판매	마케팅및판촉
영업및판매	영업프로세스
영업및판매	전문영업역량
영업및판매	특수영업
영업및판매	판매채널
사회및인적서비스	종교연구및서비스
사회및인적서비스	커뮤니티및사회사업
과학연구	기초과학
과학연구	연구방법론
과학연구	의료과학및연구
과학연구	화학및생명공학
소프트스킬	개인내적역량
소프트스킬	대인관계역량
소프트스킬	신체적역량
소프트스킬	인지적역량
운송및공급망과물류	관리및운영
운송및공급망과물류	물류
운송및공급망과물류	배달및운송

자료 : 한국숙련사전.

〈부표 2-2〉 SkillAIAssessment 코드

```

class SkillAIAssessment:
    def __init__(self, api_key: str, model: str = "gpt-4-turbo"):
        """
        AI 기반 시간 절감률 평가 클래스

        Args:
            api_key: OpenAI API 키
            model: 사용할 GPT 모델 (기본값: gpt-4-turbo-preview)
        """
        self.client = OpenAI(api_key=api_key)
        self.model = model
        self.prompt_template = """다음 단일 숙련(skill)에 대해 5인 전문가 패널로
"AI 기반 시간 절감률(%)을 3라운드 합의로 산정하라.

[평가기준]
- 복잡판단 필요성 / 반복·규칙성 / 자동화 구조화 용이성
- 점수: 0~100 점수. (0~20: 어려움 / 20~50: 일부 보조 / 50~80: 대폭 보조 /
80~100: 전면 자동화)
- 각 라운드에서 expertiser1..5가 score를 내고, 서로의 이전 의견을 반영해 조
정한다.
- 각 코멘트는 토론의 내용을 이해할 수 있도록 적당히 상세하게 작성한다.

[합의 산정]
- consensus.method = "mean_of_round3"
- consensus.score_pct = round(mean(round3의 5개 점수))
- consensus.range_pct = [min(round3), max(round3)]

[반드시 이 JSON만 출력하라. 추가 텍스트·설명 금지.]

{

```

〈부표 2-2〉의 계속

```

"skill": "{skill_name}",
"rounds": {
  "round1": {
    "expertiser1": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser2": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser3": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser4": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser5": {"score_pct": <int>, "comment": "<<=200자>"},
  },
  "round2": {
    "expertiser1": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser2": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser3": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser4": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser5": {"score_pct": <int>, "comment": "<<=200자>"},
  },
  "round3": {
    "expertiser1": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser2": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser3": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser4": {"score_pct": <int>, "comment": "<<=200자>"},
    "expertiser5": {"score_pct": <int>, "comment": "<<=200자>"},
  }
},
"consensus": {
  "method": "mean_of_round3",
  "score_pct": <int>,
  "range_pct": [<int>, <int>],
  "summary": "<<=200자>"
},
"meta": {
  "skill_id": "<slug(소문자-하이픈)>",

```

〈부표 2-2〉의 계속

```

"version": "v1.0",
"created_at": "⟨ISO-8601: YYYY-MM-DDTHH:MM:SS+09:00⟩",
"language": "ko"
}}
}"""

def call_gpt_api(self, skill_name: str, max_retries: int = 3) -> Dict[str, Any]:
    """
    GPT API 호출하여 스킬 평가 수행

    Args:
        skill_name: 평가할 스킬 이름
        max_retries: 최대 재시도 횟수

    Returns:
        API 응답 JSON 객체
    """
    prompt = self.prompt_template.format(skill_name=skill_name)

    for attempt in range(max_retries):
        try:
            response = self.client.chat.completions.create(
                model=self.model,
                messages=[
                    {"role": "system", "content": "You are an expert panel
evaluating AI automation potential. Always respond with valid JSON only."},
                    {"role": "user", "content": prompt}
                ],
                temperature=0.7,
                response_format={"type": "json_object"} # JSON 모드 활성화
            )

```

〈부표 2-2〉의 계속

```

        # 응답에서 JSON 파싱
        result = json.loads(response.choices[0].message.content)
        return result

    except json.JSONDecodeError as e:
        print(f"JSON 파싱 오류 (시도 {attempt + 1}/{max_retries}): {e}")
        if attempt == max_retries - 1:
            raise

    except Exception as e:
        print(f"API 호출 오류 (시도 {attempt + 1}/{max_retries}): {e}")
        if attempt == max_retries - 1:
            raise
        time.sleep(2 ** attempt) # 지수 백오프

    return None

def process_dataframe(self,
                      df: pd.DataFrame,
                      skill_column: str = 'skill_name',
                      output_dir: str = 'ai_assessment_results',
                      batch_size: int = 5,
                      delay_seconds: int = 1) -> pd.DataFrame:
    """
    데이터프레임의 모든 스킬에 대해 평가 수행

    Args:
        df: 스킬 목록이 포함된 데이터프레임
        skill_column: 스킬 이름이 포함된 열 이름
        output_dir: 결과 JSON 파일을 저장할 디렉토리
        batch_size: 배치 처리 크기 (API 레이트 리밋 관리)
        delay_seconds: API 호출 간 지연 시간
    """

```

〈부표 2-2〉의 계속

```

Returns:
    AI 평가 결과가 추가된 데이터프레임
"""
# 결과 저장 디렉토리 생성
os.makedirs(output_dir, exist_ok=True)

# 결과를 저장할 리스트
results = []
all_discussions = {}

# 진행률 표시를 위한 tqdm 사용
skills = df[skill_column].tolist()

print(f"총 {len(skills)}개 스킬 평가 시작...")

for i, skill in enumerate(tqdm(skills, desc="스킬 평가 진행")):
    try:
        # GPT API 호출
        result = self.call_gpt_api(skill)

        if result:
            # 합의 점수와 범위 추출
            consensus_score = result['consensus']['score_pct']
            score_range = result['consensus']['range_pct']
            summary = result['consensus']['summary']

            # 결과 저장
            results.append({
                'skill_name': skill,
                'ai_time_saving_pct': consensus_score,
                'ai_score_min': score_range[0],
                'ai_score_max': score_range[1],
            })
    except Exception as e:
        print(f"Skill {skill} evaluation failed: {e}")

```

〈부표 2-2〉의 계속

```

        'ai_assessment_summary': summary
    })

    # 전체 토론 내용 저장
    all_discussions[skill] = result

    # 개별 JSON 파일로도 저장
    skill_filename = self._create_filename(skill)
    with open(os.path.join(output_dir, f'{skill_filename}.json'), 'w',
              encoding='utf-8') as f:
        json.dump(result, f, ensure_ascii=False, indent=2)

    else:
        print(f"경고: {skill}에 대한 평가 실패")
        results.append({
            'skill_name': skill,
            'ai_time_saving_pct': None,
            'ai_score_min': None,
            'ai_score_max': None,
            'ai_assessment_summary': None
        })

    # API 레이트 리밋 관리
    if (i + 1) % batch_size == 0:
        time.sleep(delay_seconds * 2) # 배치 완료 후 추가 대기
    else:
        time.sleep(delay_seconds)

except Exception as e:
    print(f"오류 발생 - {skill}: {e}")
    results.append({
        'skill_name': skill,

```

〈부표 2-2〉의 계속

```

        'ai_time_saving_pct': None,
        'ai_score_min': None,
        'ai_score_max': None,
        'ai_assessment_summary': f"평가 오류: {str(e)}"
    ))

# 전체 토론 내용을 하나의 JSON 파일로 저장
with open(os.path.join(output_dir, 'all_discussions.json'), 'w',
encoding='utf-8') as f:
    json.dump(all_discussions, f, ensure_ascii=False, indent=2)

# 결과를 데이터프레임으로 변환
results_df = pd.DataFrame(results)

# 원본 데이터프레임과 병합
df_with_ai = df.merge(results_df, left_on=skill_column,
right_on='skill_name', how='left')

# 중복 열 제거 (skill_name이 이미 있는 경우)
if 'skill_name' in df_with_ai.columns and skill_column != 'skill_name':
    df_with_ai = df_with_ai.drop('skill_name', axis=1)

# 결과 요약 출력
print(f"\n평가 완료!")
print(f"- 성공: {results_df['ai_time_saving_pct'].notna().sum()}개")
print(f"- 실패: {results_df['ai_time_saving_pct'].isna().sum()}개")
print(f"- 평균 AI 시간 절감률: {results_df['ai_time_saving_pct'].mean():.1f}%")
print(f"- 결과 저장 위치: {output_dir}/")

return df_with_ai

```

〈부표 2-2〉의 계속

```

def _create_filename(self, skill_name: str) -> str:
    """
    스킬 이름을 파일명으로 변환 (slug 형식)

    Args:
        skill_name: 원본 스킬 이름

    Returns:
        파일명으로 사용 가능한 문자열
    """
    import re
    # 한글은 유지하되, 특수문자는 하이픈으로 변경
    filename = re.sub(r'[\w가-힣\s-]', '', skill_name)
    filename = re.sub(r'[-\s]+', '-', filename)
    return filename.lower()

```

자료: 저자 작성.

〈부표 5-1〉 OECD AI 키워드 리스트

OECD AI 키워드	한글	설명	분류 번호
K-means	K-평균	K-means는 머신러닝에서 널리 사용되는 클러스터링 알고리즘으로, 데이터 포인트를 유사성에 따라 지정된 수의 클러스터로 그룹화하여 데이터 내 패턴을 식별하는 데 도움을 준다.	1
LoRA	LoRA	LoRA(Low-Rank Adaptation)는 머신러닝 기술로, 소수의 매개변수만 조정하여 모델을 효율적으로 미세 조정할 수 있게 하여 더 빠르고 자원 소모가 적다.	1
lpboost	LPBoost	Lpboost는 부스팅 기법을 활용해 약한 분류기들을 결합하여 더 강력한 모델로 성능을 향상시키는 머신러닝 알고리즘으로, 대규모 데이터 처리와 특정 손실 함수 최적화에 특화되어 있다.	1
mlpack	mlpack	mlpack은 머신러닝 모델을 효율적으로 구축 및 배포하기 위한 다양한 알고리즘과 도구를 제공하는 빠르고 유연한 C++ 머신러닝 라이브러리이다.	1
Q-learning	Q-러닝	Q-학습은 강화 학습 알고리즘의 일종으로, 에이전트가 시행착오를 통해 보상을 극대화함으로써 주어진 상황에서 취해야 할 최선의 행동을 학습하도록 돕는다.	1
RLHF	RLHF	RLHF는 인간 피드백 기반 강화 학습(Reinforcement Learning from Human Feedback)을 의미하며, 모델이 인간 사용자로부터 지침과 피드백을 받아 더 나은 결정을 내리는 방법을 학습하는 인공지능 기법이다.	1
xgboost	XGBoost	XGBoost는 분류 및 회귀 작업에 사용되는 강력한 머신러닝 알고리즘으로, 부스팅(boosting)이라는 과정을 통해 여러 의사 결정 트리의 예측을 결합하여 예측 정확도를 향상시킨다.	1
gaussian process	가우시안 프로세스	가우시안 프로세스는 불확실한 결과에 대해 예측을 수행하기 위해 특정 분포를 따르는 무작위 변수들의 집합으로 처리하는 머신러닝 통계 기법으로, 데이터 트렌드와 패턴을 유연하게 모델링할 수 있게 한다.	1
gaussian mixture model	가우시안 혼합 모델	가우시안 혼합 모델은 복잡한 데이터 분포를 여러 개의 단순한 종형 곡선 조합으로 표현하는 기계 학습 통계 기법으로, 데이터 내 서로 다른 그룹을 식별할 수 있게 한다.	1
supervised learning	감독 학습	지도 학습은 라벨링된 데이터로 모델을 훈련시키는 인공지능 유형으로, 입력 데이터와 정답 출력이 모두 포함된 예시를 기반으로 예측이나 결정을 내리는 법을 학습한다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
reinforcement learning	강화 학습	강화 학습은 환경 내 행동에 따른 보상 또는 벌점을 받아 의사 결정을 학습하는 기계 학습의 한 유형으로, 인간이 경험을 통해 학습하는 방식과 유사하다.	1
ant colony optimisation	개미 군집 최적화	개미 군집 최적화(Ant colony optimization)는 개미의 먹이 탐색 행동에서 영감을 받은 계산 알고리즘으로, 개미가 먹이를 찾고 서로 소통하는 방식을 모방하여 복잡한 문제에 대한 효율적인 해결책을 찾는 데 사용된다.	1
decision tree	의사 결정 트리	의사 결정 트리는 인공지능 및 기계 학습에서 사용되는 플로우차트와 유사한 모델로, 다양한 입력 조건에 기반하여 가능한 여러 결과를 매핑함으로써 의사 결정을 돕는다.	1
computational intelligence	계산 지능	계산 지능은 신경망, 퍼지 논리, 진화 알고리즘 등 자연적 과정에 영감을 받은 기법을 활용해 복잡한 문제 해결에 초점을 맞춘 인공지능의 한 분야를 의미한다.	1
hierarchical clustering	계층적 클러스터링	계층적 클러스터링은 유사한 데이터 포인트를 트리 구조로 그룹화하는 방법으로, 트리의 각 수준은 데이터 간의 서로 다른 유사성 수준을 나타낸다.	1
relational learning	관계 학습	관계 학습은 개별 데이터 포인트를 독립적으로 분석하기보다 서로 다른 데이터 조각 간의 관계와 연결을 이해하고 예측하는 데 중점을 둔 기계 학습의 한 유형이다.	1
rule-based learning	규칙 기반 학습	규칙 기반 학습은 기계 학습처럼 데이터로부터 학습하기보다 사전 정의된 규칙과 논리를 사용하여 의사 결정을 내리거나 문제를 해결하는 인공지능의 한 유형이다.	1
rule learning	규칙 학습	규칙 학습은 데이터 내 패턴과 관계를 식별하여 의사 결정이나 예측에 활용할 수 있는 단순하고 해석 가능한 규칙을 생성하는 기계 학습 기법이다.	1
Gradient Boosting (Boosting Algorithms)	그라디언트 부스팅 (부스팅 알고리즘)	경사 부스팅은 일반적으로 의사 결정 트리와 같은 여러 약한 모델을 순차적으로 결합하여 강력한 예측 모델을 구축하는 기계 학습 기법으로, 각 새로운 모델이 이전 모델의 오류를 수정한다.	1
graphical model	그래픽 모델	그래픽 모델은 시스템 내 다양한 변수 간의 관계와 의존성을 그래프로 시각화하여 복잡한 데이터를 이해하고 분석하기 쉽게 만드는 표현 방식이다.	1
extreme machine learning	극한 기계 학습	익스트림 머신러닝은 고급 알고리즘과 대규모 데이터셋을 활용하여 신속하게 분석하고 예측을 수행하는, 머신러닝 모델 훈련을 위한 빠르고 확장 가능한 접근법이다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
machine intelligence	기계 지능	머신 인텔리전스는 학습, 추론, 문제 해결, 언어 이해 등 일반적으로 인간과 유사한 인지 기능이 필요한 작업을 컴퓨터 시스템이 수행할 수 있는 능력을 의미한다.	1
machine learning	기계 학습	머신러닝은 컴퓨터가 명시적으로 프로그래밍되지 않고도 데이터로부터 학습하여 작업 수행 능력을 향상시킬 수 있도록 하는 인공지능의 한 분야이다.	1
foundation models	기초 모델	파운데이션 모델은 방대한 데이터로 훈련된 대규모 AI 시스템으로, 언어 이해나 이미지 인식 등 다양한 작업에 적용될 수 있으며 보다 전문화된 애플리케이션 구축의 기반 역할을 한다.	1
multimodal-ai	다모달 AI	다중 모달 AI는 텍스트, 이미지, 오디오 등 여러 유형의 데이터 소스에서 정보를 처리하고 이해할 수 있는 인공지능 시스템을 의미하며, 이러한 입력의 조합이 필요한 작업을 수행할 수 있게 한다.	1
multi-label classification	다중 레이블 분류	다중 레이블 분류는 기계 학습 작업의 한 유형으로, 알고리즘이 데이터의 각 인스턴스에 하나의 레이블이 아닌 여러 레이블이나 범주를 할당하여 더 복잡하고 미묘한 분류를 가능하게 한다.	1
multi-layer perceptron	다층 퍼셉트론	다층 퍼셉트론은 입력 데이터를 처리하고 변환하여 예측이나 분류를 수행하는 상호 연결된 노드(또는 뉴런)의 여러 층으로 구성된 인공 신경망의 한 유형이다.	1
single-linkage clustering	단일 연결 클러스터링	단일 연결 군집화는 두 군집 사이의 거리를 서로 다른 군집 내 임의의 두 점 사이의 최단 거리로 정의하는 데이터 포인트 그룹화 방법이다.	1
deep learning	딥 러닝	딥 러닝은 다층 신경망을 활용하여 대량의 데이터를 분석하고 학습하는 인공지능의 하위 분야로, 컴퓨터가 패턴을 인식하고 인간 사고와 유사한 결정을 내릴 수 있게 한다.	1
deep belief network	딥 벨리프 네트워크	딥 신념 네트워크는 데이터 내 패턴을 인식하도록 학습하는 상호 연결된 노드의 다중 층으로 구성된 인공지능 모델 유형으로, 이미지 및 음성 인식과 같은 작업에 자주 사용된다.	1
random forest	랜덤 포레스트	랜덤 포레스트는 정확도를 높이고 과적합을 줄이기 위해 결 과값을 평균화하여 예측이나 분류를 수행하는 의사 결정 트리 집합을 사용하는 기계 학습 기법이다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
random field	랜덤 필드	랜덤 필드는 공간에 걸쳐 확률 변수가 어떻게 분포하는지 설명하는 수학적 모델로, 이미지 처리 및 공간 통계학 같은 분야에서 데이터의 패턴과 관계를 분석하는 데 자주 사용된다.	1
memetic algorithm	미메틱 알고리즘	미메틱 알고리즘은 유전자 알고리즘과 국소 탐색 기법을 결합한 최적화 기법으로, 문화 속에서 아이디어가 진화하고 적응하는 방식을 모방하여 해법을 개선한다.	1
model-compression	모델 압축	모델 압축은 성능을 유지하면서 머신러닝 모델의 크기와 복잡성을 줄여 배포 속도와 효율성을 높이는 인공지능 기술이다.	1
semi-supervised learning	반감독 학습	반지도 학습은 소량의 라벨링된 데이터와 대량의 라벨링되지 않은 데이터를 함께 사용하여 모델의 정확도와 성능을 향상시키는 머신러닝 유형이다.	1
backpropagation	백프로퍼게이션	역전파는 인공 신경망 훈련에 사용되는 방법으로, 예측 출력과 실제 출력 간의 차이를 최소화하기 위해 신경세포 간 연결 가중치를 조정하며, 오류를 네트워크를 통해 역방향으로 전파하는 과정을 포함한다.	1
Bayesian Models and Learning	베이시안 모델 및 학습	베이시안 모델 및 학습은 인공지능에서 베이지 정리를 활용하여 더 많은 증거나 정보가 확보됨에 따라 가설의 확률을 업데이트하는 통계적 접근법을 의미하며, 이를 통해 기계가 정보에 기반한 예측과 결정을 내릴 수 있게 한다.	1
vector databases	벡터 데이터베이스	벡터 데이터베이스는 벡터라는 수학적 표현을 사용하여 머신러닝 모델의 임베딩과 같은 고차원 데이터를 효율적으로 관리하고 검색하도록 설계된 특수 저장 시스템이다.	1
vector machine	벡터 머신	벡터 머신은 머신러닝에서 사용되는 알고리즘 유형으로, 다차원 공간에서 서로 다른 범주를 구분하는 최적의 경계(또는 초평면)를 찾아 데이터를 분류한다.	1
variational inference	변분 추론	변분 추론은 복잡한 확률 분포를 더 단순하고 관리하기 쉬운 형태로 변환하여 근사화하는 기계 학습 통계 기법으로, 더 빠르고 효율적인 데이터 분석을 가능하게 한다.	1
bootstrap aggregation	부트스트랩 집계	부트스트랩 집계(흔히 배깅이라고도 함)는 데이터의 서로 다른 무작위 부분 집합으로 모델의 여러 버전을 훈련시킨 후 그 예측을 결합하여 모델 정확도를 향상시키는 기계 학습 기법이다.	1
classification tree	분류 트리	분류 트리는 머신러닝에서 사용되는 의사 결정 도구 유형으로, 데이터를 트리 구조로 구성하여 항목의 특징에 기반해 범주화하거나 분류한다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
brownboost	브라운부스트	브라운부스트는 머신러닝 기법으로, 모델을 추가 데이터나 특징과 결합하여 성능을 향상시키고 종종 정확한 예측 능력을 개선한다.	1
unsupervised learning	비지도 학습	비지도 학습은 알고리즘이 특정 레이블이나 탐색 지침 없이 데이터 내 패턴을 분석하고 식별하는 인공지능의 한 유형이다.	1
generative adversarial network	생성적 적대적 네트워크	생성적 적대 신경망(GAN)은 기존 예시로부터 학습하여 사실적인 이미지, 소리 또는 기타 데이터를 생성하기 위해 서로 대립하는 두 신경망으로 구성된 인공지능이다.	1
generative-ai	생성형 AI	생성형 AI는 기존 데이터의 패턴을 학습하여 텍스트, 이미지, 음악 등 새로운 콘텐츠를 생성할 수 있는 인공지능 유형을 의미한다.	1
hidden Markov model	숨겨진 마르코프 모델	숨겨진 마르코프 모델은 상태가 직접 관측되지 않지만 관측 가능한 데이터로부터 추론할 수 있는 일련의 사건을 따르는 시스템을 분석하는 데 사용되는 통계 도구이다.	1
stacked generalisation	스택드 일반화	스택드 일반화(Stacked generalization)는 여러 모델의 출력을 최종 모델의 입력으로 사용하여 예측 정확도를 향상시키는 머신러닝 기법이다.	1
temporal difference learning	시간 차분 학습	시간 차분 학습은 강화 학습에서 사용되는 기계 학습 기법으로, 에이전트가 시간에 따라 현재 예측과 실제 결과를 비교하여 미래 보상을 예측하는 법을 학습한다.	1
adaboost	에이다부스트	에이다부스트는 여러 약한 분류기를 결합하여 강력한 분류기를 생성하는 기계 학습 기법으로, 이전 모델의 오류를 집중적으로 보완하여 예측 정확도를 향상시킨다.	1
ensemble learning	앙상블 학습	앙상블 학습은 예측이나 의사 결정 시 전반적인 성능과 정확도를 높이기 위해 여러 모델을 결합하는 인공지능 기법이다.	1
quantization	양자화	양자화는 머신러닝 모델에 사용되는 숫자의 정밀도를 낮추는 과정으로, 허용 가능한 정확도를 유지하면서 모델 크기를 줄이고 속도를 향상시키는 데 도움이 된다.	1
federated-learning	연방 학습	연방 학습은 여러 장치가 데이터를 공유하지 않고 각자의 로컬 데이터를 사용하여 협업적으로 모델을 훈련할 수 있게 하는 머신러닝 접근법으로, 프라이버시와 보안을 강화한다.	1
similarity learning	유사도 학습	유사도 학습은 항목 간의 유사성 또는 차이점을 인식하도록 모델을 훈련시키는 머신러닝 유형으로, 이미지 인식이나 추천 시스템과 같은 작업에 자주 사용된다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
genetic algorithm	유전 알고리즘	유전자 알고리즘은 자연 선택 과정을 모방한 문제 해결 방법으로, 잠재적 해결책이 세대를 거쳐 진화하여 최상의 또는 가장 효율적인 답을 찾는다.	1
artificial neural network	인공 신경망	인공 신경망은 인간의 뇌에서 영감을 받아 신경 세포의 작동 방식을 모방하여 정보를 처리하는 컴퓨터 시스템으로, 데이터로부터 학습하여 의사 결정이나 예측을 수행할 수 있다.	1
instance-based learning	인스턴스 기반 학습	인스턴스 기반 학습은 일반적인 모델을 생성하기보다 훈련 데이터의 특정 사례를 기반으로 예측을 수행하는 기계 학습 유형으로, 유사도 측정을 활용하여 새로운 입력에 가장 가까운 사례를 찾는 경우가 많다.	1
factorisation machine	인자 분해 기계	인자 분해 기계는 대규모 데이터셋 내 변수 간 상호작용을 효율적으로 포착하는 머신러닝 모델로, 추천 시스템 및 예측 분석에 주로 활용된다.	1
cognitive modelling	인지 모델링	인지 모델링은 인간의 사고 과정을 모방한 컴퓨터 시뮬레이션을 생성하여 사람들이 어떻게 생각하고, 배우고, 결정을 내리는지 이해하는 과정이다.	1
cognitive insight system	인지 인사이트 시스템	인지 통찰 시스템은 인공지능을 활용해 데이터를 분석하고 의미 있는 통찰력을 제공하여 패턴과 추세를 기반으로 더 나은 의사 결정을 내릴 수 있도록 돕는 첨단 기술이다.	1
cognitive computing	인지 컴퓨팅	인지 컴퓨팅은 복잡한 데이터 분석에서 인간의 사고 과정을 모방하는 기술을 의미하며, 기계가 보다 인간과 유사한 방식으로 이해하고 학습하며 상호작용할 수 있게 한다.	1
autoencoder	자동 인코더	자동 인코더는 데이터의 효율적인 표현을 학습하는 인공 신경망의 한 유형으로, 일반적으로 차원 축소 또는 특징 추출을 목적으로 입력 데이터를 더 작은 형식으로 압축한 후 원래 형태로 재구성한다.	1
natural gradient	자연 기울기	자연 기울기는 기계 학습에서 데이터 공간의 기하학적 구조에 따라 매개변수 업데이트 방향을 조정하여 더 효율적이고 효과적인 학습을 이끌어내는 최적화 기법이다.	1
latent variable	잠재 변수	잠재 변수는 관측 가능한 데이터나 결과에 영향을 미치는 관측 불가능한 요소나 특성으로, 숨겨진 관계를 설명하기 위해 통계 모델에서 흔히 사용된다.	1
adversarial network	적대적 네트워크	적대적 네트워크는 두 개의 경쟁 모델로 구성된 인공지능 시스템으로, 하나는 데이터를 생성하고 다른 하나는 이를 평가하여 지속적인 경쟁을 통해 생성된 데이터의 품질과 현실감을 향상시킨다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
adaptive boosting	적응형 부스팅	적응형 부스팅(AdaBoost)은 여러 약한 분류기를 결합하여 이전 분류기들의 오류를 집중적으로 활용함으로써 강력한 예측 모델을 생성하는 기계 학습 기법이다.	1
policy gradient methods	정책 기울기 방법	정책 기울기 방법은 강화 학습 기법의 일종으로, 환경으로부터 받은 보상에 따라 에이전트의 행동을 조정함으로써 의사 결정 전략(정책)을 직접 최적화한다.	1
gravitational search algorithm	중력 검색 알고리즘	중력 탐색 알고리즘은 중력의 법칙에서 영감을 받은 최적화 기법으로, 잠재적 해법을 서로를 끌어당기는 질량으로 간주하여 탐색 공간에서 이 질량들의 움직임은 시뮬레이션함으로써 문제에 대한 최상의 해법을 찾는 데 도움을 준다.	1
intelligent classifier	지능형 분류기	지능형 분류기는 학습된 패턴과 특징을 기반으로 데이터를 자동으로 분류하고 서로 다른 그룹으로 구분하는 인공지능 시스템의 일종이다.	1
support vector machine	지원 벡터 머신	지원 벡터 머신은 새로운 데이터를 정확하게 분류하기 위해 서로 다른 범주의 데이터 포인트를 분리하는 최적의 경계를 찾는 기계 학습 알고리즘의 일종이다.	1
support vector regression	지원 벡터 회귀	지원 벡터 회귀는 지정된 허용 오차 범위 내에서 오차를 최소화하는 최적의 적합 선 또는 곡선을 찾아 연속적인 결과를 예측하는 기계 학습 기법의 일종이다.	1
evolutionary algorithm	진화 알고리즘	진화 알고리즘은 자연 선택 과정을 모티브로 한 문제 해결 방법으로, 잠재적 해결책이 돌연변이, 교차, 선택 등의 기법을 통해 시간이 지남에 따라 진화하여 최상의 결과를 찾는 방식이다.	1
differential evolution algorithm	차분진화 알고리즘	차등 진화 알고리즘은 후보 해결책 집단을 활용하여 그 특성을 결합하고 변이시켜 문제를 해결하는 최적의 방안을 반복적으로 개선하는 최적화 기법이다.	1
nearest neighbour algorithm	최근 이웃 알고리즘	최근점 이웃 알고리즘은 기계 학습 및 데이터 분석에서 사용되는 방법으로, 특징 공간 내 가장 가까운 데이터 포인트를 기반으로 데이터 포인트를 분류한다.	1
recommender system	추천 시스템	추천 시스템은 사용자의 선호도와 행동 패턴을 기반으로 제품, 서비스 또는 콘텐츠를 제안하는 인공지능의 한 유형이다.	1
Kernel learning	커널 학습	커널 학습은 수학적 함수를 사용하여 데이터를 고차원 공간으로 변환함으로써 복잡한 데이터셋 내 패턴과 관계를 식별하기 쉽게 만드는 기계 학습 기법이다.	1

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
feature learning	특징 학습	특징 학습은 인공지능에서 시스템이 원시 데이터로부터 중요한 특성이나 패턴을 자동으로 식별하고 추출하여 특정 작업에 대한 이해도와 성능을 향상시키는 과정이다.	1
firefly algorithm	파이어플라이 알고리즘	반딧불 알고리즘은 자연에서 영감을 받은 최적화 기법으로, 반딧불의 깜빡이는 행동을 모방하여 해결책의 “밝기” 또는 품질에 따라 해결책을 끌어모아 문제에 대한 최상의 해법을 찾는 방식이다.	1
Fuzzy Logic and Fuzzy Systems	퍼지 논리 및 퍼지 시스템	퍼지 논리와 퍼지 시스템은 고정적이고 정확한 추론이 아닌 근사적인 추론을 다루는 인공지능 접근법으로, 컴퓨터가 불확실성을 처리하고 부정확하거나 모호한 정보를 바탕으로 결정을 내릴 수 있게 한다.	1
evaluation-metrics	평가 지표	평가 지표는 AI 모델이나 알고리즘이 의도된 작업을 수행하는 데 있어 성능과 효과를 측정하기 위해 사용되는 정량적 척도이다.	1
hebbian learning	헤비안 학습	헤비안 학습은 신경과학과 인공지능 분야의 이론으로, 뉴런들이 함께 활성화될 때 그 연결이 강화된다는 것을 제안하며, 흔히 “함께 발화하는 세포들은 함께 연결된다”는 말로 요약된다.	1
collaborative filtering	협업 필터링	협업 필터링은 유사한 사용자들의 선호도와 행동을 기반으로 사용자에게 항목을 추천하는 기술이다.	1
stochastic gradient	확률적 기울기	확률적 경사는 전체 데이터셋이 아닌 작은 무작위 부분 집합을 기반으로 매개변수를 업데이트하여 모델을 최적화하는 머신러닝 기법으로, 학습 과정의 속도를 높이는 데 기여한다.	1
diffusion models	확산 모델	확산 모델은 입자가 매질 내에서 확산되는 방식을 모방한 과정을 통해 무작위 노이즈를 점진적으로 일관된 출력으로 변환함으로써 이미지나 소리 같은 새로운 데이터를 생성하는 머신러닝 알고리즘의 일종이다.	1
regression tree	회귀 트리	회귀 트리는 머신러닝에서 사용되는 의사 결정 트리의 한 유형으로, 특징값을 기반으로 데이터를 분기(branch)로 분할하여 연속적인 수치 값을 예측하며, 최종적으로 트리의 잎(leaf)에서 최종 예측을 도출한다.	1
Caffe	카페	카페는 신경망 구축 및 훈련을 위해 설계된 오픈소스 딥러닝 프레임워크로, 특히 이미지 분류 및 처리 작업에 적합하다.	2

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
CNTK	CNTK	CNTK(Microsoft Cognitive Toolkit)는 마이크로소프트가 개발한 오픈소스 딥러닝 프레임워크로, 사용자가 머신러닝 모델을 효율적으로 구축, 훈련 및 평가할 수 있게 한다.	2
Deep learning ^{4j}	Deep learning ^{4j}	Deeplearning ^{4j} 는 자바 프로그래밍 언어를 위한 오픈소스 딥러닝 라이브러리로, 개발자가 다양한 머신러닝 작업을 위한 신경망을 구축하고 훈련할 수 있게 한다.	2
H2O.ai	H2O.ai	H2O.ai는 인공지능과 머신러닝을 활용하여 기업이 데이터를 분석하고 예측 모델을 쉽게 구축할 수 있도록 지원하는 오픈소스 소프트웨어 플랫폼이다.	2
Kaldi	Kaldi	Kaldi는 음성 인식용 오픈소스 툴킷으로, 음성 언어를 텍스트로 변환하는 모델을 개발하고 훈련하기 위한 도구와 라이브러리를 제공한다.	2
Keras	Keras	Keras는 Python에서 딥러닝 모델 구축 및 훈련을 단순화하는 오픈소스 소프트웨어 라이브러리로, 개발자가 신경망을 더 쉽게 만들 수 있도록 한다.	2
LIBSVM	LIBSVM	LIBSVM은 서포트 벡터 머신(SVM) 분류 및 회귀를 위한 도구를 제공하는 오픈소스 소프트웨어 라이브러리로, 사용자가 다양한 데이터 분석 작업을 위한 머신러닝 모델을 쉽게 구현할 수 있도록 한다.	2
MXNet	MXNet	MXNet은 머신러닝 모델의 효율적인 훈련 및 배포를 위해 설계된 오픈소스 딥러닝 프레임워크로, 특히 확장성과 다중 프로그래밍 언어 지원으로 유명하다.	2
Paddle Paddle	Paddle Paddle	PaddlePaddle은 Baidu에서 개발한 오픈소스 딥러닝 플랫폼으로, 사용자가 머신러닝 모델을 효율적으로 구축, 훈련 및 배포할 수 있게 한다.	2
Python	Python	Python은 단순성과 가독성으로 유명한 인기 있는 프로그래밍 언어로, 인공지능 애플리케이션을 포함한 소프트웨어 개발에 널리 사용된다.	2
PyTorch	PyTorch	PyTorch는 개발자가 유연하고 동적인 계산 방식을 사용하여 신경망을 쉽게 구축하고 훈련할 수 있게 하는 오픈소스 머신러닝 라이브러리이다.	2
scikit-learn	scikit-learn	Scikit-learn은 데이터 분석 및 머신러닝을 위한 간단하고 효율적인 도구를 제공하는 인기 있는 오픈소스 파이썬 라이브러리로, 사용자가 예측 모델을 구축하고 평가하기 쉽게 한다.	2

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
TensorFlow	Tensor Flow	TensorFlow는 머신러닝 모델 구축 및 훈련을 위해 구글이 개발한 오픈소스 소프트웨어 라이브러리로, 개발자가 데이터로부터 학습하고 예측을 수행할 수 있는 애플리케이션을 만들 수 있게 한다.	2
Theano	Theano	Theano는 개발자가 특히 딥 러닝 및 신경망 애플리케이션을 위해 수학 표현식을 효율적으로 정의, 최적화 및 평가할 수 있도록 하는 오픈소스 소프트웨어 라이브러리이다.	2
Torch	Torch	Torch는 주로 Lua 프로그래밍 언어를 사용하지만 이후 Python으로 PyTorch로 전환된, 딥 러닝 모델 구축 및 훈련을 위한 유연한 프레임워크를 제공하는 오픈소스 머신러닝 라이브러리이다.	2
benchmarks	벤치마크	벤치마크는 AI 시스템의 성능을 확립된 기준이나 다른 시스템과 비교 평가하기 위해 사용되는 표준화된 테스트 또는 측정 방법이다.	2
intelligent infrastructure	지능형 인프라	지능형 인프라란 교통, 에너지, 통신 네트워크와 같은 공공 서비스 및 도시 환경의 효율성, 안전성, 지속 가능성을 향상시키기 위해 데이터와 인공지능을 활용하는 첨단 시스템 및 기술을 의미한다.	2
retrieval-augmented generation	검색 강화 생성	검색 강화 생성은 텍스트 생성 능력과 데이터베이스에서 관련 정보를 검색하는 능력을 결합하여 생성된 콘텐츠의 품질과 정확성을 향상시키는 인공지능 기술이다.	3
inductive monitoring	귀납적 모니터링	귀납적 모니터링은 사전 정의된 규칙에 의존하기보다 관찰된 사례를 바탕으로 패턴을 인식하고 예측을 수행하도록 학습하는 인공지능 방법이다.	3
multi-objective evolutionary algorithm	다목적 진화 알고리즘	다중 목표 진화 알고리즘은 자연 선택의 원리를 활용하여 여러 상충되는 목표나 목적 사이의 균형을 동시에 찾는 최적화 방법의 일종이다.	3
multi-objective optimisation	다목적 최적화	다중 목표 최적화는 여러 상충되는 목표를 동시에 고려하여 이러한 목표들 사이의 균형을 이루는 최상의 해결책을 찾는 과정이다.	3
multi task learning	다중 작업 학습	다중 작업 학습은 모델이 공유된 정보를 활용하고 전반적인 성능을 향상시켜 여러 작업을 동시에 수행하도록 훈련하는 인공지능 접근법이다.	3

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
meta learning	메타 학습	메타 학습은 알고리즘이 이전 경험을 분석하고 새로운 작업에 더 효율적으로 적응함으로써 자체 학습 과정을 개선하는 방법을 학습하는 기계 학습 유형이다.	3
instruction tuning	명령어 조정	지시문 튜닝은 AI 모델이 사용자가 제공하는 특정 지시문이나 프롬프트를 더 잘 이해하고 따르도록 훈련하는 과정으로, 작업을 정확하고 효과적으로 수행하는 능력을 향상시킨다.	3
model evaluation	모델 평가	모델 평가는 특정 작업에서 인공지능 모델의 성능을 평가하는 과정으로, 일반적으로 다양한 지표를 사용하여 예측 결과와 실제 결과를 비교한다.	3
fine-tuning	미세 조정	미세 조정은 사전 훈련된 AI 모델을 가져와 특정 작업이나 데이터셋에서 성능을 개선하기 위해 작은 조정을 가하는 과정이다.	3
motion planning	모션 플래닝	모션 플래닝은 AI 시스템이 로봇이나 자율주행 차량이 장애물을 피하면서 한 지점에서 다른 지점으로 이동하는 데 필요한 일련의 동작이나 행동을 결정하는 과정이다.	3
genetic programming	유전 프로그래밍	유전적 프로그래밍은 선택과 돌연변이와 같은 진화 원리를 사용하여 특정 문제를 해결할 수 있는 컴퓨터 프로그램을 자동으로 생성하는 인공지능의 한 유형이다.	3
evolutionary computation	진화 컴퓨팅	진화 계산은 돌연변이, 교차, 선택과 같은 메커니즘을 통해 시간이 지남에 따라 해법을 진화시켜 복잡한 문제를 해결하는 자연 선택 과정을 모방하는 인공지능의 한 유형이다.	3
optimization	최적화	AI에서의 최적화는 가능한 최상의 결과를 달성하기 위해 시스템이나 모델의 매개변수나 전략을 조정하여 최대한 효과적이거나 효율적으로 만드는 과정을 의미한다.	3
prompt engineering	프롬프트 공학	프롬프트 엔지니어링은 AI 모델에 입력되는 질문이나 문장을 설계하고 개선하여 가장 정확하고 관련성 높은 응답을 이끌어내는 과정이다.	3
Julia	줄리아	줄리아는 고성능 수치 및 과학적 계산을 위해 설계된 고수준 프로그래밍 언어로, 속도와 사용 편의성으로 유명하여 데이터 과학자와 연구자들 사이에서 널리 사용된다.	4
Jupyter	주피터	주피터는 라이브 코드, 방정식, 시각화 자료 및 서술적 텍스트를 포함한 문서를 생성하고 공유할 수 있는 오픈소스 웹 애플리케이션으로, 데이터 분석 및 과학적 계산에 널리 사용된다.	4

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
MapReduce	Map Reduce	MapReduce는 작업을 두 가지 주요 작업으로 분할하여 대규모 데이터셋을 처리하고 생성하는 데 사용되는 프로그래밍 모델이다: 데이터를 정렬하고 필터링하는 매핑(mapping)과 결과를 집계하는 리듀싱(reducing) 작업이다.	4
Matlab	Matlab	Matlab은 수치 계산, 데이터 분석 및 알고리즘 개발에 사용되는 프로그래밍 플랫폼으로, 공학 및 과학 연구 분야에서 널리 활용된다.	4
R	R	R은 주로 통계 계산 및 데이터 분석에 사용되는 프로그래밍 언어이자 소프트웨어 환경으로, 데이터 과학자와 연구자들 사이에서 인기가 높다.	4
RapidMiner	Rapid Miner	RapidMiner는 사용자가 광범위한 프로그래밍 기술 없이도 데이터를 준비하고 예측 모델을 생성하며 결과를 분석할 수 있도록 지원하는 사용자 친화적인 데이터 과학 플랫폼이다.	4
Weka	Weka	Weka는 데이터 마이닝 작업을 위한 기계 학습 알고리즘 모음을 제공하는 오픈소스 소프트웨어 도구로, 사용자가 데이터를 쉽게 분석하고 시각화할 수 있게 한다.	4
high-dimensional data	고차원 데이터	고차원 데이터는 특징이나 변수의 수가 많아 분석이 복잡하고 효과적으로 처리하기 어려운 정보를 의미한다.	4
high-dimensional data and model	고차원 데이터 및 모델	고차원 데이터는 특징이나 변수의 수가 많아 분석이 복잡한 데이터셋을 가리키며, 고차원 모델은 이러한 데이터를 효과적으로 처리하고 해석하기 위해 설계된 수학적 또는 계산적 프레임워크이다.	4
data mining	데이터 마이닝	데이터 마이닝은 의사 결정에 활용할 수 있는 패턴, 추세 및 가치 있는 정보를 발견하기 위해 대규모 데이터셋을 분석하는 과정이다.	4
dynamic time warping	동적 시간 왜곡	동적 시간 왜곡은 시간이나 속도가 다를 수 있는 두 시퀀스 간의 유사성을 측정하여 데이터 포인트를 유연하게 정렬할 수 있도록 하는 데이터 분석 기법이다.	4
association rule	연관 규칙	연관 규칙은 대규모 데이터셋 내 변수 간 흥미로운 관계나 패턴을 발견하는 데이터 마이닝 기법으로, “고객이 빵을 구매하면 버터도 구매할 가능성이 높다”와 같은 ‘조건-결과’ 형태로 표현되는 경우가 많다.	4
causal-inference	인과 추론	인과 추론은 한 변수의 변화가 다른 변수의 변화와 단순히 상관관계나 연관성을 가지는 것이 아니라 직접적인 원인이 되는지 판단하는 과정이다.	4

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
dimensionality reduction	차원 축소	차원 축소는 데이터 분석 기법으로, 필수 정보를 유지하면서 변수나 특징의 수를 줄여 복잡한 데이터셋을 단순화하여 데이터 시각화 및 처리를 용이하게 한다.	4
community detection	커뮤니티 탐지	커뮤니티 탐지는 네트워크 내에서 노드(사람이나 컴퓨터 등)가 그룹 외부보다 그룹 내부에 더 밀접하게 연결된 집단이나 클러스터를 식별하는 과정이다.	4
cluster analysis	클러스터 분석	군집 분석은 특성에 따라 유사한 항목이나 데이터 포인트를 그룹화하여 데이터셋 내 패턴이나 구조를 식별하는 데 도움을 주는 통계 기법이다.	4
feature engineering	특징 공학	특징 공학은 머신러닝 모델의 성능을 향상시키기 위해 원시 데이터에서 입력 변수(특징)를 선택, 수정 또는 생성하는 과정이다.	4
feature selection	특징 선택	특징 선택은 머신러닝 모델의 성능을 향상시키기 위해 데이터셋에서 가장 중요한 변수나 속성을 식별하고 선택하는 과정이다.	4
sparse representation	희소 표현	스파스 표현은 소수의 중요한 요소나 특징만을 사용하여 데이터를 표현하는 방법으로, 복잡성을 줄이면서 분석과 처리를 용이하게 한다.	4
fastText	fastText	fastText는 Facebook이 개발한 오픈소스 라이브러리로, 단어를 고차원 공간의 벡터로 표현하여 효율적으로 단어 임베딩을 생성하고 텍스트 분류를 수행할 수 있게 한다.	5
Hugging Face Transformers	Hugging Face Transformers	허깅 페이스 트랜스포머는 자연어 처리 작업을 위한 사전 훈련된 모델과 도구를 제공하는 인기 있는 오픈소스 라이브러리로, 개발자가 인간 언어를 이해하고 생성하는 애플리케이션을 더 쉽게 구축할 수 있게 한다.	5
llm(large language model)	대규모 언어 모델(LLM)	대규모 언어 모델은 방대한 양의 텍스트 데이터를 활용하여 인간과 유사한 언어를 이해하고 생성하는 고급 인공지능 유형으로, 질문에 답하고, 에세이를 작성하며, 대화에 참여하는 등의 작업을 수행할 수 있다.	5
NLTK	NLTK	NLTK(Natural Language Toolkit)는 인간 언어 데이터 작업을 위한 인기 있는 Python 라이브러리로, 텍스트 처리, 분석, 자연어 이해 등의 작업을 위한 도구를 제공한다.	5
OpenNLP	OpenNLP	OpenNLP는 자연어 텍스트 처리를 위한 도구를 제공하는 오픈소스 소프트웨어 라이브러리로, 토큰화, 품사 태깅, 명명된 엔티티 인식 등의 작업을 가능하게 한다.	5

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
spaCy	spaCy	spaCy는 컴퓨터가 인간 언어를 효율적으로 이해하고 분석할 수 있도록 설계된, Python 기반의 고급 자연어 처리를 위한 오픈소스 소프트웨어 라이브러리이다.	5
sentiment analysis	감정 분석	감정 분석은 인공지능을 활용하여 단어 뒤에 숨겨진 감정적 어조를 판단하는 기술로, 텍스트가 긍정적, 부정적 또는 중립적인 감정을 표현하는지 식별하는 데 도움을 준다.	5
graph-neural-networks	그래프 신경 네트워크	그래프 신경망은 그래프 구조로 구성된 데이터를 분석하고 학습하도록 설계된 인공지능 모델의 한 유형으로, 노드(점)와 에지(연결선)로 구성되어 개체 간의 관계와 상호작용을 포착한다.	5
machine translation	기계 번역	기계 번역은 인간의 개입 없이 컴퓨터 소프트웨어를 사용하여 텍스트나 음성을 한 언어에서 다른 언어로 자동 번역하는 것이다.	5
link prediction	링크 예측	링크 예측은 소셜 미디어 사용자나 웹 페이지와 같은 네트워크 내 두 개체 간 연결 또는 관계 형성 가능성을 추정하는 인공지능 기법이다.	5
non negative matrix factorisation	비음수 행렬 분해	비음수 행렬 분해는 대규모 행렬을 모든 값이 비음수인 두 개의 소규모 행렬로 분해하는 수학적 기법으로, 이미지나 텍스트와 같은 데이터의 패턴이나 특징을 식별하는 데 도움이 된다.	5
dictionary learning	사전 학습	사전 학습(Dictionary learning)은 데이터에서 대표 패턴 또는 “사전” 요소의 집합을 찾아내어 원본 데이터를 효율적으로 재구성하거나 표현하는 데 활용하는 머신러닝 기법이다.	5
semantic web	시맨틱 웹	시맨틱 웹은 기계가 데이터의 의미를 이해하고 해석할 수 있도록 하여 더 스마트한 정보 공유와 서로 다른 데이터 조각 간의 의미 있는 연결을 가능하게 하는 인터넷의 확장이다.	5
neuromorphic computing	신경 모방 컴퓨팅	뉴로모픽 컴퓨팅은 인간의 뇌가 정보를 처리하는 방식을 모방한 컴퓨팅 방식으로, 학습 및 패턴 인식과 같은 작업의 효율성과 성능을 향상시키기 위해 특수 하드웨어를 사용한다.	5
neural turing machine	신경 튜링 머신	신경 튜링 머신은 신경망과 외부 메모리 자원을 결합한 인공지능 모델로, 인간의 뇌와 유사하게 정보를 저장하고 검색함으로써 복잡한 작업을 학습하고 수행할 수 있다.	5

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
speech-to-text	음성 텍스트 변환	음성 인식 기술은 말한 언어를 텍스트로 변환하여 대화, 받아쓰기 또는 오디오 녹음을 읽을 수 있는 형식으로 기록할 수 있게 한다.	5
natural language generation	자연 언어 생성	자연어 생성은 구조화된 데이터로부터 인간과 유사한 텍스트를 생성하는 데 초점을 맞춘 인공지능의 한 분야로, 컴퓨터가 일관성 있고 문맥적으로 적절한 서면 콘텐츠를 생성할 수 있게 한다.	5
natural language understanding	자연 언어 이해	자연어 이해는 컴퓨터가 인간의 언어를 의미 있고 유용한 방식으로 이해하고 해석할 수 있게 하는 인공지능의 한 분야이다.	5
natural language processing	자연 언어 처리	자연어 처리는 컴퓨터가 인간의 언어를 의미 있고 유용한 방식으로 이해하고 해석하며 응답할 수 있도록 하는 인공지능의 한 분야이다.	5
latent dirichlet allocation	잠재 디리클레 할당	잠재 디리클레 할당(LDA)은 문서 집합 내에서 단어 동시 출현 패턴을 식별하여 주제를 발견하는 데 사용되는 자연어 처리의 통계 모델이다.	5
latent semantic analysis	잠재 의미 분석	잠재 의미 분석(LSA)은 대규모 텍스트 데이터셋에서 단어와 개념 간의 관계를 분석하여 숨겨진 의미와 패턴을 발견하는 자연어 처리 기법이다.	5
long short term memory (LSTM)	장단기 기억 (LSTM)	장기 단기 기억(LSTM)은 정보를 장기간 기억하도록 설계된 인공 신경망의 한 유형으로, 음성 인식 및 언어 번역과 같은 시퀀스 관련 작업에 특히 효과적이다.	5
transfer learning	전환 학습	전이 학습은 한 작업에 대해 개발된 모델을 다른 관련 작업에 재사용하거나 적용하는 기계 학습 기법으로, 이전에 습득한 지식을 활용하여 더 효율적으로 학습할 수 있게 한다.	5
topic model	주제 모델	주제 모델은 자연어 처리에서 텍스트 집합에 존재하는 주요 주제나 테마를 식별하고 분류하는 데 사용되는 통계 모델의 한 유형이다.	5
chatbot	챗봇	챗봇은 인간 사용자와의 대화를 시뮬레이션하도록 설계된 컴퓨터 프로그램으로, 주로 텍스트나 음성을 통해 정보를 제공하거나 질문에 답변하며 작업을 지원한다.	5
text mining	텍스트 마이닝	텍스트 마이닝은 다양한 기법과 알고리즘을 활용하여 대량의 비정형 텍스트 데이터에서 유용한 정보와 통찰력을 추출하는 과정이다.	5

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
text-to-speech	텍스트 음성 변환	텍스트 음성 변환은 작성된 텍스트를 음성으로 변환하는 기술로, 컴퓨터가 자연스러운 음성으로 텍스트를 읽어줄 수 있게 한다.	5
transformers	트랜스포머	트랜스포머는 데이터를 병렬 처리하는 신경망 아키텍처의 일종으로, 특히 인간 언어의 이해 및 생성에 효과적이며 번역 및 텍스트 생성 같은 작업의 발전을 가능하게 한다.	5
OpenCV	OpenCV	OpenCV는 기계가 시각 정보를 이해하고 해석할 수 있도록 이미지 및 동영상 처리 도구와 기능을 제공하는 오픈소스 컴퓨터 비전 라이브러리이다.	6
OpenVINO	Open VINO	OpenVINO는 인텔이 개발한 오픈소스 툴킷으로, 개발자가 다양한 하드웨어 플랫폼에서 컴퓨터 비전 애플리케이션을 위한 딥러닝 모델을 최적화하고 배포하는 데 도움을 준다.	6
emotion recognition	감정 인식	감정 인식은 컴퓨터가 얼굴 표정, 목소리 톤 및 기타 단서를 기반으로 인간의 감정을 식별하고 해석할 수 있게 하는 기술이다.	6
computational pathology	계산 병리학	계산 병리학은 컴퓨터 알고리즘과 인공지능을 사용하여 조직 샘플의 의료 영상 및 데이터를 분석하고 해석함으로써 질병을 보다 정확하고 효율적으로 진단하는 데 도움을 주는 분야이다.	6
machine vision	기계 비전	머신 비전은 컴퓨터와 시스템이 인간의 눈과 뇌가 이미지를 처리하는 방식과 유사하게 세계의 시각 정보를 해석하고 이해할 수 있게 하는 기술이다.	6
brain computer interface	뇌-컴퓨터 인터페이스	뇌-컴퓨터 인터페이스는 뇌와 외부 장치 간의 직접적인 통신을 가능하게 하여 사용자가 생각만으로 컴퓨터나 다른 기계를 제어할 수 있게 하는 기술이다.	6
independent component analysis	독립 성분 분석	독립 성분 분석(ICA)은 다변량 신호를 통계적으로 독립적인 가산 성분으로 분리하는 계산 기법으로, 신호 처리 및 데이터 분석과 같은 분야에서 서로 다른 정보 소스를 분리하는 데 자주 적용된다.	6
blind signal separation	맹신호 분리	맹신호 분리는 신호 처리에서 혼합된 신호를 원본 구성 요소로 분리하는 기술로, 소스 신호나 혼합 과정에 대한 사전 지식 없이 수행된다.	6
object recognition	물체 인식	물체 인식은 컴퓨터가 이미지나 영상 내 물체를 식별하고 분류할 수 있게 하는 기술로, 인간이 보는 것을 인식하고 이해하는 방식과 유사하다.	6

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류번호
object detection	물체 탐지	객체 탐지는 컴퓨터 비전 기술로, 기계가 이미지나 영상 내의 객체를 식별하고 위치를 파악할 수 있게 한다.	6
pedestrian detection	보행자 탐지	보행자 탐지는 컴퓨터와 차량이 주변을 걷고 있는 사람을 식별하고 위치를 파악할 수 있게 하여 안전성과 내비게이션을 향상시키는 기술이다.	6
vision-transformer	비전 트랜스포머	비전 트랜스포머는 이미지를 작은 패치로 분할하고 어텐션 메커니즘을 사용하여 시각 정보를 이해하고 분석하는 인공지능 모델의 한 유형으로, 언어 모델이 텍스트를 처리하는 방식과 유사하다.	6
spectral clustering	스펙트럼 클러스터링	스펙트럴 클러스터링은 유사도 행렬의 고유값과 고유벡터를 기반으로 데이터 포인트를 클러스터로 그룹화하는 기계 학습 기법으로, 데이터 내 복잡한 구조를 식별할 수 있게 한다.	6
visual servoing	시각적 서보링	비주얼 서보잉은 카메라의 시각 정보를 활용하여 로봇이나 기계의 움직임과 위치를 실시간으로 제어하는 로봇공학 기법이다.	6
face recognition	얼굴 인식	얼굴 인식은 이미지나 영상에서 얼굴 특징을 분석하고 비교하여 개인의 신원을 식별하거나 확인하는 기술이다.	6
facial expression recognition	얼굴 표정 인식	표정 인식은 인공지능을 활용하여 얼굴 특징의 움직임과 변화를 기반으로 인간의 감정을 식별하고 해석하는 기술이다.	6
speech recognition	음성 인식	음성 인식은 컴퓨터가 말한 언어를 이해하고 처리하여 텍스트나 명령어로 변환하는 기술이다.	6
image retrieval	이미지 검색	이미지 검색은 특정 기준이나 사용자 쿼리(주로 키워드나 시각적 콘텐츠)를 기반으로 데이터베이스에서 이미지를 검색하고 찾는 과정이다.	6
image classification	이미지 분류	이미지 분류는 컴퓨터 프로그램이 이미지의 내용을 분석하여 미리 정의된 그룹으로 분류하는 과정이다.	6
image segmentation	이미지 세분화	이미지 분할은 이미지를 서로 다른 부분이나 영역으로 나누어 해당 이미지 내의 다양한 객체나 영역을 분석하고 이해하기 쉽게 만드는 과정이다.	6
image recognition	이미지 인식	이미지 인식은 컴퓨터가 디지털 이미지 내의 객체, 장면 또는 패턴을 식별하고 해석할 수 있게 하는 기술로, 인간이 보는 것을 인식하는 방식과 유사하다.	6
image processing	이미지 처리	이미지 처리는 컴퓨터 알고리즘을 사용하여 디지털 이미지를 조작하고 분석하여 품질을 향상시키거나 유용한 정보를 추출하는 기술이다.	6

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
human action recognition	인간 행동 인식	인간 동작 인식은 인공지능 기술을 활용하여 동영상이나 이미지에서 사람이 수행하는 특정 동작이나 활동을 식별하고 분류하는 과정이다.	6
gesture recognition	제스처 인식	제스처 인식은 컴퓨터와 기기가 카메라나 센서를 활용해 인간의 동작과 제스처를 해석하여 사용자 명령을 이해하고 응답할 수 있게 하는 기술이다.	6
intelligent geometric computing	지능형 기하학적 컴퓨팅	지능형 기하 컴퓨팅은 인간의 추론 및 문제 해결 방식을 모방하여 기하학적 형태와 공간적 관계를 분석, 조작, 이해하기 위해 고급 알고리즘과 인공지능 기법을 활용하는 것을 의미한다.	6
computer vision	컴퓨터 비전	컴퓨터 비전은 컴퓨터가 인간이 시각 데이터를 인식하고 처리하는 방식과 유사하게 이미지나 동영상과 같은 세계의 시각 정보를 해석하고 이해할 수 있도록 하는 인공지능 분야이다.	6
feature extraction	특징 추출	특징 추출은 머신러닝 모델의 성능 향상을 돕기 위해 원시 데이터에서 중요한 특성이나 패턴을 식별하고 선택하는 과정이다.	6
pattern recognition	패턴 인식	패턴 인식은 컴퓨터나 기계가 패턴이나 추세를 기반으로 데이터를 식별하고 분류하는 능력으로, 인간이 모양, 소리 또는 물체를 인식하는 방식과 유사하다.	6
virtual assistant	가상 비서	가상 비서는 인공지능을 활용하여 사용자에게 작업을 수행하고 서비스를 제공하는 소프트웨어 프로그램 또는 애플리케이션으로, 주로 음성 또는 텍스트 상호작용을 통해 작동한다.	7
layered control system	계층적 제어 시스템	계층적 제어 시스템은 인공지능 분야의 조직적 프레임워크로, 서로 다른 수준의 의사 결정 및 제어 기능이 독립적으로 운영되지만 복잡한 작업이나 프로세스를 효과적으로 관리하기 위해 협력한다.	7
multi-sensor fusion	다중 센서 융합	다중 센서 융합은 여러 센서의 데이터를 결합하여 환경이나 상황에 대한 보다 정확하고 포괄적인 이해를 창출하는 과정이다.	7
multi-agent system	다중 에이전트 시스템	다중 에이전트 시스템은 독립적인 개체들, 즉 “에이전트”들의 집합으로, 문제 해결이나 목표 달성을 위해 상호작용하고 협력하며, 종종 현실 세계 시나리오에서 발견되는 복잡한 행동을 시뮬레이션한다.	7
simultaneous localisation mapping	동시 위치 추적 매핑	동시 위치 추정 및 매핑(SLAM)은 로봇이나 드론과 같은 장치가 미지의 환경에 대한 지도를 생성하는 동시에 그 환경 내에서 자신의 위치를 추적할 수 있게 하는 기술이다.	7

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
unmanned aerial vehicle	무인 항공기	무인 항공기(UAV)는 탑승한 인간 조종사 없이 작동하는 항공기로, 원격 또는 자율적으로 제어되며 감시, 배송, 항공 촬영 등의 작업에 흔히 사용된다.	7
bee colony	벌 군집	벌 군집은 여왕벌, 일벌, 수벌로 구성된 구조화된 벌 공동체로, 벌집에서 함께 생활하며 먹이 수집, 새끼 돌보기, 환경 유지 등을 협력적으로 수행한다.	7
cyber physical system	사이버 물리 시스템	사이버-물리 시스템은 컴퓨터 기반 알고리즘과 물리적 프로세스를 통합하여 스마트 홈, 자율주행 차량, 산업 자동화 같은 시스템의 실시간 모니터링 및 제어를 가능하게 하는 기술이다.	7
sensor data fusion	센서 데이터 융합	센서 데이터 융합은 여러 센서의 정보를 결합하여 환경이나 상황에 대한 보다 정확하고 포괄적인 이해를 창출하는 과정이다.	7
sensor fusion	센서 융합	센서 융합은 여러 센서의 데이터를 결합하여 환경이나 대상 상태에 대한 정보의 정확성과 신뢰성을 향상시키는 과정이다.	7
swarm intelligence	스웜 지능	집단 지능은 개미나 벌 같은 사회성 생물의 집단 행동을 모방하여 분산된 협력과 소통을 통해 복잡한 문제를 해결하는 인공지능 개념이다.	7
swarm optimisation	스웜 최적화	군집 최적화는 새나 물고기 같은 사회적 유기체의 집단 행동에서 영감을 받은 문제 해결 기법으로, 여러 에이전트가 정보를 공유하고 환경에 적응하며 협력하여 최적의 해법을 찾는 방식이다.	7
agentic workflows	에이전트 워크플로우	에이전트 기반 워크플로는 인공지능 시스템 내에서 의사 결정과 행동을 독립적으로 수행할 수 있도록 하는 자동화 프로세스를 의미하며, 종종 인간과 유사한 추론 및 문제 해결 능력을 모방한다.	7
cognitive automation	인지 자동화	인지 자동화는 이해, 학습, 의사 결정과 같은 작업에서 인간의 사고 과정을 모방하기 위해 인공지능 기술을 활용하는 것을 의미하며, 이를 통해 시스템은 일반적으로 인간의 지능이 필요한 복잡한 프로세스를 자동화할 수 있다.	7
particle swarm optimisation	입자 군집 최적화	입자 군집 최적화는 새나 물고기의 사회적 행동에서 영감을 받은 계산 방법론으로, 후보 솔루션(입자) 집단이 문제 공간을 이동하며 자신과 이웃의 경험을 바탕으로 위치를 조정함으로써 최적의 해법을 찾는 데 사용된다.	7

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
self-supervised-learning	자기 감독 학습	자기 지도 학습은 인간이 제공한 레이블에 의존하지 않고 데이터 자체로부터 자체 레이블을 생성함으로써 데이터를 이해하도록 학습하는 기계 학습의 한 유형이다.	7
self-organising structure	자기 조직화 구조	자기 조직화 구조는 외부 통제 없이 스스로를 자동으로 배열하고 적응하는 시스템으로, 종종 환경 변화에 반응하여 조직이나 기능을 개선한다.	7
self-organising map	자기 조직화 지도	자기 조직화 맵은 유사한 항목들을 그들의 관계를 보존하는 방식으로 그룹화하여 복잡한 데이터를 조직화하고 시각화하는 인공 신경망의 한 유형이다.	7
autonomous weapon	자율 무기	자율 무기는 인간의 개입 없이 목표물을 식별, 선택, 공격할 수 있는 군사 장치로, 인공지능을 활용해 스스로 결정을 내린다.	7
autonomous vehicle	자율 주행 차량	자율 주행 차량은 센서 및 인공지능과 같은 기술을 활용해 인간의 개입 없이 주행하고 작동하는 자율 주행 자동차이다.	7
autonomic computing	자율 컴퓨팅	자율 컴퓨팅은 인체 기능 조절 방식과 유사하게, 인간의 개입 없이 스스로 구성, 최적화, 복구 및 보호를 수행하는 자가 관리 컴퓨팅 시스템을 의미한다.	7
obstacle avoidance	장애물 회피	장애물 회피 기술은 로봇이나 자율 시스템이 환경 내 장애물을 감지하고 우회하여 충돌을 방지하고 안전한 이동을 보장하도록 하는 기술이다.	7
collision avoidance	충돌 회피	충돌 회피는 잠재적 장애물을 감지하고 이를 피하기 위한 조치를 자동으로 취함으로써 사고를 방지하도록 설계된 기술 또는 시스템으로, 주로 차량 및 로봇 공학에 사용된다.	7
learning automata	학습 자동기	학습 자동기는 불확실한 환경에서 자신의 행동 결과로부터 학습함으로써 시간이 지남에 따라 성능을 향상시키는 적응형 의사 결정 시스템이다.	7
ambient intelligence	환경 지능	주변 지능은 우리 환경에 매끄럽게 통합되어 직접적인 상호작용 없이도 기기가 우리의 요구와 선호도에 지능적으로 반응할 수 있도록 하는 기술을 의미한다.	7
trajectory planning	경로 계획	궤적 계획은 로봇이나 차량과 같은 물체가 한 지점에서 다른 지점으로 이동할 때 속도, 장애물, 안전성 등의 요소를 고려하여 따라야 할 경로나 경로를 결정하는 과정이다.	8
trajectory tracking	경로 추적	궤적 추적은 시간에 따른 물체의 경로나 움직임을 모니터링하고 예측하는 과정으로, 경로를 정확하게 따라가기 위해 센서와 알고리즘을 사용하는 경우가 많다.	8

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
rough set	러프 세트	러프 집합은 데이터 분석에 사용되는 수학적 접근법으로, 유사성과 차이점을 기반으로 객체를 그룹화하여 불확실성과 모호성을 다루며 부정확한 정보를 처리할 수 있게 한다.	8
robot	로봇	로봇은 자동으로 작업을 수행하도록 설계된 기계로, 종종 인간이나 동물을 닮았으며 다양한 기능을 수행하도록 프로그래밍될 수 있다.	8
human-robot interaction	인간-로봇 상호작용	인간-로봇 상호작용은 사람들이 로봇과 소통하고 협력하는 방식을 의미하며, 이러한 기계가 인간의 행동과 의도를 이해하고 반응하는 데 중점을 둔다.	8
DPO	DPO	DPO(데이터 보호 책임자)는 조직이 데이터 보호 법률 및 규정을 준수하고 개인정보 및 사생활 권리를 보호하도록 책임지는 전문가이다.	9
fairness	공정성	AI의 공정성은 알고리즘과 시스템이 인종, 성별, 사회경제적 지위 등의 특성에 따른 편향이나 차별 없이 모든 개인과 집단을 평등하게 대우하도록 보장하는 원칙을 의미한다.	9
explainable-ai	설명 가능한 AI	설명 가능한 인공지능은 의사 결정 과정에 대한 명확하고 이해하기 쉬운 통찰력을 제공하도록 설계된 인공지능 시스템을 의미하며, 이를 통해 사용자가 특정 결과가 도출되는 방식과 이유를 더 쉽게 이해할 수 있게 한다.	9
trust region policy optimisation	신뢰 영역 정책 최적화	신뢰 영역 정책 최적화(Trust Region Policy Optimization, TRPO)는 강화 학습 알고리즘으로, 안정적이고 신뢰할 수 있는 학습을 보장하기 위해 작고 통제된 업데이트를 수행함으로써 에이전트의 의사 결정 정책을 개선한다.	9
safety	안전성	인공지능 맥락에서 안전성은 AI 시스템이 신뢰성 있게 작동하고 사람이나 환경에 해를 끼치지 않도록 보장하기 위해 구현된 조치 및 관행을 의미한다.	9
guardrails	안전성 가이드 라인	AI의 가드레일은 인공지능 시스템이 윤리적 경계 내에서 작동하고 해를 끼치거나 바람직하지 않은 결과를 초래하지 않도록 보장하기 위해 설계된 지침 또는 안전 조치를 의미한다.	9
human aware artificial intelligence	인간 인식 인공지능	인간 인식 인공지능은 상호작용 및 의사 결정 과정에서 인간의 감정, 의도, 사회적 맥락을 이해하고 고려하도록 설계된 AI 시스템을 의미한다.	9

〈부표 5-1〉의 계속

OECD AI 키워드	한글	설명	분류 번호
alignment	정렬	AI에서의 정렬은 인공지능 시스템의 목표와 행동이 인간의 가치관 및 의도와 조화를 이루도록 보장하는 과정을 의미한다.	9
intelligence augmentation	지능 증강	지능 증강은 기술을 활용하여 인간의 인지 능력을 향상시켜 더 나은 의사 결정을 내리고 문제를 더 효과적으로 해결하도록 돕는 것을 의미한다.	9
intelligent agent	지능형 에이전트	지능형 에이전트는 환경을 인지하고 결정을 내리며 특정 목표를 달성하기 위해 행동을 취할 수 있는 컴퓨터 프로그램 또는 시스템으로, 종종 인간과 유사한 추론 및 행동을 모방한다.	9
bias-detection	편향 탐지	편향 탐지는 AI 시스템에서 차별적 결과를 초래할 수 있는 데이터나 알고리즘 내 불공정하거나 편향된 패턴을 식별하고 분석하는 과정이다.	9

자료 : Baruffaldi et al.(2020), Squicciarini and Nachtigall(2021), Borgonovi et al. (2023) 등을 참고하여 작성.

◆ 執筆陣

- 장지연(한국노동연구원 선임연구위원)
- 심지환(동양미래대학교 교수)
- 정준호(강원대학교 교수)
- 이철승(서강대학교 교수)
- 전병유(한신대학교 교수)

인공지능 시대의 숙련

- | | |
|------------|---|
| ▪ 발행연월일 | 2025년 12월 26일 인쇄
2025년 12월 31일 발행 |
| ▪ 발 행 인 | 허 재 준 |
| ▪ 발 행 처 | 한국노동연구원
☎ 대표 (044) 287-6081 Fax (044) 287-6089
310147 세종특별자치시 시청대로 370
세종국책연구단지 경제정책동 |
| ▪ 조 판 · 인쇄 | 거목정보산업(주) (044) 863-6566 |
| ▪ 등 록 일 자 | 1988년 9월 13일 |
| ▪ 등 록 번 호 | 제2015-000013호 |

© 한국노동연구원 2025 정가 10,000원

ISBN 979-11-260-0815-5

KLI
한국노동연구원

한국노동연구원

30147 세종특별자치시 시청대로 370 경제정책동
TEL : 044-287-6083 <http://www.kli.re.kr>



ISBN 979-11-260-0815-5