



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

박사학위논문

그래프 데이터베이스를 활용한
대학생 진로 조언 플랫폼 연구



한 성 대 학 교 일 반 대 학 원

스마트융합컨설팅학과

스마트융합제품전공

김 종 울

박사학위논문
지도교수 노광현

그래프 데이터베이스를 활용한 대학생 진로 조언 플랫폼 연구

A Study on Career Advice Platform for College
Students Using Graph Database



HANSUNG
UNIVERSITY

2019년 12월 일

한 성 대 학 교 일 반 대 학 원

스마트융합컨설팅학과

스마트융합제품전공

김 종 율

박사학위논문
지도교수 노광현

그래프 데이터베이스를 활용한 대학생 진로 조언 플랫폼 연구

A Study on Career Advice Platform for College
Students Using Graph Database

위 논문을 공학 박사학위 논문으로 제출함

2019년 12월 일

한 성 대 학 교 일 반 대 학 원

스마트융합컨설팅학과

스마트융합제품전공

김 중 율

김종율의 공학 박사학위논문을 인준함

2019년 12월 일



심사위원장 _____(인)
심 사 위 원 _____(인)

심 사 위 원 _____(인)

심 사 위 원 _____(인)

심 사 위 원 _____(인)

국 문 초 록

그래프 데이터베이스를 활용한 대학생 진로 조언 플랫폼 연구

한 성 대 학 교 대 학 원
스 마 트 융 합 컨 설 톡 학 과
제 품 전 공
김 종 율

교육부의 '2017년 진로교육현황조사'에 따르면 국내 대학생들의 가장 큰 고민 사항은 진로에 대한 고민으로, 진로와 연관도가 높은 학업에 대한 고민 까지 포함할 경우 85%의 대학생들이 진로에 대해 고민하는 것으로 확인되었다.

대학생들의 진로에 대한 고민과 원인에 관한 많은 연구가 진행되고 있지만, 실제 이러한 문제를 해결하는 방안 제시는 진로 상담과 진로 교육과 같이 범위가 한정되어 있다. 본 논문은 국내 대학생들의 진로·취업 문제에 대한 정보기술을 활용한 해결방안 제시 연구이다.

연구 수행을 위한 대학생들의 요구사항은 교육부의 '2017년 진로교육현황조사' 내용 중 '진로 및 취·창업 준비를 위해 대학에서 지원받고 싶은 사항' 항목의 응답 내용과 2017년 중앙대학교의 진로의식 설문 조사에서의 '진로 결정에 어려움을 겪고 있는 이유' 항목의 응답 내용을 활용하여 학생들의 요구사항을 정의하였다.

진로 조언을 위한 모델은 실제 상담 사례들과 파슨스의 ‘특성-요인 이론’에 기반하여 개인에 대한 분석, 직업 분석, 과학적 조언을 통해 매칭하는 과정으로 조언 모델을 정의하였으며, 학생들의 요구사항과 진로 조언 이론을 기반으로 학생의 정보 분석, 직업정보 제공, 추천·상담 단계로 정보기술을 활용한 진로 조언 모델과 활용 가능한 기술 요소를 제시하였다.

진로 조언을 위해 고용노동부 워크넷 API를 통해 직업정보와 직무정보를 수집하였으며, 학생들의 요구사항인 실제 기업에 대한 채용정보 및 취·창업 준비를 위한 실무 정보의 부족으로 민간 취업정보포털에서 89개의 정보기술 관련 직업에 대한 구인공고를 수집하여 역량별로 분류하고 중요도를 계산하여 가중치를 설정하였다.

수집된 데이터들을 활용하여 학생들에게 맞춤형 조언을 제공하기 위한 설계 과정에서 각 역량 요소들과 직업, 역량과 역량의 관계들과 같이 관계에 대한 분석이 중요한 기술적 문제로 대두되었다. 이러한 관계 분석 문제 해결을 위해 관계형 데이터베이스와 그래프 데이터베이스의 관계 쿼리 효율에 대한 비교 실험과 평가를 통해 그래프 데이터베이스를 활용한 지식 그래프를 진로 조언을 위한 지식 베이스로 선정하였다.

상담 및 멘토링 기능 설계를 위해 대학생들의 자연어 질문에 대해, 자연어 처리 학습을 활용하여 챗봇 기반의 진로 조언 시스템과 입력된 개인의 역량을 기반으로 맞춤형 직업을 추천받고, 역량 개발을 지원받기 위한 e-포트폴리오 앱을 설계하였다.

본 연구를 통해 그래프 데이터베이스를 활용한 지식 베이스 설계와 구축 방법론을 제시하였으며, 제시된 지식 그래프 설계 방법론과 구축 방법론을 활용하여 고용노동부의 API에서의 직업·직무 데이터 수집과 민간채용 포털에서의 구인 정보를 수집하여 진로·취업 조언을 위한 지식 그래프를 구현하고, 방법론을 사용하지 않은 방식과의 구축속도 비교를 통해 방법론의 효율성을 증명하였다.

설계된 그래프 데이터베이스를 활용한 진로 조언 플랫폼을 실제 구현하여, 지식 그래프를 활용한 챗봇과 모바일 앱 환경의 진로 조언 시나리오를 테스트

하였으며, 정보기술 관련 대학생들을 평가단으로 선정하여 실제 시연과 테스트를 진행하였다.

평가단 설문 및 평가 결과 기업의 채용정보분석과 조언 결과(4.08), 인터페이스 만족도(4.24), 진로설계 시 도움 여부(4.24), e-포트폴리오 기능 및 활용 여부(4.52), 진로 조언 플랫폼 적극 활용 여부(4.71)로 평가된 바와 같이 진로 추천 및 역량 개발 모델의 만족도가 매우 높게 조사되었다. 또한, 기존의 모바일 또는 웹검색을 활용한 정보 검색 방식과 챗봇을 활용한 방식과의 정보 검색 시간 테스트를 통해 진로 분야와 같이 특정 기능의 경우 챗봇을 활용한 검색 효율과 만족도가 높은 것으로 평가되었다.

본 연구는 국내 대학생들의 가장 큰 고민 사항인 진로 문제를 정보기술을 활용한 해결방안을 제시한 연구로써, 관계 분석이 중요한 지식에서의 그래프 데이터베이스 기반의 지식 그래프 구축과 활용에 대한 모델을 새롭게 제시하였다.

본 연구의 성과로 다양한 직업데이터를 활용한 지식 그래프의 구축과 구축된 지식을 대학생들에게 분석·제공하여 학생의 진로 결정과 역량 개발을 지원할 수 있게 되기를 바라며, 대학 내의 학사지원시스템 및 진로·취업 통합정보시스템과 연동하여 학생 개개인의 맞춤형 지원 서비스 구축을 위한 가이드가 되기를 바란다. 특히, 그래프 데이터베이스는 구조적으로 설명할 수 있으며, 모델 설계를 위한 데이터 이외에 요구하는 데이터가 많지 않기 때문에 다양한 형태의 지능형 시스템의 구축과 활용 모델로서의 본 연구를 활용한 추가 연구를 기대한다.

【주제어】 진로 조언, e-Advisor, 챗봇, 그래프 데이터베이스, 에듀테크

목 차

제 1 장 서 론	1
제 1 절 연구 배경	1
1) 청년 실업과 일자리 문제	1
2) 밀레니얼 세대의 특성과 직업관	3
3) 대학생 진로 문제와 필요한 지원	5
4) 4차 산업혁명 시대와 직업 환경의 변화	7
5) 국가 진로지도 프로그램	8
제 2 절 연구의 필요성	10
1) 대학생 진로 문제점 및 대안	10
2) 학생과 기업의 역량 미스매치	11
3) 고용정보원의 직업정보 지원 서비스	12
4) 국내 대학들의 진로·취업 역량 강화를 위한 정보시스템	15
5) 정보기술을 활용한 국내외 대학 혁신 사례	18
제 3 절 연구의 범위와 목적	22
1) 진로 조언 플랫폼의 범위	22
2) 진로 조언 모델 기능 도출	24
3) 진로 조언 플랫폼 기본 모델 설계와 기술 요소	26
4) 연구의 방법 및 범위	29
제 2 장 이론적 배경 및 선행연구	31
제 1 절 대학생 진로 문제와 학습 동기에 관한 연구	31
1) 진로 문제 현황	31
2) 효율적인 진로 정보 검색 문제	31
3) 진로에 따른 학생의 학습 동기와 학업 만족도	32
제 2 절 챗봇을 활용한 상담 및 조언 서비스 제공 사례	34
1) 법률 및 심리 상담에서의 챗봇 활용 사례	34
2) 직업 추천 및 교육에서의 챗봇 활용 사례	35
제 3 절 진로 조언을 위한 직업정보 데이터 연구	37

1) 진로 조언 데이터 유형	37
2) 공공 직업정보 데이터	37
3) 민간 직업정보 데이터	38
제 4 절 그래프 데이터베이스(Graph Database)	39
1) 그래프 데이터베이스 개요	39
2) 그래프 데이터베이스의 특징	40
3) 사이퍼쿼리	42
4) 지식 그래프	43
제 5 절 대화형 인터페이스	46
1) 챗봇의 정의	46
2) 대학생들의 챗봇 활용	48
3) 챗봇 프레임워크	49
4) 자연어처리	50
제 6 절 진로 조언을 위한 대학생 개인 데이터	51
1) 대학생 e-포트폴리오	51
2) 대학 내부 시스템의 학생 정보 데이터	52
제 3 장 그래프 데이터베이스를 활용한 진로 조언 플랫폼 설계	54
제 1 절 진로 조언 플랫폼 설계 방향 및 모델	54
1) 진로 조언 플랫폼 설계 방향	54
2) 진로 조언 플랫폼 아키텍처 설계	56
3) 관계형 데이터베이스와 그래프 데이터베이스 비교	58
제 2 절 그래프 데이터베이스를 활용한 지식 그래프 설계	61
1) 그래프 데이터베이스를 활용한 지식 그래프 설계 방법론	61
2) 그래프 데이터베이스 기반의 지식 그래프 목표	64
3) 수집데이터 선정	65
4) 데이터 전처리	73
5) 그래프 데이터베이스를 활용한 지식 그래프 구축 방법론	76
6) 챗봇 인터페이스	80
7) 모바일 e-포트폴리오 앱	83

제 4 장	진로 조언 플랫폼 구현 및 테스트	87
제 1 절	진로 조언 플랫폼 구현 범위	87
1)	진로 조언 플랫폼 구현	87
2)	진로 조언 플랫폼 인프라 구성	88
제 2 절	진로 조언 지식 그래프 구축	92
1)	직무데이터(디렉터리 구조 API) 지식 그래프 구축	92
2)	직업정보(코드 조회 API) 지식 그래프 구축	102
3)	민간 채용정보(비정형 데이터) 지식 그래프 구축	106
제 3 절	사용자 인터페이스 구현	118
1)	챗봇 인터페이스 구현	118
2)	자연어처리 학습	122
3)	e-포트폴리오 앱 환경 구현	126
제 4 절	진로 조언 플랫폼 기능 테스트	129
1)	챗봇 인터페이스를 활용한 조언 기능 테스트	129
2)	e-포트폴리오 앱 기능 테스트	133
3)	사용자(대학생) 설문 평가	135
4)	사용자(대학생) 검색 속도 비교 테스트 및 결과	139
5)	연구의 성과	141
제 5 장	결 론	143
참 고 문 헌	145
부	록	152
ABSTRACT	156

표 목 차

[표 1-1] 중앙대학교 진로의식 설문 조사(2017년)	6
[표 1-2] 대학생과 인사담당자의 직업 역량별 비교	12
[표 1-3] 고용정보원 대학생 진로 취업지도 지원 서비스	13
[표 1-4] 진로 조언 기능별 요구사항	25
[표 1-5] 요구 기능에 대한 구현 가능 정보기술	26
[표 1-6] 진로 조언 플랫폼 기능별 기술 요소	28
[표 2-1] 챗봇의 유형별 특징	47
[표 2-2] 대학 내부의 진로·취업 관련 연동 데이터	53
[표 3-1] 진로 조언 플랫폼 주요 연구내용	54
[표 3-2] 진로 조언을 위한 지식 데이터 선정	55
[표 3-3] 진로 조언 플랫폼 아키텍처 단계별 동작 흐름	57
[표 3-4] 관계형 데이터베이스에서의 문제점	58
[표 3-5] 비교 실험을 위한 데이터베이스시스템 환경	59
[표 3-6] 데이터베이스 관계 쿼리 비교	59
[표 3-7] 관계의 증가에 따른 응답속도 비교	60
[표 3-8] 한국고용정보원 직업 사전 분류 코드 일부 항목	66
[표 3-9] 채용정보를 수집할 직업 리스트	67
[표 3-10] 직업정보 API 추출 정보	71
[표 3-11] 시나리오 테스트	75
[표 3-12] 워크넷 직업정보 지식 그래프 구축단계별 설명	78
[표 3-13] 채용정보 지식 그래프 구축 단계별 설명	78
[표 3-14] 대화 Flow 단계별 동작	83
[표 3-15] 그래프 매칭 관계 구성	85
[표 3-16] 대화 답변 절차	86
[표 4-1] 진로 조언 플랫폼 Layer 별 기능	87
[표 4-2] 진로 조언 플랫폼 하드웨어 및 소프트웨어 구조	91
[표 4-3] 디렉터리 구조에서의 데이터 수집	94
[표 4-4] 직무데이터 API 고유노드 생성 및 관계 연결	96

[표 4-5] 중복 노드 생성 및 관계 연결	97
[표 4-6] 직무데이터 API 지식 그래프 생성 노드·관계 생성 수	97
[표 4-7] 직무데이터 API 지식 그래프 검증을 위한 사이퍼쿼리	99
[표 4-8] 직무데이터 관계 연결 및 생성	101
[표 4-9] 방법론과 알고리즘에 따른 지식 그래프 생성 속도	102
[표 4-10] 각 직업 노드 생성 및 고유관계 생성	104
[표 4-11] 직업 중복 가능한 노드들과 관계 지정	105
[표 4-12] 웹프로그래머 채용정보 페이지 수집	107
[표 4-13] 역량 정보 추출을 위한 데이터 사전	108
[표 4-14] 채용 HTML 페이지에서의 직업 역량 정보 추출	109
[표 4-15] 기업 채용정보 기초데이터 생성	109
[표 4-16] 채용정보 중복제거 작업 수행	111
[표 4-17] 채용정보속성 노드 관계 생성 및 가중치 추가	112
[표 4-18] 각 역량 리스트 노드 생성 및 관계 연결	113
[표 4-19] 지식 그래프 생성 노드 목록	115
[표 4-20] 지식 그래프 생성 관계 목록	116
[표 4-21] 역량 리스트 및 노드 가중치 분석	121
[표 4-22] 역량에 대한 관련 직업 답변	121
[표 4-23] 채용정보 질문에 대한 분류와 적중률	124
[표 4-24] 직업정보 질문에 대한 분류와 적중률	125
[표 4-25] 동일한 직업정보 질문에 대한 분류와 적중률	125
[표 4-26] e-포트폴리오 앱 기능	127
[표 4-27] API 서버에 직업 역량 호출 결과	128
[표 4-28] Get으로 전달받은 값을 쿼리하여 값을 전달하는 알고리즘	129
[표 4-29] 직업정보 조언 질문 목록	129
[표 4-30] 채용정보 조언 질문 목록	131
[표 4-31] 사용자 평가단 구성 및 진로정보	136
[표 4-32] 기업의 채용정보 분석 답변 만족도	137
[표 4-33] e-포트폴리오 기능 및 활용 여부	137

[표 4-34] 챗봇 조인 결과에 대한 만족도	138
[표 4-35] 챗봇 인터페이스 만족도와 활용 여부	138
[표 4-36] 기능 구현 정도에 대한 답변	139
[표 4-37] 직업정보 검색 속도 평균	141
[표 4-38] 채용정보 검색 속도 평균	141
[표 4-39] 직업정보 답변의 만족도와 정보 검색 편의성	141



그 립 목 차

[그림 1-1] 청년고용 동향	2
[그림 1-2] 취업 희망 기업(전 학년 평균)	3
[그림 1-3] 세대의 정의	4
[그림 1-4] 대학 생활 고민 순위	5
[그림 1-5] 학생이 필요로 하는 진로·취업 조력자	6
[그림 1-6] 진로 및 취·창업 준비를 위해 대학에서 지원받고 싶은 사항	7
[그림 1-7] 학교급별 진로교육 체계	9
[그림 1-8] 워크넷 직업정보 시스템소프트웨어 개발자 지식중요도	14
[그림 1-9] 고용노동부 챗봇 고용이 시스템 개념도	14
[그림 1-10] S여대 학생경력관리시스템 시스템 구성	16
[그림 1-11] 한양대학교 빅데이터 기반의 HY-CDP 시스템 고도화 계획 ..	16
[그림 1-12] CAU e-Advisor 학사행정 서비스 예시	17
[그림 1-13] 애리조나 대학교 ‘e-Advisor’	20
[그림 1-14] 파슨스 특성-요인 이론, 진로지도의 3요소	23
[그림 1-15] 진로 조언 모델에 필요한 기능 도출	24
[그림 1-19] 진로 조언 플랫폼 기본 모델 구성	27
[그림 2-1] 그래프 데이터베이스의 구조	39
[그림 2-2] 관계형 데이터베이스와 그래프 데이터베이스의 쿼리 비교	43
[그림 2-3] 지식 그래프 예제	45
[그림 2-4] 챗봇과 대화형 인터페이스 구분	46
[그림 2-5] 메신저 앱의 사용량 변화와 채팅앱의 선호 이유	48
[그림 2-6] Microsoft Bot-Framework 구조	49
[그림 3-1] 진로 조언 플랫폼 아키텍처 설계	57
[그림 3-2] 매칭되는 관계의 증가에 따른 응답속도 비교	60
[그림 3-3] 그래프 데이터베이스를 활용한 지식 그래프 설계 방법론	62
[그림 3-4] 그래프 특성	65
[그림 3-5] 웹프로그래머 키워드 채용공고 리스트	68
[그림 3-6] 채용공고 일반 항목	68

[그림 3-7] 기업의 채용공고에서의 기업이 요구 항목 분류	69
[그림 3-8] 고용정보데이터 수집기 구조	70
[그림 3-9] 직업정보 API 요청 parameter	71
[그림 3-10] 직업 코드 '20272'에 대한 요약정보	72
[그림 3-11] 능력/지식/환경 출력 결과	72
[그림 3-12] '웹 프로그래머' 직업의 역량 노드 설계	73
[그림 3-13] 데이터 수집 지식 그래프 생성 구조	73
[그림 3-14] 해당 직무 근무경험에 대한 연관 관계 설계 예제	74
[그림 3-15] 지식 그래프 노드 구조 설계	75
[그림 3-16] 그래프 데이터베이스를 활용한 지식 그래프 구축 방법론	76
[그림 3-17] 워크넷 직업정보 지식 그래프 구축	77
[그림 3-18] 민간 채용정보의 데이터의 지식 그래프 구축	79
[그림 3-19] 디지털영상처리 전문가(직업 코드 20233) 지식 연결 형태	80
[그림 3-20] 챗봇 인터페이스 구조	81
[그림 3-21] 대화 FLOW 설계	82
[그림 3-22] 진로 조언 e-포트폴리오 기능	84
[그림 3-23] 맞춤형 직업 추천 및 진로 개발 프로세스	84
[그림 3-24] 지식 그래프에서의 답변 프로세스	86
[그림 4-1] 진로 조언 플랫폼 구조	87
[그림 4-2] 진로 조언 시스템 연구 클라우드 아키텍처	89
[그림 4-3] 지식 그래프 구축방법론의 구현 단계	92
[그림 4-4] 자료수집 및 데이터 전처리 수행	93
[그림 4-5] 워크넷 직무데이터 구조	93
[그림 4-6] 지식 그래프 구축 방법론: 그래프 생성단계	95
[그림 4-7] '이러닝 운영 학습자 관리' 직무의 지식 맵	98
[그림 4-8] 직무데이터 API 지식 그래프 검증 지식 맵	99
[그림 4-9] 직무데이터 지식 그래프 생성 로그	100
[그림 4-10] 절차와 알고리즘이 최적화된 지식 그래프 생성시간	101
[그림 4-11] 직업정보 API 지식 그래프 구축 방법론 적용	103

[그림 4-12] 코드별 직업정보 수집	103
[그림 4-13] 지식 그래프에서의 웹프로그래머 직업 호출 지식 맵	106
[그림 4-14] 채용공고 리스트에서 역량 정보 추출 절차	107
[그림 4-15] HTML 페이지에서 역량 정보 추출	108
[그림 4-16] 채용정보 데이터의 지식 그래프 구축 방법론	110
[그림 4-17] 최상위 직업 노드 연결 및 생성	112
[그림 4-18] 생성 완료된 지식 그래프 중 ‘웹프로그래머’ 지식 맵	114
[그림 4-19] 직업 속성 노드의 관계 일부	114
[그림 4-20] 최종 생성된 지식 그래프 구조	117
[그림 4-21] 메시지 처리 활동	118
[그림 4-22] 정보 검색 구분	119
[그림 4-23] 직업정보의 메인 Dialog 구성	119
[그림 4-24] 채용정보의 메인 Dialog 구성	120
[그림 4-25] LUIS에서의 채용정보 Intents 항목	122
[그림 4-26] 자격증 항목의 발화 학습 내용	123
[그림 4-27] LUIS에서의 직업정보 Intents 항목	123
[그림 4-28] 기술 및 지식요소 실체(엔티티) 학습 내용	124
[그림 4-29] e-포트폴리오 메뉴 구성	126
[그림 4-30] 개인정보 및 이력 입력	127
[그림 4-31] API 통신 구조	128
[그림 4-32] 검색 시간 및 검색 내용 만족도/편의성 테스트 방식	140

제 1 장 서 론

제 1 절 연구 배경

4차 산업혁명의 진전 속에서 빅데이터, 머신러닝, 클라우드와 같은 현대 정보기술의 발전과 학령인구의 감소와 같은 사회적 환경의 변화로 인하여 고등교육환경 역시 빠르게 변하고 있다. 현대 정보기술의 급격한 발전은 직업인이 갖추어야 할 역량의 변화와 새로운 직업의 등장에 따른 학생들이 갖추어야 할 직업 역량도 변화하고 있다.¹⁾ 대학의 기능 역시 학문 연구와 진리 탐구의 도정에서 전문적 직업인을 양성하는 직업대학으로 변해가고 있으며 이는 경제 성장률의 저하에 따른 직장의 감소와 경력자를 우대하는 기업 채용패턴의 변화에 기인한 취업률 악화 역시 하나의 요인이다.²⁾

청년실업률의 증가는 국내 대학생들에게 진로에 대한 높은 고민과 함께 미래에 대한 걱정을 안겨주고 있기에 정부 및 대학 역시 진로 문제 해결을 위한 지원 조직을 신설하고, 다양한 형태의 진로교육을 진행하고 있지만, 실제 학생들이 원하는 지원과는 차이가 있으며 참여율 역시 낮다. 대학생들의 고민과 이로 인해 발생하는 다양한 문제에 관한 연구는 계속 진행되고 있지만, 해결방안에 관한 연구는 미흡하다.

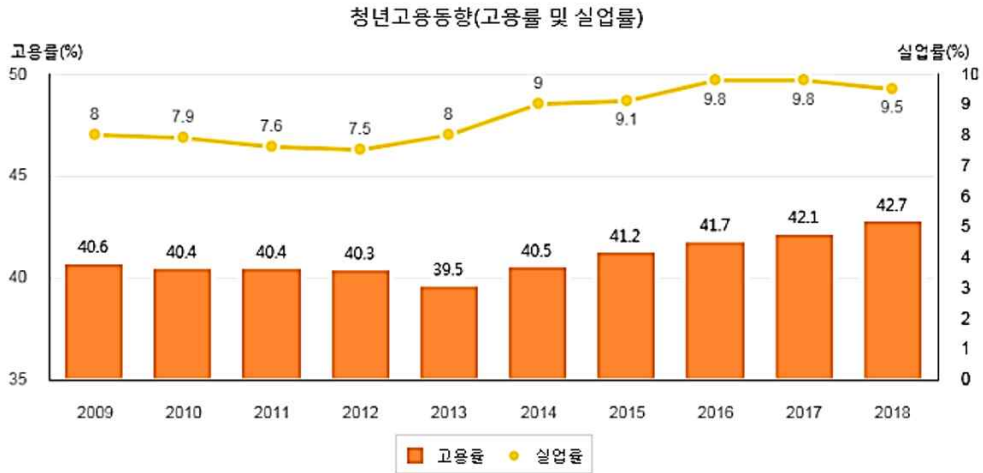
1) 청년 실업과 일자리 문제

1997년 외환위기 이후 대한민국의 가장 큰 사회 문제 중 하나는 취업난이라 할 수 있다. 그중에서 가장 문제시되는 부분은 [그림 1-1]과 같이 2018년 기준 청년실업률이 9.5%에 달하는 것과 같이 사회에 진출하고 경제적 기반을 다지기 시작해야 할 청년기의 취업난이 가장 큰 문제로 대두되고 있으며, 이와 반대로 많은 중소기업은 직무에 적합한 인력을 구하지 못하고 있는 미스

1) 이흥주. (2013). 제조업 분야의 창조적 융합. 『Korea Creative Content Agency』, 78-79.

2) 류동희. (2010). 『대학에서 전문직업인 육성을 위한 생애교육』. 서울: 서경문화사.

매칭 문제와 같은 인력 수급의 불균형 현상이 발생하고 있다.



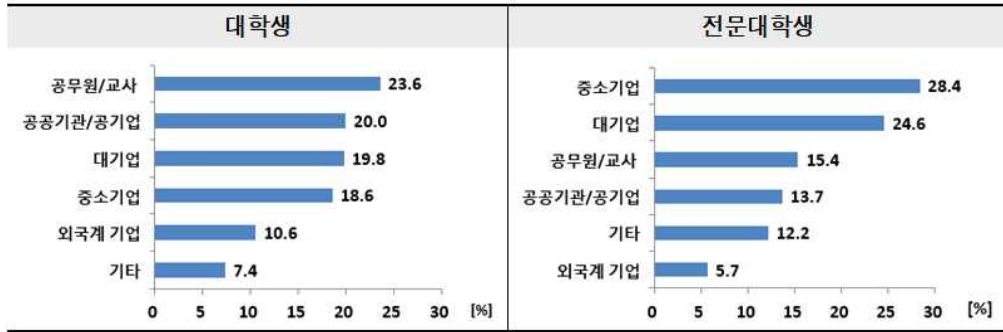
[그림 1-1] 청년고용 동향

이러한 인력수급불균형 문제는 대졸 이상 고학력자의 중소기업 기피와, 높은 임금과 고용 안정성이 보장되는 제한된 일자리를 선호하는 비율이 높기 때문에 대기 실업이 높은 상황도 원인이 된다.³⁾

[그림 1-2]와 같이 2017년 교육부의 대학 진로교육 기반 및 지원체제, 진로 및 취·창업 관련 인식과 참여 정도 등을 조사한 ‘2017년 대학진로교육현황조사’ 결과 중 ‘전 학년 평균 취업 희망기업 조사’에서 4년제 대학생의 취업 희망기업은 공무원/교사 23.6%, 공공기관/공기업 20%, 대기업 19.8%와 비교하면 중소기업은 18.6% 수준인 것으로 조사되었으며,⁴⁾ 전문대 졸업생의 경우 중소기업이 28.4%, 대기업 24.6%, 공무원 15.4%로 학력이 높을수록 좋은 일자리를 원하며, 원하는 직장에 취업하기 위한 졸업 연기나 구직활동의 지속과 같은 ‘취업 재수’를 선택하는 청년들이 증가하는 것으로 조사되었다.

3) 홍기석. (2018). 청년실업의 결정요인 연구. 『한국경제의 분석』, 24(2), 91-152.

4) 장현진. (2017). 『대학진로교육 현황조사(2017) 기초통계표』. 세종: 교육부.



[그림 1-2] 취업 희망 기업(전 학년 평균)

이러한 청년들의 중소기업 선호도가 낮은 문제를 해결하기 위하여, 정부 역시 대기업과 중소기업의 임금 격차 해소와 안정적인 사회 정착을 위해 다양한 장려금 및 세제 혜택을 지원하고 있으며 중소기업 청년고용의 증가와 장기근속자 수 증가와 같은 긍정적 효과 역시 확인되었다.⁵⁾

청년들의 중소기업 기피와 더불어, 청년들의 선호도가 높은 일자리를 가진 기업들의 채용 규모는 지속적으로 감소하고 있으며, 대학 졸업생들을 채용하더라도 현장에 적응하기 위해서는 긴 시간과 높은 비용을 들여 재교육을 진행해야 하므로 신입 직원보다는 경력직을 선호하고 있다. 이러한 신입 채용 기피는 교육과 적응 기간이 오래 걸리는 우량 일자리에서 더 두드러지게 나타나고 있으며, 사회에 진출하는 청년들의 경우 좋은 조건의 채용공고를 찾아도 경력 사항을 요구하거나 경쟁률이 매우 높을 수밖에 없다.

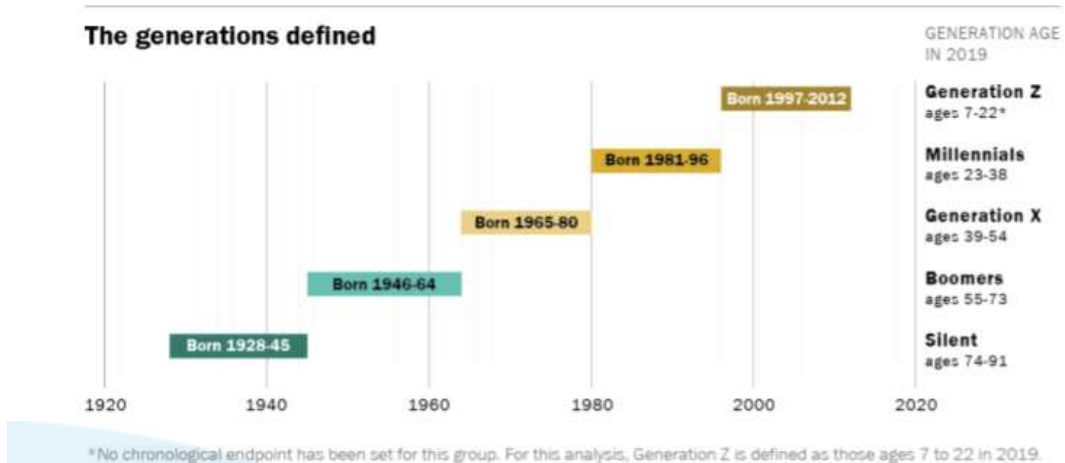
2) 밀레니얼 세대의 특성과 직업관

세대를 구별하는 경계를 나누는 것에 뚜렷한 과학적 구분은 없으나 학술적으로 많이 참고되고 있는 퓨리서치센터(Pew Research Center)의 구분에 따르면 [그림 1-3]과 같이 베이비부머, X 세대, 밀레니얼 세대, Z 세대로 나눌 수 있다.⁶⁾ 이러한 기준에 따라 분류하면, 현 대학생들은 밀레니얼 세대와 Z

5) 이성희, 유동훈. (2019). 청년추가고용장려금사업 운영현황 및 정책과제. 『노동리뷰』, 167.

6) Dimock, M. (2019). Defining generations: Where Millennials end and Generation Z begins. Pew Research Center.

세대 구분에 모두 해당하는 과도기적 세대이며, 사회에 진출하는 Z세대의 진로나 직업관 역시 밀레니얼 세대의 특성에 포함된다.



[그림 1-3] 세대의 정의

밀레니얼 세대의 대학생들은 유·소년기부터 정보통신기술의 과도기를 겪은 세대로서 정보기기의 활용력이 탁월하며 대학 진학률이 높은 특징을 가진다. 이전 세대의 노동자들은 일정 수준의 지식과 역량을 갖추면 은퇴 전까지 살아남을 수 있었으나, 제4차 산업혁명으로 부르는 산업의 변화에 새로운 기술이 빠르게 탄생하고 소멸하는 과정 속에 짧은 기간 동안 학습해야 할 역량이 크게 증가하고 있는 상황에 직면하고 있는 밀레니얼 세대 경우 대학에서 학위를 받더라도 충분히 배웠다고 여기지 못하고 있으며, 성장과 발전을 갈망하며 학습에 대한 초조함을 안고 살아가고 있다.⁷⁾

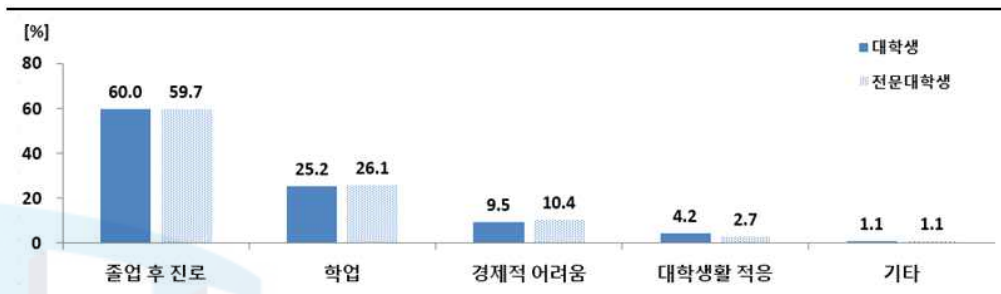
밀레니얼 세대는 베이비붐 세대보다 적극적으로 새로운 직장을 찾고 있는 비율이 높으며, 이직에 대한 거부감이 없으며, 돈보다는 개인의 발전이나 개인의 가치관을 중요시하는 특성이 있다. 이러한 특성은 조기 퇴사자의 증가, 기업의 경력직 채용 증가와 같은 채용 환경의 변화를 가져오고 있으며⁸⁾, 신입 지원자의 경우 업무 역량을 기준을 더 높여야 할 필요성이 증가하고 있다.

7) 손정은, 박재한, 김성혜, 장원섭. (2019). 초기 경력자의 자발적 이직 경험에 관한 현상학적 연구 : 학습과 성장을 중심으로. 『교육문화연구』, 25(5(B)), 641-660.

8) 박희석. (2018). 서울시 비영리스타트업 실태와 청년일자리. 『정책리포트』, (250), 1-24.

3) 대학생 진로 문제와 필요한 지원

교육부의 ‘2017년 대학진로교육현황조사’에 따르면 대학생들의 가장 큰 고민은 [그림 1-4]와 같이 ‘졸업 후 진로(대학생 60.0%, 전문대학생 59.7%)’, ‘학업(대학생 25.2%, 전문대학생 26.1%)’, ‘경제적 어려움(대학생 9.5%, 전문대학생 10.4%)’ 등의 순으로 조사되었다.⁹⁾



[그림 1-4] 대학 생활 고민 순위

2017년 중앙대에서 실시한 ‘취업·진로 의식 설문 조사’에서는 [표 1-1]과 같이 ‘자신의 능력과 역량, 흥미와 적성에 대해 잘 모르겠다’라는 응답이 53.8%, ‘직업의 종류와 세계에 대해 잘 모르겠다’라는 응답이 18.3%, ‘목표를 달성하기 위한 노력과 의지가 부족하다’라는 응답이 14.1%로 조사되었으며,¹⁰⁾ ‘노력과 의지가 부족하다’라는 학생들 역시 목표가 정해지면 그 목표를 실천하기 위해 적극적으로 노력하게¹¹⁾ 되기 때문에 약 86%의 학생들이 자기 자신에 대한 이해와 진로정보 부족이 대학생들에게 가장 지원이 필요한 분야라 할 수 있다.

9) 장현진. (2017). 『대학진로교육 현황조사(2017) 기초통계표』. 세종: 교육부.

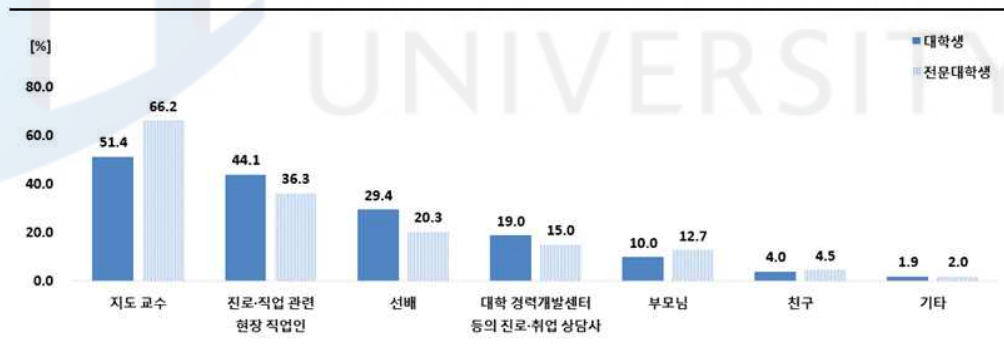
10) 김종율, 노광현. (2019). 공공 데이터를 활용한 대학생 진로 조언을 위한 그래프데이터베이스 기반 지식베이스 구축. 『전자공학회논문지』, 56(10), 41-48.

11) Bandura, A., & Simon, K. M. (1977). The role of proximal intentions in self-regulation of refractory behavior. *Cognitive therapy and research*, 1(3), 177-193.

[표 1-1] 중앙대학교 진로의식 설문조사(2017년)

진로 결정에 어려움을 겪고 있다면 그 이유는?	응답 수	비율
나의 능력과 역량에 대하여 잘 모르겠다.	1313	32%
나의 흥미와 적성에 대하여 잘 모르겠다.	915	22%
직업의 종류와 세계에 대하여 잘 모르겠다.	758	18%
목표를 달성하기 위한 노력 및 의지가 부족하다.	584	14%
내가 원하는 길이 경제적/사회적으로 안정성이 부족하다.	273	7%
기타	209	5%
내가 원하는 길과 부모님이 원하는 길이 다르다.	88	2%
합계	4140	100%

이와 같은 학생들의 진로 문제를 해결하기 위해 각 대학에서도 진로·취업 조직을 운영하고, 진로교육을 강화하고 있지만, 대부분 학생은 이러한 진로·취업 조직보다는 [그림 1-5]와 같이, 지도교수(대학생 51.4%, 전문대학생 66.2%), ‘진로·직업 관련 현장 직업인(대학생 44.1%, 전문대학생 36.3%)’의 조력을 원하는 것으로 조사된¹²⁾ 바와 같이 전공 관련 전문가로부터 진로·취업을 위한 방향 제시 및 조언을 원하고 있다.



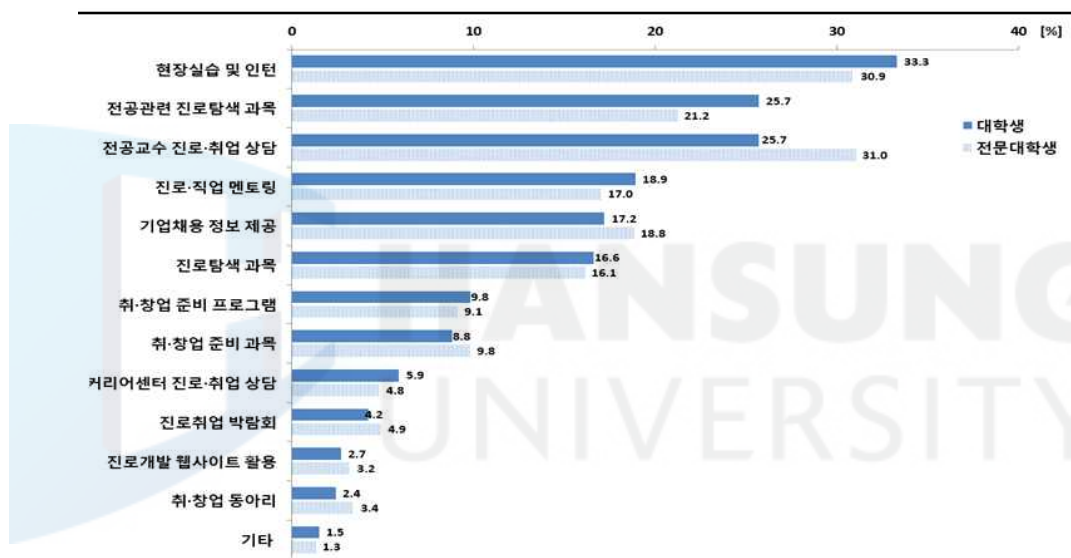
[그림 1-5] 학생이 필요로 하는 진로·취업 조력자

대학생의 진로 및 취·창업 준비를 위해 대학이 지원해 주기를 바라는 사항에 대한 조사에서는 [그림 1-6]과 같이, ‘현장실습 및 인턴십 프로그램’, ‘전공 관련 진로 탐색 과목’, ‘전공교수 진로·취업상담’ 순으로 실제 취업을 위한 정보획득의 요구가 높았다. 실제 현장실습 및 인턴 프로그램은 대학별로 ‘단기현장실습’이나 ‘기업연계형 장기현장실습(IPP)’과 같은 프로그램을 운영하

12) 장현진. (2017). 『대학진로교육 현황조사(2017) 기초통계표』. 세종: 교육부.

고 있으나 장기현장실습보다는 4주~12주 이하의 단기 현장실습의 비율이 높으며, 기업들 역시 직무교육을 위한 시간과 비용 문제로 인하여 단순 작업 위주의 실습 진행이 많으므로 실제 학생들이 원하는 현장 업무나 정보의 습득은 미비하다.¹³⁾

대부분 대학생이 취업을 목표로 하고 있지만, 실제 진로 선택과 진로 개발을 위한 정보 습득은 매우 제한적이며, 대학 역시 학생들에게 필요한 정보 제공은 부족하며, 전공 교수의 진로·취업 상담 역시 지도를 위한 시간 및 인프라의 부족과 정보 부족의 사유로 어려움이 있는 것으로 확인되었다.



[그림 1-6] 진로 및 취·창업 준비를 위해 대학에서 지원받고 싶은 사항

4) 4차 산업혁명 시대와 직업 환경의 변화

4차 산업혁명으로 변화하는 미래는 인공지능과 IoT, 클라우드, 빅데이터 등의 정보기술이 산업, 복지, 고용, 교육, 국방 등 다양한 분야와 결합하여 사회 전반의 '지능화'가 실현되는 지능정보사회가 될 것이다.¹⁴⁾ 4차 산업혁명으로 시작된 지능정보사회에서는 데이터와 지식이 기존 생산요소인 노동이나

13) 장후은, 허선영, 이종호. (2017). 대학의 현장실습 운영 실태 및 정책 과제. 『한국산학기술학회 논문지』, 18(2), 493-500.

14) 관계부처 합동. (2017). 『제4차 산업혁명에 대응한 지능정보사회 중장기 종합대책』.

자본보다 중요해지고 다양한 제품·서비스 융합으로 이중 산업간 경계가 붕괴되며, 자동화가 지적 노동 영역까지 확장되는 등 산업구조의 대대적인 변화 촉발에 따라 경제·사회 전반에 혁신적인 변화가 예상된다.¹⁵⁾ 이제 정보기술과의 융합은 컴퓨터공학과 같은 정보기술 관련 전공뿐만이 아닌 모든 대학생이 익혀야 하는 필수적인 역량으로 변해가고 있으며, 많은 대학이 비전공 학생들에게도 정보기술 역량 강화를 위한 소프트웨어 교육을 시행하고 있다.

4차 산업혁명에 따라 나타나는 직업 환경 변화로는 과학기술 간 융합과 기술혁신의 가속화로 사라지는 직업과 함께 새롭게 생겨나는 직업이 확대되는 양적 측면의 고용 변화, 플랫폼 기업의 증가와 업무환경의 변화로 인한 고용 관계 측면의 변화, 사회 격차의 증가와 같은 사회의 변화를 꼽을 수 있다.¹⁶⁾

5) 국가 진로지도 프로그램

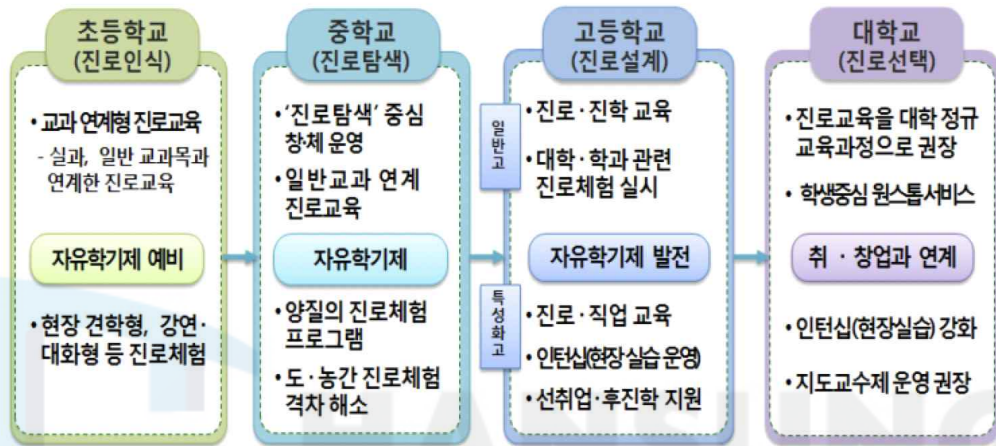
우리나라 진로교육은 변화하는 미래 직업 세계에 부응하는 창의적인 진로 결정과 개발을 돕는 역할보다는 여전히 ‘대학입시 중심의 교육’이라는 틀 속에서 수직적 상급학교 진학만을 지향하고 있으며, 창의적이고 도전적인 직업보다는 안정적이며 전통적인 직업에 대한 선호와 노동 시장에 진입 시기 등의 다양한 교육, 훈련, 고용의 문제점을 가지고 있다.¹⁷⁾ 정보기술의 발전으로 예상되는 일자리의 변화는 인공지능과 자동화의 확대에 의하여 단순 업무 일자리의 감소가 예상되며, 정보기술과 융합된 새로운 지식과 기술이 필요한 일자리는 증가할 것으로 예상된다. 빠른 기술과 사회변화에 대응하기 위해 기업 역시 직무 능력을 가진 경력직 위주의 채용 증가와 같은 고용환경의 변화를 불러오고 있으며, 이러한 변화는 신입 사원들이 경력을 쌓을 기회의 축소로 많은 대학생이 진로 문제에 직면하게 되었다. 국내의 진로교육은 2015년 「진로교육법」과 「진로교육법 시행령」 제정을 통해 진로교육 기본방향 및

15) 이상섭. (2017). 『4차 산업혁명시대의 최신 직업진로설계』. 경기: 양서원.

16) 어수봉. (2017). 『4차 산업혁명과 고용의 미래』. 서울: 경제사회발전노사정위원회.

17) 이지연. (2017). 모든 국민의 생애 전반을 지원하는 진로교육. 『서울교육』, 231.

지원체제가 확립되었으며, 기존의 학생 맞춤형 진로교육의 미흡과 대학생 및 사회적 배려 대상자에 대한 진로교육 부족과 같은 문제점을 보완하여 초등학교부터 대학과 평생학습에 이르기까지 생애 전반의 총체적 접근 안에서 학습자의 진로발달 단계별 우선 요구되는 과업들을 차근차근 준비할 수 있도록 지원하는 진로교육 체계를 [그림 1-7]과 같이 확정하였다.¹⁸⁾



[그림 1-7] 학교급별 진로교육 체계

하지만 대부분 고등학교는 여전히 대학 진학만을 목표로 하는 입시 위주의 교육으로 자신의 적성과 직업 세계에 대한 이해가 없이 진로가 결정되고, 대학에 입학하는 비율이 높으며, 이는 대학 기간의 학업의 만족도와 성취도에 부정적인 영향을 미치고 있다.¹⁹⁾ 대학생의 진로발달 수준은 전공 만족도에 유의미한 영향²⁰⁾을 미친다는 연구 결과와 같이, 고등학교에서의 진로에 대한 탐색과 자신이 희망하는 구체적인 직업의 특성을 파악하는 진로지도 중요성은 커지고 있다.²¹⁾

18) 교육부. (2016). 『진로교육 5개년 기본계획』.

19) 김봉환. (2000). 청소년 진로상담 정책의 현황과 과제. 『청소년상담연구』, 8, 146-169.

20) 윤수경, 한유경, 임소현, 김화영. (2015). 대학 및 전공 선택에 영향을 미치는 요인 분석. 『한국교육』, 42(2), 87-107.

21) ibid.

제 2 절 연구의 필요성

1) 대학생들의 진로 문제점 및 대안

국내 대학생들의 진로와 취업에 대한 고민은 수업과제와 성적, 재정 문제를 가장 큰 고민으로 조사된 미국 대학생들과 조사와는 다른 차이를 나타내고 있으며, 그 이유로는 11, 12학년에 진로지도와 취업 정보서비스를 제공하는 데 충력을 기울이는 미국과 달리 적성과 장래희망보다는 수능 성적에 따라 대학을 서열화하고 점수에 따라 대학과 전공을 선택하는 비율이 높은 제도적·사회적 문제점이 가장 큰 원인일 것이다.

대학생들의 진로에 대한 고민은 청소년기에 고민하고, 교육 시스템으로 지원받아야 할 진로 문제가 성적에 따라 부모님과 교사의 높은 영향에 의해 대학과 전공이 선택되고, 대학 입학 후 적성과 진로에 대한 고민이 시작되는 것도 이유 중 하나일 것이다. 2010년 한국고용정보원의 '대졸자 직업이동 경로 조사'에서 대졸자의 전공 불일치 비율이 50.3%로 조사되었으며, 전공 혹은 업무 분야에 따라 대학을 선택한 학생은 37.9%로 대부분 성적과 부모/교사 권유에 의해 대학과 전공을 선택한 것으로 조사되었다.²²⁾ 또한, 전공 불일치 학생들이 상대적으로 대학 생활에 더 어려움을 겪고 있으며 이는 중도탈락률을 높이는 원인이 되기도 한다.²³⁾ 이러한 문제에 대한 대안으로 대학별 진로지원조직을 운영하거나 일부 대학의 경우 트랙 제도와 같이 타 전공과의 연계를 통해 적성에 맞게 진로를 개발할 수 있는 방향을 제시하고 있으나, 개개인의 학생별 맞춤형 상담 및 지원을 위해서는 충분한 시간과 배경 데이터가 필요하며 전공에 대한 깊은 이해도를 가진 상담·조언자들이 필요하지만, 인력과 재정 등의 문제로 학생 개개인의 맞춤형 지원은 한계가 있다.

과거와 달리 오늘날의 직업은 다양한 산업이 융합된 형태로 확장되고, 발전하고 있으므로 과거와 같이 대학 기간 관련 분야를 넓게 학습하는 형태에

22) 채창균. (2013). 『대학생의 실제 전공과 희망전공 간 불일치 실태와 과제』. 세종: 한국직업능력개발원.

23) 박은선, 김광현. (2016). 학업중단요인 및 대학생활적응요인이 대학 신입생의 중도탈락의도에 미치는 영향. 『청소년학연구』, 23(10), 361-384.

서는 기업의 요구사항을 맞추기가 어렵다. 대학 역시 학생의 직업 역량 향상을 위해 특정 기술 분야에 대해 깊게 학습하거나 다양한 학문과 융합하는 세부 전공 및 트랙 제도를 운영하고 있으며, 타 대학과의 학점 공유나 정보기술의 발전으로 인하여 MOOC(Massive Open Online Course)와 같은 다양한 학습방법을 통해 지식의 습득과 진로 개발을 지원하고 있다.

진로의 개발이 다양해진 만큼 진로 개발 방법에 대한 복잡도 역시 매우 높아졌기 때문에, 자신의 진로와 진로 개발을 위한 정보를 찾고 분석하기 위해서는 많은 시간과 노력이 필요하며, 정보의 양 또한 부족하다. 학생들이 최선의 진로 선택과 개발을 지원하기 위해서는 빠르게 변화하는 직업 관련 데이터를 대학에서 실시간 분석하여 학생들이 필요로 하는 양질의 진로정보를 지속적으로 제공하여야 한다. 학생들에게 양질의 정보를 제공하기 위해서는 광범위한 형태의 정형·비정형 직업 및 구인 관련 데이터들을 실시간으로 추출·가공하여 최신의 정보를 학생들에게 제공하고, 분석된 직업 관련 기술, 자격, 태도와 같은 역량과 직업 간의 관계를 분석하여 학생의 데이터와 매칭한 맞춤형 추천 및 조언을 제공할 수 있어야 한다.

2) 학생과 기업의 역량 미스매치

[표 1-2]의 ‘대학생과 인사담당자의 직업 역량 비교 조사’²⁴⁾에서 학생들이 생각하는 취업을 위한 역량으로 외국어 능력(1순위), 구직서류 작성 능력(2순위), 구직 희망 분야 이해(3순위)를 선택했지만, 기업체 인사담당자들이 생각하는 역량 순위는 직업윤리(1순위), 도전정신(2순위), 자기 이해(3순위)와 같이 큰 차이가 있었으며, 대학생들이 취업을 위해 가장 중요하다고 생각하는 외국어 능력은 인사담당자들의 평가에서 20위로 맨 하위인 것으로 나타났다. 해당 조사에서 학생이 생각하는 역량과 기업이 요구하는 역량은 많은 차이가 있음을 확인할 수 있으며, 이런 차이는 기업의 형태와 직업에 따라 차이가 있으며, 정보기술의 발전과 사회문화에 따라 계속 변해가고 있지만, 기업이 필요한 최신의 역량과 지식에 대한 정보를 학생들에게 제공하지 못하고 있다.

24) 김은석. (2015). 『대학생과 인사담당자의 직업 역량별 비교』. 충북: 한국고용정보원.

[표 1-2] 대학생과 인사담당자의 직업 역량별 비교

구분		학생 순위	인사담당자 순위
구직 지식 군	자기 이해	4	3
	구직 희망분야 이해	3	6
	전공 지식	7	16
	외국어 능력	1	20
	구직 일반 상식	9	18
구직 기술 군	구직의사 결정 능력	10	14
	구직정보 탐색능력	11	19
	인적 네트워크 활용 능력	12	17
	구직서류 작성 능력	2	11
	구직의사 소통 능력	5	9
구직 태도군	긍정적 가치관	19	4
	도전 정신	15	2
	글로벌 마인드	14	13
	직업윤리	20	1
	직무 및 조직 몰입	17	8
구직 적용군	현장 직무 수행능력	6	15
	대인관계 능력	16	7
	문제해결 능력	8	5
	자원활용 능력	13	12
	자기관리 및 개발 능력	18	10

3) 고용정보원의 직업정보 지원 서비스

한국고용정보원은 고용지원서비스 선진화를 위해 「고용정책기본법」에 의해 2006년 설립된 고용노동부 산하 준정부기관이다. 한국고용정보원은 고용정보의 수집/분석/제공, 직업에 관한 조사·연구, 고용 관련 전산시스템의 개발·관리·운영 업무를 수행하고 있으므로²⁵⁾ 국내에서 가장 많은 직업정보를 보유하고 있으며, 공공기관의 「공공데이터 제공·이용 활성화에 관한 법률」에 의해 다양한 데이터를 오픈 API를 활용하여 등급별로 대학·직업정보제공기관·민간 기관에 제공하고 있다.

25) 한국고용정보원. <http://www.keis.or.kr/>

[표 1-3]은 고용노동부에서 지원하는 대학생 진로·취업지도 지원 서비스로 직업 심리검사, 진로 관련 교육, 콘텐츠, 일자리 정보를 대학과 대학생들을 대상으로 API와 웹사이트(<https://www.work.go.kr>)를 통해 제공하고 있다.

[표 1-3] 고용정보원 대학생 진로 취업지도 지원 서비스

No.	서비스명	서비스 내용
1	진로 탐색 지원을 위한 직업 심리검사	대학생의 자기 탐색 및 진로 방향설정 지원 시 활용 가능한 9종의 무료 직업 심리검사
2	구직역량 강화를 위한 진로·취업 지원 콘텐츠	취업 지원 교육프로그램 활용 교육 및 자료
3	워크넷 사이버진로교육	자기 주도적 진로설계 및 직업준비 지원을 위한 인터넷 강좌
4	직업 및 진로 콘텐츠	뉴스레터, 진로정보 활용 가이드
5	청년 워크넷	청년취업 지원 정책 및 청년 일자리 정보
6	사이버 진로교육센터	진로 개발을 위한 관련 강좌

정보기술의 발전으로 인하여 직업 환경은 빠른 속도로 전통적인 산업들과 융합하여 새로운 직업을 생성해 내고 있으며, 이에 따라 정보기술 관련 직업들도 다양한 세부 직무로 나뉘어 가고 있다. 그러나 대부분의 공공기관에서 실시하는 직업 관련 조사는 전통적인 표준직업분류를 기준으로 조사하고 있으며, 조사 항목도 한정되어 있어 실제 대학생들이 필요로 하는 직업정보 제공은 부족하다.

[그림 1-8]은 워크넷 직업정보에서 ‘시스템소프트웨어 개발자(프로그래머)’의 지식에 해당하는 부분을 검색한 결과이다. 가장 높은 중요도에 해당하는 지식의 내용은 ‘컴퓨터와 전자공학’과 같이 실제 학생에게 습득해야 할 기술을 안내하는 것이 아니라 범위가 매우 방대하며, 이러한 정보만으로 학생 스스로 진로를 설계하는 것은 불가능하다. ‘프로그래머’에 대한 직업정보 중 지식을 검색하는 학생이 원하는 정보는 현재 기업들이 어떤 세부 지식·기술을 필요로 하는지 안내되어야 해당 지식에 대한 진로 개발이 가능하다.

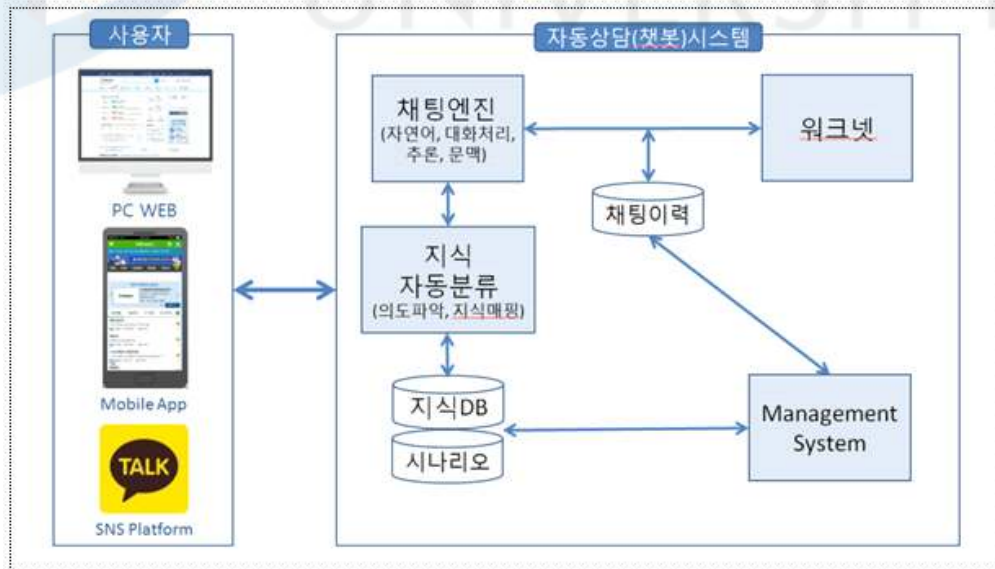
· 지식 중요도

중요도 낮음 0 100 높음

중요도	업무수행능력	설명
95	컴퓨터와 전자공학	컴퓨터의 하드웨어, 회로판, 처리장치, 반도체, 전자장비에 관한 지식
79	영어	영어를 읽고, 쓰고, 듣고 말하는데 필요한 지식
79	공학과 기술	다양한 물건을 만들고 설계하거나 서비스를 제공하기 위해 필요한 공학적인 원리, 기법, 장비 등을 실제로 적용시키는 지식
76	통신	전화기, 네트워크, 방송 등의 통신기기를 조작하고 통제하는데 필요한 지식

[그림 1-8] 워크넷 직업정보 시스템소프트웨어 개발자 지식중요도

정보기술을 활용한 컴퓨터와 사용자 간 자동질의 및 응답 시스템 기술 적용으로 효율적인 일자리 상담서비스 구현을 위해 한국고용정보원은 2019년 3월 채용정보 상담 챗봇 ‘고용이’의 시범서비스를 공개하였다. 채용정보 상담 챗봇 ‘고용이’는 ‘채용정보’, ‘개인추천’ 서비스를 제공하고 있으나 [그림 1-9]와 같이 워크넷 내부의 API를 기반으로 구축되어 제공이 가능한 정보의 한계와 특정한 몇 개의 시나리오 이외에는 답변하지 못하고 있으며, 실제 직업과 진로에 관한 지원은 아직 부족하다.



[그림 1-9] 고용노동부 챗봇 고용이 시스템 개념도

고용노동부의 대학 지원 서비스는 인지도의 부족²⁶⁾과 학생들이 필요한 기업에서 필요로 하는 역량에 대한 정보가 부족하며, 챗봇 서비스인 고용이의 경우 자연어의 인식, 메뉴 찾기의 어려움, 여러 번 반복되는 질문, 자료의 부족, 극히 일부의 대화 스크립트만 지원이 가능한 문제로 대학생들을 위한 직업정보 제공과 진로지원 서비스로는 부족하다.

4) 국내 대학들의 진로·취업 역량 강화를 위한 정보시스템

국내 많은 대학은 체계적인 직업 역량 강화와 진로 및 취업 정보 제공을 위해 진로·취업 정보시스템을 도입하여 운영하고 있다. 진로·취업 관련 정보시스템은 통합 경력관리 시스템, 학과별 진출 분야 제공 및 기업평가 시스템, 정보자원에 기반한 맞춤형 지원시스템으로 분류하기도 하나²⁷⁾ 대부분 진로·취업 통합경력관리시스템에 다양한 기능들이 통합된 형태로 운영되고 있다.

가) 진로·취업 통합경력관리시스템

통합경력관리시스템은 학생이 다양한 진로·취업 정보를 탐색하며 자기 주도적으로 경력설계·개발이 가능하고, 대학은 학생의 진로·취업 관련 활동의 체계적 관리와 맞춤형 진로상담 및 프로그램 제공에 활용할 목적으로 구축된 시스템²⁸⁾으로 학생의 경력과 역량 관리를 위한 포트폴리오, 진로 및 취·창업 교과/비교과 정보 제공 및 참여 관리, 진로지도 및 상담 관리, 취·창업 정보 제공 및 매칭 기능을 제공하며, 대학의 특성에 따라 지원 기능의 차이가 있다. [그림 1-10]은 S여대 학생경력관리시스템으로 나의 비전을 수립하고, 위크넷 자료를 활용한 직업정보를 활용한 진로정보 제공, 역량 개발을 위한 비교과 프로그램 및 e-포트폴리오, 취업 정보 제공 및 온라인 교육, 지도교수 상담, 커뮤니티로 구성되어 있으며 PC부터 스마트폰까지 다양한 환경의 기기를 지원하고 있다.

26) 이요행, 강옥희. (2010). 『대학교 직업진로지도 현황과 과제』. 충북: Employment Issues.

27) 고용노동부. (2016). 『대학경영진을 위한 진로·취업지원 전략가이드』.

28) ibid.



[그림 1-10] S여대 학생경력관리시스템 시스템 구성

나) 맞춤형 진로지원 및 학사지원을 위한 국내 대학 발전 계획

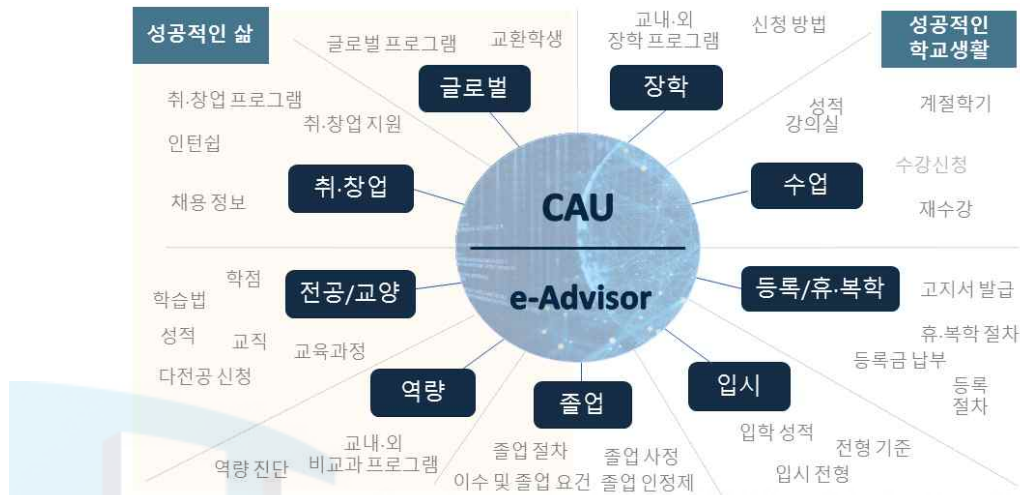
2019년 7월 공개된 한양대학교의 HY-CDP 고도화 계획에서도 4차 산업혁명 도래에 따른 빅데이터 기반의 미래혁신을 위해 학생 중심의 개인화된 시스템 구축과 구성원별 맞춤형 정보서비스 제공 계획을 확인할 수 있다. HY-CDP 고도화 계획에서는 [그림 1-11]과 같이 대학 내의 다양한 정형·비정형 데이터를 활용한 학생 맞춤형 역량 로드맵 추천, 가상 컨설턴트, 기업 맞춤형 인재 매칭, 챗봇 기반의 인터페이스를 포함하고 있다.



[그림 1-11] 한양대학교 빅데이터 기반의 HY-CDP 시스템 고도화 계획

2019년, 중앙대학교의 ‘인공지능 기술 기반의 e-Advisor 구축을 위한 BPR/ISP 용역 및 대화형 시스템 구축’ 계획에서도 [그림 1-12]와 같이 인공

지능 기술을 통한 자동질의 및 응답 시스템 기술인 대화형 시스템을 구현하여 인간과 유사한 대화기반의 상담서비스를 제공하고, 이후 교내 데이터 통합을 통한 학생 맞춤형 e-Advisor 시스템 구축계획이 공개되었다.



[그림 1-12] CAU e-Advisor 학사행정 서비스 예시

이처럼 각 대학에서는 학생의 역량 강화와 맞춤형 학생지원 서비스에 대한 필요성 증대로 고도화 계획을 수립·준비하고 있는 단계로 정보기술을 활용한 학생 진로 문제해결과 자연어 기반 맞춤형 챗봇 조인 시스템에 관한 연구의 필요성이 증가하고 있다.

국내 대학에서 도입된 학사지원 관련 챗봇 서비스들은 대부분 스크립트 방식으로 구성되어 정해진 스크립트에 대한 답변만 가능하다. 따라서 지능형 챗봇 서비스를 위한 학생 맞춤형 학사지원이나 맞춤형 진로·취업 지원 서비스 연구의 필요성이 증가하고 있으나, 머신러닝 학습에 활용할 수 있는 데이터가 매우 부족하고, 제한적이다. 따라서 머신러닝을 활용한 맞춤형 학사 또는 진로·취업 지원 서비스 도입을 위해서는 데이터 수집 전략부터 활용 연구, 데이터의 수집, 실제 학습 및 구축단계까지 진행되기까지 3년 이상의 기간이 필요할 것으로 검토된다.²⁹⁾

29) 중앙대. (2019). 『인공지능 기술 기반의 e-Advisor 구축을 위한 BPR/ISP 용역 및 대화형 시스템 구축 사업 계획』.

머신러닝 분야에서 빅데이터를 얻기 위해 데이터를 분류·라벨링하고 정제하는데 많은 인력과 비용의 문제와 각종 규제 역시 큰 장애물이다. 따라서 데이터에서 상관성 패턴을 찾아내는 학습이 아닌, 적은 데이터에서의 관계를 찾아 학습하는 머신러닝 알고리즘이 활발히 연구되고 있다.

이러한 이유로 적은 데이터만으로도 지식 그래프 모델링이 가능하고, 관계에 따라 추론·분류·매칭·유사도 분석 등에 활용과, 관계 기반 머신러닝 학습이 가능한 그래프 데이터베이스를 활용한 연구의 필요성이 요구된다.

5) 정보기술을 활용한 국내외 대학 혁신 사례

가) 대학의 혁신 사례

인공지능, 빅데이터 분석 기술의 진보로 디지털 혁명이 가속화됨에 따라, 직업 및 직종의 변화, 직업인이 갖추어야 할 새로운 역량의 등장과 직업 역량의 변화에 따라 대학은 미래사회에 적합한 인력 양성을 위한 역할 재정립과 변화가 필요한 시점이다. 대학 역시 지금까지는 대학 입학 정원이 학령인구보다 낮아 신입생의 모집에 큰 어려움이 없었다면, 이제는 대학의 생존을 위해 혁신을 통해 대학 자체의 경쟁력을 강화가 요구되는 시점이다. 특히 대학의 진로·심리 상담과 같은 학생들의 진로지원 활동과 취업률이 대학의 역량을 진단하는 지표로 평가되고 있기에 진로 부문에 대한 경쟁력 강화와 혁신의 중요성은 매우 높다.

종합대학 모델을 정착시킨 미국의 대학 역시 급격한 정보기술과 노동 시장의 변화로 인하여 전통적인 교육 시스템은 큰 위협을 받고 있다. 특히 낮은 졸업률과 입학포기, 중도탈락의 문제는 심각한 재정적 문제를 야기 시키며, 이와 같은 문제를 해결하기 위해 미국 대학들은 빅데이터 분석, 인공지능 등의 첨단 정보기술을 적극적으로 활용하여 혁신을 지속하고 있다.³⁰⁾ 국내 대학들 역시 MOOC, 플립러닝(flipped learning) 등의 새로운 교수법 도입이나 다양한 학생들의 진로·창업을 위한 다양한 프로그램을 지원하고 있으며, 정보기술을 활용한 맞춤형 지원을 위해 데이터 수집과 분석을 위한 고도화 계획

30) 김종율, 노광현. (2019). 공공 데이터를 활용한 대학생 진로 조언을 위한 그래프데이터베이스 기반 지식베이스 구축. 『전자공학회논문지』, 56(10), 41-48.

을 수립하고 있다.

대부분의 국내 대학은 학부생 중심으로 운영되고 있으므로 연구보다는 인력 양성의 목적이 더 크며, 사회에서 요구되는 역량을 지닌 인력 양성을 위해서는 학생 개개인에 대한 정확한 진로 조언과 진로 개발을 지원하기 위한 연구가 필요하다.

나) 애리조나 대학교 ‘e-Advisor’ 사례

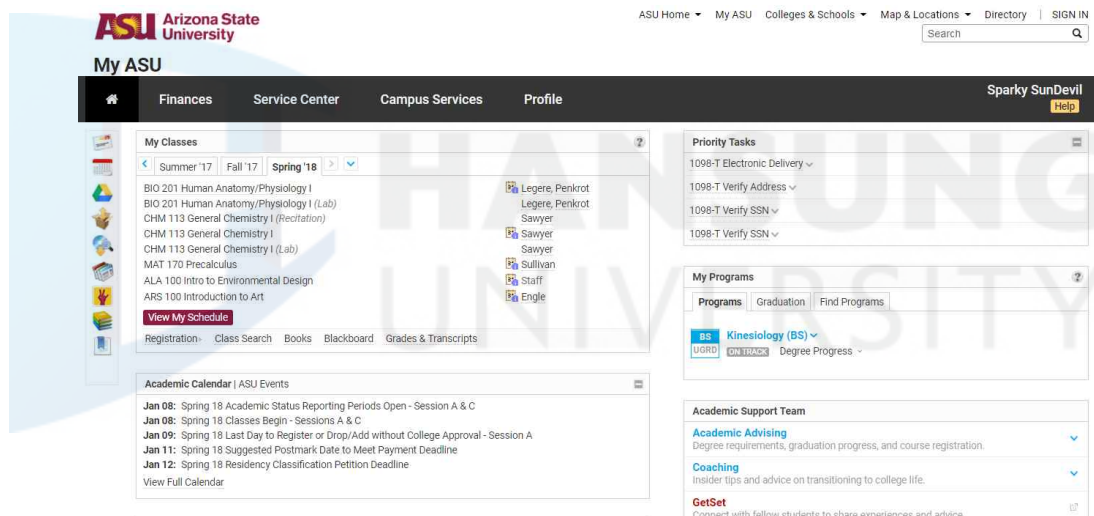
해외 대학들의 경우 대학의 다양한 문제에 대해서 현대 정보기술을 적극적으로 활용하여 문제해결의 성과를 보이고 있다. 해외 대학의 혁신 사례로 대표적인 대학은 유에스뉴스앤드월드리포트 선정 2015년부터 2017년까지 가장 혁신적인 대학 1위에 선정된 애리조나 대학의 ‘e-Advisor’가 대표적이다. ‘e-Advisor’란 정보기술을 활용하여 대학과 대학생들의 가장 큰 문제점인 진로상담과 수강 신청 문제해결을 중심으로 학업 및 학습 관리를 지원하는 소프트웨어를 말하며, 학생지원 부서를 통한 학업 지원의 한계 문제를 정보기술을 활용하여 해결하기 위해 설계되었다.

전통적인 상담 시스템의 복잡성과 비효율성 문제를 컴퓨터 기술과 데이터 분석 기법의 발전에 따라 학생 개인별 성공 패턴의 체계적인 분석이 가능해짐에 따라 이를 활용하여 성공적인 학업 진행을 예측하는 기준 개발을 통해 개별 학생의 성과를 예상되는 성공 패턴과 매핑하고, 학생들의 진행 과정에 따라 ‘e-Advisor’는 현재 이수 과정과 추가 이수해야 할 과정을 결정하여 제안하며, 전공 변경을 위한 최적의 경로를 제안한다.³¹⁾

‘e-Advisor’는 [그림 1-13]과 같이 학생들의 학업 성취에 관한 피드백을 제공하고, 학생 개개인의 관심사와 역량에 적합한 전공을 안내해주며, 대학에 상주한 전문 컨설턴트는 이 프로그램을 통해 졸업 요건들을 사전에 모니터링하여 피드백을 제공하고, 학생들은 학사일정 및 졸업 요건 충족을 위해 필요한 정보를 개개인의 성장과 진도에 맞춤형 피드백을 받을 수 있게 되었다. 또한, 학생들의 출석, 학습 진행 상황, 과제 제출 정보를 분석하여 조기 알람과 필요한 경우, 수강 취소와 대체 교과들을 제안하기도 한다.³²⁾

31) 김종율, 노광현. (2019). AI 를 활용한 대학생 진로 조언 시스템 모델 및 데이터 수집과 융합에 대한 연구. 『디지털융복합연구』, 17(2), 177-185.

전공 선택에 어려움이 있는 학생을 위해 모바일 기반의 맞춤형 진로·학과 추천 애플리케이션인 ‘me3’를 통해 홀랜드의 직업적 성격유형에 기반의 학생에게 적합한 진로 및 전공 데이터를 제안하고 있으며, 학생들의 기초 수학 능력의 약화와 이로 인하여 관련 전공 심화 과목 이수에 어려움이 있다는 문제를 해결하기 위한 적응형 학습을 도입하여 학생의 학습 행동을 분석해 학생의 해당 항목별 이해 정도와 학습방법 등을 파악하고, 분석된 데이터를 기반으로 학생별로 맞춤형 학습을 지원한다.³³⁾ 적응형 학습 도입 이후 애리조나 대학교는 기초 수학 과정 통과 학생 비율이 64%에서 75%로 증가하고, 중도 포기 학생 비율이 16%에서 7%로 감소했고 조기 수료 학생 비율이 한 명도 없었던 0%에서 45%로 증가했다는 성과를 공개했다.³⁴⁾



[그림 1-13] 애리조나 대학교 ‘e-Advisor’

다) 조지아주립대학교 ‘Pounce’ 사례

미국의 대학에 지원하는 과정은 재정지원 신청서 작성, 최종 고등학교 성적 증명서 제출, 예방 접종, 대출 및 수업료 납부와 같은 여러 가지 어렵고

32) Page, L. C., & Gehlbach, H. (2017). How an artificially intelligent virtual assistant helps students navigate the road to college. AERA Open, 3(4).

33) 김진숙. (2017). 『학습자 맞춤형 학습 실현을 위한 지능정보기술』, 한국교육신문.

34) Phillips, E. D. (2013). Improving advising using technology and data analytics. Change: The Magazine of Higher Learning, 45(1), 48-55.

복잡한 서류와 환경적인 문제로 입학에 포기하는 서머멜트(Summer-Melt) 문제를 해결하기 위하여 AI가 주도하는 메시징 플랫폼인 'Pounce'를 도입하였다. Pounce는 각 학생이 어디에서 어려움을 겪고 있는지 그리고 어디서 도움이 필요한지 이해하기 위해 입학 전 프로세스에 대한 대학 내부 자료를 분석한다. 이 정보를 사용하여 모든 사람에게 일반 메시지를 보내는 대신 도움이 필요한 학생들에게만 연락할 수 있도록 개인화된 텍스트 메시지를 생성하며, 자연어처리 기술을 활용하여 24시간 다양한 문의에 대한 원활한 상담을 지원한다. 또한, 딥러닝 알고리즘 기술을 기반으로 학생들과 상호 작용하면서 더 스마트하게 성장하고 있다.³⁵⁾ 실험 연구 결과에 따르면 Pounce의 지원을 받은 학생들은 필요한 사전 프로세스를 완료했고 지원받지 않은 학생들보다 훨씬 더 높은 비율로 등록했다. 그 결과, 조지아주립대학교는 고등학교 졸업과 대학 입학 사이인 여름 동안의 학생들의 동기 부여가 사라지는 현상(Summer-Melt)인 진학 포기율이 21% 감소하였다.³⁶⁾

라) 베를린공과대학 'Alex' 사례

베를린공과대학의 'Alex'는 AI를 사용하여 학생들의 커리큘럼 검색, 수강 신청 지원과 같은 복잡하고 시간이 많이 소요되는 문제를 해결하기 위하여 도입되었다. 학생들은 온라인 시간표를 검색하는 대신 수업 시간, 강의 대상자 및 응시할 시험과 같은 일반적인 질문을 Alex에 입력하고 봇은 자연스러운 문장을 생성하여 답변하며, 후속 질문을 통해 더 세부적인 정보를 제공할 수 있게 되었고, 챗봇 도입을 통해 기존의 온라인상에서의 복잡한 정보 검색의 시간을 단축할 수 있게 되었다.³⁷⁾

35) Williamson, B. (2018). The hidden architecture of higher education: building a big data infrastructure for the 'smarter university'. International Journal of Educational Technology in Higher Education, 15(1), 12.

36) Page, L., & Gehlbach, H. (2018). How Georgia State University used an algorithm to help students navigate the road to college. Harvard Business Review.

37) Cortez, M. B. (2018). Universities Deploy Chatbots to Aid Students in the Admissions Process and Beyond. EdTech.

제 3 절 연구의 범위와 목적

1) 진로 조언 플랫폼 범위

가) 진로 조언의 정의

100년 전, 미국에서 주 교통수단이 마차에서 자동차로 바뀌던 2차 산업혁명기에 프랭크 파슨즈(Frank Parsons)는 보스턴 시민청에 직업국을 세워서 방황하는 구직자들에게 체계적인 진로상담을 제공하였다.³⁸⁾ 파슨즈는 특성-요인 이론에서 “감독이 배우에게 가장 적합한 역할과 각본을 제시해 주면 배우는 각본대로 열심히 연기함으로써 성공을 보장받을 수 있다는 것처럼 진로상담사는 개인의 재능과 가장 잘 맞는 직업을 연결해 주는 역할을 해야 한다.”고 하였다.³⁹⁾

진로지도는 파슨스의 직업지도로부터 시작되었으며, 개인의 적성을 분석하여 그에 맞는 직업을 선택하는 것을 주요 내용으로 삼고 있다. 파슨스의 직업지도는 ‘첫째, 자신의 능력과 흥미 등에 대한 이해’, ‘둘째, 직업에 관한 각종 정보의 제공’, ‘셋째, 직업선택을 올바르게 하도록 돕는 과정’으로 구성되어 있다. 미국의 국립직업지도협회(The Vocational Guidance Association)에서는 “직업지도란 어떤 개인이 일정한 직업을 선택하고 그 직업에 임하는 준비를 하며, 또한 취업하여 향상, 발전하는 것을 돕는 과정이고, 직업지도의 주요 과업은 각 개인이 장래 계획을 세우고 또 자신의 직업에 만족하게 적용할 수 있도록 결정과 선택을 도와주는 것”이라고 정의하고 있다.⁴⁰⁾

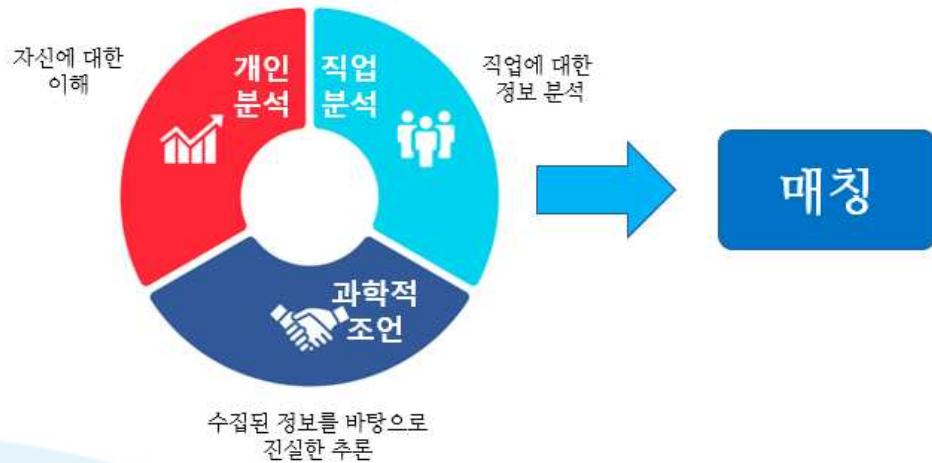
파슨스는 [그림 1-14]와 같이 개인 분석, 직업 분석, 과학적 조언을 통한 매칭을 주장하였는데, 이는 자신의 강점과 약점을 포함한 개인적 성향을 충분히 이해하고, 주어진 직업에서의 성공조건 및 보상과 승진에 관한 정보를 알아야 하며, 입수한 정보를 바탕으로 선택과정에서 ‘진실한 추론’을 해나가야 한다는 것이다.⁴¹⁾ 이러한 특성-요인 이론은 정서적인 불안으로 불확실한 직업선택과 같은

38) Zytowski, D. G. (2001). Frank Parsons and the progressive movement. The career development quarterly, 50(1), 57.

39) J. G. Maree. (2019). 『진로구성상담의 원리와 실제』, 서울: 학지사.

40) Katz, M. (1966). A Model of Guidance for Career Decision-Making. Vocational Guidance Quarterly, 15(1), 2-10.

문제점이 있지만, 개념이 간결한 이유로 많은 직업상담자가 활용하고 있기 때문에 연구 수행을 위한 진로지도 이론으로 선정하였다.



[그림 1-14] 파슨스 특성-요인 이론, 진로지도의 3요소

나) 대학에서의 진로 조언

국내에서의 진로지도는 주로 초등학교와 중학교에서 적극적으로 진행되고 있으며, 성적과 입시 위주의 경쟁과 교육으로 인하여 학교급별로 올라갈수록 진로 교육에 대한 만족도가 낮아지고, 장애 요인이 되고 있다.⁴²⁾

대학에서의 진로지도는 이미 전공이 결정된 이후이기 때문에 자신의 전공에 대한 세부 진로 방향에서 자신의 능력과 적성, 흥미 등을 찾고 자신에게 적합한 진로와 진로를 개발하기 위해 지원하는 방향으로 진행된다. 예를 들어, 공학 관련 학과에 입학하는 신입생들은 학과를 선택한 다양한 이유가 있으며 많은 학생이 공학 분야에 관심이 있지만, 특정 공학 분야를 명확하게 선택하는 것은 어렵다. 수년에 걸쳐 공학 분야, 특히 정보기술 관련 직무는 더 광대하고 깊어졌기에 많은 세부 직업들이 있으며, 이러한 세부 직업들은 학생들에게 다양한 선택을 제공할 수 있게 된다. 학교에서의 조언은 학생들이 세부 직업선택과 같은 중요한 결정을 내리도록 지속해서 지원해야 하며, 학생이 세부 직업선택과 같은 주요 결정을 내

41) 조만업, 정은선. (2014). 『직업 상담 실무』. 서울: 정훈사.

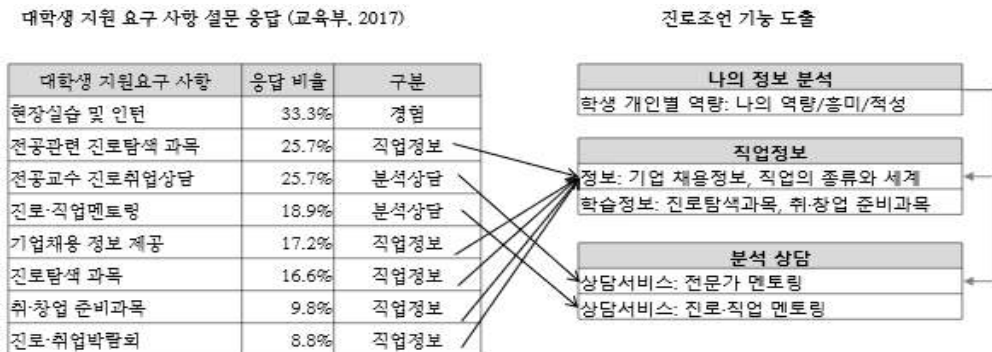
42) 장현진. (2015). 『진로교육의 현황과 과제』. 세종: 한국직업능력개발원.

리는 것은 학생의 미래 직업 목표뿐만 아니라 경우에 따라 학생의 미래에도 영향을 줄 수 있으므로 진로 조언 서비스 제공은 학생들과의 성공적인 관계를 구축하고 학생들의 흥미와 참여를 유도해야 한다.

정보기술을 활용한 진로지도 역시 그 과정은 동일하지만, 다양한 데이터와 관계 분석을 통해 대학생들에게 최적화된 정보를 제공하고, 효율적이고 올바른 선택을 지원하며, 정보를 통해 학생들이 자율적으로 진로 목표를 결정하고, 목표 달성을 위한 역량 개발을 지원하는 정보시스템이라 할 수 있다.

2) 진로 조언 모델 기능 도출

진로 조언은 현대 정보기술을 활용하는 것 이외의 목적과 정의는 전통적인 직업지도와 동일하다. 대학생들의 요구사항 분석을 위해 교육부의 ‘2017년 대학 진로교육현황조사’에서 ‘진로 및 취·창업 준비를 위해 대학에서 지원받고 싶은 사항’으로 응답한 대학생의 ‘진로지원 요구사항’을 기능별로 매핑하여 진로 조언 모델이 갖추어야 할 기능을 [그림 1-15]와 같이 진로를 위한 학습정보와 진로·취업 정보에 해당하는 ‘직업정보’, 전공 관련 전문 조언 서비스인 ‘분석 상담’, 학생의 역량 정보를 활용한 맞춤형 직업 추천을 위한 ‘나의 정보 분석’으로 도출하였다. 도출된 기능은 파슨스의 직업지도에서 정의한 자신의 능력과 흥미에 대한 이해, 직업정보의 제공, 직업선택 과정, 직업에 임하는 준비를 위한 결정과 선택을 도와주는 진로지도의 특성-요인 이론과 일치한다.



[그림 1-15] 진로 조언 모델에 필요한 기능 도출

[표 1-4]는 도출된 진로 조언 플랫폼 기능에서 각각의 기능에 대한 세부 요구 사항이며, 요구사항을 기반으로 정보기술로 구현해야 할 범위와 수집해야 할 데이터 및 구현에 최적화된 기술을 분석한다.

[표 1-4] 진로 조언 기능별 요구사항

기능	요구사항
나의 정보 분석	· 기술·지식, 역량, 자격증, 성향, 학과 등의 학생 개인의 상태 데이터
직업정보·진로 개발 정보	· 기업 채용정보 분석 · 직업의 종류와 세계 · 능력과 역량 정보 · 흥미와 적성 정보 · 연봉, 미래 전망 · 진로 탐색 과목 · 취·창업준비과목
분석 상담	· 학생의 다양한 질문에 대한 자연어 질문에 대한 답변 기능 · 학생의 정보를 기반으로 한 직업 추천 · 역량 정보와 직업과의 관계 분석을 통해 학생이 진로를 선택하고 개발하기 위한 정보 제공

[표 1-3]에서 기능별 요구사항기능 대한 정보기술을 활용한 구현 방안을 찾는다면 [표 1-5]와 같이 학생 이력 관리 프로그램, 빅데이터 수집·전처리 기술, 지식 베이스·그래프, 대화형 앱·웹 인터페이스와 같은 기술로 구현할 수 있다.

‘나의 정보 분석’은 실제 학생의 특성을 확인하기 위한 직업심리검사 또는 개인의 역량 정보를 활용한 분석이 필요하다. 직업심리검사의 경우 고용정보원에서 각 대학에 공공데이터로 제공하고 있으며, 학생의 개별 역량 정보는 학교 내부의 학사정보시스템이나 e-포트폴리오에서 데이터를 보유하고 있으나 개인정보 보호법상 민감 정보에 해당하는 데이터가 많기 때문에 활용에 한계가 있다. 때문에, 매칭되는 데이터를 통한 추천 시나리오를 테스트하기 위해서는 별도의 e-포트폴리오 형태의 프로그램이 필요하다.

‘기업 채용정보’의 경우 실제 상담 사례나 선행연구⁴³⁾들을 통해 기업들이 채

43) 천영민. (2015). 『채용 경향 변화 분석 및 이를 활용한 취업진로지도 방안 연구』. 충북: 한국 고용정보원.

용정보 분석을 통해 기업들의 가장 최근의 트렌드를 반영한 채용정보의 중요성과 이를 통한 체계적인 진로지도의 필요성이 부각 되었으며, ‘진로탐색과목’과 ‘취·창업 준비 과목’ 역시 학생들이 필요로 하는 부분은 실제 취업 시장에서 기업이 원하는 역량 정보로, ‘진로·취업 박람회’에서의 요구는 실제 현장 전문가의 상담 및 멘토링(조언)에 대한 부분이며 청년들의 채용박람회 참여는 조사결과기업의 채용 정보를 얻는 목적(46%)이 큰 것으로 확인되었지만 실제 만족도는 높지 않았다⁴⁴⁾, ‘직업의 종류와 세계’는 직업정보 데이터, 경제적 안정성은 직업별 임금 분석과 미래 전망 데이터가 필요하다.

‘분석 상담’에서 ‘전문가의 멘토링’, ‘진로·직업 멘토링’의 경우 취업 해당 직업 전문가의 상담과 멘토링을 원하는 것이기 때문에 전공 관련 전문가 수준의 지식 베이스 구축의 필요성과 이러한 지식베이스를 통한 지식의 전달과 조언 알고리즘의 역할이 중요시된다.

[표 1-5] 요구 기능에 대한 구현 가능 정보기술

기능	요구사항	구현 가능 정보기술
학생 개인의 역량 데이터	학생 정보 관리 프로그램	· 학생 이력 관리 프로그램(e-포트폴리오)
기업 채용 및 직업정보 데이터	데이터 수집·가공·저장	· 빅데이터 수집 기술 · 빅데이터 전처리 기술
축적된 진로 관련 지식의 저장 및 관리	다양한 경로를 통해 저장된 지식의 저장	· 지식 베이스, · 지식 그래프
상담 및 조언	사용자와의 인터페이스	· 대화형 인터페이스 · 웹·앱 인터페이스

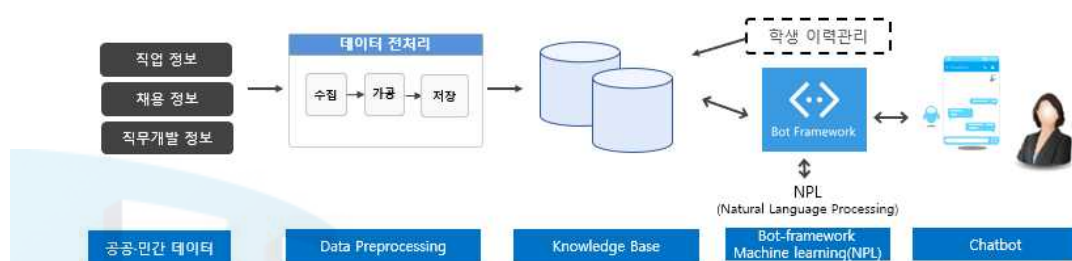
3) 진로지원 플랫폼 기본 모델 설계와 기술 요소

가) ‘진로 조언’ 플랫폼 기본 모델

진로 조언 플랫폼은 다양한 경로를 통해 직업정보와 조언 데이터를 수집, 가공하여 데이터베이스에 저장하고 사용자 인터페이스를 통해 최적화된 답변과 조언을 제공할 수 있는 시스템이다. 이러한 시스템 구성에서는 학생들에게 필요한 직업정보 및 조언을 위한 데이터들의 수집과 수집된 데이터들로 만들어진 지식 베

44) 박홍성, 이상현, 전시홍. (2014). 『청년 채용박람회 운영실태』. 청년위원회.

이스 구축이 중요하며, 구축된 지식 베이스를 활용하여 학생들에게 질문을 받고, 조언을 제공할 수 있는 사용자 인터페이스, 사용자의 요구에 대한 정확한 정보를 쿼리 하여 답변을 제공해 줄 수 있는 알고리즘으로 분류할 수 있다. 사용자 인터페이스 최근 상담, 조언 시스템의 챗봇 인터페이스 도입이 증가하고 있으며, 다수의 연구를 통해 효과성이 연구⁴⁵⁾되었기 때문에 대표적인 대화형 인터페이스인 챗봇으로 사용자 인터페이스를 구성하고, 정보의 수집과 지식 베이스 구축, 추천 및 조언 알고리즘 영역으로 구분하여 [그림 1-16]과 같이 기본 모델을 설계하였다.



[그림 1-16] 진로 조언 플랫폼 기본 모델 구성

나) 진로 조언 플랫폼 기본 모델 기술 요소

플랫폼 설계를 위하여 직업정보 수집, 지식 베이스 구축, 진로 조언, 챗봇 단계에 대한 기능별 기술 요소는 [표 1-6]과 같다. 먼저 공공기관의 오픈 데이터를 분석하여 조언에 필요한 데이터를 API를 활용하여 수집하고, 학생들에게 필요한 직업정보를 제공해 주기 위하여 각 직업의 채용정보를 검색·수집하여 데이터 전처리 과정을 거쳐 ‘기술 요소’, ‘성향’, ‘우대사항’, ‘학과’와 같은 구인공고 내의 역량 정보를 분류한다.

수집·분류된 정보를 진로 및 직업 상담에 전문적인 분야의 지식 집합체인 지식 베이스를 구축한다. 지식 베이스 구축에서 문제가 되는 부분은 직업정보에 해당하는 다양한 역량에는 서로 연관성이 있으며, 추천과 같은 정확한 정보 제공을 위해서는 이러한 연관 관계 분석이 선행되어야 한다. 지식 베이스에 가장 많이 활용되는 관계형 데이터베이스 관리시스템의 경우 이러한 관계 분석을 위해서는

45) 문용식. (2018). 『인공지능 기반의 챗봇 서비스의 국내외 동향분석 및 발전 전망』. 대구: 한국정보화진흥원.

테이블마다 많은 조인 작업이 복잡성과 함께 성능의 문제를 일으킴으로써 최근 이러한 관계 모델의 빅데이터 분석에 많이 활용되고 있는 그래프 모델을 활용하여 그래프 데이터베이스 기반의 지식 베이스(지식 그래프, Knowledge Graph)를 설계·구축한다.

그래프 데이터베이스 기반의 지식 베이스를 활용하여 학생들에게 직업정보와 조언을 제공하기 위해 현대 학생들에게 가장 친숙한 인터페이스인 대화형 인터페이스인 챗봇 시스템과 학생의 역량 데이터 저장과 역량 기반 매칭 테스트를 위한 모바일 e-포트폴리오 앱 인터페이스를 설계·구현한다.

[표 1-6] 진로 조언 플랫폼 기능별 기술 요소

구현 항목	기술 요소	설명
직업정보 수집	<ul style="list-style-type: none"> · Restful API · HTML Crawling 	<ul style="list-style-type: none"> · 고용노동부 직업정보 수집(API) · 채용정보포털에서 기업들의 채용정보 수집 · 채용정보에서의 직업역량 추출
지식 베이스 구축	<ul style="list-style-type: none"> · 그래프 데이터베이스 · 관계형 데이터베이스 	<ul style="list-style-type: none"> · 수집된 직업정보의 저장 · 직업 역량의 연관 관계 구성 설계 · 설계된 관계 항목 저장 · 저장된 직업, 역량 관계 분석
진로 조언	<ul style="list-style-type: none"> · 추천 알고리즘 · 질의 분석 알고리즘 · 정보 매칭 알고리즘 · 답변 생성 알고리즘 	<ul style="list-style-type: none"> · 맞춤형 데이터 기반의 직업 추천 · 학생의 자연어 질의에 대한 질의 분석·정보 매칭 · 매칭된 정보를 활용한 답변 생성
상담 및 조언	<ul style="list-style-type: none"> · 챗봇 인터페이스 · 자연어처리 	<ul style="list-style-type: none"> · 챗봇 인터페이스 구성 · 조언을 위한 대화 흐름 구성 · 자연어처리(질문 의도/실체 분석)
e-포트폴리오	<ul style="list-style-type: none"> · 모바일 앱 · API 연동 	<ul style="list-style-type: none"> · 개인 역량 정보 · 역량 기반 맞춤형 직업 추천 · 내부 챗봇 인터페이스 · 역량 개발 정보

다) 진로 조언 기능

대학생을 위한 직업 및 진로 조언은 정보화, 세계화 시대에서 전문직업인으로 사회적인 역할을 담당하고 자아실현을 성취하도록 도와주는 주요한 과업 중의

하나일 뿐만 아니라,⁴⁶⁾ 최근 대학의 복잡해진 학부·트랙 제도 등으로 인해 그 필요성이 크게 대두되고 있다.

경제 성장률의 저하와 청년 실업의 증가, 기업의 경력자 우대와 같은 채용패턴의 변화에 따라 대학 역시 학생들의 성공적인 취업을 위한 직업 역량을 강화하기 위해 노력하고 있다. 대학에서의 진로교육은 학생들의 성격, 흥미, 가치관, 능력, 지식, 환경, 교육, 자격, 훈련 등을 고려하여, 해당 전공 분야에 알맞은 각종 취업 정보의 제공과 멘토링 등의 지원을 통해 자신의 진로를 현명하게 선택하고, 진로를 개발을 위한 기틀을 마련해 주어야 한다. 이에 따라, 실제 진로지도의 목표인 자신에 대한 이해와 직업에 대한 각종 정보 제공을 통한 직업의 선택, 그리고 선택한 직업에 임하는 준비과정을 조언하기 위해서는 학생의 정보와 지식 베이스 정보와의 매칭을 통해 최적화된 진로정보를 제공해야 하며, 자연어를 활용한 학생의 진로 문제 상담에 명확한 질의 내용을 분석하여 지식 베이스에서 정보를 찾고, 이렇게 찾은 정보를 활용하여 학생에게 조언할 수 있어야 한다.

라) 챗봇 기능

인공지능 기술과 빅데이터 기술의 발전으로 상담, 각종 지원업무, 정보 검색 등과 같은 정보를 찾고, 상담을 받는 업무들에서의 챗봇의 활용이 크게 증가하고 있다. 챗봇은 밀레니엄 세대와 Z세대 대학생들에게 매우 친숙한 인터페이스이며, 이용을 위한 별도의 학습 과정을 거치지 않기에 빠르고 편리하게 이용할 수 있는 장점과 대화로 모든 정보가 전달되는 형태로 다양한 상담 및 조언 분야에서 연구·활용되어 활용성이 검증된 챗봇을 활용하여 학생과의 상담·조언 인터페이스를 구현한다.

4) 연구의 방법 및 범위

본 연구는 정보기술을 활용하여 대학생들의 가장 큰 고민인 진로 문제 해결을 위한 연구로 공공, 민간의 진로·직업 데이터를 수집하여 그래프 데이터베이스 기반의 지식 베이스인 지식 그래프를 설계·구축하고 대화형 인터페이스인 챗봇과

46) 이대용, 류동희, 김인환, 고태용. (2011). 한국대학생의 진로의식에 관한 연구. 『취업진로연구』, 1(2), 41-68.

맞춤형 데이터와 진로 개발을 위한 e-포트폴리오 시스템을 활용하여 대학생들에게 맞춤형 진로정보와 조언 제공을 목적으로 한다. 대학생들의 가장 큰 고민인 진로 문제를 정보기술을 활용해 해결해 나가기 위해서는 정형·비정형 데이터를 분석하여 진로 조언을 위한 데이터의 수집과 수집된 데이터와 직업(직무)과의 역량의 관계를 분석하고, 이를 가공하여 학생들이 필요한 진로정보를 생성하여야 한다.

학생들의 개별 데이터의 경우 「개인정보 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」 등의 관계 법률에 따라 실제 학생 데이터의 활용을 통한 연구는 한계가 있기 때문에, 모바일 e-포트폴리오 앱을 개발하여 학생들의 보유 역량에 따른 맞춤형 직업 추천과 진로를 개발하기 위한 활용 모델 제시와 밀레니엄 세대와 Z세대 학생들에게 친숙한 인터페이스인 챗봇을 활용하여 자연어 질문에서의 의도와 실체를 판별하여 지식 그래프와의 매칭으로 진로 관련 정보를 제공할 수 있는 플랫폼 구축을 목적으로 한다.

정보기술의 급격한 발전과 함께 다양한 산업, 환경, 공공 등과 융합이 가속화되고 있다. 빠르게 변화하는 경제·기술 환경 속에서 학생들이 검색엔진을 통해 수집할 수 있는 진로 관련 정보는 한정되어 있으며, 오래된 정보가 대부분이다.

이러한 진로 관련 문제 해결을 위해 기업의 구인 데이터에서 수집·분석된 역량 정보는 학생들의 진로 선택과 진로 개발을 위한 중요도가 높은 정보이며, 관계 분석을 통해 역량과 직업 간의 관계를 분석하여 학생들에게 조언함으로써 학생들의 최선의 선택을 지원하기 위한 플랫폼의 설계와 구현, 테스트를 통해 대학생들의 진로·취업 문제 해결에 기여하고자 한다.

제 2 장 이론적 배경 및 선행연구

제 1 절 대학생 진로 문제와 학습 동기에 관한 연구

1) 진로 문제 현황

전통적인 직업들과 달리 현대의 직업들은 매우 다양하게 세분되어있고 이렇게 세분된 직업들은 다시 기술, 지식 등의 수많은 관계와 밀접하게 엮여있기 때문에 대부분 대학생은 자신에게 맞는 진로 관련 정보를 찾거나 해당 직무에 종사하는 전문가들에게 조언을 구하는 것은 매우 힘든 것이 현실이다.

현 대학들이 도입하고 있는 진로·취업 지원센터 역시 정형화된 심리검사나 진로설계를 위한 프로그램의 수행 이외에 실제로 대학생들이 원하는 전공 관련 전문 상담과 학생 맞춤형 진로지원은 어렵다. 이러한 문제점으로 인하여 많은 대학생은 진로에 대한 충분한 고민과 정보가 부족한 상태에서 대학을 졸업하고 취업 시장에 내몰리게 되면서 공시생의 증가, 졸업 연기, 조기 퇴사와 같은 사회 문제의 원인이 되기도 한다.⁴⁷⁾ 기업의 인사담당자들을 대상으로 한 조사에서, 지난 1년간 퇴사자 중 약 49% 절반가량이 입사한 지 1년 미만인 신입 사원인 것으로 드러난 것과 같이 진로의 문제와 고민은 대학 재학 기간에서뿐만 아니라 졸업 이후로도 생애 전 주기에 걸쳐 계속 이어지고 있다.⁴⁸⁾

2) 효율적인 진로 정보 검색 문제

대학생은 초·중등학생과 달리 진로에 대해 스스로 정보를 찾고, 분석하며, 결정이 가능한 나이와 지식을 갖추었다고 볼 수 있다. 이러한 이유로 초·중등 교육과 같은 진로지도라는 개념이 아닌 학생이 진로를 선정하고, 선정될 진로

47) 김지경, 이상호, 이성용. (2016). 『대학생 졸업유예 실태 및 지원 방안 연구』. 한국청소년정책연구원 연구보고서.

48) 손정은, 박재한, 김성혜, 장원섭. (2019). 초기 경력자의 자발적 이직 경험에 관한 현상학적 연구 : 학습과 성장을 중심으로. 『교육문화연구』, 25(5(B)).

를 개발할 수 있도록 맞춤형 정보 제공과 조언이 가장 효율적인 방법일 것이다.⁴⁹⁾ 하지만 기술의 발전과 직업 환경의 변화에 따라 취업을 위해 갖추어야 할 직업정보 역시 새로운 지식과 같은 역량들이 빠르게 변화하고 있으며, 검색된 정보에서 이를 반영한 유용한 정보를 찾기는 매우 힘들다. 또한, 현대의 밀레니엄 세대와 Z세대 학생들의 경우 특히 대면상담에 부담감⁵⁰⁾을 가지는 경우가 많으므로 이에 대한 보완도 검토해야 한다. 이러한 문제 해결 방안으로 국내외 대학들의 학생들에게 친숙한 챗봇의 도입과 활용 사례가 증가하고 있으며, 기업에서 인력 채용 시 필요로 하는 역량 정보와 요구사항들을 대학생들에게 조언함으로써 역량 개발을 통한 사회 진출을 준비하기 위한 대학 내부의 노력과 정보의 수집·가공·분석·활용의 중요성이 커지고 있다.

3) 진로에 따른 학생의 학습 동기와 학업 만족도

대부분 학생의 학업을 위한 동기는 취업·진학 등을 통해 직업을 갖기 위해서이다. 따라서 학생의 수강 과목에 대한 주요 목표 이해를 통해 학습 과정 전반에 걸쳐 학생의 만족도를 향상하게 시키는 것이 중요하며, 선택한 과정이 미래의 직업과 관련이 있을 때 학생들은 더 큰 만족을 느끼게 된다.⁵¹⁾ 학생의 만족도는 학습 활동뿐만 아니라 학습이 원하는 목표를 달성하는 데 어떻게 기여하는지에 대한 포괄적인 이해를 통해 달성된다.

학생이 자신의 진로에 대한 업무 역량과 환경의 이해에 따라 명확한 목적을 가지고 학업을 진행하기 위해 증거기반의 접근 방식이 필요하다.⁵²⁾ 이러한 증거기반의 접근은 실제 회사에서 필요한 업무 역량과 내가 학습하고 있는 업무 역량의 일치성을 분석하여 보여주는 것이 가장 효과적일 것이다. 진

49) 이대용, 류동희, 김인환, 고태용. (2011). 한국대학생의 진로의식에 관한 연구. 『취업진로 연구』, 1(2), 41-68.

50) 오지은. (2011), 『성공적인 상담을 위한 커뮤니케이션의 첫 걸음』. 서울대학교 교수학습개발센터

51) Shury, J., & IFF Research Ltd. (2017). Planning for Success: Graduates' Career Planning and Its Effect on Graduate Outcomes. Department for Education.

52) McIlveen, P., Brooks, S., Lichtenberg, A., Smith, M., Torjul, P., & Tyler, J. (2011). Career development learning frameworks for work-integrated learning. In Developing Learning Professionals. Springer.

출 가능한 직업에 대한 이해와 학습 과목을 통해 얻은 기술이 진로에 필요한 기술과 어떻게 일치하는지에 대한 인지를 통해 학생들이 선택한 학습 과정의 개인 및 직업적 이점을 예측할 수 있다면 학습 동기와 성과를 높일 수 있다.⁵³⁾

취업 정보 포털의 온라인 구인 광고는 이러한 데이터가 구인 및 수요와 밀접한 관련이 있으므로 인기 있는 구직 모델이 되었다. 학생, 졸업생, 고용주, 연구원, 대학 및 커리큘럼 개발자는 온라인 구직 데이터를 오늘날 구직 시장의 역동성을 탐구하기 위한 실용적인 데이터로 보고 있으며, 온라인 구인 광고는 특정 직무에 대한 다양한 유형의 기술에 대한 상대적 수요를 보여줄 수 있다. 실시간 취업 시장 지표는 의료 및 정보기술과 같은 신규·급변하는 산업의 인력 요구와 교육 커리큘럼을 조정하는 데 특히 유용할 수 있다.⁵⁴⁾



53) Crosling, G., Heagney, M., & Thomas, L. (2009). Improving student retention in higher education: Improving teaching and learning. Australian Universities' Review, The, 51(2), 9.

54) Carnevale, A. P., Jayasundera, T., & Repnikov, D. (2014). Understanding online job ads data: a technical report. Georgetown University, McCourt School on Public Policy, Center on Education and the Workforce.

제 2 절 챗봇을 활용한 상담 및 조언 서비스 제공 사례

1) 법률 및 심리 상담에서의 챗봇 활용 사례

대화형 인터페이스를 활용한 진로 조언 기술은 고객 상담 및 지원 업무와 유사하다. 기존의 정보시스템에서 구현과 제공이 힘들었던 분야 중 하나는 상담업무로 컴퓨터가 사람의 자연어를 알아듣고, 정확히 이해하여 최적의 정보를 찾아서 문의한 사람에게 답변하는 시스템은 오래전부터 연구되어 왔지만, 실제 정보를 찾는 데 도움을 주는 전문가 시스템에서 크게 발전하지 못하고 있었다. 그러나 최근 컴퓨팅 파워의 증가, 빅데이터의 증가로 인하여 머신러닝 알고리즘이 빠르게 발전하게 되면서 자연어처리 기술도 크게 발전하였다. 이러한 기술의 발전에 따라 인공지능을 활용한 다양한 형태의 인공지능을 활용한 상담업무가 도입되고 있다. 챗봇은 고객지원 상담 이외에도 우울증 치료 챗봇(Woebot), 법률 상담(DoNotPay)과 같은 전문적인 상담의 영역에까지 확장되어가고 있으며 진로 조언과 유사한 구인·구직 서비스에서도 인공지능 챗봇을 활용한 서비스가 성과를 보이고 있다.

미국인들이 겪는 질병에서 우울증은 상위권을 차지하고 있고, 미국 대학생의 50%는 불안 또는 우울증을 겪고 있다.⁵⁵⁾ 오늘날 정신건강 분야의 의료체계는 비용이 매우 높을 뿐 아니라, 정신과를 다녀온 기록이 남는다는 점도 부담되어 청년들에게 적합하지 않다. 이러한 문제를 해결하기 위해 Alison Darcy 팀은 대학생 자원자들에게 몇 가지 형태의 치료를 시도해 본 결과, 챗봇이 가장 효율적이라는 사실을 발견했다.⁵⁶⁾ 2017년 의학 저널에 발표한 이들의 논문에 따르면 워봇이 2주간 학생들의 우울증 증상을 효과적으로 낮췄다. 이런 효과가 있는 이유는 대화야말로 자연스럽게 감정적 지지를 얻는 방법이며 사람들이 자신의 문제를 이야기하고 대화를 통해 치료해 나가기 때문이다.⁵⁷⁾

55) Deroma, V. M., Leach, J. B., & Leverett, J. P. (2009). The relationship between depression and college academic performance. *College Student Journal*, 43(2), 325-335.

56) Skjuve, M., & Brandtzæg, P. B. (2018). Chatbots as a new user interface for providing health information to young people. *Youth and news in a digital media environment - Nordic-Baltic perspectives*.

57) 테크M. (2018). AI 거물 앤드류 응이 개발한 우울증 완화 챗봇. 『MIT테크놀로지리뷰』.

MIT 심리치료 로봇 ‘올리’와 같이 인지적 영역에 대한 효과성 검증과 함께 학습자의 흥미, 태도, 감정 등 정신적 영역의 효과성을 검증하기 위한 연구도 다수 수행되었다. 이러한 연구에서 봇(Bot) 활용 교육이 수업 흥미도, 만족도, 학습몰입도 등 학습자의 학습 태도 향상에 긍정적인 효과가 있음이 보고되었다.⁵⁸⁾ 연구를 통해서도 사람의 감정을 이해하기도 하고, 감정의 표현을 통해 봇에게 감정이입과 공감(Artificial Empathy)을 느끼게 된다. 챗봇은 인공 친구의 개념으로 접근하여 감정적 공감대를 형성할 수 있으며, 이를 통해 친숙한 대화가 가능하다.

이처럼, 상담·조언·고객지원 등 다양한 커뮤니케이션이 필요한 업무에 인공지능 기술이 연구·활용되고 있다.⁵⁹⁾

2) 직업 추천 및 교육에서의 챗봇 활용 사례

미국의 유명 구직사이트인 퍼스트잡(First job)은 인공지능 기반의 챗봇 Mya를 도입하여 채용과정의 75%를 자동화하여 채용 효율을 무려 38%나 증대시켰다. 미국 구직자의 약 85%는 지원 후 결과 통보를 듣지 못한 채 무엇이 문제인지 모른 상태에서 다시 지원서를 넣고 있으며, 채용기업의 약 71%는 적합한 인력을 찾지 못하고 있다고 한다. Mya는 구직자와의 채팅으로 기업 문화, 연봉, 정책 등 구직자가 원하는 조건에 맞는 기업을 추천해 준다. 게다가 구직자가 특정 회사에 지원했다가 탈락하면 Mya는 어떤 부분이 부족했는지 설명해주고 추가 질문을 통하여 다시 조건과 자격에 맞는 기업들을 추천해 준다.⁶⁰⁾

Mya 덕분에 구직자와 기업 인사담당자 모두 지원에 드는 비용을 줄이고 면접에 더욱 집중하여 최적의 인재를 뽑을 수 있다. 퍼스트 잡은 Mya를 통하여 구직 참여율이 150%나 늘어났고 통보 수신율은 3배나 확대되는 등 고용 창출에 크게 이바지하였으며, 덕분에 구직자와 기업 인사담당자 모두 지원에 드는 비용을 줄이고 면접에 더욱 집중하여 최적의 인재를 선발할 수 있게 되었다.⁶¹⁾

58) 서순식. (2016). 『로봇 활용 SW교육 효과성 검증 연구』. 대구: 한국교육학술정보원.

59) 박동아. (2017). 인공지능 기반 대화형 공공 행정 챗봇 서비스에 관한 연구. 『멀티미디어학회 논문지』, 20(8), 1347-1356.

60) Frey, C. B., M. A. Osborne. (2017). The future of employment: how susceptible are jobs to computerisation?. 『Technological Forecasting and Social Change』, 114, 254-280.

온라인 학습 상담에서의 챗봇 사례로 2012년 개설한 "인공지능 소개" 과정에
서 시작된 Udacity 플랫폼은 다양한 컴퓨터 과학 분야에 수십 개의 강좌와 나노
학위 프로그램을 제공하기 위해 성장하였다. 하지만, 이러한 코스의 확장으로 인
하여 학생들에게 방대한 카탈로그를 탐색해야 자신의 기술과 경험에 맞는 것을
찾을 수 있게 되었으며, 이러한 문제를 해결하기 위해서는 수천 명의 사용자에게
개인화된 상담을 제공해야 하는 것과 같이 회사의 용량 및 예산을 초과하는 인적
자원이 필요하게 된다. Udacity는 이 문제를 해결하기 위해 AI 챗봇을 도입하였으
며, AI 챗봇은 까다로운 명령 형식에서 벗어나 편안하고 대화 방식으로 사용자와
상호 작용하여 학생들은 수강생들이 수동으로 코스를 찾지 않고 교수나 상담사와
이야기할 때와 비슷한 경험을 할 수 있게 되었다. AI 챗봇은 Udacity에서 학생들
에게 올바른 코스나 나노 학위를 찾도록 유도하거나 강좌 비용이 얼마나 들 것인
가, 혹은 그것을 완료하는 데 시간이 얼마나 걸릴 것인가와 같은 지속적인 질문에
대한 답변과 목표를 달성하기 위한 올바른 행동 방침을 조언함으로써 진로 개발
과 상담을 돕고 있다.⁶²⁾

61) KT경제연구소. (2017). 『한국형 4차 산업혁명의 미래: KT경제경영연구소가 찾아낸 미래 한국
의 7가지 성장전략』. 서울: 한스미디어.

62) Iftene, A., & Vanderdonckt, J. (2016). Moocbuddy: a chatbot for personalized learning
with moocs. In RoCHI-International Conference on Human-Computer Interaction.

제 3 절 진로 조언을 위한 직업정보 데이터 연구

1) 진로 조언 데이터의 유형

학생 맞춤형 조언 서비스 제공을 위해서는 직업과 밀접하게 연관된 능력, 자격증, 성격 및 태도, 적성, 전공과 같은 다양한 역량들을 분석하여 지식 그래프를 구축하고, 대학생들의 개인화된 데이터를 구축된 지식 그래프에 매칭하여 최적화된 정보를 찾을 수 있어야 한다.

이러한 직무와 연관 관계 데이터 수집을 위해서는 공공기관들의 오픈 데이터뿐만 아니라 대학, 민간 직업정보제공기업들이 보유하고 있거나 공개하고 있는 다양한 데이터들을 분석하여 직업에 밀접하게 연관된 다양한 요소들을 분석하고 이를 활용하여 지식 그래프를 구축하여야 한다. 다양한 기술의 발전과 더불어 빠르게 변화하고 있는 복잡화된 현대 사회 속에서 신규 직업들의 등장과 직업들과의 연관 요소들 역시 매우 다양하고 서로 간에 밀접하게 연결되고 있으며, 이에 대한 관계 분석의 중요성이 커지고 있다.

진로 조언 플랫폼을 이용하는 대학생의 역량 데이터를 지식 그래프에 매칭하여 학생 개개인에게 최적화된 직무를 찾고, 직무를 위해 부족한 부분과 개발 방법에 대한 조언을 위해서는 진로 조언에 필요한 최적화된 데이터가 필요하다.

2) 공공데이터의 정의 및 필요성

공공데이터는 공공부문에서 업무를 통해 만들어낸 수많은 데이터를 기업체와 시민이 이용, 접근, 재사용할 수 있게 함으로써 정부는 신뢰성과 투명성을 향상하고 가치를 창출하며 국민 참여 및 국민의 의사결정에 더 많은 도움을 주기 위한 데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하는 광 또는 전자적 방식으로 처리된 자료 또는 정보를 말한다.⁶³⁾

63) Global ICT R&D Policy Trends, (2018), 『4차 산업혁명 시대에 대응한 주요국의 공공데이터 개방 현황 분석』, 대전: 정보통신기술진흥센터.

제4차 산업혁명은 데이터의 시대이다. 지능정보기술을 바탕으로 지능화·자동화·자율화를 추구하는 4차 산업혁명에서 데이터는 필수요소이자 핵심 자원이다. 대규모 데이터를 보유하고 활용하는 기업이 시장 혁신을 주도하고 있고, 국가사회의 전 영역에서 급속하게 부상하는 인공지능과 로봇도 데이터를 기반으로 한다.⁶⁴⁾ 이러한 추세에 따라 미국을 비롯한 서구 선진국들은 미래 경쟁력을 좌우하는 데이터의 중요성을 인식하고 데이터 산업 활성화를 위한 전략 수립과 투자확대 등 데이터 패권경쟁에 본격 돌입하고 있으며,⁶⁵⁾ 대한민국 역시, 지능형 정부를 구현과 데이터 경제 육성, 데이터 활용 사회적 가치 실현을 위하여 법률제정, 제도 개선 등을 본격적으로 추진하고 있다.⁶⁶⁾

3) 민간 직업정보 데이터의 수집

공공부문에서 공개하고 있는 공공데이터와 달리 민간데이터는 일반적으로는 민간기업이 보유하고 있는 데이터를 공개하는 것을 말하지만, 민간 기관에서 수집할 수 있는 데이터를 통칭하여 민간데이터라 부를 수 있다.

학생들에게 실제 기업에서 필요로 하는 다양한 직무 관련 역량의 채용정보 데이터를 수집하기 위해 고용정보원 워크넷 채용정보과 민간 채용정보 포털에서 '웹프로그래머' 키워드로 검색한 결과를 비교했을 때 민간 채용정보 포털에서 검색된 구인공고가 10~20배 더 많았다. 대부분의 민간 취업 정보사이트들도 API를 제공하지만, 실제 기업의 구인 정보에 해당하는 사항은 별도의 링크에 HTML 형태로 존재하며, 표준화가 되어있지 않다. 따라서 직업/직무별 민간 구인 정보에서 기업이 요구하는 역량을 분석하기 위해서는 각 구인공고에서 직무 관련 역량 정보에 해당하는 항목을 분리하여 별도의 자연어 처리 또는 별도의 매핑 사전을 활용하여 원하는 직업별 역량 정보를 추출해야 한다.

64) 윤상오, 현지우. (2019). 공공데이터 개방정책의 실태분석 및 개선방안에 관한 연구: 공공데이터 포털의 국가중점데이터 개방 사례를 중심으로. 『한국공공관리학보』, 33(1), 219-247.

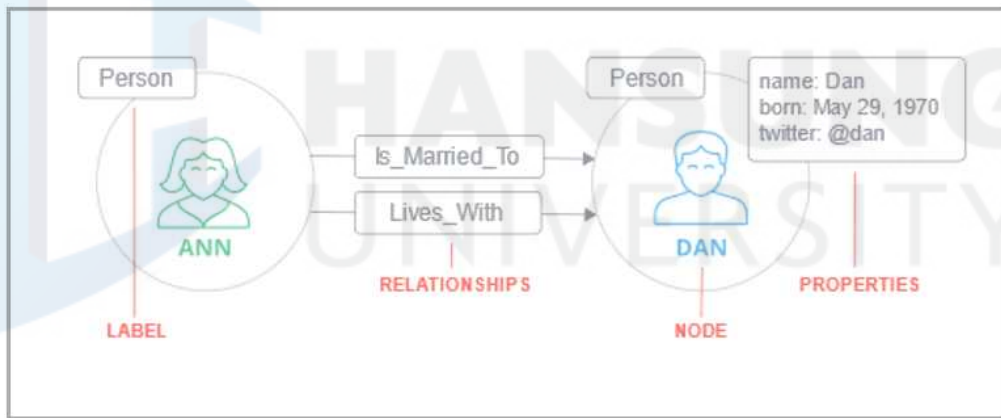
65) 관계부처 합동. (2018). 『데이터 산업 활성화 전략』.

66) 관계부처 합동. (2019). 『2019년도 공공데이터 제공 및 이용 활성화 시행계획』.

제 4 절 그래프 데이터베이스(Graph Database)

1) 그래프 데이터베이스 개요

그래프 데이터베이스(Graph Database)는 빅데이터 시대에 접어들며, 정형 데이터⁶⁷⁾ 위주의 처리만 가능한 관계형 데이터베이스(Relational Database)를 보완하기 위해 등장한 NoSQL 계열의 데이터베이스로 그래프 데이터베이스는 [그림 2-1]⁶⁸⁾과 같이 사람, 기업, 계정과 같이 개체에 해당하는 (Node), 객체 간의 관계를 표시해 주는 관계(Edge), 노드와 관계에서 데이터의 추가 세부 정보를 저장하는 속성(Property), 성질이 비슷한 개체들을 묶어주는 묶음(Label)으로 구성되어 있다.



[그림 2-1] 그래프 데이터베이스의 구조

노드(node)는 그래프의 엔티티이며, 속성(key-value pair)을 원하는 수만큼 보유할 수 있다. 노드에는 레이블로 태그를 지정할 수 있으며 도메인에서 각기 다른 역할을 나타낸다. 노드 레이블은 특정 노드에 메타 데이터를 첨부하는 역할을 할 수도 있다.

관계는 두 개의 노드 간에 명명된 의미론적 관련 연결을 나타낸다. 관계에

67) 정해진 특정 조건에 맞게 저장된 데이터

68) Neo4j, <https://neo4j.com/>

는 항상 방향, 유형, 시작 노드 및 종료 노드가 있으며, 노드와 마찬가지로 속성 정보를 가질 수 있다. 대부분의 관계에는 가중치, 비용, 거리, 등급, 시간 간격 또는 강점과 같은 정량적 속성이 있다. 그래프 데이터베이스에서는 이러한 관계가 효율적으로 저장되므로 두 개의 노드가 성능 저하 없이 모든 유형의 관계를 공유할 수 있으며, 관계는 특정 방향으로 저장되지만, 관계는 항상 어느 방향으로나 효율적으로 탐색할 수 있다.

2) 그래프 데이터베이스 특징

가) 현실 세계의 관계를 반영한 데이터베이스

그래프 이론은 18세기에 오일러에 의해 개척되었고, 그 이후로 수학, 사회학, 인류학 등의 다양한 영역에서 활발하게 연구되고 개선되어왔다. 2000년대 이후 그래프 이론과 그래프 사고가 정보 관리에서 연구되고, 적용되면서 그래프 데이터베이스는 소셜 네트워킹, 생물 정보학, 지리 공간 등의 영역에서 중요한 문제를 해결하는 데 도움이 되었다. 페이스북, 구글, 트위터와 같은 그래프 기술을 기반으로 하는 회사의 대규모 상업적 성공과 데이터의 폭발적 증가로 인한 관계 분석의 중요성이 커짐에 따라 그래프 데이터베이스의 연구와 활용이 증가하고 있다.

관계형 데이터베이스에서 링크 정보나 연결 정보를 찾으려면 한 테이블에 있는 값을 이용해 다른 테이블에 있는 값을 찾는 조인 작업을 수행하여야 한다. 이러한 조인 작업은 데이터베이스 크기와 테이블 수가 많아질수록 시간과 자원이 많이 소모되지만, 그래프 데이터베이스에서는 조인을 수행하는 대신 노드에서 노드까지 관계로 연결하여 표현하기 때문에 간단한 구조와 빠른 작업 처리가 가능하다.⁶⁹⁾

나) 복잡한 관계 질의에 대한 성능 문제 해결

우리는 연결된 세상에 살며, 주위에 풍부하게 연결된 도메인이 존재하고 있으며 기본적으로 관계를 포용하는 데이터베이스만이 연결을 효율적으로 저

69) Sullivan, D. (2015). NoSQL for mere mortals. Addison-Wesley Professional.

장, 처리 및 쿼리할 수 있다. 관계형 데이터베이스는 비효율적인 JOIN 작업을 통해 쿼리시 관계를 계산하지만, 그래프 데이터베이스는 모델의 데이터와 함께 관계를 저장한다.

복잡한 질의는 원하는 데이터를 얻기 위해 여러 가지 복잡한 조인 형태의 연산으로 구성되어 있다. 이러한 연산은 모든 관계형 데이터베이스 관리자가 알고 있듯이, 사용되는 테이블 간에 가능한 모든 결합의 수를 모두 구해야 하기 때문에 높은 시스템 자원의 소비가 필요한 작업이다. 관계형 데이터베이스 관리시스템에서 2~3개의 테이블 사이에 1개 또는 2개의 조인을 수행하는 것은 괜찮을 수 있지만 사용할 테이블의 개수가 늘어날수록 조인 가능한 경우의 수가 기하급수적으로 늘어난다는 것이 문제가 되는 이유다. 이처럼 관계형 시스템에서 복잡한 질의는 쉽게 해결할 수 없는 문제이다.⁷⁰⁾

그러나 그래프 데이터베이스에서는 조인 작업을 수행할 필요가 없다. 그래프 데이터베이스에서는 일반적으로 인덱스 탐색을 이용해 시작 노드를 찾은 후 색인 없는 인접성 특성을 사용해 연결 관계((노드)-[:관계]->(노드))를 통해 한 노드에서 다음 노드로 이동한다. 이 경로를 따라 이동하는 모든 홑은 사실상 조인 작업과 동일하다. 이러한 유형의 쿼리를 패턴 매칭 쿼리(Pattern matching query)라고 부른다. 패턴을 지정한 후, 이 패턴을 하나 또는 그 이상의 시작점에 고정하고 이 패턴이 일치하는 항목을 찾는다. 그래프 데이터베이스는 시작 노드 주위를 회전하면서 연결된 패턴이 일치하는지를 판단하는 도구이기 때문에, 일치하지 않는 패턴은 무시되며, 시작 노드에 연결되지 않은 패턴은 고려하지 않는다. 대부분 그래프에서 모든 노드가 다른 모든 노드와 연결된 것은 아니므로 쿼리 성능은 데이터 집합의 크기와 독립적이며, 다른 데이터베이스들과 달리 노드의 데이터 집합 크기가 증가하더라도 성능상의 저하가 크게 발생하지 않는다.

다) 4차 산업혁명에서의 그래프 데이터베이스의 활용

빅데이터는 요소로 크기, 속도, 다양성, 복잡성 이란 특성이 있지만, 이러한 빅데이터 분석에서 복잡성을 간과하는 경우가 의외로 많다. 비정형데이터

70) Baton, J., & Van Bruggen, R. (2017). Learning Neo4j 3. x: Effective data modeling, performance tuning and data visualization techniques in Neo4j. Packt Publishing Ltd.

의 폭발적인 증가와 복잡성으로 인한 데이터 간의 관계에 대한 분석의 필요성이 증가하고 있지만, 기존 관계형 데이터베이스의 한계로 인하여 이러한 관계 분석에 최적화된 새로운 데이터베이스의 필요성 또한 커지게 되었다. 가트너는 2019년 10대 데이터 및 분석 기술 트렌드⁷¹⁾에 그래프 분석을 선정하면서, “그래프 분석은 조직, 사람 및 거래와 같은 관심 있는 엔티티 간의 관계를 탐색할 수 있는 일련의 분석 기술로 그래프 데이터 저장소에서는 데이터 사일로⁷²⁾를 가로지르는 복잡한 상호관계를 통해 데이터를 효율적으로 모델링, 탐색 및 쿼리할 수 있다. 이러한 그래프 데이터베이스는 전문 기술이 필요한 부문이라 현재까지 채택이 제한적이었으나, 데이터 전반에 걸쳐 복잡한 질문을 제기할 필요성이 대두됨에 따라, 빠르게 성장할 것이다.”라고 발표하였다.

4차 산업혁명은 다양한 정보들이 서로 연결되고 지능화된 초연결사회를 불러왔으며, 이로 인하여 연결된 복잡한 상호관계 분석의 중요성이 커짐에 따라 그래프 데이터베이스를 활용한 연구와 활용 사례 역시 증가하고 있다.

3) 사이퍼쿼리

사이퍼(Cypher)는 그래프 데이터에 대한 직관적인 질의를 위하여 개발된 오픈소스 언어로 그래프의 효율적인 질의 및 업데이트를 허용하는 선언형 그래프 질의어이다. Cypher는 그래프를 검색하는 방법보다는 그래프에서 검색하는 것에 초점을 맞춘 선언적 쿼리 언어로 기존 질의어로는 그래프 데이터 구조를 효율적으로 다룰 수 없으므로 Neo4j는 새로운 질의 언어인 사이퍼를 개발하였다. 사이퍼는 선언적 질의어로, (NODE1)-[:RELATION]->(NODE 2)와 같이 찾고 싶은 패턴을 선언하면 된다.

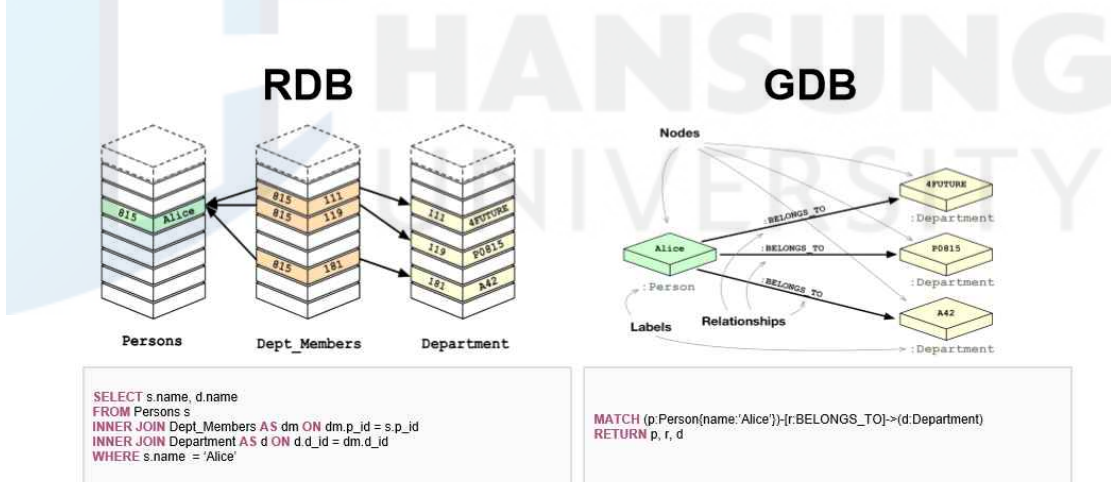
Neo4j의 Cypher 쿼리 언어는 패턴을 기반으로 한다. 이 패턴은 기본 그래프 구조를 일치시키는 데 사용되므로, 여기서는 추천 엔진 구축과 같은 추가 처리를 위해 패턴을 사용할 수 있다. 패턴 매칭(Pattern matching) 사이퍼는 패턴 일치 질의어다. 패턴 매칭을 통해 데이터 세트의 여러 개체 간 복잡

71) 가트너. <https://www.gartner.com/>

72) 단일 부서의 관할 하에 있어 나머지 구성원이 고립되는 결과를 불러오는 고정 데이터 보관소 또는 저장소

한 관계를 아주 쉽게 스케치하고 그릴 수 있다. 일반적으로 사람의 두뇌는 패턴을 이용해 쉽게 작업하려는 경향이 있으므로 패턴으로 작업하는 데 매우 익숙하다. 사이퍼는 상대적으로 단순하지만, 매우 강력한 언어로 복잡한 데이터베이스 쿼리들을 쉽게 표현할 수 있으므로 데이터베이스의 관계에 대한 더 쉬운 접근이 가능하다.

[그림 2-2]에서 관계형 데이터베이스는 외래 키 제약 조건을 사용하는 관계형 데이터베이스의 사람과 및 부서 테이블 간 JOIN 테이블로, 데이터 관계를 쿼리하기 위해서는 테이블 간에 수많은 JOIN이 포함된다. 이러한 JOIN 작업은 JOIN 할 테이블의 잠재적으로 색인화된 행의 기본 키와 외래 키를 일치시켜 쿼리시 계산되기 때문에 쿼리가 증가함에 따라 더 많은 컴퓨팅 자원을 소모하게 되지만, 그래프 데이터베이스 모델의 각 노드에는 다른 노드와의 관계를 나타내는 관계 레코드 목록을 직접 포함하고 있으므로 연결된 노드에 직접 액세스할 수 있으므로 복잡한 쿼리나 계산이 필요하지 않다.⁷³⁾



[그림 2-2] 관계형 데이터베이스와 그래프 데이터베이스의 쿼리 비교

4) 지식 그래프

일반적으로 그래프 데이터베이스 기반의 지식 베이스를 지식 그래프라 부리며, 지식 그래프는 관련 있는 정보와 지식을 서로 연결해놓은 것을 의미하

73) Hunger, M., Boyd, R., & Lyon, W. (2016). The Definitive Guide to Graph Databases for the RDBMS Developer. Neo Technology.

며, 정보만 있는 관계형 데이터베이스와 달리 실제 개체, 이벤트, 상황 또는 추상적 개념과 같은 개체 간의 상호연결된 모음을 나타낸다.

지식 베이스를 구축하는 데 있어 가장 많이 활용되고 있는 관계형 데이터베이스의 경우 관계를 분석하기 위해서는 테이블 조인으로 인해 2-depth 이상의 관계 데이터를 가져오기 위해서는 복잡한 쿼리와 많은 시간이 소요되는 구조적 한계로 인하여 관계의 분석이 중요한 소셜네트워크, 교통망, 단백질의 상호 작용과 같은 연관 관계 저장과 분석에 최근 그래프 데이터베이스 기반의 지식 그래프의 연구와 활용이 증가하고 있다.

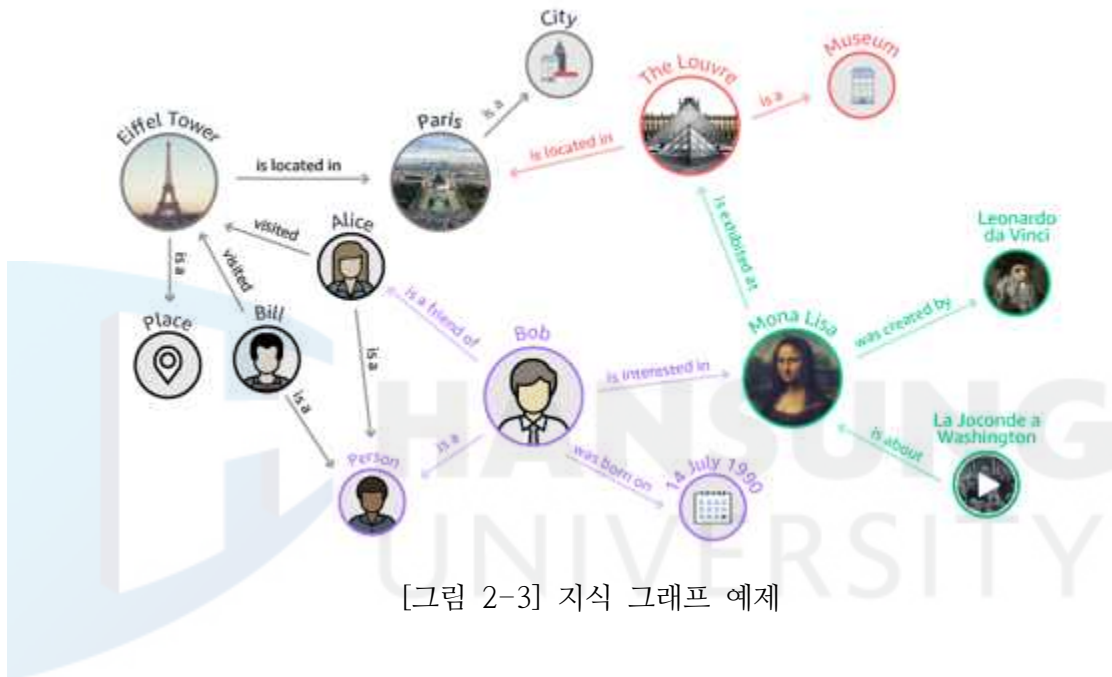
지식 그래프는 개체와 개체 간의 상호관계를 사람이 이해할 수 있는 형태로 설명하는 서로 연결된 사실 집합으로, 데이터 관계를 사용하여 인접 정보를 수집하고 통합하여 새로운 지식을 도출하게 된다.

지식 그래프로 가장 대표적인 서비스는 사용자들이 검색하는 단어를 이해하고 그 단어와 연관성이 있는 정보를 같이 보여주는 구글의 지식 그래프로 구글 지식 그래프는 5억 개 이상의 개체(인물이나 지명 등) 외에도 35억 개 이상의 사실과 개체 간의 관계를 포함하고 있으며, 동음 이의어 검색, 요약정보 제공, 관련 정보의 제공을 지원하고 있다. 구글의 지식 그래프 이외에도 지식 그래프는 외부의 공급망 위험 평가나 식별, 자연어처리, 의사결정 지원을 위해 활용되며, 특히 입력 데이터의 특징을 관계 그래프로 생성하여 분류 결과에 연결된 노드로부터 분류 근거 확인하는 설명 가능한 인공지능(eXplainable AI, XAI) 모델로 연구되고 있다.

[그림 2-3]과 같이 지식 그래프는 사용자가 모나리자에 관심이 있는 경우, 레오나르도 다빈치의 다른 작품 또는 루브르 박물관에 있는 다른 작품을 찾도록 지원할 수 있으며, 데이터베이스에 저장된 모든 데이터에 의미를 부여함으로써 사용자의 입력에 정확한 결과를 제시한다.⁷⁴⁾ 즉, 모든 데이터에 랜드마크, 유명인, 도시, 스포츠, 팀, 건물, 장소, 예술작품 등 인간이 이해하는 개념을 부여하고, 특정 데이터와 연관된 정보를 연결하여 검색의 범위를 정확하게 한정할 수 있으며, 정보를 그래프 모델에 저장하고 그래프 쿼리를 사용해 사용자가 상호연결성이 높은 데이터 세트를 손쉽게 탐색하도록 지원할 수

74) AWS. <https://aws.amazon.com/ko/neptune>

있다. 지식 그래프는 서로 다른 유형의 것들을 체계적으로 연결하여 기존의 관계형 데이터베이스와 같이 행과 열의 테이블이 아닌 노드와 링크로 구성된 네트워크에서 지식을 인코딩하고, 이를 통해 사람과 기계는 사물에 관한 사실에 대한 의미론적 네트워크가 동적으로 확장되어 데이터 통합, 지식 발견 및 심층 분석에 사용할 수 있다.⁷⁵⁾



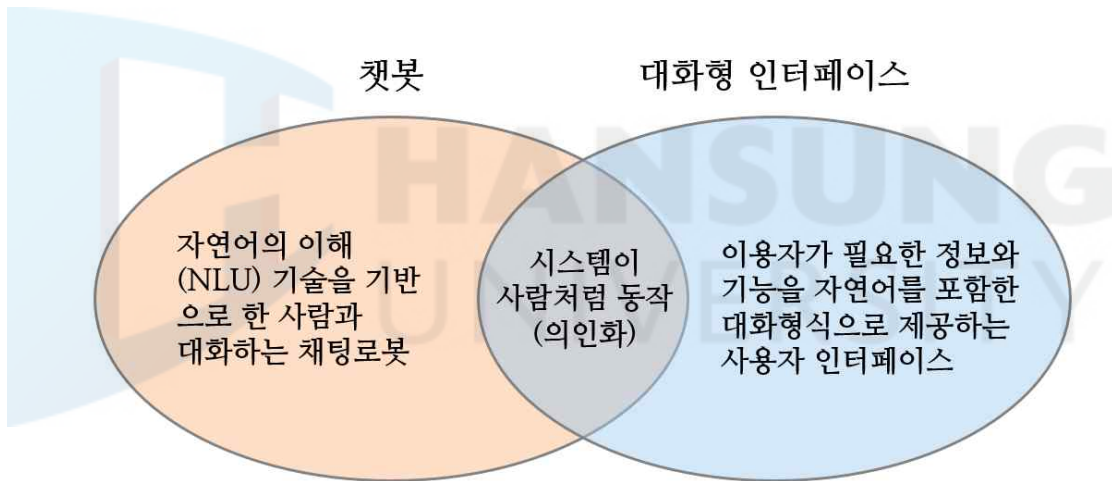
[그림 2-3] 지식 그래프 예제

75) Andreas Blumauer, (2018). Knowledge Graphs – Connecting the Dots in an Increasingly Complex World. Semantic Web Company.

제 5 절 대화형 인터페이스

1) 챗봇의 정의

사용자 인터페이스는 다양한 형태로 발전되어 왔으나, 대표적으로 명령어 기반의 CLI, 그래픽 기반의 GUI, 최근 챗봇과 결합하여 대화형 인터페이스로 발전하고 있다. 챗봇과 대화형인터페이스를 엄격하게 구분하기는 어려우나 [그림 2-4]와 같이 챗봇이 자연어를 기반으로 사람과 대화하는 채팅 로봇에 중점을 둔다면, 대화형인터페이스는 이용자가 필요한 서비스(정보, 기능 등)를 대화 형식으로 제공하는 사용자 인터페이스로 정의할 수 있다.



[그림 2-4] 챗봇과 대화형 인터페이스 구분

챗봇이란, 채팅과 로봇의 합성어로, 대표적인 대화 에이전트에 해당하며 사람과 문자 또는 음성 대화를 통해 질문에 답하거나 다양한 관련 정보를 제공하는 인공지능 기반의 커뮤니케이션 소프트웨어를 지칭한다.⁷⁶⁾

챗봇은 단순한 문제나 행정적 문제에 대한 질의응답은 물론이고 학습자가 궁금해하는 내용과 학습자원을 즉시 피드백할 수 있다는 장점이 있으며 딥러닝,

76) 강민정. (2018). 사용목적에 따라 선호하는 챗봇의 성격에 관한 연구. 『한국콘텐츠학회논문지』, 18(5), 319-329.

자연어처리 기술 등 인공지능 기반기술이 발전과 함께 가운데 사용자가 체감할 수 있는 인공지능 서비스인 '챗봇'이 상담에서부터 대화형 인터페이스까지 전 산업 분야로 확산되고 있는 중이다.

챗봇은 서버에서 인공지능을 통해 사용자의 질의를 분석한 뒤 앱 또는 웹으로 서비스를 제공하며, 챗봇은 개발 방식에 따라 질의 응답형 챗봇과 인공지능(AI) 기반의 상담 챗봇으로 구분할 수 있다.

[표 2-1]과 같이 질의 응답형 챗봇은 입력 텍스트의 문맥에 맞추어 정해진 규칙에 따라 정해진 응답을 제공하는 단방향 정보 전달 방식으로 사용자 맞춤형 정보가 아닌 일반적인 포괄적 정보를 제공하며, 정보획득 측면에서 사용자가 자신에게 적합한 정보를 위해 스스로 질문을 만들고 답변을 찾는 과정을 반복하게 된다. 인공지능 기반 방식은 최근 급격히 발전하고 있는 머신러닝을 활용하여 쌍방향 정보 교류와 맞춤형 정보 제공이 가능하며, 정보획득 측면에서 가이드에 따라 원하는 정보에 빠르게 접근할 수 있다. 질의 응답형 챗봇은 구현하기 쉽지만, 미리 정의된 문장에서 조금만 변형이 되면 응답을 제대로 하지 못하는 경우가 많으며, 인공지능 기반의 챗봇은 학습데이터 부족과 모델 생성의 어려움으로 인하여 질의 응답형 방식과 혼합하여 구현하는 하이브리드 방식이 주로 많이 쓰이고 있다.⁷⁷⁾

[표 2-1] 챗봇의 유형별 특징

구분	질의 응답형 챗봇	인공지능 상담 챗봇
대화방식	- 단방향 정보전달 : 사용자가 질문에 대한 답변만 가능하며 추가 보완은 어려움	- 쌍방향 정보 교류 : 대화 에이전트가 주도적으로 사용자와의 질문-답변 과정을 반복하여 부족한 정보를 보완
정보제공	- 포괄적 정보제공 : 일반적인 정보만 제공하기 때문에 자신에게 맞는 정보인지 스스로 판단 필요	- 맞춤형 정보 제공 : 대화를 통해 사용자의 상황정보를 인식하여 사용자에게 적합한 정보 제공
정보획득	- 사용자 스스로 질문 주도 : 사용자는 자신에게 적합한 정보를 얻기 위해 스스로 질문을 만들고 답변을 찾는 과정을 반복	- 대화를 통한 전문가의 도움 : 지식이 없는 사용자도 대화 에이전트의 가이드에 따라 원하는 정보에 빠르고 쉽게 접근

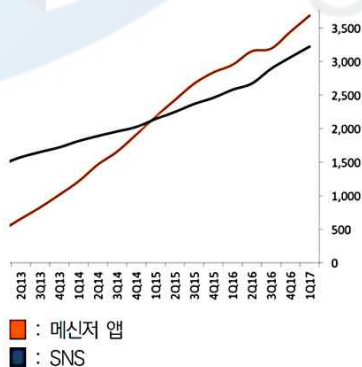
77) 문용식. (2018). 『인공지능 기반의 챗봇 서비스의 국내외 동향분석 및 발전 전망』. 대구: 한국정보화진흥원.

2) 대학생들의 챗봇 활용

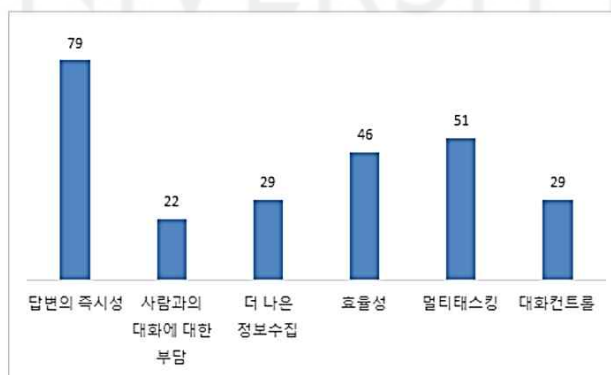
밀레니얼 세대 이후 1997년도에서 2012년대에 태어난 세대인 Z세대가 대학에 입학하면서 새로운 세대로 주목받고 있다. 이들의 가장 큰 특징은 태어나면서부터 디지털을 접한 세대라는 점이다. 스마트폰과 PC, 태블릿을 자연스럽게 접하며 자란 이 세대는 이상주의적 경향이 있는 이전 세대에 견주어 개인적이고 합리적이다. 또한, 디지털 네이티브라는 특성에 걸맞게 인터넷, 이메일, 실시간 채팅, 인터넷 게임, 랩 음악을 좋아하고 SNS를 적극적으로 하는 것이 특징이다. [그림 2-5]와 같이 젊은 세대가 채팅을 선호하는 이유가 자기 모습을 자기 마음대로 나타낼 수 있고, 시간을 소유하며, 내용을 편집할 수 있기 때문이다.⁷⁸⁾

챗봇은 공유 사회 패러다임에서 서비스들을 연결해 주는 매개체로 사용되며 인간의 생활양식에도 커다란 변화를 가져오고 있다. 즉, 정보를 찾기 위해 직접 검색하거나 통화하기보다는 메신저 기반 챗봇으로 원하는 정보를 얻을 수 있게 됨에 따라 정보획득의 패러다임이 변화하고 있다.

〈 메신저앱과 SNS 사용량 변화 추이 〉



〈 라이브 채팅 앱을 선호하는 이유 〉



[그림 2-5] 메신저 앱의 사용량 변화와 채팅앱의 선호 이유

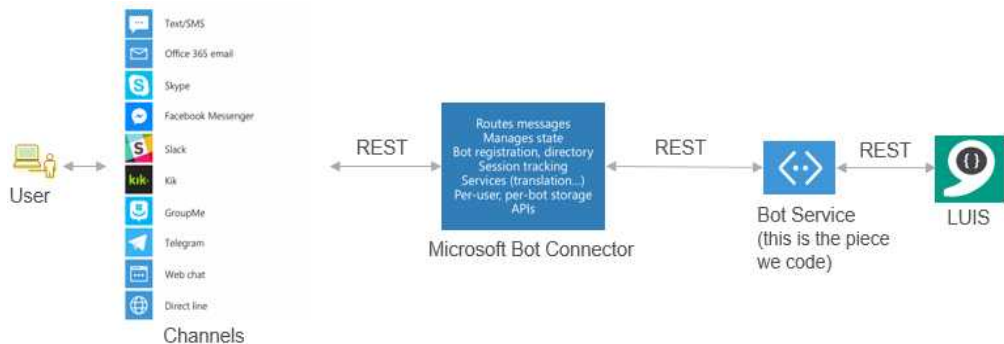
78) 문용식. (2018). 『인공지능 기반의 챗봇 서비스의 국내외 동향분석 및 발전 전망』. 대구: 한국정보화진흥원.

3) 챗봇 프레임워크(Microsoft Bot Framework)

봇 서비스를 개발하기 위하여 다양한 형태의 챗봇 라이브러리가 개발되어 제공되고 있으며, 대표적인 챗봇 플랫폼으로는 Microsoft Bot-Framework, ChatScripts, Facebook의 Bot-for-Messenger 등이 있다. 마이크로소프트의 챗봇 프레임워크인 Bot-Framework은 .NET과 Node.js 언어를 지원하며 클라우드 기반의 봇 포털을 통해 다양한 메신저들과 연결을 할 수 있으며, Bot-Framework SDK(Software Development Kit)를 오픈소스로 제공하고 있으며, 클라우드 환경에서 빠른 테스트, 배포가 가능한 마이크로소프트사의 소프트웨어 기반 봇 개발 프레임워크 이다.

Bot-Framework은 챗봇을 개발하기 위한 다양한 기술적 사항을 프레임워크로 제공하여 개발 과정의 문제 해결을 지원하는 도구와 개발 환경을 제공한다.

[그림 2-6]과 같이 Microsoft의 Bot-Framework에서는 Bot-Channel을 통해 다양한 인터페이스를 제공하고 있다. 국내 대학생들의 친구들과의 대화를 위해 주로 이용하고 있는 페이스북 메신저나 글로벌 점유율이 높은 라인 등의 다양한 채널을 활용할 수 있으며, 직접적인 웹 채팅 채널도 보유하고 있으므로 웹사이트 내의 진로 조언을 위한 채팅 채널의 구성이나 e-포트폴리오와 같은 모바일 앱과 효율적으로 연동할 수 있다.



[그림 2-6] Microsoft Bot-Framework 구조

4) 자연어처리(Natural Language processing)

자연어처리란 우리가 일상생활에서 사용하는 언어인 자연어에서 의미를 분석하여 컴퓨터가 처리할 수 있도록 하는 일을 말하며, 음성 인식, 번역, 사용자의 감성 분석, 텍스트 분류 작업, 질의응답 시스템에서 주로 활용되는 기술이다.⁷⁹⁾

챗봇은 단계형 대화에서부터 자유 형식 대화까지 다양한 대화 스타일을 사용할 수 있으며, 일반적인 챗봇의 장점은 자유 형식 대화에서 더 효율적이고 정확한 사용자의 의도를 찾거나 조언을 위해 시나리오나 선택창을 활용하여 사용자의 정확한 의도 파악과 추가적인 조언 서비스 제공을 통해 효율성을 높일 수 있다. 일반적으로 사용자와 챗봇 간 상호 작용은 대부분 자유 형식이며, 챗봇 인터페이스에서 사용자의 자연어를 이해하기 위한 자연어처리(자연어 이해) 과정을 통해 사용자가 원하는 것을 파악하고 적절한 응답과 동작을 수행한다.

마이크로소프트의 자연어처리 서비스 LUIS(Language Understanding Intelligent Service)는 일반 언어 텍스트를 챗봇이 인식하고 실행할 수 있는 개체로 번역하는 자연어처리 API이다. 딥러닝 기반 엔진을 활용하여 사전에 제시된 문장 기반의 학습 결과로 사용자의 의도를 파악하며, 대화 분석을 통해 사용자가 무엇을 원하는지 이해할 수 있도록 발화 의도(Intent)와 사용자가 입력한 문장 중에 의미 있는 요소인 실체(Entity)를 판별한다.

일반적으로 채널이 사용자의 자연어 메시지를 Bot-Connector에 전송하면, Bot-Connector는 그 메시지를 챗봇에 전달한다. LUIS는 메시지를 발화 의도와 실체로 판별하여 전달하며, 의도는 사용자가 이루고자 하는 목표에 연결되며, 실체는 챗봇이 의도의 매개변수로 필요로 하는 발화 텍스트 내부의 실체에 연결된다. 챗봇은 이러한 의도와 실체를 기반으로 로직을 처리해서 어떤 행동을 취할지 그리고 사용자에게 어떤 응답을 보낼지를 결정하게 된다.

79) 전창욱, 최태균, 조중현. (2019). 『텐서플로와 머신러닝으로 시작하는 자연어처리』. 위키북스.

제 6 절 진로 조언을 위한 대학생 개인 데이터

1) 대학생 e-포트폴리오

전통적으로 포트폴리오는 미술, 건축, 사진, 문학 등의 창작활동을 통해 만들어진 작품들을 조직적으로 모아놓은 작품집을 일컫는 용어로 사용되었다.⁸⁰⁾ 이러한 관점을 토대로 교육 분야에서의 포트폴리오는 학습자들이 무엇을 알고 수행할 수 있는지를 표현하고 이를 평가받기 위해 체계적으로 수집된 다양한 자료들의 모임을 말한다. 구체적으로 기술하자면, 포트폴리오는 시간의 경과에 따른 학습자의 학습 과정, 목표, 노력, 태도, 교육 활동, 성과물 등을 의도적으로 모아둔 것으로, 이를 통하여 교육 목표 달성 정도를 확인하고 학습 개선에 필요한 조언 등 학습 과정 및 결과를 관리하는 기능을 포함하고 있다. 또한, 포트폴리오는 자신의 학습 과정에 대한 성찰의 도구, 교수-학습 활동에 필요한 다양한 정보를 제공하는 도구, 학습자의 적극적 참여를 유도하는 도구 등으로 알려져 있다.⁸¹⁾

e-포트폴리오는 특정의 영역, 활동, 역량에서 장기간에 걸쳐 생성된 개인의 학습과 성장의 과정, 그리고 이에 대한 성찰을 보여주는 각종 디지털 자료를 특정 의도에 맞춰 구조화한 시스템을 의미한다.⁸²⁾

대학에서의 e-포트폴리오는 과거의 학습, 경험, 대외활동에 대한 학점인정과 포트폴리오를 통하여 학습자 스스로 자신의 학습을 주도해 나갈 수 있도록 지원해 주는 시스템으로 직업교육 및 훈련 분야에서의 포트폴리오는 학습자 중심, 결과 중심의 평가 목적으로 사용되어 과거의 학습경험에 대한 인증, 자격증, 지속적인 자기계발을 위한 객관적 자료수집에 활용된다. 대학에 따라 e-포트폴리오는 진로 문제 해결을 위한 접근으로 직업 심리검사와 같은 진로 설계, 역량 진단과 개발, 수업 이수명세, 수상명세 등의 학생 이력에 관한 내

80) 박인우. (2008). 『e-포트폴리오의 교육적 활용 방안 연구 : 초등학교급을 중심으로』. 대구: 한국교육학술정보원.

81) Wolf, K., & Dietz, M. (1998). Teaching portfolios: purposes and possibilities. Teacher Education Quarterly, 25(1), 9-22.

82) 김영애. (2010). 『입학사정관제의 효율적 운영을 위한 e-포트폴리오 설계 연구』. 대구: 한국교육학술정보원.

용을 포함하고 있으며, 상담사 또는 지도교수와의 매핑을 통한 진로지도를 지원하고 있으므로 진로 조언 플랫폼에서 학생의 개인화된 매칭 데이터로 활용하기에 최적화되어 있는 데이터이다.

하지만 대학의 e-포트폴리오의 문제점 중 하나는 연속성이다. 대학에서 배운 지식으로 평생 고용을 보장받았던 과거와 달리 현대 시대는 빠른 기술과 환경의 변화로 인하여 대학 이후로도 평생 계속 새로운 지식과 기술을 학습해야 하는 시대이다. 하지만 대부분의 대학 e-포트폴리오는 대학 생활에 최적화되어 설계되었기 때문에 졸업 이후의 활용에 한계가 있다.

평생학습에 맞춰 취업 이후로도 퇴사와 지속적인 진로 변경이 요구되는 현실에 맞춰 대학 이후로도 계속 학습이 이루어지고 e-포트폴리오에서의 역량과 경력이 업데이트되어야 하며, 취업 이후로도 필요시 진로 조언을 통한 진로 변경이 가능해야 한다. 자기계발, 기술 및 지식의 성장과 같은 개인의 역량 강화는 개개인의 경쟁력뿐만 아니라 직장 생활에서의 중요한 경쟁 요소가 된다. e-포트폴리오와 동시에 평생학습의 개념을 생각하면, e-포트폴리오의 중요성은 시간이 갈수록 더 중요해질 것이다.⁸³⁾

2) 대학 내부 시스템의 학생 정보 데이터

최근 에듀테크 분야에서는 학습자들의 학습데이터를 수집·분석하는 학습분석(Learning Analytics) 기술이 주목받고 있다. 학습분석은 학습자로부터 발생하는 다양한 데이터 분석을 통하여 효과적인 학습 모델을 구축하는 것으로, 학습자의 성적, 학점, 행동 등 다양한 학습자 데이터를 수집·분석과 이를 학습하여 맞춤형 교육을 가능하게 한다.⁸⁴⁾

학습분석 형태로, 애리조나주립대학, Austin Peay State University, University of Florida 등의 대학에서는 학업 성적 관리 프로그램과의 연동을 통해 학생들의 교육 데이터를 활용, 학생 개개인의 학업 성과 및 적성에 맞는 전공 지도를 추천해 주고, 그와 관련된 수강 과목을 제안해 주고 있다. 국내

83) Sullivan, R. (2018). Exploring Emerging Technologies for Lifelong Learning and Success. The State University of New York.

84) 정상현. (2017). 『‘대학 4.0’과 학습분석...맞춤형 고등교육 앞당긴다』. 서울: 한국대학신문.

대학들 역시 학습분석을 통한 맞춤형 지도 서비스 제공을 위해 대학 내 데이터의 활용 가능성을 분석하여 수집 및 연구하고 있는 단계이다.

[표 2-2]와 같이 대학 내부의 진로·취업 관련 연동 데이터들을 학사 관련 시스템, LMS 시스템에서부터 취업, 봉사 활동과 같이 다양한 시스템들과 연계를 통한 맞춤형 진로 지원이 가능하다.

또한, 이러한 데이터들을 지식 그래프와 연계하여 성공적으로 취업에 성공한 선배들의 역량 정보들의 매칭을 통한, 진로 방향설정 지원, 역량 개발을 위한 학습 로드맵을 추천과 같이 개인의 저장된 데이터를 기준으로 최선의 선택을 위한 정보 제공과 조언이 가능하다.

[표 2-2] 대학 내부의 진로·취업 관련 연동 데이터

구분	연동 정보
통합역량관리시스템	<ul style="list-style-type: none"> · 학생의 이력 및 경력 · 역량 개발 현황 · e-포트폴리오 · 진로·취업 관련 활동 · 진로상담 내역 · 비교과 관리
학사 관련 시스템	<ul style="list-style-type: none"> · 사용자 정보 · 교과 정보 · 이수/학점 정보
취업 관련 시스템	<ul style="list-style-type: none"> · 인턴십/기업/채용정보
봉사 관련 시스템	<ul style="list-style-type: none"> · 봉사 활동 정보
LMS 시스템	<ul style="list-style-type: none"> · 수강 정보, 과제물 정보 · 토론, 글 게시 횟수
저장소	<ul style="list-style-type: none"> · 졸업학생 정보

제 3 장 그래프 데이터베이스를 활용한 진로 조언 플랫폼 설계

제 1 절 진로 조언 플랫폼 설계 방향 및 모델

1) 진로 조언 플랫폼 설계 방향

가) 연구내용

본 연구는 국내 대학생들의 가장 큰 문제점인 진로 문제를 해결하기 위해 ‘2017년 대학진로교육현황조사’에서 대학생들이 요구사항들을 분석하여 현대 정보기술을 활용한 진로 조언 플랫폼을 설계하고, 설계된 진로 조언 플랫폼의 실제 구현을 통해 정보기술을 활용한 대학의 진로 문제점 해결 방안 제시와 학생들의 진로 목표 결정과 진로 개발을 지원하기 위한 상담과 진로 정보 제공 및 역량 관리 기능을 제공하는 플랫폼 연구로 연구의 주요 내용은 [표 3-1]과 같으며, 직업정보 데이터, 지식 그래프 구축, 챗봇, e-포트폴리오로 구성된 진로지도 서비스 제공을 위한 플랫폼 설계와 구축이다.

[표 3-1] 진로 조언 플랫폼 주요 연구내용

설계 항목	설명	구현 내용
직업정보 데이터	학생들의 진로 조언을 위한 높은 품질의 진로 관련 지식 데이터	· 고용노동부 공공 데이터 · 기업 구인 데이터 수집
지식 그래프 구축	진로 조언을 위한 지식 그래프	· 수집된 직업 역량 내에서의 그래프 설계
챗봇	학생의 질문에 대한 상담 및 조언	· 24시간 진로상담이 가능한 챗봇 인터페이스
e-포트폴리오	맞춤형 직업 추천 및 진로 개발 지원	· 개인 역량 정보 기록과 직업 매칭 및 관리

나) 변화가 빠른 직무 조업을 위한 지식 수집

정보기술 분야와 같이 변화가 빠른 직업에 대한 조업을 위해서는 현재의 기술 트렌드와 기업이 필요로 하는 인재상과 같은 데이터를 빠르게 수집할 수 있어야 하며, 학생들에게 진로 조업을 위한 정보 제공을 위해서는 해당 직업군의 현장조사를 통한 연봉, 기술, 자격증, 학과 등의 데이터와 미래 전망을 예측할 수 있는 데이터 역시 필요하다.

진로 조업에 최적화된 데이터를 찾기 위해 정부·지자체·공공기관이 보유한 공공데이터들과 민간 취업 정보 포털들에서 수집할 수 있는 데이터들을 분석하여 [표 3-2]와 같이 한국고용정보원 워크넷 직업·직무정보 데이터를 선정하였으며, 민간 취업 정보 포털에서 기업들이 구인공고 중 정보기술 직업들에 대한 채용공고 데이터를 조업을 위한 실제 채용정보 데이터로 선정하였다.

[표 3-2] 진로 조업을 위한 지식 데이터 선정

구분	데이터 내용	업데이트 주기
워크넷 직무정보 API	NCS 기반의 직무데이터로 각 직무의 내용, 지식 및 기술, 수행 태도, 관련 자격증, 관련 전공, 관련 훈련 정보	업데이트 느림
워크넷 직업정보 API	직업 설명, 하는 일, 교육/자격/훈련, 임금/직업만족도/전망, 능력/지식/환경, 성격/흥미/가치관, 업무 활동 등	매년 조사 실시, 데이터 업데이트는 느림 (1~5년)
민간채용정보 포털	기업이 채용하고 있는 직무에 대한 학력, 나이와 같은 일반 데이터와 해당 직무의 직원을 채용하기 위한 기술, 학력, 우대사항, 자격증, 성향 등	필요할 때 즉시 수집, 데이터 업데이트 즉시 가능

(1) 워크넷 직업정보 API

공공데이터인 워크넷 직업정보의 경우, 국내 대표적인 830여 개 직업에 대해 고용정보원 미래직업연구팀이 매년 실시하는 재직자 조사 결과를 바탕으로, 각 직업의 업무와 업무를 위한 교육, 자격, 특성 등과 같은 종합적인 데이터 사전을 구축하고 있다. 워크넷 직업정보 데이터, 연구보고서 및 실제 수집데이터를 확인하여 업데이트 주기를 분석결과 항목별로 1~4년 정도의 간

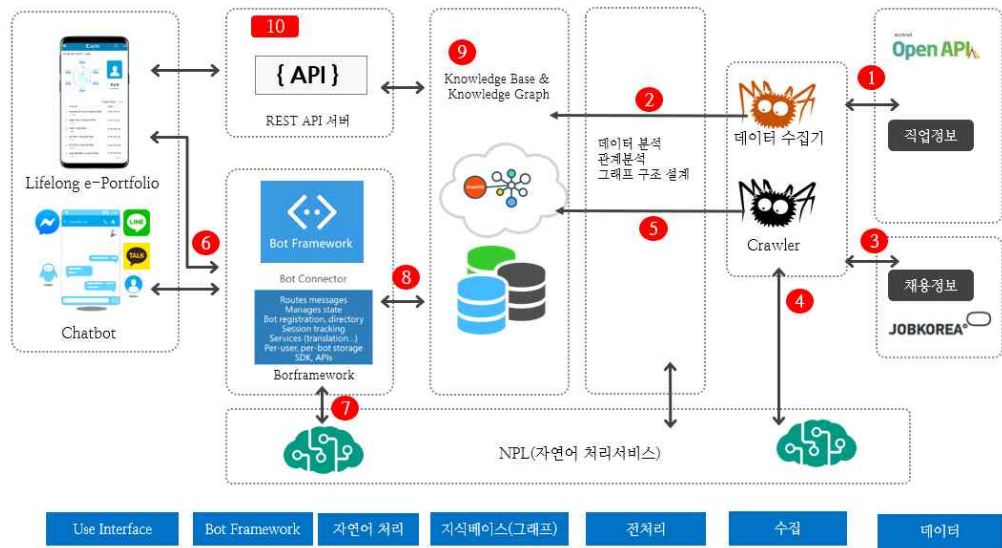
격이 있는 것으로 확인되며, 학생들에게 조언이 필요한 일자리 전망, 연봉, 지식의 가중치와 같이 실제 조언에 필요한 데이터를 다수 포함하고 있지만, 일부 항목들의 경우 5년 이상의 오래된 정보들도 확인되었다.

(2) 기업의 채용 데이터 수집

민간 채용정보 포털의 경우, 실제 기업들이 원하는 채용공고에 올린 해당 직무를 수행과 요구사항에 적합한 인력을 채용하기 위한 다양한 역량 데이터들이 포함되어 있으므로, 현재 취업 시장에서의 기술, 자격, 성향, 나이, 경력, 우대사항들과 같은 학생들에게 필요한 정보의 추출이 가능하나, 이러한 구인 정보들은 대부분 형식과 구조가 정해져 있지 않은 HTML 형태의 비정형데이터로 실제 활용을 위해서는 데이터의 수집·분석 단계와 전처리 과정을 거쳐 실제 활용이 가능한 형태의 지식 데이터를 생성한다. 이러한 기업의 채용정보는 학생들이 실제 취업 시장 진입을 위한 현실 데이터로 실제 학생들의 진로 결정과 취업 시장 진입을 위한 효과적인 조언 데이터로 활용 가능하며, 필요할 때 실시간으로 데이터를 수집하여 최근의 채용정보를 업데이트할 수 있는 장점이 있다.

2) 진로 조언 플랫폼 아키텍처 설계

진로 조언 플랫폼은 선행연구에서 구현하고 검증했던 아키텍처 구조를 기반으로 진로 조언 데이터의 변경과 추가, 데이터 수집과 수집된 데이터의 전처리를 위한 시스템, 실제 학생의 데이터 저장과 목표 설정 및 관리를 할 수 있는 e-포트폴리오 시스템을 추가하여 [그림 3-1]과 같이 진로 조언 플랫폼 아키텍처를 설계하였으며, 각 단계에 대한 동작 흐름에 대한 설명은 [표 3-3]과 같다.



[그림 3-1] 진로 조언 플랫폼 아키텍처 설계

[표 3-3] 진로 조언 플랫폼 아키텍처 단계별 동작 흐름

동작	단계명	설명
1	직업정보 API 수집	API 활용 고용정보원 워크넷 직무데이터 수집
2	데이터 전처리	직업정보데이터 분석/설계하여 그래프 데이터베이스에 저장
3	채용정보 크롤링	민간 채용정보 크롤링
4	자연어 분석	비정형 문서에서의 중요 키워드 추출
5	데이터 전처리	추출된 키워드 분석/설계를 통한 그래프 데이터베이스에 저장
6	조언 요청	사용자 인터페이스를 통한 조언 요청
7	자연어 분석	조언 요청 내용 분석 (의도, 실체(엔티티))
8	그래프 쿼리	매칭, 가중치 계산을 활용한 지식 쿼리
9	그래프 데이터베이스	요청한 지식에 대한 쿼리 내용 출력
10	API 서비스	e-포트폴리오의 사용자 역량을 그래프 데이터에 매칭한 결과값(직업 정보) 전달

3) 관계형 데이터베이스와 그래프 데이터베이스 비교

가) 관계형 데이터베이스에서의 관계 분석 문제점

그래프 데이터베이스는 2000년 중반부터 연구되기 시작하여 2010년 중반에 들어서면서 빅데이터에서의 복잡한 관계 저장·분석을 위한 데이터베이스시스템으로 알려지기 시작했기 때문에 관련된 연구가 매우 적었기 때문에 진로조언을 위한 지식 베이스로 가장 먼저 테스트를 진행한 데이터베이스는 관계형 데이터베이스였으나 관계 기반의 지식 베이스를 구축과 관계 기반의 분석에 활용하기 위한 테스트 시 [표 3-4]와 같은 문제점이 발견되었다.

[표 3-4] 관계형 데이터베이스에서의 문제점

문제점	설명
관계 Depth 증가에 따른 효율 저하	· 관계(JOIN)가 늘어날수록 쿼리 속도 증가 · 쿼리문의 복잡도 및 컴퓨팅 자원 소모 증가
지식의 확장 어려움	· 서로 다른 종류와 형식을 가지는 지식을 확장하기 위해서는 두 데이터베이스를 연결해 주기 위한 별도의 복잡한 프로세스가 필요
관계 분석의 어려움	· 복잡한 JOIN과 쿼리문으로 인해 관계 분석의 복잡도 증가
시각화의 어려움	· 데이터 간의 관계를 분석하기 위한 시각화가 어려우며, 별도의 작업과 소프트웨어 추가 필요

나) 성능 비교를 위한 데이터베이스 구성

그래프 데이터베이스와 관계형 데이터베이스의 성능 비교를 위하여 [표 3-5]와 같이 두 개의 데이터베이스 환경을 클라우드 환경에 구축하고, 고용노동부 직무데이터 사전을 호출·전처리 과정을 거쳐 양쪽의 데이터베이스에 저장하였다. 비교 대상인 두 개의 상품의 차이로 인하여 인프라 환경은 다르지만, 마이크로소프트 Azure 클라우드 내에서는 비교 대상의 하드웨어가 모두 최저 사양이며, 비용 역시 큰 차이가 없으므로 해당 하드웨어 구성을 비교 대상으로 구성하였다.

[표 3-5] 비교 실험을 위한 데이터베이스시스템 환경

관계형 데이터베이스 환경 구성	그래프 데이터베이스 환경 구성
·시스템 구성: Azure SQL Server ·시스템 사양: DTU기반 Basic 모델	·시스템 구성: Container Instances ·시스템 사양: 1 vCPU, 1.5G RAM

다) 관계형 데이터베이스와 그래프 데이터베이스의 비교 실험

실험은 [표 3-6]과 같이 5가지의 역량에 대해 관계가 매칭되는 직업정보를 찾는 방식으로 진행하였으며, 각각의 쿼리문은 점선의 박스로 표시하였다. 실제 그래프 데이터베이스의 사이퍼 쿼리의 경우 관계 구조만 알고 있으면 쿼리가 가능했지만, 관계형 데이터베이스에서는 복잡한 JOIN 과정이 별도로 필요했다.

[표 3-6] 데이터베이스 관계 쿼리 비교

[관계 쿼리 비교]

기술적 성향, 컴퓨터공학 전공, 정보보안기사 자격증, 정보화사업 기술, CISA 대비반 훈련을 받은 학생에게 매칭되는 직업은?*

The screenshot shows a SQL Server query window with a complex JOIN query. The query is as follows:

```

SELECT
FROM
ABIT
LEFT JOIN ABITATT ON ABIT.ID = ABITATT.ABITID
LEFT JOIN ATT ON ABITATT.ATTID = ATT.ID
LEFT JOIN ABITMAJOR ON ABIT.ID = ABITMAJOR.ABITID
LEFT JOIN MAJOR ON ABITMAJOR.MAJORID = MAJOR.ID
LEFT JOIN ABITCRFITEM ON ABIT.ID = ABITCRFITEM.ABITID
LEFT JOIN CRFITEM ON ABITCRFITEM.CRFITEMID = CRFITEM.ID
LEFT JOIN ABITKNWSKI ON ABIT.ID = ABITKNWSKI.ABITID
LEFT JOIN KNWSKI ON ABITKNWSKI.KNWSKIID = KNWSKI.ID
LEFT JOIN ABITTRNG ON ABIT.ID = ABITTRNG.ABITID
LEFT JOIN TRNG ON ABITTRNG.TRNGID = TRNG.ID
WHERE
ATT.Value = '기술적'
AND
MAJOR.Value = '컴퓨터공학'
AND
CRFITEM.Value = '정보보안기사'
AND
KNWSKI.Value = '정보화사업'
AND
TRNG.Value = '보안' CISA 대비반

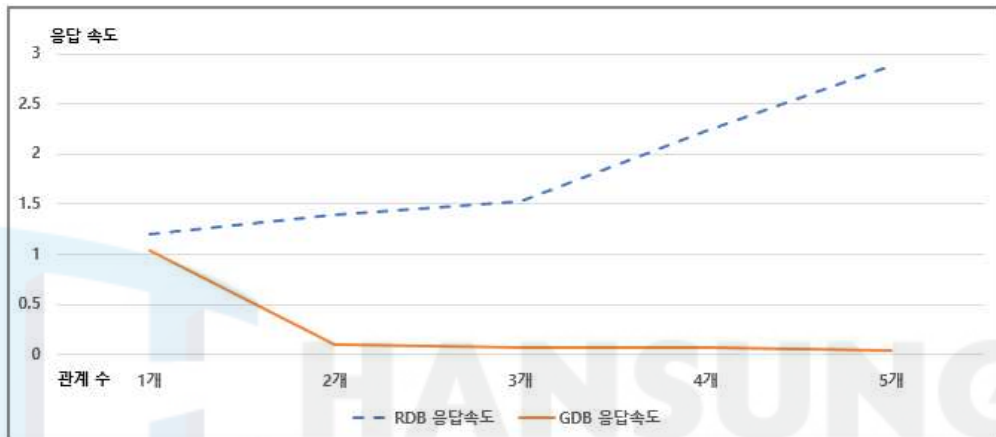
```

The execution results on the right show the following details:

- Neo4j/3.5.13
- ver address: cf-neo4j-md-eastasia.azurecontainer.io:7687
- ry: MATCH (n:Abit)-[:ATT]->({att{Name:'기술적'}),(n)-[:MAJOR]->({major{Name:'컴퓨터공학'}),(n)-[:CRFITEM]->({crfitem{Name:'정보보안기사'}),(n)-[:KNWSKI]->({knwsk{Name:'정보화사업'}),(n)-[:TRNG]->({trng{Name:'보안'} CISA 대비반)}) RETURN n
- mary: [{"statement": [{"text": "MATCH (n:Abit)-[:ATT]->({att{Name:'기술적'}),(n)-[:MAJOR]->({major{Name:'컴퓨터공학'}),(n)-[:CRFITEM]->({crfitem{Name:'정보보안기사'}),(n)-[:KNWSKI]->({knwsk{Name:'정보화사업'}),(n)-[:TRNG]->({trng{Name:'보안'} CISA 대비반)})inRETURN n", ...
- ponse: [{"keys": [...
- reaming 1 records after 1 ms and completed after 43 ms.

[그림 3-2]. [표 3-7]은 역량을 매칭 개수에 따른 응답속도의 실험 결과이다. 그래프 데이터베이스의 경우 매칭되는 대상의 숫자가 증가할수록 오히려 더 빠른 성능을 보여주고 있으며, 매칭 데이터의 수와 관계없이 일정한 응답속도를 보여주고 있다.

관계형 데이터베이스의 경우 JOIN 되는 개수가 증가할수록 성능과 함께 응답속도가 증가하고 있음을 확인할 수 있다. 이러한 비교는 클라우드의 환경이나 시스템의 성능에 의해서도 차이가 있으나, MySQL과 그래프 데이터베이스를 비교한 Chad Vicknair의 연구⁸⁵⁾, Alexandra Martinez의 연구⁸⁶⁾에서도 조인 작업이 필요한 쿼리에서는 그래프 데이터베이스의 성능이 더 우수한 것으로 실험을 통해 확인되었다.



[그림 3-2] 매칭되는 관계의 증가에 따른 응답속도 비교

[표 3-7] 관계의 증가에 따른 응답속도 비교

관계 쿼리 수	1개	2개	3개	4개	5개
RDB 응답속도	1.206	1.402	1.524	2.230	2.876
GDB 응답속도	1.044	0.094	0.077	0.076	0.043

85) Vicknair, C., Macias, M., Zhao, Z., Nan, X., Chen, Y., Wilkins, D. (2010). A comparison of a graph database and a relational database: a data provenance perspective. In Proceedings of the 48th annual Southeast regional conference (p. 42). ACM.

86) Martinez, A., Mora, R., Alvarado, D., Lopez, G., Quiros, S. (2016). A Comparison between a Relational Database and a Graph Database in the context of a Personalized Cancer Treatment Application. In Proceedings of the 10th Alberto Mendelzon International Workshop on Foundations of Data Management.

제 2 절 그래프 데이터베이스를 활용한 지식 그래프 설계

1) 그래프 데이터베이스를 활용한 지식 그래프 설계 방법론

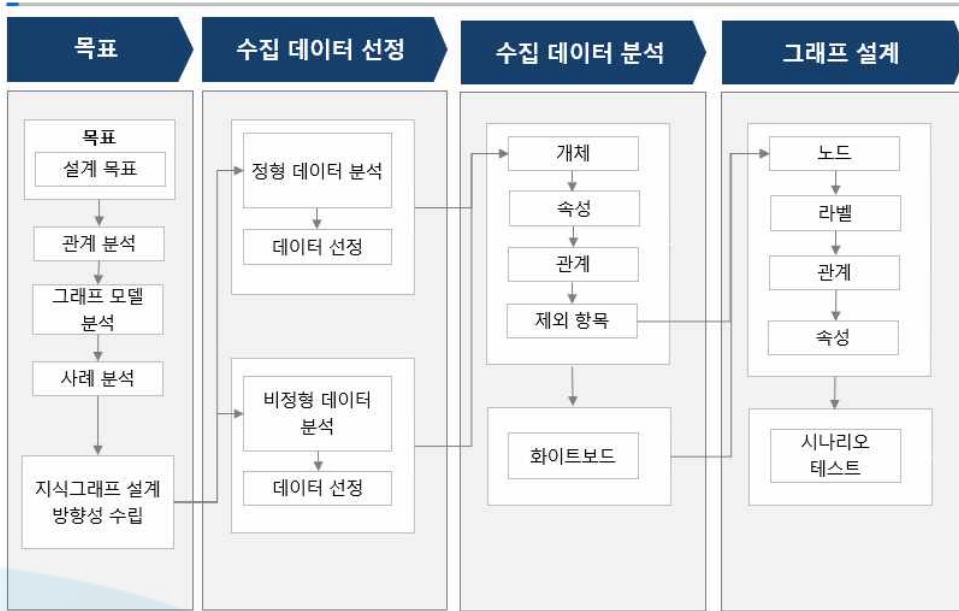
최근 데이터의 크기가 증가하고 복잡도가 커짐에 따라 산업계와 학계에서의 그래프 데이터베이스의 관심도가 증가하고 있으나 국내에서의 그래프 데이터베이스에 관한 연구는 아직 초기 단계이며, 해외 연구를 통해 제시된 일부 방법론들 역시 활용 목적과 연구 방향에 따라 차이로 활용할 수 없었다.

선행연구와 본 연구를 시행하는 과정에서의 절차, 최적화, 문제점과 해결 사항들을 기반으로 정형데이터와 비정형 데이터를 수집하여 “그래프 데이터베이스를 활용한 지식 그래프 설계 방법론”(이하 설계 방법론)과 “그래프 데이터베이스를 활용한 지식 그래프 구축 방법론”(이하 구축방법론)을 제시하였으며 연구 과정의 데이터 분석부터 지식 그래프 생성까지의 전 단계에서 연구된 설계 방법론과 구축방법론을 활용하였다.

[그림 3-3]과 같이 그래프 데이터베이스를 활용한 지식 그래프 설계 방법론은 그래프 데이터베이스 활용을 위한 목표에서부터 수집데이터 선정, 분석, 그래프 설계 프로세스로 제시하였다.

가) 설계 목표

데이터 모델링은 추상화 프로세스이다. 비즈니스 혹은 사용자의 목표 및 요구사항에 따라 데이터 구조의 설계 방식은 달라지며, 특히 그래프 데이터베이스에서는 데이터 구조 및 방법이 크게 달라지기 때문에 실제 이를 통한 관계 분석을 통해 어떤 그래프 모델이 적합할지에 대해 선정해 나가야 한다. 또한, 목표와 비슷한 형태의 성공 사례를 찾을 수 있다면 해당 성공 사례를 참고할 수 있다. 만약, 목표에서 관계 분석이 매우 적거나 간단한 형태의 관계 분석이라면 그래프 데이터베이스보다는 다른 형태의 데이터베이스를 선정해야 한다. 이러한 설계 목표 분석을 통해 지식 그래프 설계 방향성이 수립되어야 한다.



[그림 3-3] 그래프 데이터베이스를 활용한 지식 그래프 설계 방법론

나) 수집 데이터 선정

수집데이터는 일반적인 형태로 데이터의 성격에 따라 정형데이터와 비정형데이터로 분류하였다. 정형데이터는 API를 통해 수집되거나 DB, CSV 등의 형태로 제공되는 데이터로 활용을 위한 분류가 되어 있기 때문에 데이터의 구조분석과 역량과 같은 속성의 추출이 쉽다. 비정형 데이터는 문서, 음성 등과 같이 규칙이 정해져 있지 않은 데이터로 데이터를 활용하기 위해서는 별도의 가공을 거쳐야 하며, 필요한 데이터를 찾거나 수집된 데이터에서 관계를 설계하기도 비교적 어렵다. 하지만 정형데이터보다는 비정형데이터 형태로 존재하는 데이터의 종류가 더 많으며, 가치 있는 유용한 데이터가 많으므로 지식 그래프 구축을 위한 지식 데이터로 검토되어야 한다.

다) 수집데이터 분석

관계를 분석하기 위해서는 사람, 사물, 직업, 성격, 지식 등과 같은 개념 또는 정보의 단위에 해당하는 객체에 해당하는 분류와 이러한 객체의 집합을 분류해야 한다. 정형데이터들은 이러한 객체들이 일정한 순서대로 분류되어

있으므로 확인 및 분류가 쉽지만, 비정형 데이터들은 각 속성을 묶어 별도의 분류 과정이 필요하다. 그래프 데이터베이스의 가치는 이러한 관계 간의 분석에서 비롯되기 때문에 연결되는 관계가 없는 개체는 제외하는 것이 바람직하다. 모든 개체에는 하나 이상의 관계가 존재해야 하며, 관계의 이름은 관계를 표현하는데 가장 알맞은 이름을 사용해야 한다.

데이터에서의 제외 항목은 그래프 데이터베이스에서 불필요한 항목이다. 관계가 없는 데이터나 혹은 속성 정보가 너무 많은 것과 같이 그래프 데이터와 알맞지 않은 형태의 데이터들은 데이터들의 특성에 적합한 별도의 데이터베이스로 분리하는 것이 바람직하다.

그래프 데이터베이스는 화이트보드 친화적인 데이터베이스이기 때문에 분석된 개체, 속성, 관계에 해당하는 구조를 스케치해 보는 것이 설계단계에 들어가기에 앞서 가장 효율적이며, 문제점을 빠르게 파악하여 보완할 수 있다.

라) 그래프 설계

분석된 데이터를 그래프 기반으로 설계 작업을 수행해야 한다. 이전 단계에서 화이트보드에 데이터와 관계에 대해 스케치를 했었다면, 그 스케치된 노드와 관계 구조를 기반으로 그래프를 설계할 수 있다.

노드는 중복되지 않는 고유노드와 중복되는 노드로 나누어지며, 일반적인 대표 노드에 해당하는 고유노드를 제외하면 대부분의 관계되는 노드들은 중복이 된다. 예를 들어 직업 분석에서 직업에 해당하는 노드는 중복되지 않지만, 직업에 관련된 다양한 역량 정보들은 중복이 되며, 이러한 관계를 분석하고자 할 때는 중복노드를 제거하고, 해당 노드를 보유하고 있었던 노드들과의 관계를 연결해야 한다.

최종적으로 설계된 그래프 구조에서 구축 목표에 대한 시나리오 중 일부를 호출하는 방법의 테스트를 통해 설계 과정에서의 문제점을 발견하여 보완할 수 있다.

2) 그래프 데이터베이스 기반의 지식 그래프 목표

가) 설계의 목표

지식 그래프 설계의 목표는 연구의 목적인 학생들에게 진로정보를 조언해 주기 위한 관련 지식의 모음이다. 따라서 학생들의 질문에 대한 원활한 답변이 가능하고, 분석된 연관 정보들을 조언함으로써 유용성과 활용성을 극대화해야 한다. 이를 위한 지식 그래프 설계의 목표는 ①~⑥과 같다.

- ① 챗봇으로 문의한 학생의 직업 관련 역량 문의에 대한 답변
- ② 역량 정보를 기반으로 한 직업 관련 정보 분석
- ③ e-포트폴리오 내의 데이터를 활용한 맞춤형 조언
- ④ 실제 비정형 채용정보를 수집하여 진로 조언을 위한 그래프 설계
- ⑤ 채용정보의 직업과 연결된 역량 정보를 활용하여 학생들이 직업개발을 위해 필요한 역량 추출
- ⑥ 정형데이터인 워크넷의 직업정보 API를 수집·분석하여 지식 그래프 설계

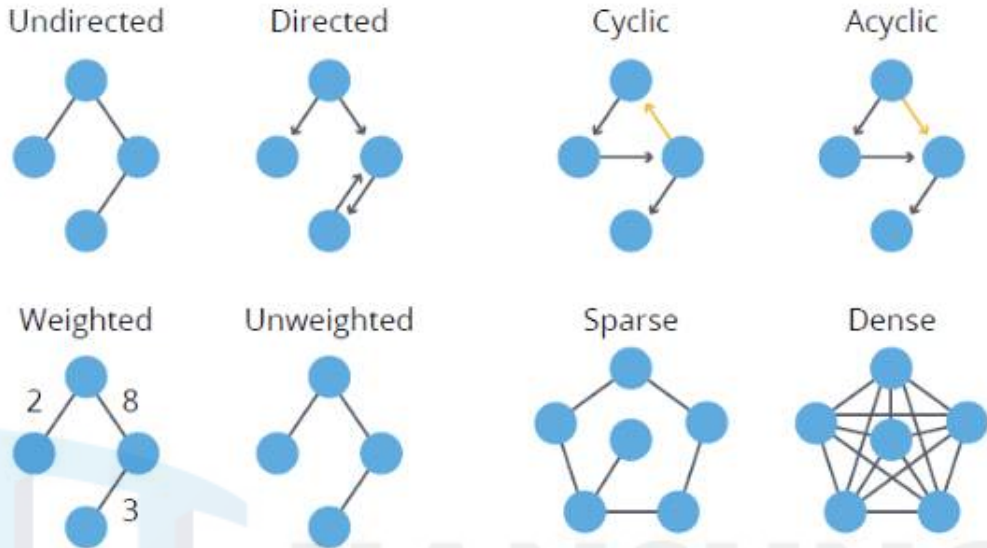
나) 관계 분석

직업과 직업 관련 역량과의 관계를 맺는 과정에서 너무 많은 관계가 직접한 개에 노드에 연결되는 형태는 쿼리 속도, 시각화의 문제, 복잡성의 증가와 같은 문제를 일으킬 수 있다. 워크넷 직업정보 API의 경우 직접 연결되는 역량 노드가 100개가 넘기 때문에 직접 역량 정보를 직업 노드에 연결하는 것은 바람직하지 않다. 따라서 각각의 역량 정보의 단위마다 분류 노드를 삽입하였으며, 채용정보 분석과 워크넷 직업정보와의 노드를 분리하여 사용자의 질문에 따라 관계를 쿼리할 수 있도록 설계한다.

다) 그래프 모델 분석

그래프는 목표에 따라 [그림 3-4]와 같이 다양한 특성의 그래프를 구성할 수 있다. 직업 관련 역량 노드는 직업 노드에 포함되며, 포함되는 관계를 화살표로 구성할 수 있기 때문에 방향을 지정하는 Directed 특성과 학생들에게 도움이 되는 조언 정보를 제공해 주기 위한 Weighted 특성을 가지도록 설계

한다. 또한, 직업정보 데이터 중 일부는 Unweighted 특성으로 설계하는 것과 같이 다양한 그래프 특성이 혼합된 형태의 구조가 일반적이다.



[그림 3-4] 그래프 특성

라) 지식 그래프 설계 방향성 수립

지식 그래프는 실제 개체, 이벤트, 상황 또는 추상적 개념과 같은 개체의 상호 연결된 모음을 나타낸다. 그래프에는 정보와 관계, 해당 정보 간의 연결이 있으며, 지식 그래프에 데이터가 추가되면 추가된 데이터 역시 기존 데이터와의 관계가 연결되어야 한다. 관계가 연결되지 않은 지식은 그래프 데이터베이스에서 활용성이 떨어지며, 관계가 끊어져 있는 지식의 묶음 역시 비효율적이며, 데이터베이스 구조만 더 복잡하게 만들 수 있다.

진로 조언을 위한 지식 그래프에서의 학생들이 방향을 설정하거나 목표로 하는 직업이 고유노드가 되며, 직업에 다양한 관련 관계와 정보들이 계속 연결되는 모델로 설계한다.

3) 수집데이터 선정

가) 기업의 채용정보 데이터 수집

대학생들의 진로 조언을 위한 직업 관련 지식 데이터로 워크넷 직무사전의 경우 대학생들에게 실제 필요한 정보는 현장 전문가나 지도교수와 같은 전공과 직업에 이해도가 높은 전문가의 진로 조언을 위한 지식 데이터로는 부족하다. 가장 큰 이유는 빠르게 변화하는 특성 정보와 실제 기업이 원하는 기술, 우대사항과 같은 인재상에 대한 데이터 역시 업데이트가 빠르게 진행되어야 하지만 실제로 업데이트 주기가 늦기 때문에 조언을 위한 데이터로 활용하기는 부족하다. 실제 기업들이 채용을 원하는 인재상에 대한 정보는 기업들이 실제 채용공고를 등록하는 민간 취업 정보 포털이 가장 많이 보유하고 있으므로 가장 대표적인 두 개 포털에서 해당 직무에 대한 구인을 원하는 기업들이 실제 등록한 채용 데이터를 수집한다.

(1) 수집 채용 직업분류

연구 수행을 위한 직업분류 및 직업명은 [그림 3-5]와 같이, 한국고용정보원의 직업 사전 분류 코드를 기준으로 하였으나, 전통적 직업에 맞춰져 있으므로 정보기술이 속해있는 정보기술 관련 업종의 많은 직업이 분류되어 있지 않기 때문에 민간 취업 정보 포털에 있는 분류를 기준으로 ‘클라우드컴퓨팅개발자’, ‘인공지능연구’ 등과 같은 직업 사전 분류 코드에 포함되어 있지 않은 직업 46개를 추가하여 신규 직업들을 추가하여 [표 3-9]와 같이 총 89개의 정보기술 분야 직업을 선별하였다.

[표 3-8] 한국고용정보원 직업 사전 분류 코드 일부 항목

중분류명	직업코드	소분류명
정보통신관련직	20120	통신공학기술자
	20128	통신장비기술자
	20220	시스템 소프트웨어 개발자(프로그래머)
	20231	응용소프트웨어개발자
	20232	컴퓨터프로그래머
	20233	디지털영상처리전문가
	20234	가상현실 전문가
	20236	게임 프로그래머

[표 3-9] 채용정보를 수집할 직업 리스트

문제점	설명
직업 사전 분류 코드 정보통신 관련직 분류 직업(43개)	웹 기획자, 게임기획자, 애니메이션기획자, 컴퓨터 하드웨어 기술자, 통신공학기술자, 통신기기·장비 기술자, 통신망 운영 기술자, 인공위성 개발원, 통신기술 개발자, 통신장비기술자, RFID시스템개발자, 컴퓨터 시스템 설계 및 분석가, 시스템 소프트웨어 개발자(프로그래머), 응용 소프트웨어개발자, 컴퓨터프로그래머, 디지털영상처리전문가, 가상현실 전문가, 음성처리전문가, 게임 프로그래머, 모바일앱개발자, 데이터베이스개발자, 네트워크 관리자, 네트워크엔지니어, 컴퓨터보안전문가, 웹엔지니어, 웹프로그래머, 웹 마스터(운영자), 정보시스템 운영자, 통신장비기사, 방송송출장비기사, 컴퓨터시스템감리전문가, IT 기술지원 전문가, 정보통신건설턴트, MIS전문가(경영정보시스템개발자), 영상 및 관련장비설치 및 수리원, 통신 및 관련장비설치 및 수리원, 증강현실 전문가, IT 테스터 및 IT QA 전문가, 네트워크 프로그래머, 방송·통신·인터넷 케이블 설치·수리원, 모바일 애플리케이션 개발자, 빌딩정보 모델링(BIM)전문가, 빅데이터분석가
민간 채용정보 포털에서 추가 수집한 정보통신관련직 분류 직업(46개)	지능로봇연구개발자, 지능형교통시스템(ITS)연구원, 모바일게임QA전문가, 게임시나리오작가, 프로그래머, 게임마케터, 게임운영자, 게임사운드크리에이터, 게임프로그래머, 게임그래픽디자이너, 게임기획자, 컴퓨터프로그래머, 컴퓨터보안전문가, 컴퓨터하드웨어기술자, 클라우드컴퓨팅개발자, 네트워크시스템분석가 및 개발자, 사물인터넷(IOT)개발자, 스마트폰애플리케이션개발자, 빅데이터분석가, 감성인식개발자, 전자공학기술자, 게임 QA, IOT 개발, 임베디드 개발, 시스템 엔지니어, 클라우드 시스템 엔지니어, 생체인식 개발, GIS 개발, 웹퍼블리셔, 인공 지능 연구, 인공지능 개발, 블록체인 개발, 클라우드 엔지니어, 스마트 시티, 서버 엔지니어, 시스템엔지니어, 데이터 분석, 빅데이터 전문가, GIS DB, 빅데이터 엔지니어, 프론트엔드개발, 백엔드개발, 금융 개발, 인공지능 개발자, 자연어처리, 빅데이터 개발

(2) 채용공고 수집항목

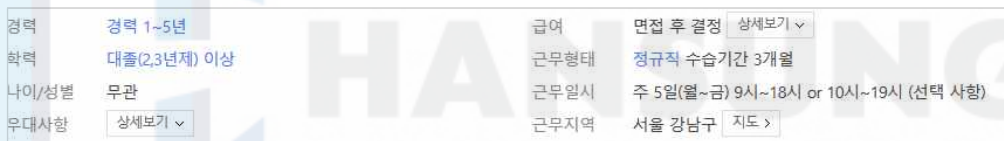
선정한 직업명 키워드로 민간 취업 정보 포털에서 채용정보 검색을 하면, 해당 직업(직무)을 채용하는 기업들의 공고 리스트가 출력된다, [그림 3-5]는 ‘웹프로그래머’ 키워드로 검색된 결과로 총 3,571건의 채용공고 리스트를 확인할 수 있다.

각 채용공고에서 [그림 3-6]과 같이 경력, 학력, 나이, 급여 등의 일반적인 정보는 각 민간 취업 정보 포털의 API 서비스를 통해 수집할 수 있지만, 실제 기업이 채용하는 세부 역량 정보의 경우 정형화되어 있는 양식이 아닌, 각 기업이 직접 제작하거나, 혹은 민간 취업 정보 포털에서 제공하는 에디터를 활용하여 직접 작성한 비정형 문서에 포함되어 있으며, 이러한 정보는 기

업들이 자율적으로 작성한 문서이기 때문에 양식이나 표현방식, 약어 등과 같은 다양한 변수를 추가로 고려해야 하는 문제점 또한 존재하였다.

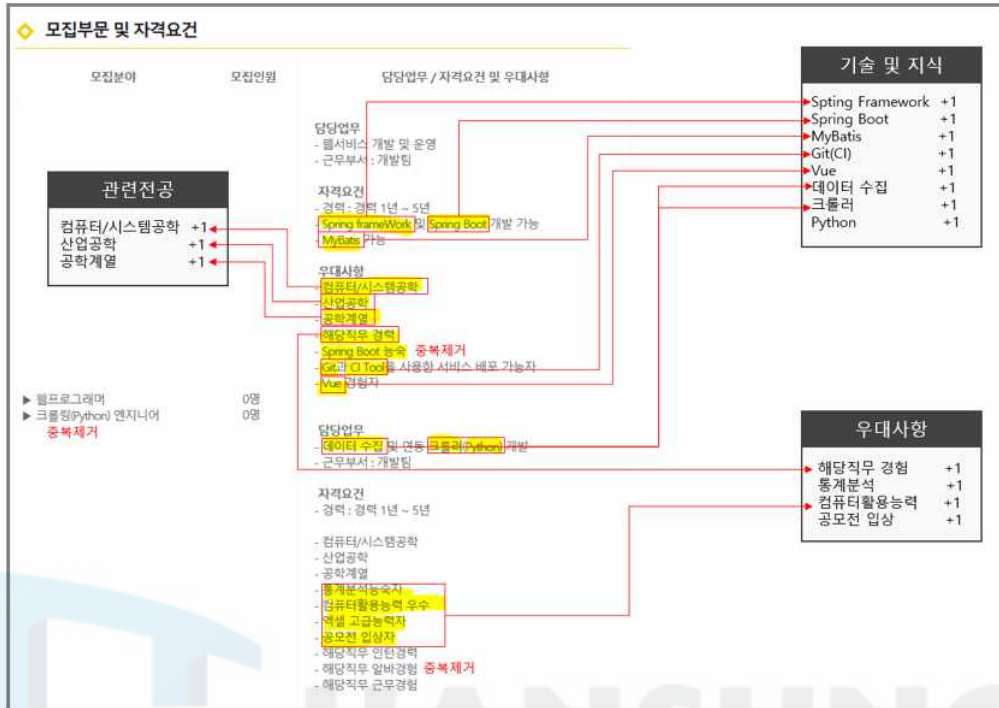


[그림3-5] 웹프로그래머 키워드 채용공고 리스트



[그림3-6] 채용공고 일반 항목

[그림 3-7]과 같이 일반적으로 모집 부분 및 자격요건에 기술 및 지식에 해당하는 항목, 관련 전공, 우대사항과 같은 기업이 필요로 하는 내용에 대한 데이터를 찾을 수 있으며, 이러한 데이터 추출을 위해 HTML 태그를 제거하고 채용정보 데이터 수집을 위해 필요한 텍스트를 분류하여 추출 후, 해당 항목(개체)을 요구하는 요구 기업의 총 숫자를 합산하여 관계(Edge)의 가중치 값으로 할당한다면 가중치 계산을 통해 실제 학생들에게 해당 역량의 중요도를 조언해 줄 수 있게 되어 학생들이 진로 목표 결정과 진로 개발을 위한 유용한 조언 정보로 활용할 수 있다.



[그림 3-7] 기업의 채용공고에서의 기업이 요구 항목 분류

민간 취업 정보 포털에서의 해당 직무를 위한 기업의 자격, 우대사항, 기술 항목을 수집하기 위한 고용데이터 수집기를 [그림 3-8]과 같이 설계하였으며, 고용정보데이터 수집기는 민간 취업 정보 포털에서 자료를 수집하여 API를 통한 일반정보 수집과 HTML 페이지를 수집하여 태그 제거, 텍스트 분류, 키워드 매칭(혹은 형태소 분류), 중복제거와 가중치 집계를 통해 모든 채용정보를 반복하여 수집/분석하여 그래프 데이터 기반의 지식 베이스(지식 그래프)를 만들기 위한 직업별 채용정보 자료를 생성한다.

수집기를 통해 기업들의 채용정보 HTML 페이지를 분석하여 해당 기업이 필요한 역량을 분석하면 직업별 CSV(Comma-Separated Values) 형식의 분석 데이터를 생성할 수 있으며, 이렇게 생성된 직업별 채용공고 데이터를 통해 학생은 특정 직업 및 직무를 채용하는 기업들의 활용 언어와 기술사항, 취업을 위해 준비해야 할 사항, 관련 자격증 및 성향과 같이 진로 목표 결정과 진로 개발을 위해 역량 요소들을 조언받을 수 있게 된다.



[그림 3-8] 고용정보데이터 수집기 구조

나) 워크넷 직업데이터 수집

기업의 채용정보를 효율적으로 수집하기 위하여 민간 채용정보 포털 통해 수집된 구인 정보는 직업과 직업 관련 기술 및 지식, 학과, 자격증, 우대사항과 같은 노드 정보와 여기에 연결된 관계와 관계에 연결된 역량(가중치(채용기업 수))이 있으므로 바로 그래프 데이터베이스에 저장하여 활용할 수 있다. 하지만, 노드나 관계의 추가 정보에 해당하는 속성에 해당하는 정보가 없으므로 관계에 대한 분석이 아닌, 해당 직업이나 기술에 대한 설명 등의 속성에 해당하는 정보가 부족하여 추가 정보 제공에는 한계가 있다. 실제 해당 직업에 대한 세부적인 조언을 위해서는 연봉, 전망, 설명 등의 데이터가 추가되어야 한다. 채용정보 수집 단계에서 직업사전 분류코드에 해당하는 키워드를 활용하여 수집된 채용정보의 경우 동일한 코드명을 가진 워크넷의 직업정보와의 관계 연결을 통한 지식의 확장으로, 직업에 대한 더 자세한 추가 정보의 제공과 함께 분석할 수 있는 관계의 범위를 더 확장할 수 있다.

워크넷의 직업정보 수집은 직업 사전 분류 코드의 정보기술에 해당하는 정보통신 분류를 기준으로 수집하였으며, [그림 3-9]와 같이 파라미터에 따라 정보 호출이 가능하며, 호출되는 내용은 [표 3-10]과 같다.

dtlGb	String	Y	직업정보 상세구분을 지정합니다. - 1 : 요약 - 2 : 하는 일 - 3 : 교육/자격/훈련 - 4 : 임금/직업만족도/전망 - 5 : 능력/지식/환경 - 6 : 성격/흥미/가치관 - 7 : 업무활동
-------	--------	---	---

[그림 3-9] 직업정보 API 요청 parameter

[표 3-10] 직업정보 API 추출 정보

Key	설명	수집 정보
1	요약	요약 <jobSum>
2	하는일	직업 코드<JobCd>, 직업명<JobNm>, 직무개요<JobSum>, 수행직무<execJob>
3	교육/자격/훈련	학력분포<edubg>, 전공학과분포<schDpt>, 관련자격명<relCertList>
4	임금/직업만족도/전망	비정형 문서에서의 중요 키워드 추출
5	능력/지식/환경	추출된 키워드 분석/설계를 통한 그래프 데이터베이스에 저장
6	성격/흥미/가치관	사용자 인터페이스를 통한 조언 요청
7	업무활동	조언 요청 내용 분석 (의도, 실제 분석)

[그림 3-10]은 실제 API를 통해 호출한 ‘웹프로그래머’, 직업 코드‘20727’에 대한 출력 결과이다.

요약과 하는 일의 경우 해당 직업 노드의 직업 속성 정보로 저장하여 학생의 질문에 대한 답변이 가능하도록 설계하며, 교육/자격/훈련, 임금/직업만족도/전망, 능력/지식/환경, 업무 활동의 경우 그래프 데이터베이스의 관계 분석을 활용하여 해당 직업뿐만이 아니라 관련 있는 모든 요소(역량 등) 간의 관계를 분석하여 학생들이 필요한 정보를 분석하여 제공해 줄 수 있다.

능력/지식/환경 구분에서는 업무 수행능력 중요도<jobAblStatus>, 업무수행능력 수준<jobAblLvlStatus>, 지식중요도<knwldgStatus>, 지식수준<knwldgLvlStatus>, 업무환경<jobEnvStatus>을 출력하며, 각각의 수준에 대한 중요도

는 [그림 3-11]과 같이 중요도(0:낮음 ~ 100:높음)로 구분된다. 각각의 노드에 대한 중요도는 관계의 가중치 데이터로 활용하여 해당 직업의 중요도에 대한 정보 제공과 함께, 학생의 개별 역량과의 데이터 매칭 시 중요도에 따른 추천 알고리즘을 통해 더 유용한 조언 정보 제공이 가능하다.

```
<jobSmlNm>웹프로그래머</jobSmlNm>
<jobSum>웹상에서 각종 자료들을 보여줄 수 있도록 웹 프로그래밍 언어를 이용하여 프로그램을 설계하고 작성한다.</jobSum>
<way>웹프로그래머로 일하려면 전문대학 및 대학교의 컴퓨터공학과, 전자학과, 인터넷(정보)공학과 등에서 웹 구축에 필요한 프로그래밍언어를 비롯해 서버, 네트워크, 데이터베이스 등에 대해 공부하거나 프로그래밍, 디자인 등으로 과정이 나누어져 있는 사설교육기관에서 강의를 이수해 이론과 기술을 익히는 것이 좋다. 웹 기획과 개발을 위해서는 웹 언어뿐만 아니라 SQL 등의 데이터베이스는 물론 홈페이지(나도웹메이커, 드림위버), 그래픽소프트웨어(포토샵, 일러스트레이터, 플래시, 플래시) 등의 지식과 능력을 갖추고 있으면 취업과 업무수행에 유리하다. 웹사이트를 구축하는 데 필요한 언어가 쉬웠기 때문에 사설 교육기관을 통해서도 입직할 수 있으며 채용 시 학력보다는 경력을 우선시하기 때문에 많은 실습을 통해 경험을 쌓는 것이 중요하다.</way>
<relMajorList>
  <majorCd>8</majorCd>
  <majorNm>공학교육과</majorNm>
</relMajorList>
<relMajorList>
  <majorCd>1</majorCd>
  <majorNm>교육학과</majorNm>
</relMajorList>
<relMajorList>
  <majorCd>23</majorCd>
  <majorNm>응용소프트웨어공학과</majorNm>
</relMajorList>
<relMajorList>
  <majorCd>22</majorCd>
  <majorNm>컴퓨터공학과</majorNm>
</relMajorList>
<sal>조사년도:2018년, 임금 하위(25%) 3241만원, 평균(50%) 3682만원, 상위(25%) 4836만원</sal>
<jobSatis>64.1</jobSatis>
<jobProspect>증가(42%) 현상유지(23%) 감소(35%)</jobProspect>
<jobAbil>기술 설계/건설/기술 분석/설치/학습능력</jobAbil>
<knowl>컴퓨터와 전자공학/디자인/영어/공학과 기술/산수와 수학</knowl>
<jobEnv>4차산업 도구, 기술 사용, 클라우드시스템/앞에서 근무/자동화 정도/갈등 상황/이메일 이용하기</jobEnv>
<jobChr>혁신/분석적 사고/신뢰성/책임성/공공함</jobChr>
<jobIntrst>탐구형(Investigative)/관습형(Conventional)</jobIntrst>
<jobVals>선 제활동/타인에 대한 영향/자율/다양성/성취</jobVals>
<jobActvImprtn>컴퓨터 업무/창조적 생각/정보의 의미 해석/새로운 지식의 습득, 활용/정보 작성, 기록</jobActvImprtn>
<jobActvLvl>컴퓨터 업무/창조적 생각/새로운 지식의 습득, 활용/목표, 전략 수립/정보의 의미 해석</jobActvLvl>
```

[그림 3-10] 직업 코드 ‘2072’에 대한 요약정보

<jobAbil>			
<jobAbilStatus>	Number	업무수행능력 중요도 : 중요도(0:낮음 ~ 100:높음)	</jobAbilStatus>
<jobAbilNm>	String	업무수행능력 중요도 : 업무수행능력	</jobAbilNm>
<jobAbilCont>	String	업무수행능력 중요도 : 설명	</jobAbilCont>
</jobAbil>			

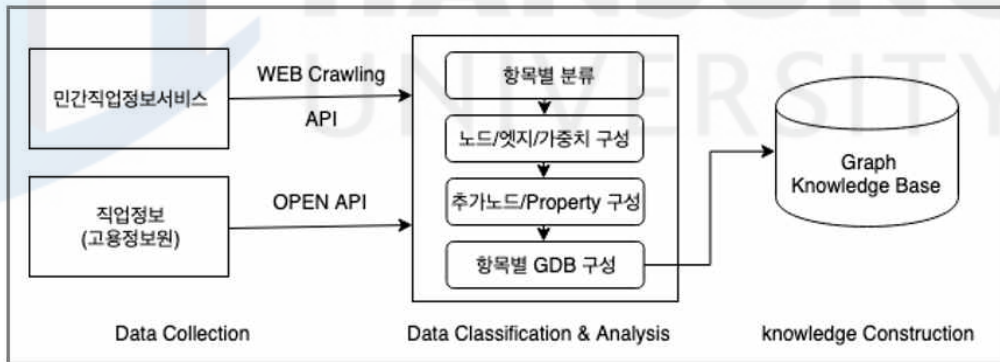
[그림 3-11] 능력/지식/환경 출력 결과

이렇게 수집되는 데이터는 [그림 3-12]와 같이 ‘웹 프로그래머’ 직업의 역량에 속하는 노드의 정보로 설계되며, 수행능력 및 역량의 점수는 역량과 노드 간의 가중치로 설정한다.



[그림 3-12] ‘웹 프로그래머’ 직업의 역량 노드 설계

[그림 3-13]과 같이 API 또는 크롤링을 통해 수집된 데이터는 항목별로 분류하여 노드, 관계, 가중치를 구성하고 관계되는 노드와 속성 정보들을 추가하여 진로 조언을 위한 지식 그래프를 생성한다.



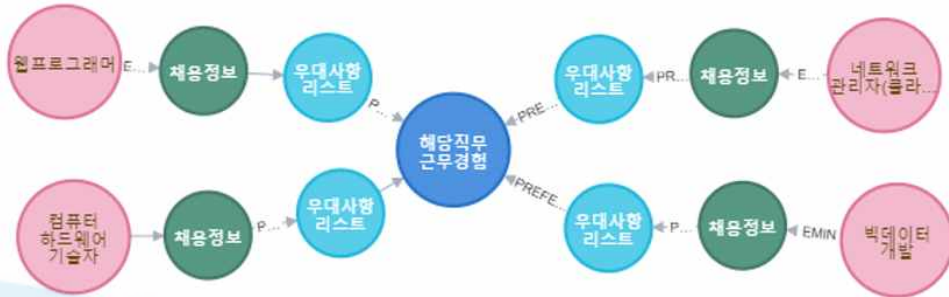
[그림 3-13] 데이터 수집 지식 그래프 생성 구조

4) 데이터 전처리

가) 데이터의 관계 설계

그래프 데이터베이스의 가장 큰 장점은 데이터들과의 관계 분석이다. 관계형 데이터베이스와 달리 데이터 간의 관계 구성이 설계 시 고려되어야 한다. [그림 3-14]와 같이 ‘해당 직무 근무경험’이라는 역량 노드는 ‘웹프로그래머’,

‘응용프로그램개발자’, ‘서버엔지니어’ 등의 대부분의 다른 직업의 채용 우대 사항에 포함된다. 따라서 모든 대부분의 역량 노드들은 해당 역량과의 관계를 가지는 직업과 관계가 연결되어야 하며, 이러한 관계를 활용하여 학생의 보유 역량과 매칭하여 학생에게 적합한 직업 추천이 가능하며, 가중치에 따른 적합도와 중요도를 계산할 수 있게 된다.



[그림 3-14] 해당 직무 근무경험에 대한 연관 관계 설계 예제

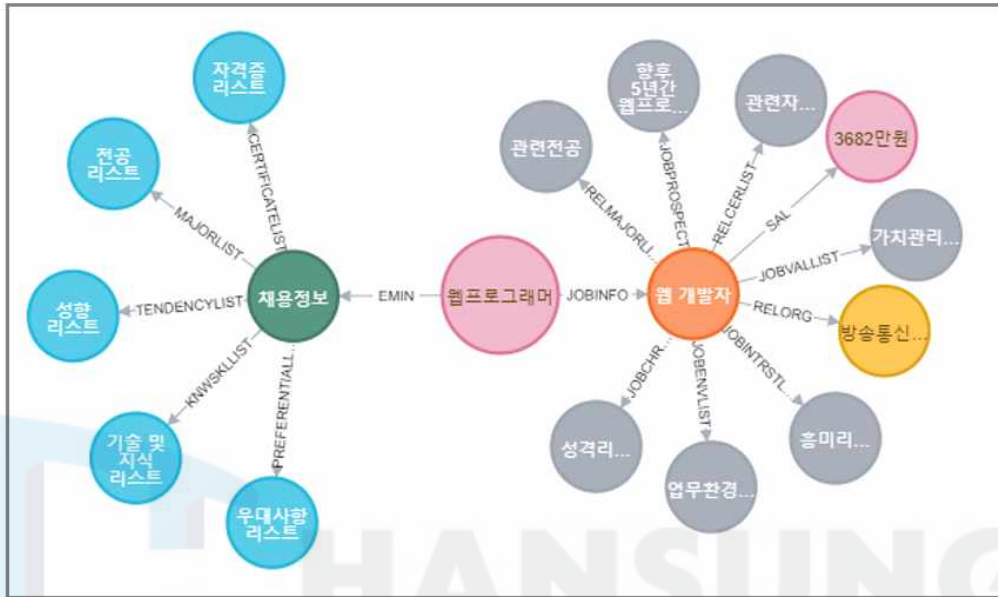
나) 그래프 노드의 설계

학생들의 진로 조언의 다양한 직업정보가 중요한 만큼 직업의 관계 노드들을 해당 직업에 연결하게 되면, 수십~수백 개의 관계 노드가 한 개의 노드와 연결되게 된다. 이러한 경우 쿼리 구조는 간단해질 수 있어도, 분류가 어려워지고 연결의 복잡성이 늘어남에 따라 시각화의 어려움과 쿼리 성능에도 영향을 미칠 수 있다. 때문에 [그림 3-15]와 같이 각 직업과 수집된 데이터 사이에 구분을 위한 분류 노드를 추가함으로써 구조의 간결성과 시각화의 이점을 모두 만족시킬 수 있게 된다.

다) 시나리오 테스트

시나리오 테스트는 [표 3-11]과 같이 설계된 노드와 관계를 기준으로 질문에 대한 답을 찾는 과정이다. 특정한 질문들에 대하여 가상의 쿼리문과 같은 형태의 답을 찾는 과정을 통해 설계의 문제점은 없는지, 노드의 방향에 오류는 없는지, 답을 찾는 과정에서 더 필요한 관계는 없는지를 점검하는 과정이라고 볼 수 있다. 테스트를 수행하기 위해서는 지식 베이스 생성 목적에 따라 질문을 준비하고 이에 대한 노드 및 관계를 연결하여, 이에 대한 점검을

수행한다. 이러한 시나리오 테스트는 후에 구현을 위한 설계에서 어떤 점을 고려하여 구현할 것인가에 대하여 사전 검토하고 설계의 문제점을 최소화할 수 있다.



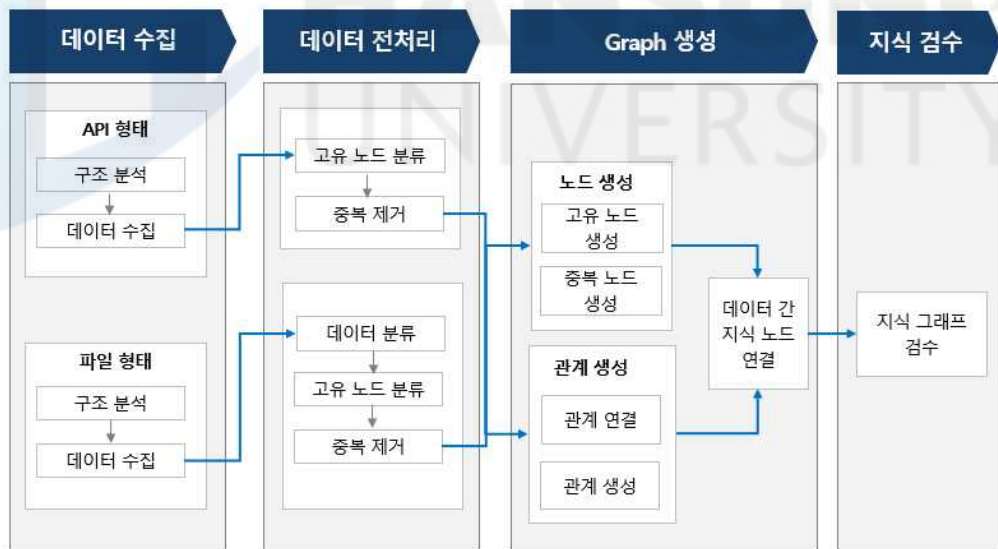
[그림 3-15] 지식 그래프 노드 구조 설계

[표 3-11] 시나리오 테스트

No	질문	답변 과정
1	웹프로그래머는 어떤 자격증이 있어야 해?	①자격증-자격증리스트-채용정보-직업 ②자격증-관련자격증-직업정보-직업
2	진취적인 성격에 어울리는 직업은 뭐야?	①진취적-성향리스트-채용정보-직업 ②진취적-성격리스트-직업정보-직업
3	중국어 잘하는 사람을 우대해 주는 직업은?	중국어-우대사항-채용정보-직업
4	Java 잘하는 사람에게 좋은 직장은?	JAVA-기술및지식리스트-채용정보-직업
5	컴퓨터공학 전공, 정보처리기사 자격증, 긍정적인 나에게 어울리는 직업은?	컴퓨터공학, 정보처리기사, 긍정적-리스트-채용정보-직업
6	웹프로그래머는 전망은?	웹프로그래머-직업정보-향후 전망-전망노드
7	웹프로그래머는 무슨 일을 해?	웹프로그래머-직업정보(속성정보)

5) 그래프 데이터베이스를 활용한 지식 그래프 구축 방법론

그래프 데이터베이스를 활용한 지식 그래프 설계 방법론에 따라 설계된 그래프는 [그림 3-16]의 지식 그래프 구축 방법론의 절차에 따라 실제 구축을 진행한다. 지식 베이스 설계와 달리 실제 지식 베이스를 구축의 경우 단지 구현의 영역이기 때문에 절차를 간과하기 쉽다. 그러나 관계형 데이터베이스와 달리 지식 그래프의 경우 노드의 생성뿐만이 아니라 노드 간의 관계들을 연결해야 하며, 이러한 관계 정보 분류와 연결을 위해 관계형 데이터베이스보다 복잡한 추가 절차가 필요하다. 지식 그래프의 구축을 위해서는 수집 자료의 분석, 데이터 수집, 고유노드 분류, 중복제거, 고유노드와 중복노드의 생성, 관계 연결, 관계 생성, 데이터 간 지식 노드 연결이라는 순서로 지식 베이스 생성이 진행되고, 최종적으로 구축된 지식 베이스를 지식 맵과 사이 퍼쿼리를 활용하여 검수하는 절차로 진행된다.



[그림 3-16] 그래프 데이터베이스를 활용한 지식 그래프 구축 방법론

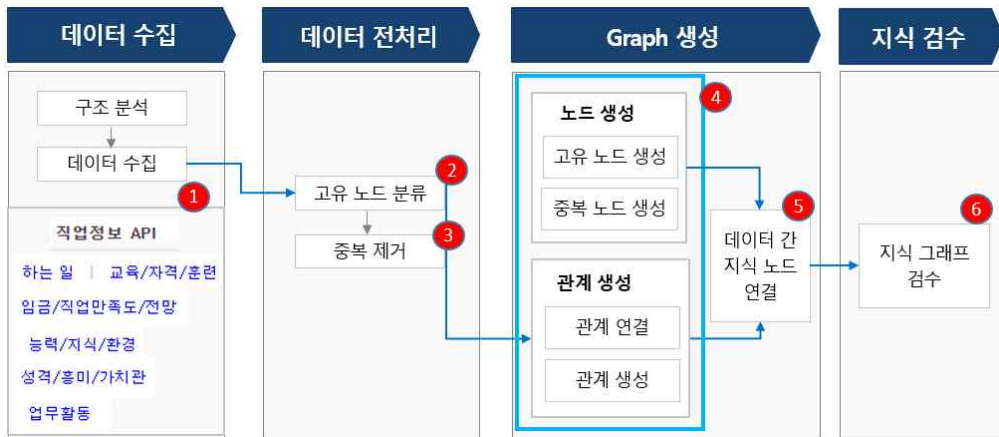
제시하는 지식 그래프 방법론은 실제 디렉터리 구조 API, 코드호출 API, CSV 파일 등 다양한 유형의 데이터를 수집하여 그래프 데이터베이스를 생성하는 과정에서의 가장 효율적인 방법을 제시한 것이며, 구축 프로세스가 정리

되지 않았던 초기의 설계 문제점과 실제 생성 과정에서 시행착오가 있었던 부분을 고려하여 설계와 구축 방법론을 제시하였으며, 실제 제시된 방법론을 활용한 연구 과정에서의 디렉터리 구조 API, 코드호출 API, CSV 파일 형태의 데이터를 활용한 지식 그래프 설계와 구축을 통해 활용성을 증명하였다.

가) 고용노동부 직업정보 API 수집을 통한 지식 그래프 생성

워크넷 직업정보 API에서는 하는 일, 교육/자격/훈련, 임금/직업만족도/전망, 능력/지식/환경, 성격/흥미/가치관, 업무 활동에 관한 6가지 분류의 직업 데이터를 받아 각각의 역량 데이터와 직업 간의 연관 관계를 분석하여 그래프 데이터베이스를 설계한다.

워크넷 직업정보 API는 846개의 한국 직업이 있지만, 연구를 위해서 정보통신 분류 43개와 기타 민간 채용정보 포털에서 수집한 46개를 합하여 89개의 정보기술 관련 직업의 정보를 수집하고 이를 분류하여 그래프 데이터베이스에 저장한다. 수집된 데이터는 연관 관계가 설정되어 있지 않기 때문에 그대로 저장할 경우 직업마다 중복되는 노드들이 생성되고, 같은 이름을 가진 다수의 중복노드로 인하여 관계의 분석이 매우 복잡하고 어려워진다. 따라서 중복되는 노드들을 분석하여 중복노드의 제거 및 노드 간의 관계를 연결해야 하며, [그림 3-17], [표 3-12]의 프로세스에 따라 직업정보 지식 그래프를 생성한다.



[그림 3-17] 워크넷 직업정보 지식 그래프 구축

[표 3-12] 워크넷 직업정보 지식 그래프 구축단계별 설명

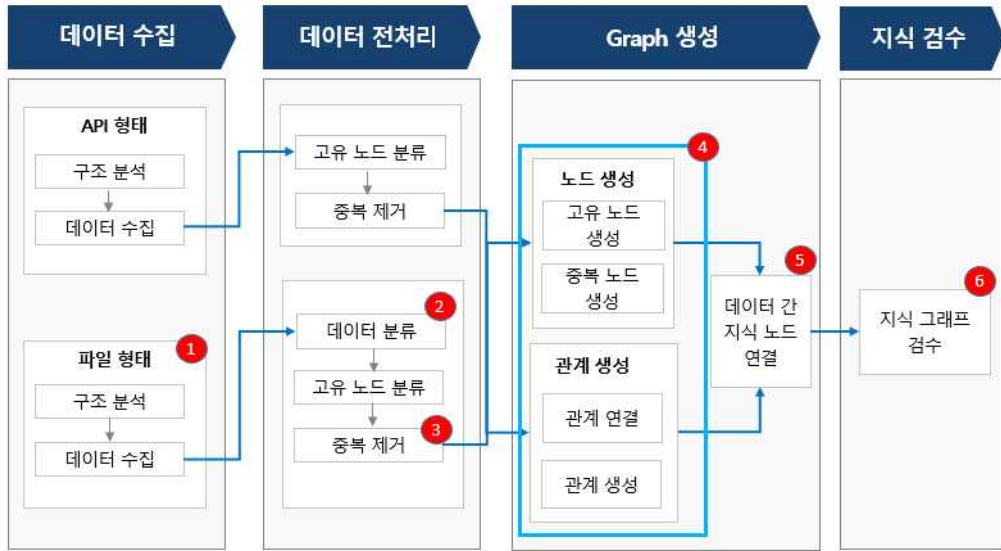
생성단계	설명
1	수집 직업 리스트를 한 줄씩 반복하여 직업에 대한 6가지 검색 타입에 의한 API 정보를 수집한다
2	직업별로 고유노드와 중복 가능한 노드들과의 관계를 추가할 수 있는 노드로 분류한다.(분류노드 추가 구성) (중복가능 노드: 관련 전공, 관련 자격증 등의 직업 역량 노드)
3	중복가능 노드들은 중복제거 후 합하여 노드를 생성한다.
4	각 직업에 대한 최상위 노드 생성과 그와 연결된 하위노드 및 관계를 순서대로 연결한다.
5	4번의 생성이 끝난 후, 중복가능 노드들과의 관계를 연결한다.
6	지식 그래프를 지식 맵과 사이퍼쿼리를 활용하여 검수한다.

나) 민간 채용정보 지식 그래프 생성

CSV 형태의 데이터는 API 형태의 데이터와 달리 데이터가 명확하게 분류되지 않은 경우가 많으므로 먼저 데이터에 대한 분류 작업이 필요하다. 민간 채용정보 데이터의 경우 직업에 관련된 역량 데이터를 수집했기 때문에 [그림 3-18], [표 3-13]과 같이 데이터 분류, 고유노드 분류, 중복제거 절차를 거쳐 채용정보 지식 그래프를 생성한다.

[표 3-13] 채용정보 지식 그래프 구축 단계별 설명

생성 단계	설명
1	· 직업별 파일을 하나씩 오픈
2	· 파일형 데이터의 문자열 분리를 통해 배열을 구성하고 직업별 하나의 리스트에 통합
3	· 파일의 중복가능 노드들에 대한 중복제거 작업을 수행
4	· 각 직업에 대해서 최상위 노드 생성 및 그와 연결된 하위노드 및 관계를 순서대로 생성
5	· 생성된 노드들과 중복가능 노드들과의 관계 연결
6	· 생성된 지식 그래프에 대한 지식 맵과 쿼리 활용 검수



[그림 3-18] 민간 채용정보의 데이터의 지식 그래프 구축

다) 지식의 확장 및 연결

연구 수행을 위한 데이터는 고용정보원의 직업정보 API 데이터와 직접 민간채용정보 포털에서 수집한 직업별 채용정보 데이터이며, 서로 다른 구조와 유형을 하고 있다. 지식 그래프에서는 서로의 관계가 연결되어야 지식 그래프의 가치가 높아지며 실제 조언할 수 있는 지식의 깊이도 더 증가한다. 한국고용정보원의 워크넷 직업정보 API는 직업 사전 분류 코드에 해당하는 직업에 대한 정보만 가지고 있으므로 수집된 채용정보의 전체 노드를 연결할 수는 없지만, 연결이 가능한 유사한 형태의 직업에 대해서는 관계를 연결하여 지식의 영역을 확장할 필요가 있다.

웹프로그래머를 부르는 이름도 ‘웹개발자’, ‘웹프로그래머’, ‘web프로그래머’와 같이 상이하므로 직업 이름을 기준으로 연결하는 것은 적절하지 않다. 고용정보원의 직업정보 API는 직업분류 코드별로 API를 호출하여 해당 직업에 대한 정보를 제공하는 방식으로 구성되어 있으므로 채용정보 데이터에 동일한 직업 코드를 추가해 준다면 두 개의 지식 간 연결이 가능하다.

지식의 연결을 위해 동일한 직업이나 유사성이 매우 높은 직업의 경우 직업명으로 저장된 파일명의 이름을 매칭되는 직업 코드로 변환하였으며, 실제 CSV 파일을 읽어 들일 때 해당 직업 코드를 확인하여 직업 코드가 있는 경

우 [그림 3-19]와 같이 지식을 서로 간의 지식을 연결하도록 설계하였다.

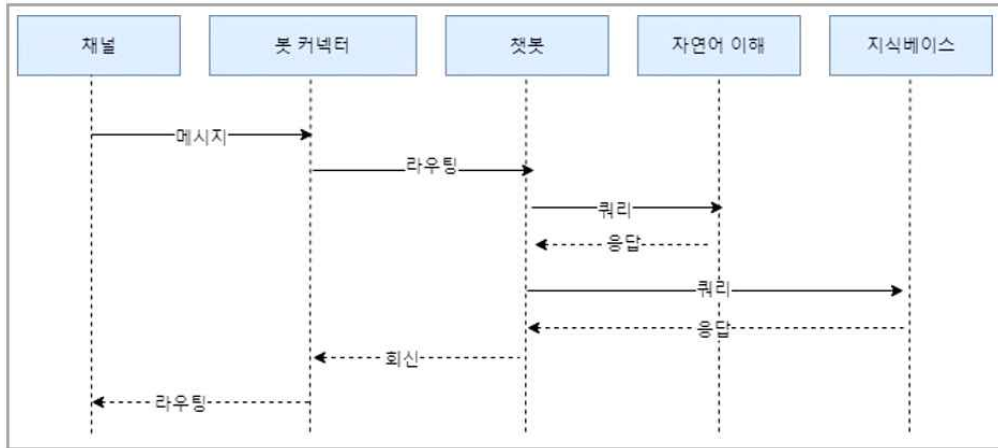


[그림 3-19] 디지털영상처리 전문가(직업 코드 20233) 지식 연결 형태

지식의 연결은 지식 그래프의 활용성을 크게 높일 수 있다. [그림3-19]와 같이 연결된 지식에 대학 졸업생 중에 디지털영상처리 전문가로 진로를 정하거나 취업을 한 선배의 역량 등의 지식을 추가하여 조언해 줄 수 있는 지식을 확장할 수 있으며, 디지털영상처리 전문가의 직업 훈련과정이나, 채용정보와 연결하여 역량 획득을 지원하거나, 실제 채용공고와 연결하여 디지털영상처리 전문가를 채용하는 기업과 학생의 역량을 기반으로 맞춤형 취업 추천도 가능하다. 이러한 지식의 확장을 통해 진로 조언 플랫폼은 더 다양하고 깊은 단계에서의 진로 조언과 학생들에게 더 다양한 정보 제공이 가능하다.

6) 챗봇 인터페이스

[그림 3-20]과 같이 챗봇 인터페이스는 채널<->봇커넥터<->챗봇의 구조로 구성된다. 웹 또는 봇-애플리케이션을 통해 사용자가 입력한 메시지는 봇-커넥터(Bot-Connector)를 통해 챗봇으로 전달된다. 전달된 사용자의 메시지를 자연어 이해 과정을 통해 정확한 의도와 실체를 분석 후 지식 베이스에 정보를 찾아 챗봇에서 사용자에게 이해가 쉬운 형태로 가공하여 봇-커넥터(Bot-Connector)를 거쳐 사용자에게 답변할 수 있도록 구성한다.



[그림 3-20] 챗봇 인터페이스 구조

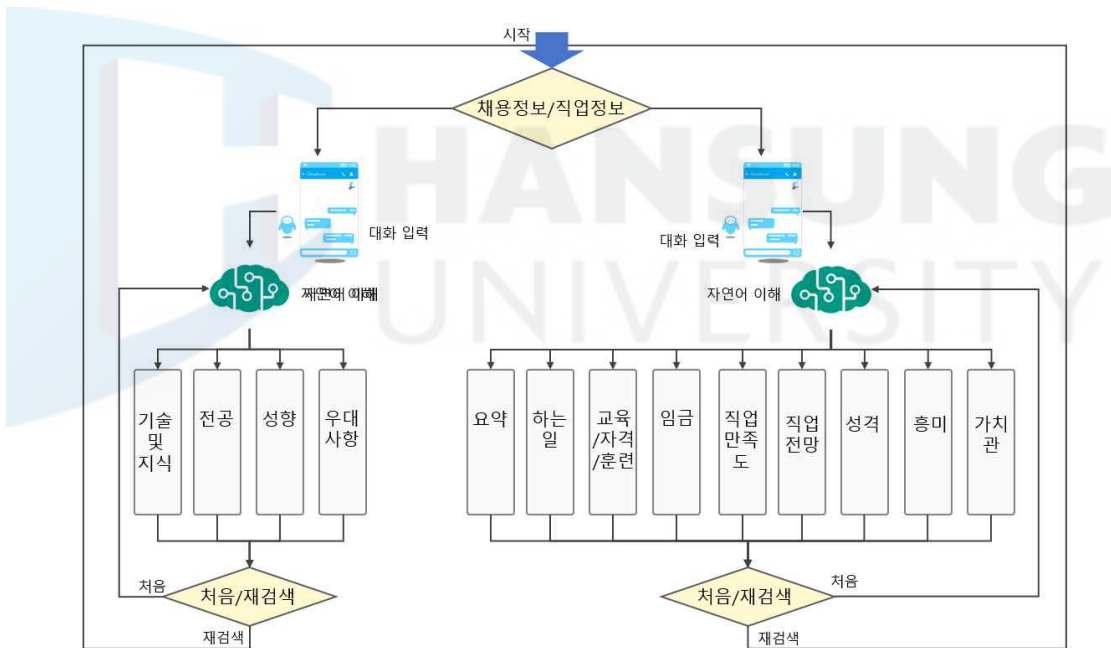
연구에서의 챗봇은 학생들이 익숙한 형태의 인터페이스를 활용하여 직업에 대한 다양한 정보를 질의하고 질의에 대한 정보 제공과 조언을 위한 인터페이스 역할을 제공하고 있으며, 국내외의 다양한 챗봇 프레임워크 중에서 Azure Cloud 기반의 연구 환경에 최적화된 오픈소스인 마이크로소프트의 Bot-Framework를 활용하여 구현하며, 자연어의 이해는 Azure Cognitive 서비스인 LUIS를 활용하여 발화 의도와 실체(엔티티)를 학습시켜 자연어처리 환경을 구현한다.

형식이 없는 대화에서 문제시되는 것은 명확하지 않은 질문이다. 실제 직업에 대한 질문에서 직업명이 없이 기술이나 자격증과 같은 역량 정보를 문의한다면 명확한 정보를 제공하기 위해서 챗봇은 원하는 직업명을 문의하는 프로세스를 추가하여 명확한 정보를 제공해야 한다. 또한, 직업에 대한 명확한 질문 “웹프로그래머 직업을 위한 자격증 알려줘”라는 질문이라면 LUIS에서 정확한 의도와 엔티티 결과를 받아 정보를 추출하기 위한 쿼리문을 지식 그래프에 매칭하여 정보를 수신하고 이를 재가공하여 어떤 자격증이 웹프로그래머 직업에서 어떤 중요도를 지니는지에 대해 답변을 제공해야 한다.

워크넷 직업정보 API에서 수집한 역량과 취업 정보 포털에서 비정형 데이터인 채용정보를 수집하여 얻은 역량에는 동일한 역량이 존재한다. 직업정보 API에 포함된 항목인 자격증, 기술, 전공, 흥미와 같은 역량과 채용정보에서

수집한 전공, 성향, 자격증, 기술 및 지식 항목 중 비슷한 내용은 자연어 분석을 활용하여 사용자의 정확한 의도를 파악하는 것은 매우 난해하며, 비슷한 의도의 다른 표현이라 오히려 자연어 이해 능력을 떨어뜨릴 수 있다.

중복되지만 다른 정보의 경우 이용자에 따라 중요도가 차이 날 수밖에 없다. 때문에 [그림 3-21]과 같이 이용자가 문의 사항을 입력하기에 앞서 채용 정보를 선택할 것인지, 직업정보를 선택할 것인지를 직접 선택 후 문의 사항을 입력하도록 질문 프로세스를 설계하였다. 입력된 문의 사항은 자연어처리 서비스인 LUIS를 통해 의도와 실체(엔티티) 평가 정보를 수신하고, 이 정보를 기준으로 해당 노드 및 관계 정보를 지식 그래프에서 쿼리하여 이용자에게 제공하도록 구성하였으며, 각각의 단계별 기능은 [표 3-14]와 같다.



[그림 3-21] 대화 FLOW 설계

[표 3-14] 대화 Flow 단계별 동작

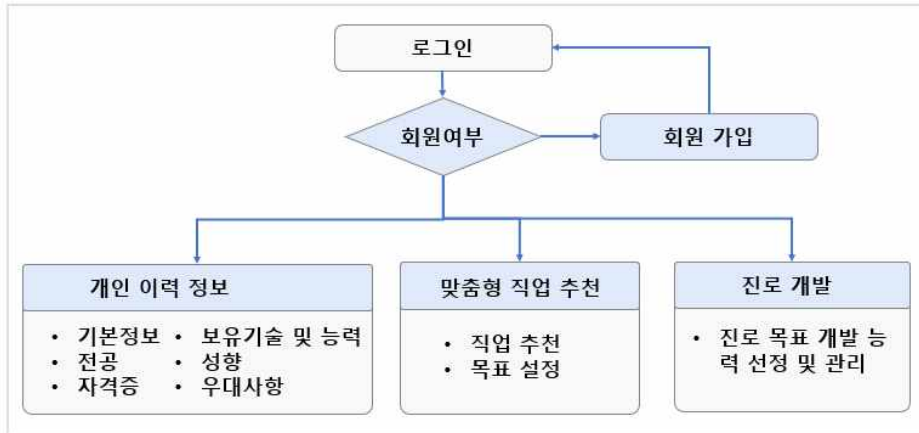
No	단계	동작
1	사용자 질의 입력	사용자 질의를 입력
2	질문 유형 선택	채용정보와 직업정보 중 조언받기를 원하는 정보 선택
3	사용자 대화 입력	사용자의 문의 사항 또는 관심 사항에 대한 자연어 입력
4	LUIS(자연어 이해)	자연어 이해를 통해 발화 의도와 실체(엔티티) 평가
5	발화 의도와 실체 확인	발화 의도와 실체(엔티티)가 명확한지 확인하고 명확하지 않은 경우 챗봇은 추가 질문을 통해 발화 의도와 실체(핵심어) 판별
6	의도와 실체에 따른 분류	질문별, 질의어와 가중치 계산이 상이하므로 발화 의도에 따른 대화 프로세스로 전달
7	그래프 쿼리 생성	지식 그래프에 의도와 실체를 매칭하여 결과 값 호출
9	가중치 계산	역량 관계에서의 가중치 추출 및 계산

7) 모바일 e-포트폴리오 앱

스마트폰의 보급이 늘어남에 따라 학생들에게 가장 친숙하고 즉시 이용이 가능한 서비스의 형태는 모바일 앱 서비스이다. 이에 따라 대학들의 학생지원 서비스 역시 PC를 기반으로 설계한 웹브라우저 방식에서 모바일 앱 서비스를 추가 지원하고 있다.

학생들에게 맞춤형 정보를 제공해 주기 위해서는 개별 학생의 역량 정보가 필요하지만, 이러한 데이터들의 경우 「개인정보 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」에 의한 비식별화 등의 절차가 필요하므로 실제 학생의 데이터를 사용하여 연구를 진행하는 것은 무리가 있다.

학생에게 역량에 기반한 맞춤형 직업 추천 정보와 학생이 선택한 진로 목표관리 테스트를 수행하기 위해 챗봇과 함께 e-포트폴리오 앱을 추가 개발하여 조언 플랫폼으로서의 챗봇의 장점과 학생 개인의 역량 정보 관리, 진로 개발, 목표관리, 챗봇 조언 기능을 통합한 모바일 e-포트폴리오 앱을 [그림 3-22]와 같이 추가 구현하였다.



[그림 3-22] 진로 조언 e-포트폴리오 기능

e-포트폴리오의 기능은 [그림 3-23]과 같이 맞춤형 진로 조언을 위한 ‘개인 이력 정보’, ‘진로정보’, ‘조언 서비스(챗봇)’, ‘목표관리’로 구성된다.

The screenshot displays the e-portfolio system interface with several key sections:

- 로그인 (Login):** Includes fields for ID and password, and buttons for login, sign up, and social media logins (Facebook, Google).
- 개인 이력 정보 관리 (Personal Information Management):** A form for entering personal details like name, age, gender, major, and email.
- 경력 사항 (Career History):** A section for listing work experience with fields for company name, position, and dates.
- 주제별 역량 (Thematic Competency):** A section for managing competencies related to specific themes, with a list of competencies and a '자세히 보기' (View Details) button.
- 추천직업 (Recommended Job):** A section for managing recommended jobs, with a list of jobs and a '자세히 보기' (View Details) button.
- 목표직업: 웹프로그래머 (Target Job: Web Programmer):** A section for setting career goals, including a table for skill levels and a '직업 목표 설정' (Set Career Goal) button.
- 목표관리 (Goal Management):** A section for managing goals, with a list of goals and a '목표 관리' (Goal Management) button.

[그림 3-23] 맞춤형 직업 추천 및 진로 개발 프로세스

‘개인 이력 정보’의 경우 기업들의 고용정보와 워크넷 직업정보와 매칭되는 정보 위주로 학력, 경력, 보유 기술 및 능력, 자격증 항목으로 설계하였으며 보유 기술 및 능력에 개인별 현재 등급을 설정하도록 구성하여 지식 그래프의 관계 노드의 가중치와 대입하여 개인에게 적합한 정보를 제공할 수 있도록 설계하였다.

가) 데이터와 노드 매칭을 통한 정보 검색

개인화된 데이터를 활용한 맞춤형 직업 추천의 경우 데이터에 해당하는 노드와 필요한 관계만으로 지식 그래프의 정보를 쿼리할 수 있다.

[표 3-15]의 경우 사용자의 데이터와 노드의 매칭을 통한 정보의 추출 쿼리이다. 하지만 이러한 쿼리의 경우 단순히 매칭되는 정보만 제공이 가능한 한계점이 있으므로 효율적인 진로 조언을 위해서는 노드의 가중치를 활용하여 최적화된 점수를 계산하는 방식이 추가로 검토해야 한다.

[표 3-15] 그래프 매칭 관계 구성

[그래프 쿼리]

```
// 직업, 기술 관계에서의 직업 문의
(직업)-[:관계]-(기술) Return 직업

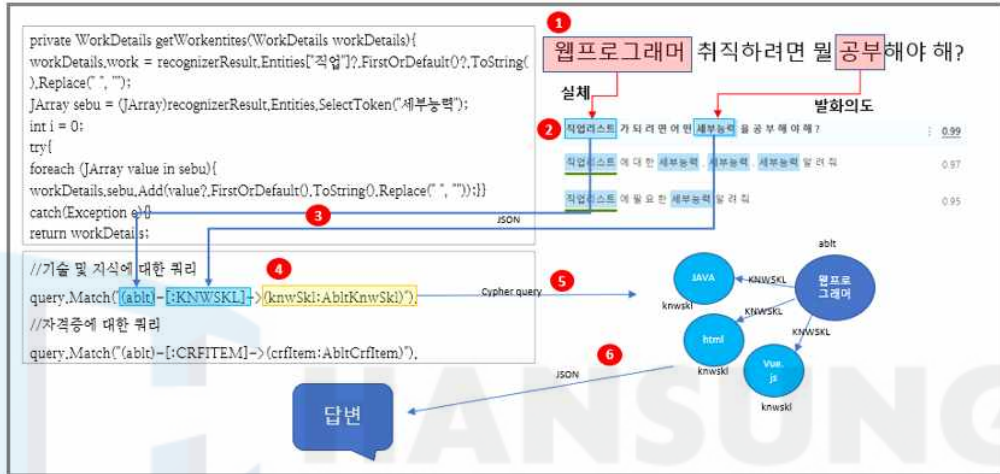
// 여러개의 관계와 노드로 구성된 구조에서의 관계 2 문의
(노드)-[:관계1]-(노드)-[:관계2]-(노드) Return 관계2

//직업의 채용정보에 있는 리스트 중 기술및지식 목록과 가중치 문의
(직업)-[:채용정보]-(기술및지식 리스트)-[가중치:기술및지식 관계]-(기술
및지식 노드) Return 기술및지식, 기술및지식 관계 가중치
```

나) 지식 그래프에서의 자연어 대화 답변 프로세스

[그림 3-24]는 사용자가 챗봇을 통해 질문한 “웹프로그래머 취직하려면 뭘 공부해야 해?”라는 문의에 대한 응답 과정에 대한 구현 시나리오이며, [표 3-16]은 사용자의 챗봇 질문에 대한 그래프 데이터베이스에서의 대화에 대한 답변 절차로, “웹프로그래머”는 직업으로 학습되어 직업 노드를 “웹프로그래

며”로 지정하고, “공부”는 이에 해당하는 기술및지식으로 평가되어 관계 노드로 분류되었으며, 각각의 웹프로그래머(직업)과 관계(기술및지식)에 해당하는 노드 정보를 지식 그래프에 쿼리하여 수신된 정보를 사용자에게 제공하게 된다. 이처럼 지식 그래프는 기존의 관계형 데이터베이스로 구축한 지식 베이스와는 달리 관계 분석이 필요한 다양한 서비스에 활용될 수 있다.



[그림 3-24] 지식 그래프에서의 답변 프로세스

[표 3-16] 대화 답변 절차

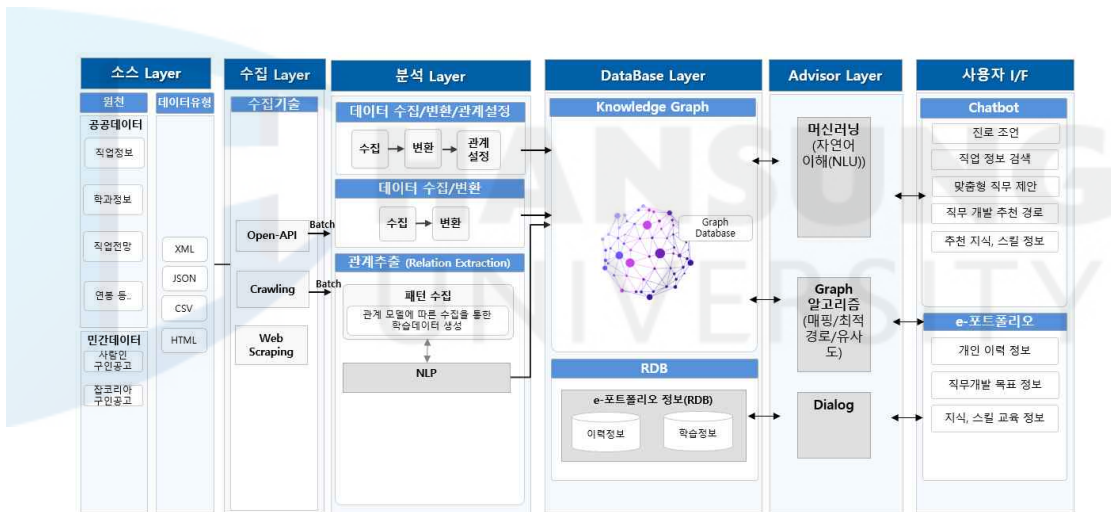
No	단계	동작
1	질문에서의 형태소 분류	질문에서의 형태소를 분류한다.
2	질문의 의도와 실체 분류	질문에서 의도와 실체의 분류 (의도와 실체는 별도의 학습 진행)
3	직업이름 매칭	직업으로 판정된 형태소에 대해 직업에 해당하는 노드 매칭
4	기술 및 지식 관계 매칭	사용자의 질문은 공부라는 문의이기 때문에 지식 및 지식 관계에 매칭
5	지식 그래프 쿼리	(웹프로그래머)-[:기술및지식]-(?) 쿼리
6	지식 그래프에서 관계 노드 답변	(웹프로그래머)-[:기술및지식]-(?) 에 해당하는 모든 노드 답변 관계의 가중치를 활용하여 답변의 품질 향상 가능

제 4 장 진로 조언 플랫폼 구현 및 테스트

제 1 절 진로 조언 플랫폼 구현 범위

1) 진로 조언 플랫폼 구현

제3장에서 설계된 진로 조언 플랫폼 모델 구현을 위한 플랫폼 구조는 [그림 4-1]과 같이 소스 Layer, 수집 Layer, 분석 Layer, DB Layer, 조언 Layer, 사용자 인터페이스와 같이 6개의 Layer로 구분하였으며, 각 Layer 별 기능 설명은 [표 4-1]과 같다.



[그림 4-1] 진로 조언 플랫폼 구조

[표 4-1] 진로 조언 플랫폼 Layer 별 기능

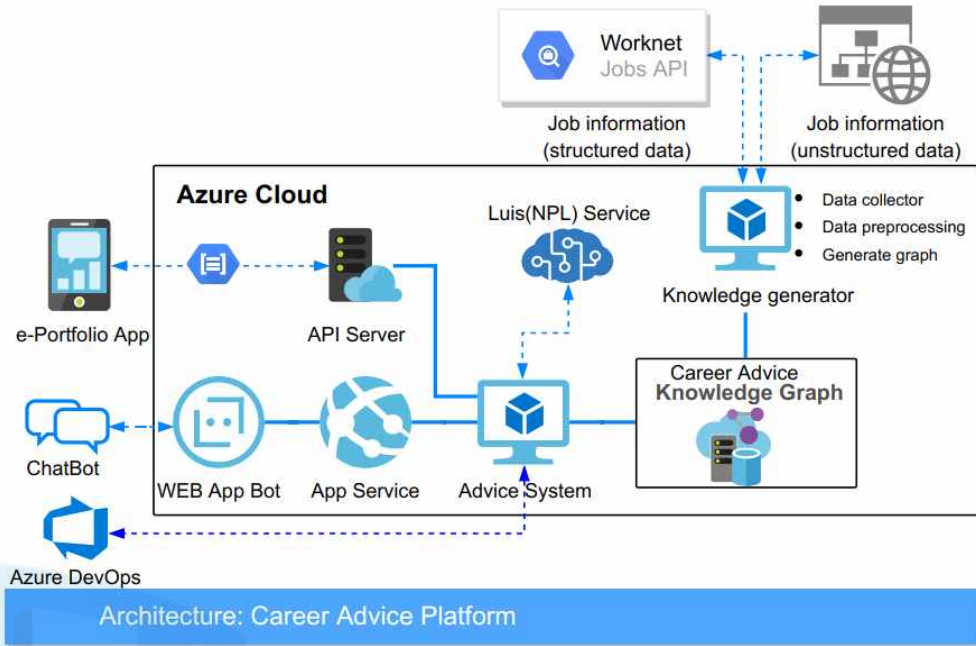
No	단계	동작
소스 Layer	원천 데이터	· 공공데이터: 워크넷 직업정보 API · 민간데이터: 채용정보 포털의 채용정보
	데이터 유형	· 공공데이터 : JSON, XML · 민간데이터: CSV, HTML
수집 Layer	OpenAPI	· 공공데이터 표준 API

	Crawling	· 주기적인 공공, 민간 데이터 수집
	Scraping	· 특정 이벤트에 따른 데이터 수집 진행
분석 Layer	변환	· 수집된 데이터를 분석하여 노드와 관계 요소 분류
	관계 설정	· 직업과 직업 역량 노드에 대한 관계 분류
	관계 추출	· 비정형 문서에서의 노드와 관계 추출
데이터베이스 Layer	지식 그래프	· 그래프 데이터베이스 환경의 지식 그래프
	관계형 DB	· 관계형 데이터베이스 환경의 지식 베이스
조언 Layer	자연어처리	· 사용자 대화에서 의도 분석
	그래프 알고리즘	· 지식 베이스에서 최적의 직업정보 및 조언 정보를 검색
	대화	· 직업 조언 대화 흐름 개발
사용자 인터페이스	챗봇	· 대화형 인터페이스를 활용하여 학생의 진로 관련 질문들에 대한 상담 및 조언
	모바일 앱	· 개인의 이력 정보와 역량 기반 직업 추천 및 목표 설정을 위한 앱 서비스

2) 진로 조언 플랫폼 인프라 구성

가) 플랫폼 인프라 클라우드(하드웨어) 구성

효율적인 연구 수행을 위하여 연구 수행 플랫폼은 [표 4-2]와 같이 Microsoft의 클라우드 인프라(IaaS, infrastructure As A Service) 서비스인 Azure 환경에서 구성하였으며, 데이터 수집 및 처리, API 시스템은 Azure App Service를 그래프 지식 베이스 구축은 Azure Container, 챗봇은 Web-App-Bot을 서비스를 활용하여 플랫폼 인프라 환경을 구성하였다. 클라우드 기반의 연구 환경은 인프라의 빠른 구축과 배포를 할 수 있으며, 인프라 자원을 연구 단계에 따라 효율적으로 활용할 수 있는 장점이 있다.



[그림 4-2] 진로 조언 플랫폼 연구 클라우드 아키텍처

(1) 지식 생성기(Knowledge Generator)

고용노동부 워크넷 오픈 API에서 수집되는 직업·직무 정보인 정형데이터와 취업 정보 포털에서 수집된 비정형데이터를 수집하며, 수집된 데이터의 분류·그래프 모델링, 그래프 생성 기능을 수행한다.

조언을 위한 데이터를 수집하여 지식을 만드는 과정을 수행하는 지식생성기 역할로 설정된 직업정보를 지정한 규칙에 따라 수집·전처리하여 지식 그래프 내 정보에 대한 업데이트를 수행으로써 조언 정보의 품질을 높이기 위한 시스템이다.

(2) 진로 조언 지식 그래프

Neo4j 그래프 데이터베이스로 구성된 지식 그래프이며, 진로 조언을 위한 직업과 직업에 관한 기술, 우대사항, 자격증, 성향과 같은 역량 정보와 연봉, 만족도와 같은 직업정보들과의 관계가 저장되어 있다.

지식 그래프는 챗봇, e-포트폴리오와 같은 사용자 인터페이스와 연동되어

쿼리문을 통해 요청한 관계에 대한 정보를 제공한다.

(3) 조언 시스템

챗봇, e-포트폴리오 앱을 통해 사용자가 질의한 진로 및 직업정보를 제공하고, 사용자 정보를 지식 그래프에 매칭하여 정보를 추천하는 기능을 가진 시스템으로 사용자의 대화를 자연어 이해 서비스를 통해 의도와 실체로 분류하고 지식 그래프와의 매칭을 통해 진로 조언 기능을 수행하는 시스템이다.

(4) 챗봇 (Bot-Framework)

학생들에게 가장 친숙한 인터페이스인 인터페이스를 활용하여 시간과 장소와 관계없이 간편하게 진로정보를 문의하고, 조언을 받기 위한 인터페이스로, 마이크로소프트 Azure 환경의 Bot-Framework와 Azure 환경의 자연어처리 시스템인 LUIS를 활용하여 시스템을 구성하였다.

(5) e-포트폴리오 App

학생의 상태 및 역량 정보를 담고 있는 포트폴리오로 자신이 쌓은 역량을 등록하고 등록한 역량 정보를 기준으로 관계가 높은 순으로 직업을 추천해 줄 수 있으며, 챗봇과 연결된 대화형 조언 기능, 역량 개발 및 관리 기능을 제공한다.

(6) LUIS (자연어처리 서비스)

대화형 진로 조언 기능 구현을 위해 사용자의 질문 의도와 엔티티를 찾아내기 위한 자연어처리서비스로 각 진로 관련 다양한 데이터들을 의도와 실체(엔티티)로 나누어 학습을 수행한다. 학습된 모델은 조언 시스템에서 활용된다. 사용자의 질문을 의도와 실체로 적중률을 평가하여 분류한다.

(7) Azure DevOps

연구 프로젝트에 대한 칸반(Kanban) 기반 관리와 애자일 프로세스를 활용하여 진행 절차를 관리하고, 생성된 소스의 형상관리와 문서관리를 지원하기 위한 시스템이다.

나) 개발 환경 및 소프트웨어 구조

[표 4-2]는 진로 조언 플랫폼 연구를 위한 소프트웨어 및 개발 환경으로 플랫폼의 범위가 큰 관계로 각각의 영역에 적합한 인프라 구조와 개발 언어를 활용하여 클라우드 서비스에 최적화된 개발 환경과 소프트웨어 구조 설계를 통해 연구를 진행하였다.

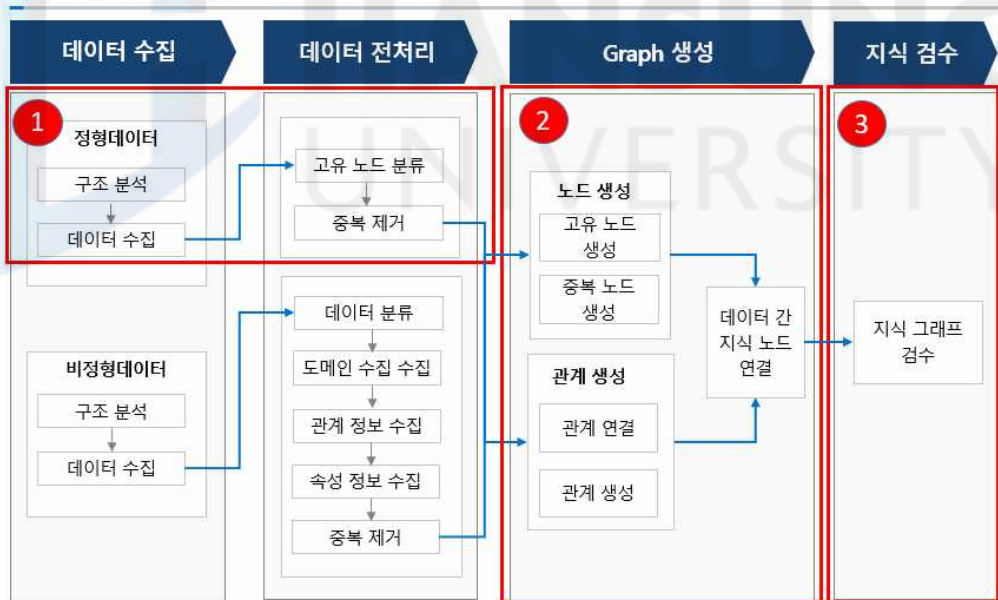
[표 4-2] 진로 조언플랫폼 하드웨어 및 소프트웨어 구조

구분	내용
민간 데이터 수집	<ul style="list-style-type: none"> · Python · windows
공공 데이터 수집	<ul style="list-style-type: none"> · Stack : .NET · .NET Framework version : V.47 · Windows
데이터 분석	<ul style="list-style-type: none"> · Stack : .NET · Azure Web App
Graph Database (직업정보)	<ul style="list-style-type: none"> · Azure Container Service <ul style="list-style-type: none"> - Properties : 1 CPU cores,1.5 Gib Memory - OS Type : Linux - Container image : neo4j:latest
Graph Database (직무정보)	<ul style="list-style-type: none"> · Azure Container Service <ul style="list-style-type: none"> - Properties : 1 CPU cores,1.5 Gib Memory - OS Type : Linux - Container image : neo4j:latest
e-포트폴리오	<ul style="list-style-type: none"> · FireBase-Spark · Flutter Framework
API 서버	<ul style="list-style-type: none"> · Azure App Service · 개발 언어 :C#
Bot Service	<ul style="list-style-type: none"> · Web App bot · 개발 언어 :C#
Advice System	<ul style="list-style-type: none"> · Azure App Service · 개발 언어 :Python

제 2 절 진로 조언 지식 그래프 구축

1) 직무데이터(디렉터리 구조 API) 지식 그래프 구축

워크넷 직무 사전 API는 정형데이터 형태로 직업정보를 제공하기 때문에 [그림 4-3]과 같이 ① 각 직업 코드별 직업정보를 수집하여 중복 데이터를 제거하기 위한 전처리 작업을 수행하는 부분 ② 인덱스와 직업의 기술, 전공과 같은 중복 가능한 직업 역량 노드를 생성하고, 각 직업에 해당하는 고유노드와 중복노드를 생성한다. ③ 각 관계의 연결과 관계의 생성, 데이터 간 지식 노드 연결을 통해 지식 그래프가 생성되고 최종 생성된 그래프 구조를 시각적으로 확인하여 잘못된 연결이나 혹은 관계가 연결되지 않은 노드가 있는지 확인한다.



[그림 4-3] 지식 그래프 구축방법론의 구현 단계

가) 직무데이터 API의 자료수집 및 전처리

지식 그래프 구축방법론에 따라 [그림 4-4]의 절차에 따라 데이터 수집

및 전처리를 수행한다.

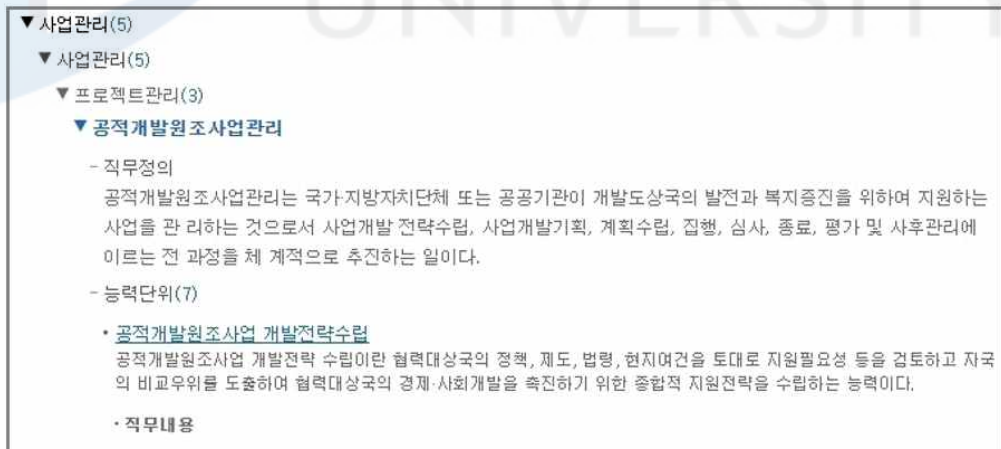


[그림 4-4] 데이터 수집 및 데이터 전처리 수행

(1) 구조분석

워크넷 직무데이터 API의 분류체제는 [그림 4-5]와 같이 대분류→중분류→소분류→세분류→능력단위로 이어지는 디렉터리 구조로 되어있다.

따라서 대분류를 수집 후, 대분류 코드를 기준으로 중분류, 중분류 내의 소분류, 소분류 내의 세분류, 세분류 내의 능력 단위로 반복하여 수집해야 한다.



[그림 4-5] 워크넷 직무데이터 구조

(2) 데이터 수집

[표 4-3]과 같이 각 분류별 NcsCD(NCS 코드 번호)를 기준으로 대분류,

중분류, 소분류, 세분류, 능력 단위의 데이터를 수집하여, 해당하는 분류에서는 각각의 분류 별 역량 정보들을 수집하여, Data 변수에 child 형태로 데이터를 디렉터리 형태의 데이터를 저장하였다.

[표 4-3] 디렉터리 구조에서의 데이터 수집

[디렉터리 구조에서의 데이터 수집]

Ncs코드별로 대분류>중분류>소분류>세분류>직무역량 데이터 수집

```
var dataAndAbtlLists = (await openApiHelper.모든대분류리스트가져오기())
{
  Data = 대분류데이터
  Child = (대분류와 NcsCd코드가 동일한 중분류 데이터 가져오기)
  {
    Child = (중분류와 NcsCd코드가 동일한 소분류 데이터 가져오기)
    {
      Child = (소분류와 NcsCd코드가 동일한 세분류 데이터 가져오기)
      {
        Child = (소분류와 NcsCd코드가 동일한 각 직무역량 데이터 가져오기)
      } } } }
}
```

(3) 고유노드 분류

대분류, 중분류, 소분류, 세분류, 능력 단위에 해당하는 노드들의 경우 1:N 관계를 띄는 고유노드 형태이므로 중복제거 작업은 필요하지 않다. 디렉터리 구조로 저장된 데이터들을 각 분류별로 호출하여 그래프에 저장하기 위해 앞에 문자열 배열이 있으면 모두 하나의 배열인 것처럼 합해주는 SelectMany 메서드를 활용하여 대분류, 중분류, 소분류, 세분류, 능력 단위별 수집된 데이터를 합하고, 이를 리스트 형태로 저장하는 평면화 작업을 수행하였다.

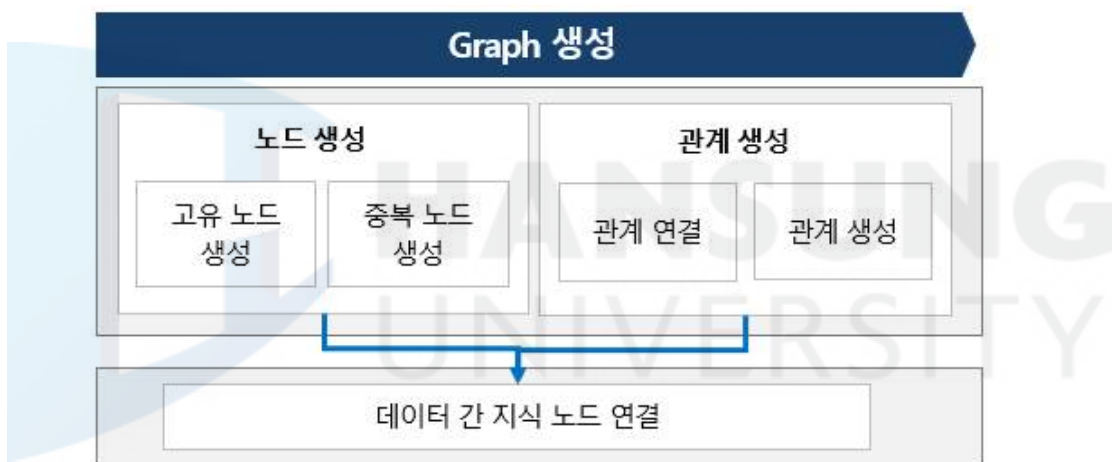
(4) 중복 노드의 제거

직무 능력에 대한 각각의 역량 데이터의 경우 각각의 직무 능력 데이터에 중복되어 있다. 예를 들어 정보통신 전공의 경우 관계형 데이터베이스의 경우 정보통신에 관련된 모든 직업의 정보통신 관련 직업의 관련 학과에 중복되어 저장되어 있었다. 관계를 저장하고 이러한 관계를 분석해야 하는 그래프 데이

터베이스에서는 노드의 중복을 제거, 하나의 노드에 관계를 가지는 노드의 정보들을 저장해야 하므로, 정보통신 전공은 한 개의 노드를 가지며, 모든 정보통신 관련 직업과 관련학과라는 이름으로 관계가 연결된다. 따라서 list에서의 중복제거를 수행하는 ‘Distinct()’ 메서드를 이용하여 중복을 제거 후 리스트로 반환하였다.

나) 직무데이터 API의 데이터 수집 및 전처리

지식 그래프 구축방법론에 따라 [그림 4-6]의 절차에 따라 노드와 관계를 생성하고, 각각의 데이터간 지식을 연결한다.



[그림 4-6] 지식 그래프 구축 방법론: 그래프 생성단계

(1) 고유노드 생성

워크넷 직무데이터와 같이 고유노드가 디렉터리 구조로 구성된 경우 쓸모 없는 노드들을 버리거나, 혹은 각 디렉터리 구조에 따른 관계를 형성해 주어야 한다. 각 분류별 관계도 지식 그래프에서 활용할 수 있으므로 [표 4-4]와 같이 고유노드로 생성하였으며, 디렉터리별 구조로 연결된 분류 별로 [:PARENT_OF] 관계를 생성하였다. 분류별 관계는 일반적으로 관계를 분석해야 하는 이후의 단계와 다르므로 생성과 함께 관계 연결하도록 구성하였다. 고유 코드로 분류되는 대분류의 NcsCd 코드와 중분류의 NcsCd 코드가 같은 항목

들에 해당하는 노드를 불러들여 관계를 생성한다. 이처럼 대분류 -> 중분류 -> 소분류 -> 세분류 -> 능력 단위까지 동일한 관계를 수행하여 고유노드 생성을 진행하였다.

[표 4-4] 직무데이터 API 고유노드 생성 및 관계 연결

[고유노드 생성 및 관계 연결 쿼리]

```
.Match("(대분류데이터노드:Data)", "(중분류데이터노드:Data)")
.Where((GraphDB.Data 대분류 NcsCd값이 같은 대분류데이터노드)
.AndWhere((GraphDB.Data 중분류데이터 NcsCd값이 같은 노드)
.Create("(대분류데이터노드)-[:PARENT_OF]->(중분류데이터노드)")
```

대분류 노드 생성

{ 중분류 데이터들 루프

{ 중분류 데이터 생성

{ (대분류)-[:PARENT_OF]->(중분류) 관계 생성

{.Create("(대분류데이터노드)-[:PARENT_OF]->(중분류데이터노드)")

{ 소분류 데이터들 루프

{ 소분류 데이터 생성

{ (중분류)-[:PARENT_OF]->(소분류) 관계 생성

{ 세분류 데이터들 루프

{ 세분류 데이터 생성

{ (소분류)-[:PARENT_OF]->(세분류) 관계 생성

{ 능력단위 데이터들 루프

{ 능력단위 데이터 생성

{ (세분류)-[:PARENT_OF]->(능력단위) 관계 생성

(2) 중복 노드와 관계 생성

워크넷 직무정보 API에서 능력 단위 이하의 노드들은 한 개의 직무역량 노드들이 다수의 직무와 관계가 연결된다. 때문에 [표 4-5]와 같이 해당 직무 내용(수행태도, 기술및지식, 자격증, 학과, 훈련과정)을 ForEach를 활용하며 반복하면서 직무 내용을 생성하고, 각각의 세부 역량과의 관계를 생성하였으며, 각 세부 역량들도 ForEach를 활용하여 연결된 관계를 확인하여 관계 생성을 반복하여 생성하였으며, 지식 그래프 구축 방법론의 프로세스와의 차이점을 비교하기 위하여 일반적인 ForEach 방식으로 중복 노드 관계 생성과 관계 연결을 구현하였다.

[표 4-5] 중복 노드 생성 및 관계 연결

[중복 노드 생성 및 관계 연결 쿼리]	
.Merge("(직업스킬노드:지식및기술 { Name: {knwSkil} }"))	
.OnCreate()	
.Set("abltnwSkilNode = {newAbltnwSkil}")	
직무내용들 루프	
abltnwSkilList?.ForEach(knwSkil =>	
{	
역량 노드 생성 및 관계 연결 생성	
지식 및 기술 노드 생성	
(능력단위)-[:관계]->(역량) 관계 생성	
역량 노드들 전체 반환	
}	

다) 지식 그래프 생성

지식 그래프가 생성이 완료되면, [표 4-6]과 같이 생성된 노드와 관계의 개수를 확인하여 정상적으로 지식 그래프가 생성되었는지를 확인할 수 있다.

[표 4-6] 직무데이터 API 지식 그래프 생성 노드·관계 생성 수

분류	노드	설명	합계
노드	Ablt	직무능력	12,979
	AbltnwSkil	지식 및 기술	25,388
	AbltMajor	관련 전공	1,959
	Data	분류(대·중·소·세 분류)	1,290
	AbltAtt	관련 훈련	1,826
	AbltcrfItem	수행태도	47
	AbltcrfItem	관련 자격증	1,631
	AbltJobCont	직무내용	11,014
	합계		56,034
관계	ATT	관련 수행태도 관계	124,866
	CRFITEM	관련 자격증 관계	39,468
	MAJOR	관련 전공 관계	21,943
	KNWSKL	관련 지식·기술 관계	128,447
	PARENT_OF	분류 관계	1,290
	JOBCONT	직무내용 관계	23,893
	합계		342,758

(1) 지식 맵을 활용한 구조 검수

[illegible]

[그림 4-7] ‘이러닝 운영 학습자 관리’ 직무의 지식 맵

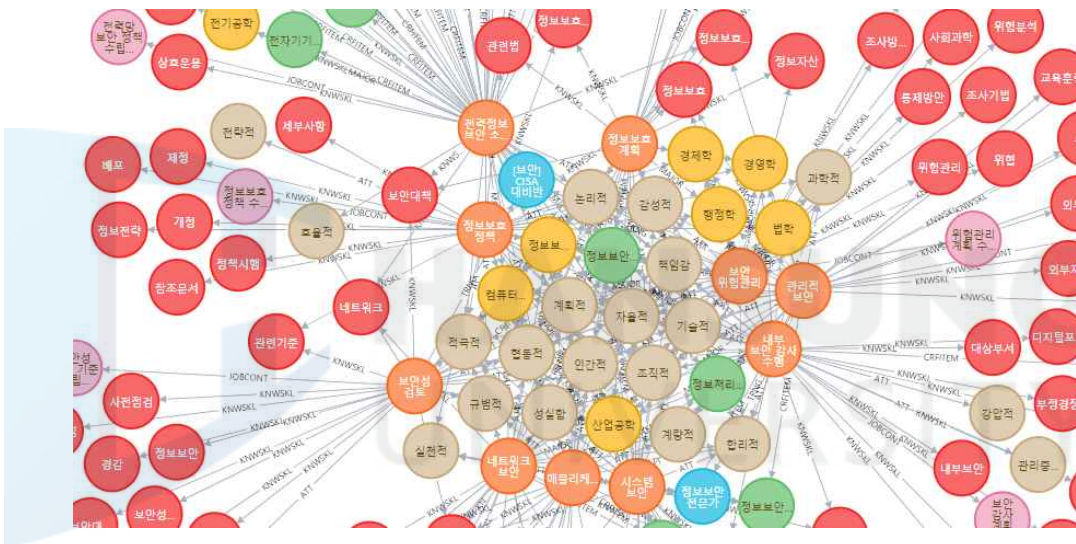
실제 구축된 지식 그래프를 검증하기 위해 [표 4-7]과 같이 기술적인 수행태도와, 컴퓨터공학 전공, 정보보안기사와 매칭되는 직무를 검색하고, 해당 그래프의 관계를 쿼리 하였으며, 결과는 [그림 4-8]과 같이 수집된 지식을 기반으로 설계된 정상적인 지식 그래프를 확인할 수 있다.

[표 4-7] 직무데이터 API 지식 그래프 검증을 위한 사이퍼쿼리

[기술적 수행태도, 컴퓨터공학 전공, 정보보안기사 자격증 소지자의 능력단위 가져오기]

```

MATCH p=(n:Ablt)-[]->()
WHERE (n)-[:ATT]->({Name:'기술적'})
AND (n)-[:MAJOR]->({Name:'컴퓨터공학'})
AND (n)-[:CRFITEM]->({Name:'정보보안기사'})
RETURN p
    
```



[그림 4-8] 직무데이터 API 지식 그래프 검증 지식 맵

마) 문제점 및 해결 사항

(1) 지식 그래프 구축 문제점

일반적인 방법으로 워크넷 API를 활용한 지식 그래프 구축에서 가장 큰 문제는 지식 그래프 구축에 너무 긴 시간이 소요되었다. [그림 4-9]와 같이 04:29:20에 실행된 프로그램이 17:26:13에 종료되었기 때문에 총 12시간 56분 52초가 소요되었으며, 이 중 노드와 관계의 생성이 12:54:05초로 가장 큰 시간 소요가 있었다.

```

2019-11-26 04:31:54,062 [19] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Completed(Elapsed:00:00:00.0623181)
2019-11-26 04:31:54,062 [8] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Started
2019-11-26 04:31:54,174 [10] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Completed(Elapsed:00:00:00.1106410)
2019-11-26 04:31:54,175 [8] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Started
2019-11-26 04:31:54,274 [12] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Completed(Elapsed:00:00:00.0980683)
2019-11-26 04:31:54,274 [14] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Started
2019-11-26 04:31:54,338 [13] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Completed(Elapsed:00:00:00.0622544)
2019-11-26 04:31:54,338 [19] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Started
2019-11-26 04:31:54,441 [10] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Completed(Elapsed:00:00:00.1016009)
2019-11-26 04:31:54,442 [13] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Started
2019-11-26 04:31:54,587 [12] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Completed(Elapsed:00:00:00.1439594)
2019-11-26 04:31:54,588 [19] DEBUG DataFetcher.Program [(null)] - GetAbtlListByNcsCd Started
Create Data, Abtl Nodes and Relations Started
Create Data, Abtl Nodes and Relations Ended(Elapsed:12:54:05.1057413)
DataFetcher Program Ended(Running Duration:12:56:52.4258257)
2019-11-26 04:32:07,952 [7] DEBUG DataFetcher.Program [(null)] - Create Data, Abtl Nodes and Relations Started
2019-11-26 17:26:13,059 [7] DEBUG DataFetcher.Program [(null)] - Create Data, Abtl Nodes and Relations Ended(Elapsed:12:54:05.1057413)
2019-11-26 17:26:13,061 [1] INFO DataFetcher.Program [(null)] - DataFetcher Program Ended(Running Duration:12:56:52.4258257)

```

[그림 4-9] 직무데이터 지식 그래프 생성 로그

(2) 문제점 해결방안

디렉터리 구조의 반복문으로 인한 데이터 관계 매칭 및 생성에 의한 것으로 지식 그래프 구축방법대로 관계의 연결과 생성을 별도로 구성하고 ForEach대신 목록의 요소가 있는 행에서 변수의 데카르트 곱을 수행하는 UNWIND를 활용하여 관계를 구성한다.

(가) index 값 추가

지식 그래프 생성 속도와 함께 쿼리 속도 향상을 위해 중복 제거된 노드들에 index값을 추가하였으며, index는 지식 베이스 생성이 아닌 지식 베이스 쿼리 속도를 향상하기 위해 설정한다.

(나) 고유·중복 노드 생성

지식 베이스 ‘구축방법론’의 절차에 따라 Data 노드, 직무능력 노드, 직무내용 노드, 지식 및 기술 노드, 수행태도 노드, 관련 자격증 노드, 관련 전공 노드, 관련 훈련 노드들을 생성한다. 상위 과정에서 중복제거가 완료되었기 때문에 전 단계의 인덱스가 추가된 목록으로 고유노드와 중복노드를 모두 생성한다.

(다) 관계 연결 및 생성

[표 4-8]과 같이 지식 베이스 구축 방법론에 의한 관계를 우선 생성해 주

며, UNWIND와 index 값을 활용하여 각각의 관계를 연결하며 노드들 간의 관계를 생성하였다.

[표 4-8] 직무데이터 관계 연결 및 생성

[역량 노드 생성 및 관계 연결 쿼리]

```
client.Cypher
    .Unwind(역량 관계정보, "역량 관계 정보")
    .Match("(source:직업 (직업과 기술 및 리스트 관계모두 매칭))")
    .Create("(직업)-[: 관계]->(역량 노드)")
    .ExecuteWithoutResults();
```

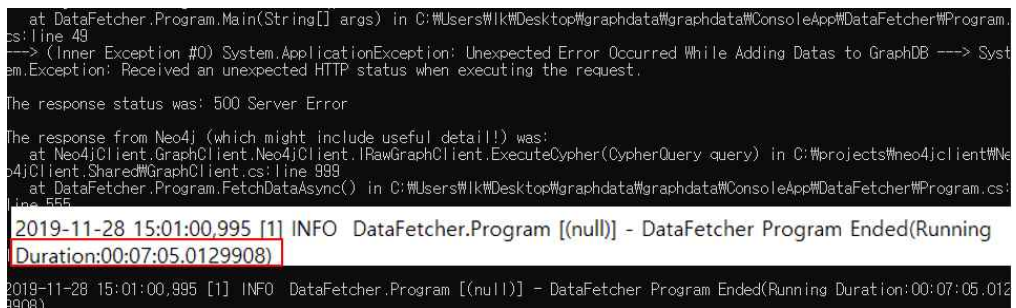
PARENT_OF 관계들 생성: Data to Data

PARENT_OF 관계들 생성2: Data to Abt

```
{
    JOBCONT 관계들 생성
    KNWSKL 관계들 생성
    ATT 관계들 생성
    CRFITEM 관계들 생성
    MAJOR 관계들 생성
    TRNG 관계들 생성
}
```

(라) 지식 그래프 생성

지식 그래프 구축 방법론의 절차와 알고리즘을 최적화를 통해 지식 그래프를 다시 생성하였으며, 그 결과는 [그림 4-10]과 같이 이전의 방식보다 큰 폭의 속도 향상이 있었다.



```
at DataFetcher.Program.Main(String[] args) in C:\Users\Wlk\Desktop\graphdata\graphdata\ConsoleApp\DataFetcher\Program.cs:line 49
--> (Inner Exception #0) System.ApplicationException: Unexpected Error Occurred While Adding Datas to GraphDB --> System.Exception: Received an unexpected HTTP status when executing the request.
The response status was: 500 Server Error
The response from Neo4j (which might include useful detail!) was:
at Neo4jClient.GraphClient.Neo4jClient.IRawGraphClient.ExecuteCypher(CypherQuery query) in C:\projects\neo4jclient\Neo4jClient\Shared\RawGraphClient.cs:line 999
at DataFetcher.Program.FetchDataAsync() in C:\Users\Wlk\Desktop\graphdata\graphdata\ConsoleApp\DataFetcher\Program.cs:line 555
2019-11-28 15:01:00,995 [1] INFO DataFetcher.Program [(null)] - DataFetcher Program Ended(Running Duration:00:07:05.0129908)
2019-11-28 15:01:00,995 [1] INFO DataFetcher.Program [(null)] - DataFetcher Program Ended(Running Duration:00:07:05.0129908)
```

[그림 4-10] 절차와 알고리즘이 최적화된 지식 그래프 생성시간

(2) 지식 그래프 생성시간 비교

[표 4-9]와 같이 방법론과 알고리즘에 따라 생성 지식 베이스 생성 속도는 큰 차이를 보이며, 동일한 데이터와 동일한 컴퓨팅 자원을 활용하여 생성된 지식 그래프 생성 비교를 통해 방법론에 따른 절차와 해당 절차의 구현을 위한 최적화된 알고리즘 적용은 지식 그래프 구축을 위한 중요한 요소임을 확인할 수 있었다.

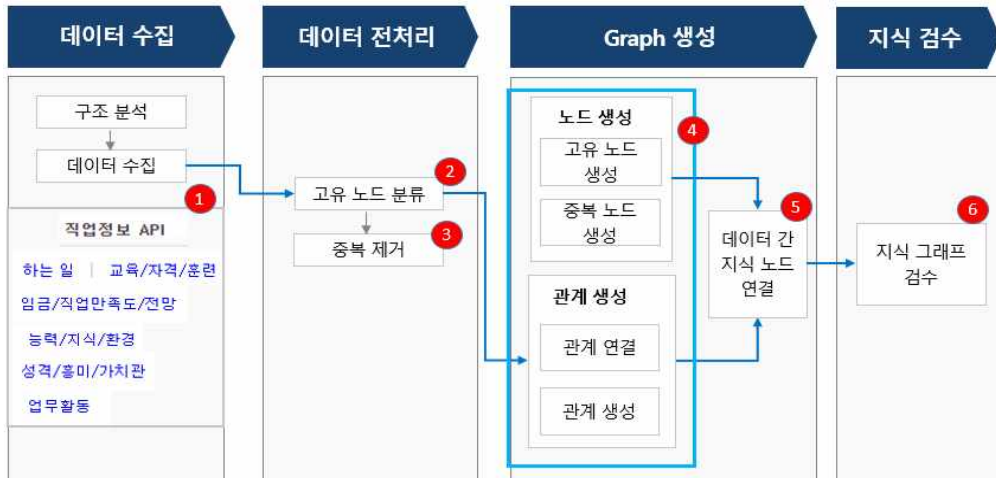
[표 4-9] 방법론과 알고리즘에 따른 지식 그래프 생성 속도

구현	일반 구현 방식	방법론과 알고리즘 변경 방식
결과	<pre> er.Program [null] - GetAbtlIstByNcs0d Started er.Program [null] - GetAbtlIstByNcs0d Completed(Elapsed:00:00:00.0000000) er.Program [null] - GetAbtlIstByNcs0d Started er.Program [null] - GetAbtlIstByNcs0d Completed(Elapsed:00:00:00.0000000) er.Program [null] - GetAbtlIstByNcs0d Started er.Program [null] - GetAbtlIstByNcs0d Completed(Elapsed:00:00:00.0000000) nd Relations Started nd Relations Ended(Elapsed:12:54:05.105741) (Running Duration:12:56:52.4258257) </pre>	<pre> at Ncs0dClient.GraphClient.Ncs0dClient.TrwGraphClient.Executed at Ncs0dClient.SharedGraphClient.cs:line 989 at DataFetcher.Program.FetchDataAsync() in C:\Users\Wilk\Desktop\ line 555 2019-11-28 15:01:00.995 [1] INFO DataFetcher.Program Duration:00:07:05.0129908) 2019-11-28 15:01:00.995 [1] INFO DataFetcher.Program [null] - [9908) </pre>
생성 시간	12:56:52	00:07:05

2) 직업정보(코드 조회 API) 지식 그래프 구축

가) 지식 그래프 구축 방법론 적용

워크넷 직업정보 API는 직무데이터와 동일하게 정형데이터지만, 디렉터리별 분류 방식이 아닌, 특정 코드를 기준으로 파라미터값에 따라 하는일, 교육/자격/훈련, 임금/직업만족도/전망, 능력/지식/환경, 성격/흥미/가치관, 업무활동 데이터를 수집할 수 있다. 워크넷 직업정보 API를 활용한 지식 베이스 구축 역시 [그림 4-11]과 같이 지식 그래프 구축 방법론에 따라 데이터 수집, 중복제거, 노드 생성, 관계 생성 절차로 지식 그래프를 생성한다.



[그림 4-11] 직업정보 API 지식 그래프 구축 방법론 적용

나) 직업정보 API 수집

[그림 4-12]와 같이 직업정보 코드별로 워크넷 직업정보 API에서 ‘하는 일’, ‘교육/자격/훈련’, ‘임금/직업만족도/전망’, ‘능력/지식/환경’, ‘성격/흥미/가치관’, ‘업무활동’ 정보를 수집한다.

//api에서 받아온xml을 바탕으로 각 변수에 내용 저장

```
job = 데이터분류.하는일(await apiHelper.GetXml(uri + (int)SearchType.work), job);
job = 데이터분류.교육/자격/훈련(await apiHelper.GetXml(uri + (int)SearchType.edu), job);
job = 데이터분류.임금/직업만족도/전망(await apiHelper.GetXml(uri + (int)SearchType.sal), job);
job = 데이터분류.능력/지식/환경(await apiHelper.GetXml(uri + (int)SearchType.abil), job);
job = 데이터분류.성격/흥미/가치관(await apiHelper.GetXml(uri + (int)SearchType.chr), job);
job = 데이터분류.업무활동(await apiHelper.GetXml(uri + (int)SearchType.act), job);
```

//각 직업 관련정보를 담은 클래스는 XmlNode를 이용하여 생성자 초기화가 진행

// 검색 조건 "하는일"

public static Job 하는일(XDocument xml, Job job)

```
{
    var node = xml.Element("jobsDo");
    job.JobCd = node.Element("jobCd")?.Value; // 직업코드
    job.JobNm = node.Element("jobSmclNm")?.Value; // 직업 소분류명
    job.JobSum = node.Element("jobSum")?.Value; // 직무 개요
    job.ExecJob = node.Element("execJob")?.Value; // 수행 직무
    job.RelJobList = node.Elements("relJobList")?.Select(x => new RelJob(x)).ToList(); //직업코드, 직업명
    return job;
}
```

[그림 4-12] 코드별 직업정보 수집

다) 중복 데이터 제거 및 전처리 작업 시작

워크넷 직업정보의 경우 직업별로 직업 역량 정보들이 별도로 구성되어 있다. 그러나 관계 분석을 위해서는 역량 정보들이 별도로 구성되는 방식이 아닌, 하나의 역량 정보가 해당 역량 정보를 가지고 있는 모든 노드에 연결되어야 한다. 이러한 작업 수행을 위해서 중복을 할 수 있는 노드들인 관련 전공 리스트, 관련 자격증 리스트, 관련 정보처 리스트, 업무수행 능력 중요도, 업무수행 능력 수준, 지식 중요도, 지식 수준, 업무환경, 성격 리스트, 흥미 리스트, 가치관 리스트, 업무활동 중요도 리스트, 업무활동 수준 리스트에 대한 중복 데이터를 제거하는 전처리 작업을 수행한다.

라) 중복가능 노드들에 대한 생성 및 각 직업 노드 생성

중복가능 노드들에 대한 생성은 전처리 작업에서 중복 데이터를 제거한 각 직업 역량들의 리스트(Unwind(all역량 리스트, "역량"))에서 중복 가능한 노드들에 대한 생성 작업을 수행한다.

최상위 노드인 각 직업 노드들은 ‘요약’, ‘하는일’, ‘교육/자격/훈련’, ‘임금/직업만족도/전망’, ‘능력/지식/환경’, 성‘격/흥미/가치관’, ‘업무활동’과 같은 상세 구분으로 분류되어 있다. [표 4-10]과 같이 각 직업 노드 생성과 함께 이러한 직업분류 노드를 만들고 고유관계를 생성한다.

[표 4-10] 각 직업 노드 생성 및 고유관계 생성

[각 직업 노드 및 연결되는 관계 생성]
<pre>//직업과 역량 노드 생성 .Create(\$"(job:직업{Property})") .Create(\$"(jobInfo: 직업정보 구분{Property})") .Create(\$"(sal: 임금{Property})") .Create(\$"(jobSatis: 직업만족도{Property})") .Create(\$"(edubg: 학력분포{Property})") .Create(\$"(schDpt: 전공학과분포{Property})") //노드들 간의 관계 생성 .Create("(jobRelCert:JobRelCerList{Name:W\"관련자격증W\"})") .Create("(jobInfo)-[:RELCERLIST]->(jobRelCert)") ... 관련전공, 업무환경리스트, 성격리스트, 흥미리스트, 가치관리스트 생성</pre>

마) 직업 중복 가능한 노드들과 관계 지정

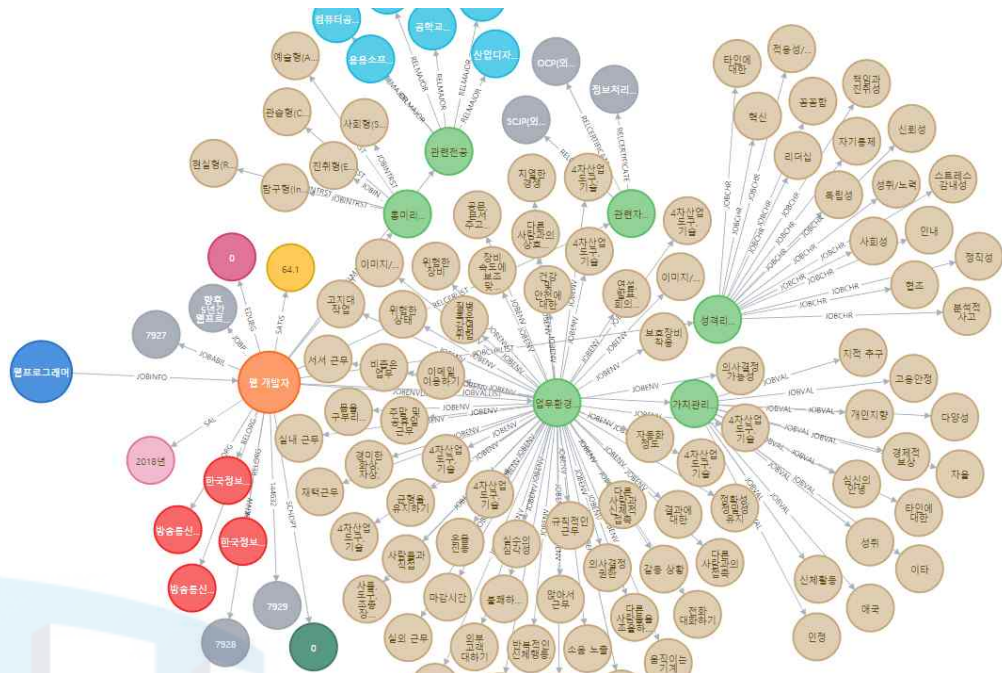
[표 4-11]과 같이 중복 가능한 직업 역량 노드들과 직업과의 관계를 지정하는 절차로 직업정보 API에서 수집된 각 직업 역량들의 중요도를 각 직업에 연결하기 위한 관계 가중치 값으로 지정하여 관계를 형성한다. 중요도와 같은 가중치 값이 존재하지 않는 노드는 가중치 값이 없이 관계만 생성한다.

[표 4-11] 직업 중복 가능한 노드들과 관계 지정

```
[직업 중복 가능 노드들과의 관계 지정 및 생성]
.Match($"(직업정보 관계 매칭)")
.Unwind(중복 가능 리스트 관계, "node")
.Create("(중복가능 리스트 생성)")
.Create("(직업정보)-[:관계{상태 및 가중치}]->(중복가능 리스트)")
.Create("(직업정보)-[:관련일자리전망{상태}]->(일자리전망 리스트)")
.Create("(직업정보)-[:관련수행능력{상태}]->(수행능력리스트)")
.Create("(JobInfo)-[:관련지식중요도{상태}]->(지식중요도 리스트)")
.Create("(job)-[:관련지식수준{상태}]->(지식수준중요도 리스트)")
.Create("(JobInfo)-[:관련업무환경{상태}]->(JobEnvList,JobNm)")
.Create("(직업정보)-[:관련성격{상태}]->(성격중요도 리스트)")
.Create("(직업정보)-[:관련흥미{상태}]->(흥미중요도리스트)")
.Create("(JobInfo)-[:관련가치관{상태}]->(직업가치관리스트)")
.Create("(직업정보)-[:관련활동중요도{상태}]->(업무활동중요도 리스트)")
.Create("(직업정보)-[:관련활동수준{상태}]->(업무활동 수준 리스트)")
.Create("(직업정보)-[:관련 자격증]->(자격증 리스트)")
.Create("(직업정보)-[:관련전공]->(전공 리스트)")
.Create("(직업정보)-[:관련정보처]->(관련정보처리리스트")
```

바) 지식 그래프 검수

[그림 4-14]는 생성된 지식 그래프에서 정상적으로 구축된 지식 베이스에서 MATCH (n:Job{JobNm:'웹프로그래머'})를 쿼리하여 '웹프로그래머' 관련 노드가 정상적으로 생성되었는지 구조 검수를 시행하였으며, 워크넷 직업정보 API를 수집하여 생성한 그래프는 설계대로 구현된 지식 그래프가 생성되었음을 확인할 수 있었으며 생성된 구조와 설계 구조상의 문제점을 지식 맵을 통해 확인할 수 있다.



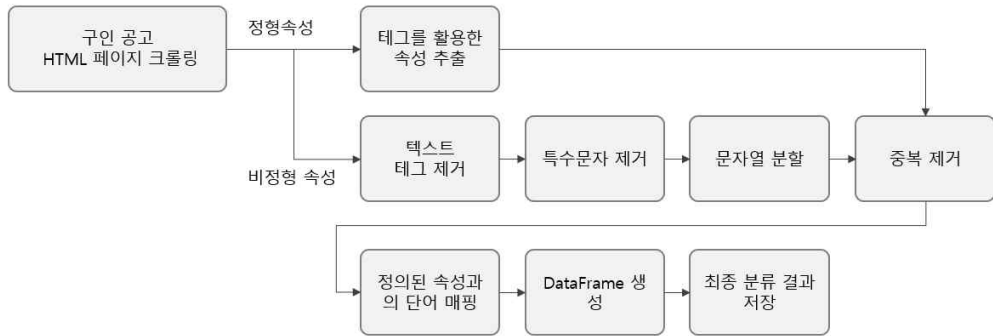
[그림 4-13] 지식 그래프에서의 웹프로그래머 직업 호출 지식 맵

3) 민간 채용정보(비정형 데이터) 지식 그래프 구축

가) 분류 사전 활용 채용정보 데이터 추출

실제 학생들에게 가장 도움이 되는 데이터는 기업들이 실제 인력을 채용하는 채용광고이다, 따라서 기업들의 실제 채용공고를 분석하면 기업이 찾는 인력의 역량 정보를 수집할 수 있으며, 실제 현업에서 필요하거나 우대하는 기술, 성향, 자격증, 학과 등과 같은 다양한 세부 정보 수집이 가능하다. 많은 수의 기업들이, 직업정보제공 서비스의 정형화된 양식이나 고유명사보다는 기업마다 각기 다른 이름으로 부르고 있으며, 이를 그대로 공고에 올리기 때문에 유사어나 동의어 역시 분류하여 상세 역량을 추출해야 한다.

실제 수집된 채용광고 링크에서 정형·비정형 문서별로 역량 정보를 [그림 4-14]와 같이 추출한다.



[그림 4-14] 채용공고 리스트에서 역량 정보 추출 절차

(1) 채용공고 크롤링

기업들의 채용공고에서 실제 기업이 필요로 하는 인재 역량을 추출할 수 있는 정보는 기업별 자체적으로 작성한 채용정보가 포함된 HTML 페이지 링크이다. 채용공고 크롤링의 첫 번째 단계로 직업별 채용공고가 있는 채용공고 페이지 링크 정보를 [표 4-12]와 같이 수집하였다.

[표 4-12] 웹프로그래머 채용정보 페이지 수집

No.	‘웹프로그래머’ 수집 채용공고 링크
1	http://www.jobkorea.co.kr/Recruit/GI_Read/29821802
2	http://www.jobkorea.co.kr/Recruit/GI_Read/29786711
3	http://www.jobkorea.co.kr/Recruit/GI_Read/29762048

(2) 역량 정보 추출을 위한 단어 사전 작성

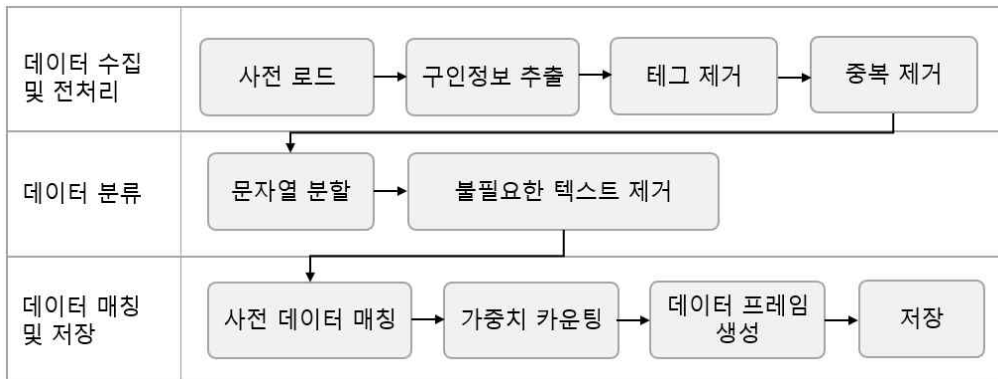
비정형 문서에서 역량 정보 추출을 위한 방법의 하나로 단어 사전을 활용한 역량 정보 추출 방법을 구현하였다. 채용공고와 채용 내용에 대해 작성하는 단어 형식이 없으므로 [표 4-13]과 같이 CSV 파일을 활용하여 동의어 사전과 구분별 단어 사전을 통해 매칭하여 실제 역량 정보를 추출하도록 구성하였다. 스프레드시트 방식의 장점은 학습의 필요성이 없고 단순화되어있는 장점으로 분류 항목 및 동의어의 추가가 쉽다.

[표 4-13] 역량 정보 추출을 위한 데이터 사전

기술 및 지식	우대사항	자격증	학과	항목 명
파이썬	영어	정보처리기사	전기/전자공학	새로운기술
머신러닝	영어	리눅스마스터	컴퓨터공학	주체적
인공지능	일어	CCNA	시스템공학	디테일
딥러닝	근무경험	PMP	수학	꼼꼼함
텐서플로우	운전가능자	MCP	물리학	커뮤니케이션
자연어처리	관련 자격증	MCSE	통계학	협업
데이터마이닝	관련 전공	ISMS	전자계산학	긍정적
React	공모전	OCJP	인지과학	박은
Java	포트폴리오	정보통신기사	공학계열	아이디어
챗봇	스타트업	사무자동화	산업공학	논리적
...
총 223개	총 27개	총 10개	총 13개	총 27개

(3) HTML 페이지에서 역량 정보 추출 및 기초데이터 생성

HTML 페이지에서 역량 정보 수집은 크롬 브라우저를 제어하기 위해 Selenium 라이브러리를 활용하였으며, [그림 4-15]와 같이 데이터 추출 및 전처리 -> 데이터 분류 -> 데이터 매칭 및 저장 단계로 구분하며, [표 4-16]과 같이 Body의 HTML 전체를 추출하여 전처리 후 사전매칭을 통한 데이터 프레임 생성 및 [표 4-15]의 형태로 최종 기업들이 원하는 채용정보 데이터를 생성한다.



[그림 4-15] HTML 페이지에서 역량 정보 추출

[표 4-14] 채용 HTML 페이지에서의 직업 역량 정보 추출

HTML 페이지에서의 정형·비정형 역량 정보 추출	
soup.find("body")	// body 내용 전체 추출
remove_tag	// 태그 제거
remove_symbol(origin_str)	// 텍스트, 특수문자 제거
origin_str.split()	// 문자열 분할
list(set(split_str))	// 중복제거
for sheet in 데이터사전	// 각각의 추출 소스 LOOP == 매칭
base.역량구분[ori_str] += 1	// 역량 가중치 카운팅
pd.DataFrame((base.skills.ke	// DataFrame 생성
ys(),(base.skills.values()))	
make_CSV(merge_df)	// 직업별 채용정보 데이터 생성
* 직업별로 분류된 CSV 형식의 채용정보 데이터 생성	

[표 4-15] 기업 채용정보 기초데이터 생성

기술 및 지식	가중치	우대사항	가중치	성향	가중치
Java	72	포트폴리오	39	커뮤니케이션	20
PHP	51	관련 자격증	23	성장	11
SQL	41	프로젝트관리	13	발전	10
유지보수	39	영어	9	주체적	8
JSP	30	문서능력	6	책임감	7
Spring	2	프리젠테이션	4	적극적	6
챗봇	17	컴퓨터활용	3	성실	4
HTML	18	스타트업	3	새로운기술연마	2
jQuery	15	PL	3	리더쉽	2
챗봇	17	통계분석	1	창의적	1
JavaScript	13	중국어	1	협업	1

나) 민간 채용정보 데이터 지식 그래프 구축

(1) 지식 그래프 구축방법

진로 및 직업 관련 정보 수집 시 공공데이터만으로도 정형화되어 있는 가치 있는 자료를 수집하여 활용할 수 있지만, 실제 학생들이 필요로 하는 데이터는 쉽게 구할 수 있는 데이터들이 아닌 비정형 데이터 형태로 존재하는 다양한 정보들을 활용할 수 있어야 한다.

[표 4-16] 채용정보 중복제거 작업 수행

[데이터 전처리]

각각의 노드에 대한 중복 제거 작업 수행

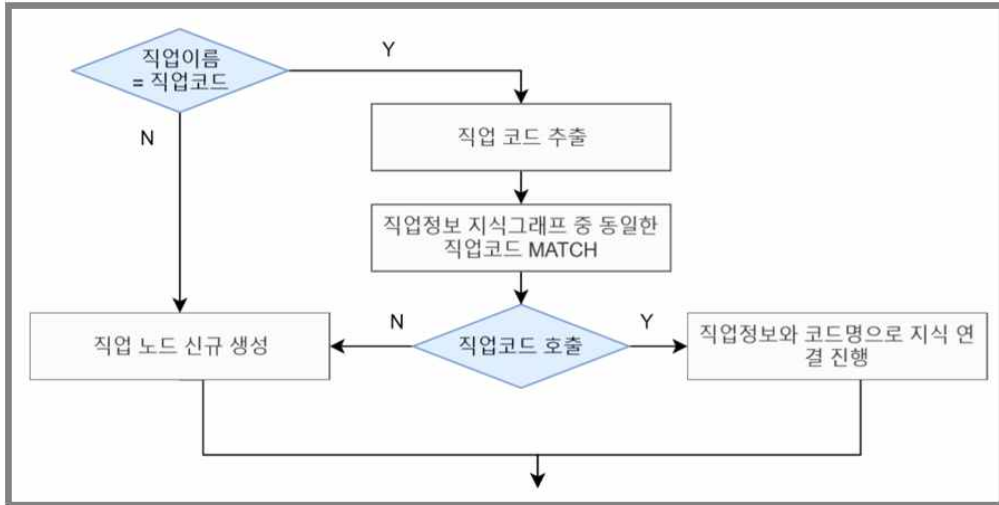
```
Foreach 역량 in List
{
    // 각 역량 별 중복 제거 작업 수행 후
    *기술 및 지식 노드 중복 제거 List 저장 (1:n 관계 생성)
    우대사항 노드 중복 제거
    자격증 노드 중복 제거
    전공 노드 중복 제거
    성향 노드 중복 제거
}
*지식및기술 노드 = all_지식및기술_리스트.Distinct().Select.ToList();
```

(4) 중복가능 노드 생성

채용정보 데이터의 경우 고유노드가 없이 직업 관련 속성 정보만 있으므로 [표 4-16]과 같이 중복가능 속성 노드들인 기술 및 지식, 우대사항, 자격증, 전공, 성향에 대한 노드를 생성한다.

(5) 최상위 지식 노드 연결 및 생성

CSV 파일을 생성할 때 표준 직업 코드와 같은 이름을 가진 직업은 파일 이름을 코드명으로 설정하였다. 그 때문에 [그림 4-17]과 같이 파일 이름이 코드 형태였던 경우 이미 직업 노드가 지식 그래프에 존재하기 때문에 해당 직업 노드에 채용정보로 그래프 생성하며, 파일 이름이 직업 이름일 때 관계를 구성할 직업 노드가 없으므로 최상위 직업 노드를 신규로 생성한다.



[그림 4-17] 최상위 직업 노드 연결 및 생성

(6) 데이터 간 지식 노드의 연결

[표 4-17]과 같이 채용정보 노드 및 최하위 노드 연결을 위한 각 중간 단계의 노드와 관계를 생성한다.

[표 4-17] 채용정보 속성 노드 관계 생성 및 가중치 추가

[속성 노드 생성 및 가중치 추가]

채용정보 속성 노드 관계 생성 및 가중치 추가

query

```

.Match("(채용정보)-[:기술리스트관계]->(지식및기술리스트 노드)")
.Unwind(job.지식및기술, "node")
.Match("(a:지식및기술{Name:지식및기술이름})")
.Create("(지식및기술리스트)-[:관계{Weight:Count}]->(a)")
{
  //기타 직업 역량 정보(우대사항, 자격증, 전공, 성향)도 동일한
  중복가능 속성 노드 생성 수행
}
  
```

(7) 채용정보 속성 노드 관계 생성 및 가중치 추가

[표 4-18]과 같이 각 직업 속성 노드와 연결 관계를 구성하고, 연결 관계마다 구인을 가중치로 추가하여, 추후 진로 조언에서 중요도를 산출할 수 있

도록 구성한다.

[표 4-18] 각 역량 리스트 노드 생성 및 관계 연결

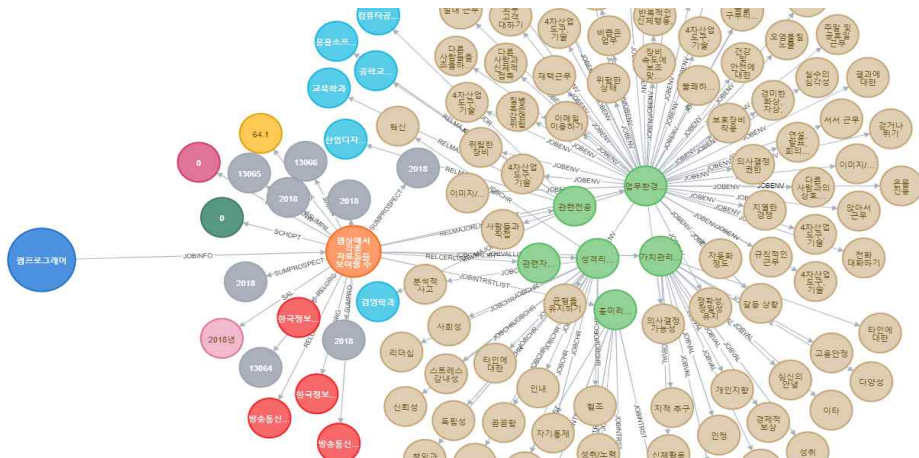
[역량 노드 생성 및 관계 연결]

채용정보 노드와 역량 노드 연결을 위한 중간 노드 생성 및 관계 생성

```
.Create("{Name:채용정보}")
.Create("{Name:기술 및 지식 리스트}")
.Create("{Name:우대사항 리스트}")
.Create("{Name:자격증 리스트}")
.Create("{Name:전공 리스트}")
.Create("{Name:성향 리스트}")
.Create("(n)-[:채용정보관계]->(채용정보)")
.Create("(채용정보)-[:지식및기술리스트 관계]->(지식및기술List 노드)")
.Create("(채용정보)-[:우대사항리스트 관계]->(우대사항List 노드)")
.Create("(채용정보)-[:자격증리스트 관계]->(자격증List 노드)")
.Create("(채용정보)-[:전공 리스트 관계]->(전공List 노드)")
.Create("(채용정보)-[:성향 리스트 관계]->(성향List 노드)")
```

(8) 생성 완료된 지식 그래프

[그림 4-18]은 생성 완료된 지식 그래프에서 ‘웹프로그래머’ 직업을 호출한 지식 맵이다. 위의 프로그래밍 절차와 같이 [직업]-[jobinfo]-[직업정보]의 하위노드에 생성된 노드들을 확인할 수 있으며, 우측의 중복 가능한 노드인 업무환경리스트, 가치관리리스트, 성격리스트, 흥미리스트, 관련전공, 관련자격증 노드에서 세부 노드를 확인할 수 있다. [그림 4-19]는 직업 속성 노드 중 ‘4차산업도구기술’의 확장 노드를 호출하였으며, 다양한 직업과 속성들과 연관된 관계를 확인할 수 있다.



[그림 4-18] 생성 완료된 지식 그래프 중 ‘웹프로그래머’ 지식 맵



[그림 4-19] 직업 속성 노드의 관계 일부

다) 최종 생성 지식 그래프 검수

(1) 생성 노드 및 관계의 수

(가) 직업정보와 구인 정보 분석 생성 노드의 수

연구 수행을 위해 변화가 빠른 정보기술 직업으로 데이터를 수집했으며, 일부 직업 리스트의 경우 구분이 어려울 정도로 채용공고가 적었기 때문에 의미 있는 68개의 직업에 대한 직업 관련 노드를 생성하였으며, [표 4-19]와 같이 속성 노드와 워크넷 직업정보에서 1901개의 노드, 채용정보 분석에서 637개의 노드가 생성되어 총 2538개의 노드가 생성되었다.

[표 4-19] 지식 그래프 생성 노드 목록

분류	No	노드	설명	합계
Worknet 직업정보	1	Job	직업	104
	2	JobStatusNmCont	업무수행능력	328
	3	JobSumProspect	평균직업만족도	340
	4	Edubg	학력분포	68
	5	JobAbil	업무수행능력	68
	6	JobAct	업무활동	68
	7	JobChrList	성격 리스트	68
	8	JobEnvList	업무환경	68
	9	JobInfo	직업정보	68
	10	JobIntrstList	흥미 리스트	68
	11	JobKnw	지식	68
	12	JobProspect	향후 5년 전망	68
	13	JobRelCerList	자격증 리스트	68
	14	JobRelMajor	관련 전공	68
	15	JobSatis	직업만족도	68
	16	Sal	연봉	68
	17	SchDpt	전공학과분포	68
	18	JobValList	가치관 리스트	68
	19	RelCertificate	자격증	37
	20	RelMajor	학과	36
	21	RelOrg	관련정보처	36
	직업정보 노드 합계			1,901
채용정보 분석	1	EmIn	채용정보	59
	2	KnwSkl	기식및기술	169
	3	KnwSkIList	지식및기술 리스트	59
	4	Major	학과	33
	5	MajorList	전공리스트	59
	6	Preferential	우대사항	24
	7	PreferentialList	우대사항 리스트	59
	8	Tendency	직업 성향	22
	9	TendencyList	성향 히스트	59
	10	Certificate	자격증	35
	11	CertificateList	자격증 리스트	59
	채용정보 분석 노드 합계			637
생성 노드 합계			2,538	

(나) 직업정보와 구인 정보 분석 생성 관계의 수

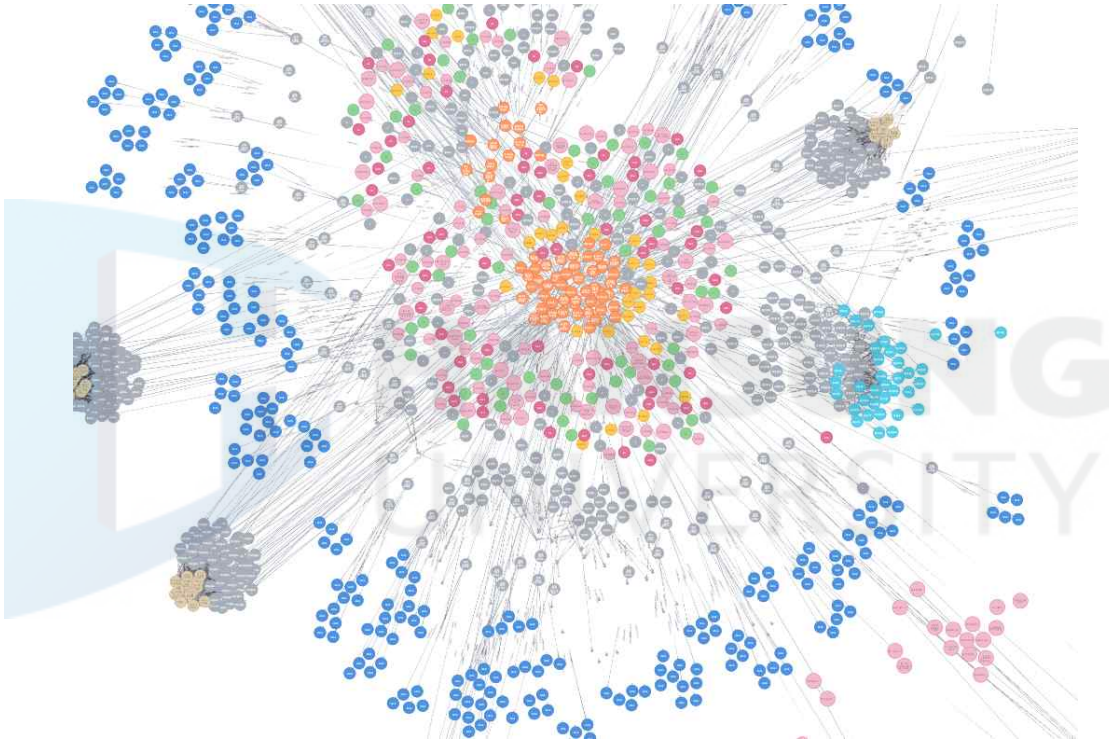
그래프 데이터베이스 분석에 중요성이 높은 노드와의 관계는 [표 4-20]과 같이 워크넷 직업정보에서 34,851개, 구인 정보 분석에서 3,889개가 생성되어 총 39,760개의 관계가 형성되었다.

[표 4-20] 지식 그래프 생성 관계 목록

분류	워크넷 직업정보 관계		채용정보 분석 관계	
No	노드	생성 수	설명	합계
1	ABIL	5,806	EMIN	59
2	ABILLVL	5,800	KNWSKL	2,225
3	KNWLDG	4,356	KNWSKLLIST	59
4	KNWLGLVL	4,352	MAJOR	210
5	JOBACTVIMPRTNC	4,182	MAJORLIST	59
6	JOBACTVLVL	4,182	TENDENCY	363
7	JOENV	3,011	TENDENCYLIST	59
8	JOBCHR	1,056	CERTIFICATE	174
9	JOBVAL	852	CERTIFICATELIST	59
10	RELMAJOR	375	PREFERENTIAL	563
11	SUMPROSPECT	340	PREFERENTIALLIST	59
12	JOBINTRST	318		
13	RELCERTIFICATE	135		
14	RELOGR	86		
15	EDUBG	68		
16	JOBABIL	68		
17	JOBACT	68		
18	JOBCHRLIST	68		
19	JOENVLIST	68		
20	JOBINFO	68		
21	JOBINTRSTLIST	68		
22	JOBKNW	68		
23	JOBPROSPECT	68		
24	JOBVALLIST	68		
25	RELCERLIST	68		
26	RELMAJORLIST	68		
27	SAL	68		
28	SATIS	68		
29	SCHDPT	68		
워크넷 직업정보 관계 합계		34,851	채용정보분석 관계 합계	3,889

(2) 생성된 지식 그래프의 지식 맵

직무정보 API는 별도의 지식 그래프에 저장하여 실험을 수행했으며, 직무정보 API와 직업정보 API는 동일한 연결 코드가 존재하지 않기 때문에 최종 생성된 지식 그래프는 직업정보 API와 채용정보수집 데이터를 합하여 구성하였다. [그림 4-20]은 최종 생성된 그래프의 지식 맵이며, 이러한 지식 맵 검증을 통해 지식 그래프에서의 관계 연결 상태와 문제점을 파악할 수 있다.

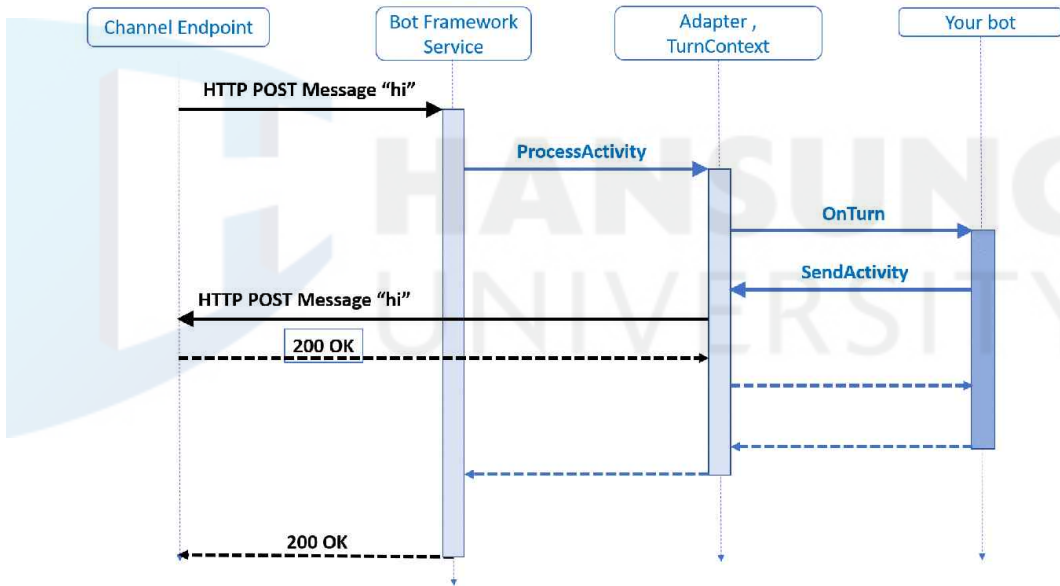


[그림 4-20] 최종 생성된 지식 그래프 구조

제 3 절 사용자 인터페이스 구현

1) 챗봇 인터페이스 구현

Microsoft Bot-Framework에서의 메시지 활동 처리는 [그림 4-21]과 같으며, 실제 Bot Framework Service와 사용자 인터페이스는 설정 이외에는 기본적으로 제공되기 때문에 비교적 빠른 대화형 서비스 구현이 가능하며, Microsoft Cognitive 서비스인 LUIS 자연어처리 서비스를 활용하여 대화형 서비스를 구현하였다.



[그림 4-21] 메시지 처리 활동

가) 정보 선택 구분

사용자가 채팅창에 처음 정보를 입력하거나 대기 시 직업정보와 채용정보 중에서 원하는 정보를 선택할 수 있도록 선택 창을 [그림 4-22]와 같이 구성하였으며, 사용자의 선택에 따라 원하는 분야의 메인 Dialog에서 대화를 처리하도록 구성하였다.

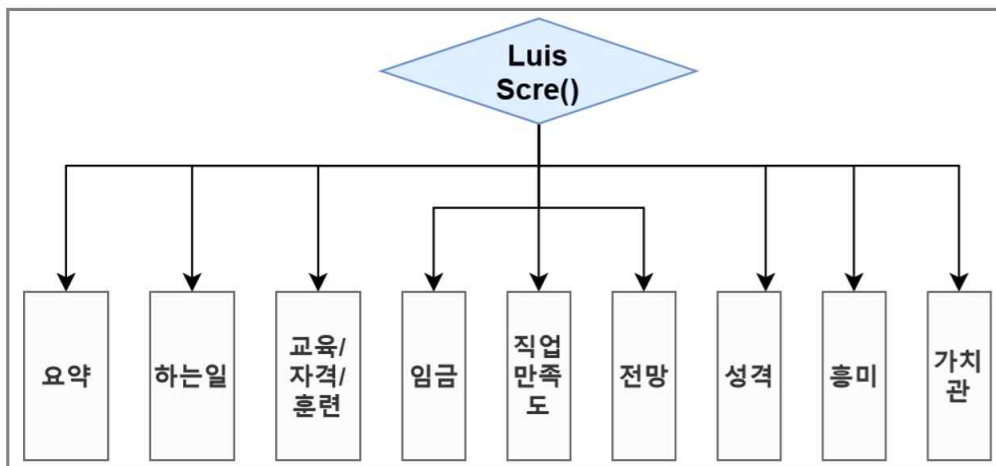
[그림 4-22] 정보 검색 구분

(나) 자연어처리 및 분류

자연어처리에서는 실제 사용자의 의도와 엔티티를 자연어처리에 학습된 정보를 기준으로 분류하여 직업과 속성 정보를 각각의 직업정보와 채용정보의 메인 Dialog로 전달한다.

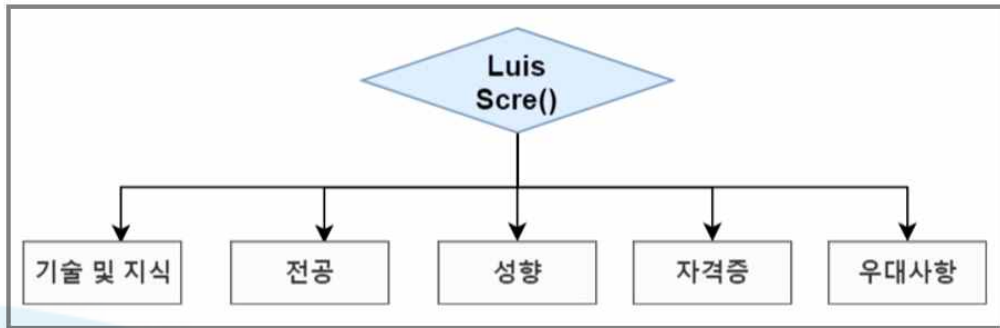
다) 직업정보/채용정보의 메인 Dialog

직업정보의 메인 Dialog는 자연어처리에서 분류된 직업정보와 속성 정보를 전달받아 [그림 4-23]과 같이 각각의 분야 해당 항목 처리 프로세스로 전달된다.



[그림 4-23] 직업정보의 메인 Dialog 구성

[그림 4-24]의 채용정보의 메인 Dialog 역시 동일한 작업을 수행하며, 하나의 항목에서 처리하지 않고 각각의 항목으로 분류한 이유는 가중치 값에 대한 계산 및 활용 정보가 차이가 있으며, 데이터에 따라 표현과 답변을 달리 해줘야 하는 부분이 있기 때문이다.



[그림 4-24] 채용정보의 메인 Dialog 구성

라) 채용정보의 기술 및 지식 항목 Dialog

직업 항목의 Dialog는 전달된 값을 기준으로 사이퍼 쿼리문을 수행한다. 해당 항목의 매칭 값은 '(직업)-[:채용정보]->()-[:지식및기술 리스트]-()-[REL:지식및기술]-(지식및기술 항목)'이며 반환되는 값은 직업, 지식 및 기술, 관계의 가중치 값이다.

(마) 쿼리 된 역량 리스트 및 가중치 분석

[표 4-21]과 같이 검색된 전체 기술 및 지식 역량 및 가중치 값의 수집을 수행하면서 리스트 형태로 저장하며, 가중치 노드의 개수를 구하고 전체 가중치 노드를 합산하여 저장하며, 직업 역량 노드와 가중치 노드에 대한 정렬을 수행한다.

[표 4-21] 역량 리스트 및 노드 가중치 분석

[챗봇에서의 역량 리스트 쿼리]

```
foreach (var 직업 in result){    // 반복문(해당 직업 노드)
    기술및지식=a.KnwSkl.ToList(); // 지식 및 기술-> 리스트로 저장
    가중치 = a.weight.ToList();    // 가중치 -> 리스트로 저장
    foreach (var bb in knwSkl)    // 연결된 역량 정보 전체 반복
    { 채용정보.가중치[i++]        // 가중치 합산
        foreach (선택 역량노드전체) // 연결된 역량 노드의 가중치 총 합산
        { sum += w.Weight; }
        가중치의 합 = sum;
    }
}
```

(바) 결과 메시지 및 가중치 계산

사용자에게 전달되는 결과 메시지는 [표 4-22]와 같이 사용자에게 전달된다. 전체 연결된 노드에 대해 반복 수행하면서 관련된 직업 이름과 가중치를 출력하며 가중치 값은 해당 노드의 역량 합산 값에서의 해당 역량 가중치 평균값으로 계산하여 백분율로 출력한다.

[표 4-22] 역량에 대한 관련 직업 답변

[챗봇에서의 역량에 대한 답변과 가중치 평균값 계산]

```
foreach (var job in jobList)    // 전체 연결 노드 반복
{
    print(역량 이름)            // 역량 이름
    print(가중치)                // 가중치 값 *
    foreach (weight in job.Node) // 연관된 노드 가중치 값에 대한 평균
    {                             계산
        역량노드 수 += weight.Name;
        역량노드 수 +=
        (역량가중치/역량합산*100);
    }
}
```

2) 자연어처리 학습

지식 그래프의 지식은 채용정보와 직업정보로 나뉘어 있으며 각각의 지식이 서로 틀리지만, 비슷한 부분이 있으므로 자연어처리 모델을 2개 구성하여 가장 높은 적중률을 기준으로 조언을 할 수 있도록 구성하였다.

가) 발화 의도(Intent) 학습

[그림 4-25]는 채용정보의 LUIS에서 발화 의도 목록이다. 기술 및 지식, 성향, 우대사항, 자격증, 전공으로 나뉘 266개의 발화 구문을 만들어 학습을 진행하였으며 세부 학습 구문은 [그림 4-26]과 같다.

[그림 4-27]은 직업정보에 대한 LUIS 발화 의도 목록이다. 가치관, 교육/자격/훈련, 성격, 요약, 임금, 전망, 직업만족도, 하는일, 흥미 분류로 나뉘 총 845개의 발화 구문을 만들어 학습을 진행하였다. 그러나 LUIS의 특성상 데이터 크기가 증가하면서 과적합으로 인한 적중률 감소 문제 역시 발견되었다.

발화 의도는 [그림 4-26]과 같이 이용자가 질문할 수 있는 항목에 대해 질문 리스트를 만들고 그 리스트에 해당하는 값을 학습시키는 방식으로 진행하였다.

Intents ?	
+ Create new intent + Add prebuilt domain intent	
Search intents ...	
<input type="checkbox"/> Name ^	Labeled Utterances
None	15
기술 및 지식	50
성향	50
우대사항	50
자격증	50
전공	51

[그림 4-25] LUIS에서의 채용정보 Intents 항목

전공

This intent has a number of unclear predictions. [Apply filters.](#)

Labelled entities: None

Edit Reassign intent Add as pattern Delete Search Filter View options

<input type="checkbox"/> Example utterance	Score ?
<input type="text" value="Enter an example of what a user might say and hit Enter."/>	
전공 요소 출신 이 학원 강사 나 교 사 말 고 가 질 만 한 직 업 추 천 좀 .	0.94
전공 요소 와 전공 요소 과 를 복 수 전 공 했 는 데 두 전 공 을 모 두 살 리 면 서 할 만 한 직 업 이 뭐 가 있 어 ?	0.98
전공 요소 단 일 전 공 으 로 어 느 분 야 로 취 업 되 는 지 알 려 주 세 요 .	0.99
전공 요소 과 취 업 어 디 로 됨 ?	0.99
전공 요소 과 나 오 면 어 떤 직 업 을 가 질 수 있 을 까 요 ?	0.95
기 술 및 지 식 요 소 전 문 가 가 되 고 싶 은 데 전공 요소 과 가 관 련 이 있 을 까 요 ?	0.98
전공 요소 과 가 갈 수 있 는 영 화 관 련 직 업	0.96
전공 요소 과 랑 전공 요소 과 의 점 점 이 되 는 직 업 이 뭐 가 있 을 까 요 ?	0.98

[그림 4-26] 자격증 항목의 발화 학습 내용

Intents ?

Create new intent Add prebuilt domain intent

<input type="checkbox"/> Name ^	Labeled Utterances
None	95
가치관	94
교육/자격/훈련	95
성격	94
요약	92
임금	96

[그림 4-27] LUIS에서의 직업정보 Intents 항목

나) 엔티티(실체) 학습

[그림 4-28]과 같이 질문에 대한 실체 쪼는 핵심어에 해당하는 엔티티의 경우 적중률을 높이기 위해서 리스트 타입으로 328개의 기술 및 지식 요소,

성향 요소, 우대사항 요소, 자격증 요소, 전공 요소, 직업 이름 요소에 대한 학습을 수행하여 지식 그래프에 있는 노드와 관계에 대한 자연어 질문에 답변할 수 있게 자연어처리 모델을 배포하였다.

기술 및 지식 요소

Entity type: List

Values

Add new sublist ... Import values Search ... ☐ Exact match

☐ Normalized Value Synonyms ?

Python	파이썬 ×
golang	Golang × GoLang ×
C#	Csharp ×
Java	자바 × JAVA ×

[그림 4-28] 기술 및 지식요소 실체(엔티티) 학습 내용

다) 학습 결과 테스트 수행

(1) 채용정보 질문에 대한 분류 및 적중률

[표 4-23]과 같이 ‘웹마스터는 어떤 기술이 필요해’라는 질문에 intants:기술 및 지식, Entities:직업이름(웹마스터)로, ‘웹프로그래머 우대사항은 뭐야’라는 질문에 intants:우대사항, Entities:직업이름(웹프로그래머)과 같이 높은 점수로 정확히 분류 되었다.

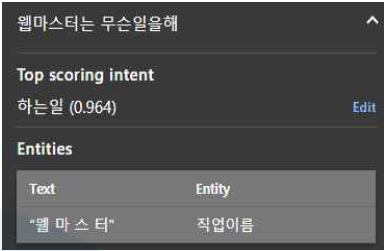

[표 4-23] 채용정보 질문에 대한 분류와 적중률

질문	웹마스터는 어떤 기술이 필요해	웹프로그래머 우대사항은 뭐야
의도/엔티티 분류		
적중률	0.964	0.957

(2) 직업 정보 질문에 대한 분류 및 적중률

[표 4-24]와 같이 ‘웹마스터는 무슨 일을해’라는 질문에 intents:하는일, Entities:직업이름(웹마스터)로, ‘웹프로그래머 우대사항은 뭐야’라는 질문에 intents:우대사항, Entities:직업이름(웹프로그래머)과 같이 높은 점수로 정확히 분류 되었다.


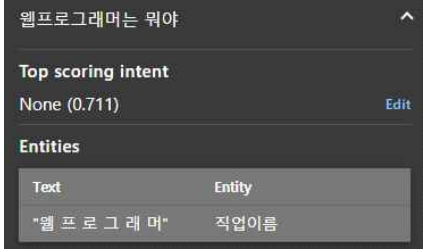
[표 4-24] 직업정보 질문에 대한 분류와 적중률

질문	웹마스터는 무슨일을해	웹프로그래머 우대사항은 뭐야
의도/엔티티 분류		
적중률	0.864	0.957

(3) 동일한 질문에 대한 2개의 LUIS 서비스 응답 결과

채용정보와 직업정보를 학습시킨 2개의 LUIS 서비스에 동일한 질문을 던져 보았다. [표 4-25]과 같이 점수가 높은 직업정보로 Intent:요약, Entities:직업이름(웹프로그래머) 답변 결과를 응답하며 챗봇은 해당 결과를 기반으로 지식 그래프에 쿼리하여 답변을 생성한다.

[표 4-25] 동일한 직업정보 질문에 대한 분류와 적중률

LUIS	직업정보	채용정보
질문	웹프로그래머는 뭐야?	웹프로그래머는 뭐야?
의도/엔티티 분류		
적중률	0.934	0.711

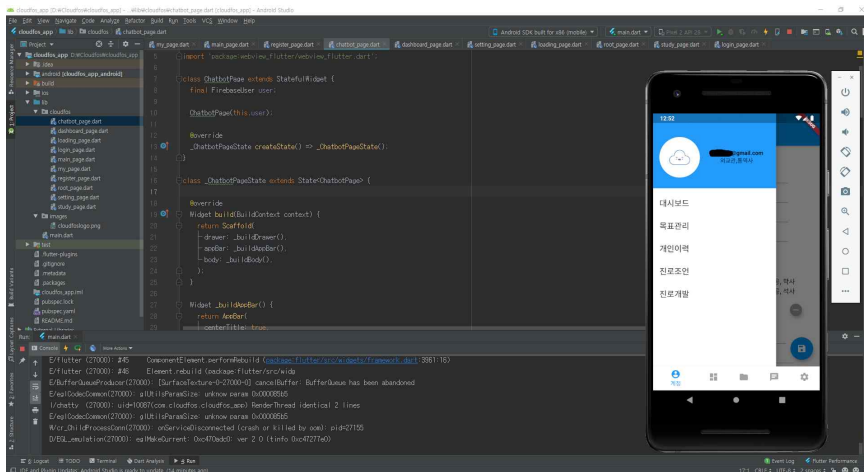
3) e-포트폴리오 앱 환경 구현

가) e-포트폴리오 구현 환경

Flutter는 Google에서 개발한 크로스 플랫폼 모바일 앱 개발 프레임워크이다. 모바일 환경은 크게 Android와 iOS로 나누어져 있으므로 하나의 코드를 여러 플랫폼에서 사용할 수 있게 해주는 기술인 크로스 플랫폼과 디자인 엔진 탑재를 통한 UI의 구현이 간편한 장점과 이용자 수가 적은 Dart 언어의 사용으로 지원되는 플러그인이 매우 적은 문제가 있지만, 본 연구에서는 크로스 플랫폼과 사용자 데이터 매칭을 위한 간단한 e-포트폴리오 모바일 앱 개발을 목표로 하고 있으므로 Flutter를 활용하여 간단히 사용자 관리, 개인 이력 및 역량 입력, 입력된 데이터를 활용한 직업 추천 기능과 iframe을 활용하여 챗봇 기능을 구현하였다.

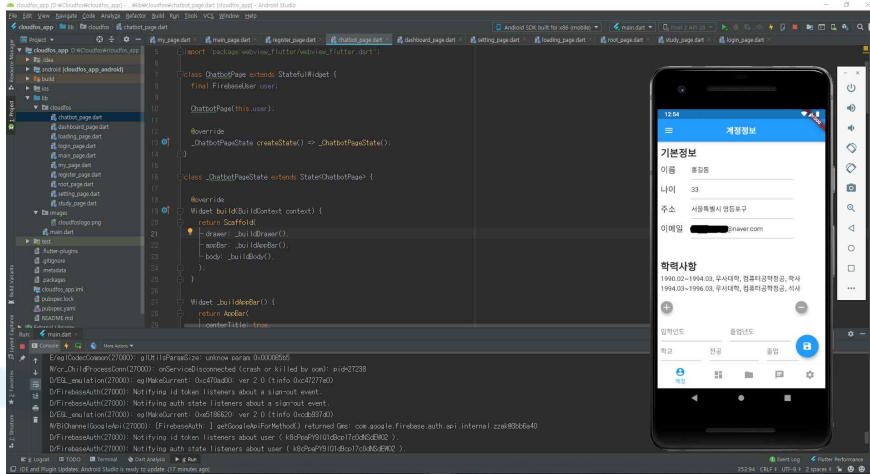
나) e-포트폴리오 구현 및 기능

[그림 4-29]는 목표관리, 개인이력, 진로 조언 챗봇, 진로 개발이라는 좌측과 하단메뉴, 상단에 이용자의 계정 정보와 목표 직업을 표시할 수 있도록 구현하였다.



[그림 4-29] e-포트폴리오 메뉴 구성

[그림 4-30]은 개인정보 및 사용자 역량 입력 메뉴이며, 여기서 사용자의 개인 이력과 학력, 기술, 자격증, 성향과 같은 역량을 입력하고, 역량의 관리와 필요시 사용자의 역량을 바탕으로 한 매칭되는 직업 추천 기능을 제공하도록 개발되었으며, 구현된 기능은 [표 4-26]과 같다.



[그림 4-30] 개인정보 및 이력 입력

[표 4-26] e-포트폴리오 앱 기능

구분	설명
챗봇	대학생들에게 가장 친숙한 대화형 인터페이스를 활용한 진로 조언 및 진로 관련 정보 제공
이력 정보	자신의 현재 상태를 입력함으로써 진로 개발을 위한 매칭 데이터로 활용
진로개발	진로 목표를 찾기 위한 정보 제공
목표관리	현재 목표와 목표 달성을 위한 세부 항목들을 표시
대시보드	직무개발 목표와 현재 상태를 시각적으로 표현하고, 이를 통해 학생의 진로 개발을 지원

구현된 e-포트폴리오에서는 실제 그래프 데이터베이스에 대한 직접 쿼리를 위한 드라이버를 제공하고 있지 않기 때문에 데이터 매칭을 위한 별도의 API 서버를 구현하여 데이터 매칭을 통한 추천 기능을 구현하였다.

다) API 서버 구현

[그림 4-31]과 같이 API 서버는 외부에서 그래프 데이터베이스를 환경과

정보 전달을 위한 메서드는 GetJobmatch로 지정하며 추천 직업 검색을 위한 HTTP 요청 쿼리와 결과는 [표 4-27]과 같다.



[표 4-27] API 서버에 직업 역량 호출 결과

JSON

```
1 [
2 {
3     "JobNm": "컴퓨터보안전문가",
4     "count": 4
5 },
6 {
7     "JobNm": "빅데이터 전문가",
8     "count": 4
9 },
10 {
11     "JobNm": "인공지능 개발",
12     "count": 4
13 },
14 {
15     "JobNm": "프론트엔드개발",
16     "count": 3
17 },
18 ]
```

- 128 -

여 15개만 출력되도록 구현하였다.

[표 4-28] Get으로 전달받은 값을 쿼리하여 값을 전달하는 알고리즘

[전달받은 값을 리스트를 활용해 관계된 직업 노드 호출]

```
.With($" {list} as abil")
.Match("(n:Job)-[*]->(c)")
.Where("c.Name in abil")
.Return(n => new { JobNm = n.As<Job>().JobNm, count = n.Count() })
.OrderByDescending("COUNT(n)")
```

제 3 절 진로 조언 플랫폼 기능 테스트

1) 챗봇 인터페이스를 활용한 조언 기능 테스트

가) 직업정보 조언 기능 테스트 결과

챗봇의 직업정보 지식 조언 기능 검증을 위해 [표 4-29]와 같이 기능별 질문을 추출하여 질문과 답변 내용을 테스트하였다.

[표 4-29] 직업 정보 조언 질문 목록

[직업정보에 대한 질문 목록]

- (1) 어떤 흥미를 가져야 웹프로그래머가 될 수 있어? (흥미)
- (2) 웹프로그래머는 뭐야? (요약)
- (3) 웹프로그래머는 어떤일을 해? (하는일)
- (4) 웹프로그래머의 전망은 어때? (전망)
- (5) 웹프로그래머의 연봉은 얼마야? (연봉)
- (6) 웹프로그래머의 만족도는 어때? (만족도)

<p>(1) 어떤 흥미를 가져야 웹프로그래머가 될 수 있어?</p>	<p>(2) 웹프로그래머는 뭐야?</p>
<p>어떤 흥미를 가져야 웹프로그래머가 될 수 있어?</p> <p>Just now</p> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>-----관련 흥미 목록-----</p> <p>사회형(Social) - 다른 사람들을 후원시키고, 발달시키고, 치료해주기 위한 활동들을 선호하고, 자로나 도구 혹은 기계를 포함하는 명확하고 체계적인 활동을 싫어하는 경향이 있다. 이러한 경향성은 사회적 및 교육적 능력에서의 능력을 획득하게 하는 반면에 조직, 기계, 논리, 전기 및 기술적인 능력들에서는 결함을 보여준다. - 37점</p> <p>진취형(Enterprising) - 조직적인 목표나 경제적인 이익을 얻기 위한 다른 사람과의 상호작용 활동을 선호하고, 관찰적이고, 상징적이며, 체계적인 활동을 싫어하는 경향이 있다. 이러한 경향성은 리더십, 대인관계 능력 및 설득적인 능력을 획득하게 하는 반면에 과학적인 능력에서 결함을 보여준다. - 75점</p>	<p>웹프로그래머는 뭐야?</p> <p>Just now</p> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>요약 설명 : 웹상에서 각종 자료들을 보여줄 수 있도록 웹 프로그래밍 언어를 이용하여 프로그램을 설계하고 작성한다.</p>
<p>(3) 웹프로그래머는 어떤 일을 해?</p>	<p>(4) 웹프로그래머의 전망은 어때?</p>
<p>웹프로그래머는 어떤 일을 해?</p> <p>Just now</p> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>하는 일 : -웹상에 올릴 자료의 성격과 형태에 대한 요구를 파악하고 분석한다. -프로그래밍언어(Php, Asp, Java, Jsp 등)를 이용하여 프로그램을 코딩한다. -웹상에서 테스트한 후 문제점을 확인하고 수정한다. -기존에 개발된 프로그램을 유지 및 보수한다. -웹디자이너와 업무를 협의한다.</p>	<p>웹프로그래머의 전망은 어때?</p> <p>Just now</p> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>전망 설명 : 향후 5년간 웹프로그래머의 고용은 증가할 것으로 전망된다. 「중장기 인력수급 수정전망2015~2025」(한국고용정보원, 2016)에 따르면, 웹 및 멀티미디어기획 및 개발자는 2015년 약 16천 명에서 2025년 약 19천 명으로 향후 10년간 약 3천 6백 명(연평균 2.1%) 증가할 것으로 전망된다. 블로그와 소규모 사업자의 사이트가 늘어나는 등 인터넷 웹사이트 수가 급증하고 있다. 또한, 집에서 홈네트워크를 이용해 각종 민원 업무를 보거나 교육정보, 생활정보를 얻는 등 인터넷 웹사이트를 통한 정보교환, 금융거래, 상거래 등도 향후 더욱 가속화될 것이다. 이처럼 웹사이트 이용자가 늘어남에 따라 인터넷을 이용한 다양한 응용프로그램과 콘텐츠의 개발 필요성이 커지면서 웹프로그래머의 수요가 늘어날 것으로 보인다. 기업이 마케팅이나 홍보수단으로 웹을 활용하는 것도 일반화되고 있어 대기업이나 공공기관의 웹을 구축하는 SI업체에서의 웹프로그래머 고용도 꾸준할 것이다. 또한, 기존의 웹을 주기적으로 보완·개편하는 사례가 늘어남에 따라 관련 전문가의 고용에 긍정적 영향을 미칠 것으로 보인다. 그러나 업체에서는 실력 있는 경력자를 원하고 있어 신규인력의 취업 난이도 예상된다.</p> <p>감소 - 6%</p>
<p>(5) 웹프로그래머의 연봉은 얼마야?</p>	<p>(6) 웹프로그래머의 만족도는 어때?</p>
<p>웹프로그래머의 연봉은 얼마야?</p> <p>Just now</p> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>상위 25% : 4836만원 평균 : 하위 25% : 3241만원 조사기준년도 - 2018년</p>	<p>웹프로그래머의 만족도는 어때?</p> <p>Just now</p> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>만족도 : 64.1</p>

나) 채용정보 조연 기능 테스트

챗봇의 채용정보 지식 조연 기능 검증을 위해 [표 4-30]과 같이 각 기능별 질문을 추출하여 질문과 답변 내용을 테스트하였다.

[표 4-30] 채용 정보 조연 질문 목록

[채용정보에 대한 질문 목록]



- (1) 웹프로그래머는 어떤 기술을 가져야 해? (기술 목록)
- (2) java기술을 가진 사람은 어디에 취업할 수 있어 (직업)
- (3) 진취적인 성격은 어느 직업에서 뽑아가? (성격)
- (4) 웹프로그래머는 어떤 사람을 우대해줘? (우대사항)
- (5) 중국어 잘하는 사람을 우대해주는 기업은? (우대사항)
- (6) 웹프로그래머 자격증 (자격증)

(1) 어떤 흥미를 가져야 웹프로그래머가 될 수 있어?	(2) 웹프로그래머는 뭐야?
<p>어떤 흥미를 가져야 웹프로그래머가 될 수 있어?</p> <p>웹프로그래머로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>-----관련 흥미 목록-----</p> <p>사회형(Social) - 다른 사람들을 훈련시키고, 발달시키고, 치료해 주기 위한 활동을 선호하고, 자료나 도구 혹은 기계를 포함하는 명확하고 체계적인 활동을 싫어하는 경향이 있다. 이러한 경향성은 사회적 및 교육적 능력에서의 능력을 획득하게 하는 반면에 조작, 기계, 농경, 전기 및 기술적인 능력들에서는 결함을 보여준다. - 37 점</p> <p>진취형(Enterprising) - 조직적인 목표나 경제적인 이익을 얻기 위한 다른 사람과의 상호작용 활동을 선호하고, 관찰적이고, 상징적이며, 체계적인 활동을 싫어하는 경향이 있다. 이러한 경향성은 리더십, 대인관계 능력 및 설득적인 능력을 획득하게 하는 반면에 과학적인 능력에서 결함을 보여준다. - 75 점</p>	<p>웹프로그래머는 뭐야?</p> <p>웹프로그래머로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>요약 설명 : 웹상에서 각종 자료들을 보여줄 수 있도록 웹 프로그래밍 언어를 이용하여 프로그램을 설계하고 작성한다.</p>

<p>(3) 웹프로그래머는 어떤 일을 해?</p>	<p>(4) 웹프로그래머의 전망은 어때?</p>
<div> <div>웹프로그래머는 어떤 일을 해?</div> <div>Just now</div> <div> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>하는 일 : -웹상에 올릴 자료의 성격과 형태에 대한 요구를 파악하고 분석한다. -프로그래밍언어(Php, Asp, Java, Jsp 등)를 이용하여 프로그램을 코딩한다. -웹상에서 테스트한 후 문제점을 확인하고 수정한다. -기존에 개발된 프로그램을 유지 및 보수한다. -웹디자인과 업무를 협의한다.</p> </div> </div>	<div> <div>웹프로그래머의 전망은 어때?</div> <div>Just now</div> <div> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>전망 설명 : 향후 5년간 웹프로그래머의 고용은 증가할 것으로 전망된다. 「중장기 인력수급 추정전망2015~2025」(한국고용정보원, 2016)에 따르면, 웹 및 멀티미디어기획 및 개발자는 2015년 약 16천 명에서 2025년 약 19천 명으로 향후 10년간 약 3천 6백 명(연평균 2.1%) 증가할 것으로 전망된다. 블로그와 소규모 사업자의 사이트가 늘어나는 등 인터넷 웹사이트 수가 급증하고 있다. 또한, 집에서 홈네트워크를 이용해 각종 민원 업무를 보거나 교육정보, 생활정보를 얻는 등 인터넷 웹사이트를 통한 정보교환, 금융거래, 상거래 등도 향후 더욱 가속화될 것이다. 이처럼 웹사이트 이용자가 늘어남에 따라 인터넷을 이용한 다양한 응용프로그램과 콘텐츠의 개발 필요성이 커지면서 웹프로그래머의 수요가 늘어날 것으로 보인다. 기업이 마케팅이나 홍보수단으로 웹을 활용하는 것도 일반화되고 있어 대기업이나 공공기관의 웹을 구축하는 si업체에서의 웹프로그래머 고용도 꾸준할 것이다. 또한, 기존의 웹을 주기적으로 보완·개선하는 사례가 늘어남에 따라 관련 전문가의 고용에 긍정적 영향을 미칠 것으로 보인다. 그러나 업체에서는 실력 있는 경력자를 원하고 있어 신규인력의 취업난이 예상된다.</p> <p>감소 - 6%</p> </div> </div>
<p>(5) 웹프로그래머의 연봉은 얼마야?</p>	<p>(6) 웹프로그래머의 만족도는 어때?</p>
<div> <div>웹프로그래머의 연봉은 얼마야?</div> <div>Just now</div> <div> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>상위 25% : 4836만원 평균 : 하위 25% : 3241만원 조사기준년도 - 2018년</p> </div> </div>	<div> <div>웹프로그래머의 만족도는 어때?</div> <div>Just now</div> <div> <p>웹프로그래머 으로 검색한 결과는 총 1건으로 다음과 같습니다.</p> <p>직업 이름 : 웹프로그래머</p> <p>만족도 : 64.1</p> </div> </div>

2) e-포트폴리오 앱 기능 테스트

e-포트폴리오 앱에서의 메뉴 구성과 학생 자신의 역량 정보를 입력하여 적합한 직업을 추천받는 맞춤형 추천 기능에 대한 기능 테스트 수행과 챗봇을 활용한 조언 기능 테스트를 수행하였다.

(1) 전체 메뉴	(2) 자신의 역량 입력 1/2
	

(3) 자신의 역량 입력 2/2	(4) 맞춤형 직업 추천
	
(5) 챗봇을 활용한 진로 조언 1/2	(6) 챗봇을 활용한 진로 조언 2/2
	

3) 사용자(대학생) 설문 평가

가) 설문 평가 개요

개발된 ‘진로 조언’ 플랫폼의 활용성은 개개인의 주관적인 평가에 의존할 수밖에 없다. 따라서 실제 사용자 계층인 정보통신 분야 대학생 21명의 평가단을 모집하고, 실제 진로 조언 플랫폼을 활용하여 50개의 문항에 대한 정보 검색을 통해 정보 검색의 활용성, 편의성, 맞춤형 검색, 진로 조언을 위해 활용 의향을 평가하였다.

설문은 조사대상자가 응답할 범위를 다양화함으로써 타당도를 높이고 상대적으로 안정된 서열구조로 응답할 수 있도록 만든 리커트척도를 활용하였으며, 스마트폰 및 PC를 활용한 정보 검색과 진로 조언 플랫폼의 비교를 통해 정보검색 평균 속도와 편의성에 대한 평가를 추가하였다.

나) 평가단 선정

개발된 진로 조언 플랫폼은 정보기술 전공 대학생을 대상으로 하므로 평가단의 구성은 연구의 성격을 대표할 수 있는 서울 소재 정보기술 전공 대학생 21명을 모집하였다.

연구를 위한 평가단 선정은 서울 소재 4년제 대학의 대나무숲을 활용하여 모집 공고를 등록했으며, 모집 공고에는 연구 제목과 목적, 연구 장소, 참여 조건 및 혜택, 연구 참여에 요구되는 작업 및 소요시간을 명시하였다. 신청 후 실험에 참여한 전원에게 연구에 대한 사전 설명 후 서면 동의를 구하였다.

다) 평가단의 구성과 진로에 대한 조사

[표 4-31]과 같이 평가단은 진로에 고민이 많은 3학년생 위주로 설문을 진행하였으며 총 21명의 평가단을 구성하여 진로 조언 플랫폼을 직접 이용해 보면서 설문을 진행하였으며, 조사 결과 76%에 해당하는 평가단이 아직 진로를 결정하지 못한 것으로 확인되었다. 졸업 후 진로에 대한 답변으로는 85.7%의 학생이 취업을 선택하였다.

[표 4-31] 사용자 평가단 구성 및 진로 정보

학년 구분	2학년	3학년	4학년	전체
구분(백분율)	2(9.5%)	14(66.7%)	5(23.8%)	21(100%)
진로결정 여부	예	아니오	고민 중	전체
구분(백분율)	4(19%)	1(4.8%)	16(76.2%)	21(100%)
졸업 후 진로	취업	진학	창업	전체
구분(백분율)	18(85.7%)	3(14.3%)	0	21(100%)

라) 진로 조언 플랫폼의 기능 만족도 평가

[표 4-32]와 같이 진로 조언 플랫폼의 기능 만족도에 대한 평가로는 도움(57.1%), 매우도움(28.6%) 순이었으며, 도움이 되지 않는다(2%)의 의견도 있었다.

직업의 종류와 세계에 대한 답변은 도움(38.1%), 보통(33.3%), 매우도움(28.6%) 순으로 조사되었으며, 현실과 다른 표준직업 코드로 분류된 직업명을 통해 데이터를 수집했기 때문에 직업 검색이 효율적이지 않았던 문제점으로 보통으로 평가된 응답이 많았다.

진로 개발을 위한 학습정보 답변 만족도에서는 도움(57.1%), 매우도움(33.3%), 보통(9.5%) 순으로 실제 기업의 채용정보를 수집한 답변을 통해 학생들이 진로 개발을 위한 학습정보를 얻을 수 있는 것으로 확인되었다.

취업 준비 과목 관련 답변 만족도로는 이전의 진로 개발을 위한 학습정보와 비슷한 유형이었기 때문에 동일하게 답변되었다.

챗봇을 활용한 취업 준비 과목 관련 전문가 멘토링 기능에서는 보통(47.6%), 도움(42.9%), 매우도움(9.5%) 순으로 현 챗봇으로는 멘토링과 같은 기능 구현으로는 부족한 것으로 확인되었다.

챗봇을 활용한 진로·직업 멘토링 관련 답변으로는 도움(42.9%), 보통(38%), 매우도움(19%) 순으로 도움이 된다는 의견이 많았으며, 별도로 도움이 되지 않는다는 의견(4.8%)도 있었다.

[표 4-32] 기업의 채용정보 분석 답변 만족도

구분	매우 도움	도움	보통	도움이 되지 않음
채용정보 분석 답변 만족도	6(28.6%)	12(57.1%)	1(4.8%)	2(9.5%)
직업의 종류와 세계 답변 만족도	6(28.6%)	8(38.1%)	7(33.3%)	0(0%)
진로 개발을 위한 학습정보	7(33.3%)	12(57.1%)	2(9.5%)	0(0%)
취업 준비과목 관련 답변	7(33.3%)	12(57.1%)	2(9.5%)	0(0%)
챗봇을 활용한 취업 준비과목 전문가 멘토링 기능	2(9.5%)	9(42.9%)	10(47.6%)	0(0%)
챗봇을 활용한 진로·직업 멘토링 기능	4(19%)	9(42.9%)	7(33.3%)	1(4.8%)

마) e-포트폴리오 기능 만족도 평가

[그림 4-33]의 e-포트폴리오 기능의 경우 매우도움(33.3%), 도움(57.1%), 보통(9.5%)으로 e-포트폴리오를 활용한 나의 역량 관리와 역량 기반 맞춤형 직업 추천 기능에 대한 만족도가 높았다.

e-포트폴리오에서 직업 추천을 기반으로 개발해야 할 역량을 목표로 설정하고, 역량 달성을 위한 강좌 추천과 자신이 목표로 하는 직업에 자신이 어느 정도에 있는지에 대한 위치를 알 수 있는 기능이 추가된다면 76.2%가 매우 도움이 될 것이라고 답변하였으며, 61.9%가 적극 활용, 8%가 활용할 것이라 응답한 것과 같이 학생들은 자신들의 현 수준과 역량 개발에 대한 요구가 높았다.

[표 4-33] e-포트폴리오 기능 및 활용여부

구분	매우 도움	도움	보통
e-포트폴리오 기능(현재)	7(33.3%)	12(57.1%)	2(9.5%)
e-포트폴리오 기능 (진로 개발 기능 완료 후)	16(76.2%)	4(19%)	1(4.8%)
e-포트폴리오 적극 활용 여부 (진로 개발 기능 완료 후)	13(61.9%)	8(38.1%)	0(0%)

바) 챗봇을 활용한 진로·직업 정보에 대한 만족도

(1) 챗봇 조언 결과에 대한 만족도

[표 4-34]와 같이 조언 결과에 대한 만족도는 만족(57.1%), 매우만족(28.6%), 보통(14.1%) 순으로 확인되었다.

채용정보 결과의 경우 보통(52.4%), 매우도움(23.8%), 보통(14.3%) 순이며, 직업정보 결과의 경우 도움(42.9%), 매우도움(38.1%), 보통(19%) 순으로 조사된 것처럼 채용정보 분석에 대한 만족도가 조금 더 높았다.

[표 4-34] 챗봇 조언 결과에 대한 만족도

구분	매우만족	만족	보통	도움이 되지 않음
진로 조언의 결과 만족도	6(28.6%)	12(57.1%)	3(14.3%)	0(0%)
채용정보 결과 만족도	5(23.8%)	11(52.4%)	3(14.3%)	2(9.5%)
직업정보 결과 만족도	8(38.1%)	9(42.9%)	4(19%)	0(0%)

(2) 진로설계 시 도움과 적극 활용 여부

[표 4-35]와 같이 챗봇을 활용한 진로 조언 인터페이스에 대한 만족도에 서는 도움(51.1%), 매우도움(33.3%), 보통(9.5%)로 챗봇을 활용한 인터페이스 만족도는 높은 것으로 확인되었으며, 진로 조언 플랫폼이 진로설계 시 도움이 될 수 있는지에 대한 답변으로는 도움(52.4%), 매우도움(33.3%), 보통(14.3%)로 조사 되었다. 진로 조언 플랫폼의 지식 확장과 같은 모든 기능이 완성된다면 적극적으로 활용할 생각에 대한 질문에는 적극 활용(71.4%), 활용(28.6%) 순으로 조사된 것과 같이 학생들을 위한 챗봇을 활용한 진로 조언 플랫폼의 추가적인 연구와 도입이 필요한 것으로 나타났다.

[표 4-35] 챗봇 인터페이스 만족도와 활용 여부

구분	매우도움	도움	보통
챗봇 인터페이스 만족도	7(33.3%)	12(57.1%)	2(9.5%)
진로 설계 시 도움 여부	7(33.3%)	11(52.4%)	3(14.3%)
진로 조언 플랫폼 적극 활용 여부 (플랫폼 완성 후)	15(71.4%)	6(28.6%)	0(0%)

사) 기능 구현 정도에 대한 답변

[표 4-36]과 같이 직업정보 제공기능의 경우 잘 구현되었다(66.7%), 매우 잘 구현되었다(28.6%), 구현되었다가(4.8%) 순으로 답변하였으며, 기업 채용 정보 분석과 진로·직업 상담의 경우도 위의 직업정보 제공기능과 동일하게 구현되었다고 답변이 이루어진 바와 같이 기능 구현은 잘 되었다는 응답이 높았다.

[표 4-36] 기능 구현 정도에 대한 답변

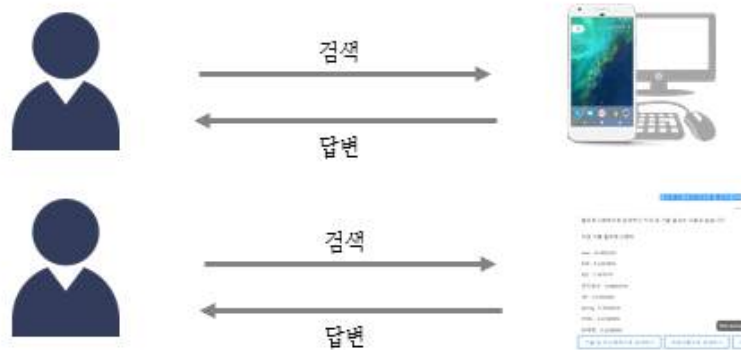
구분	매우 잘 구현되었다	잘 구현되었다	구현되었다
직업정보 제공 기능 구현정도	6(28.6%)	14(66.7%)	1(4.8%)
채용정보 제공 기능 구현 정도	6(28.6%)	14(66.7%)	1(4.8%)
진로·직업 상담 구현 정도	7(33.3%)	12(57.1%)	2(9.5%)

4) 사용자(대학생) 검색 속도 비교 테스트 및 결과

정보 검색 시간 및 검색 만족도 평가는 [그림 4-32]와 같이 12개의 주어진 진로(직업) 정보를 사용자의 모바일 웹과 진로 조언 플랫폼을 번갈아 가며 사용하여 정보를 검색하고 검색 시간과 만족도를 기재하는 방식으로 평가를 시행하였으며, 실험 참가자는 위의 설문 조사 평가단 21명이 설문이 끝난 후 테스트에 참여하였다.

기존 방식은 직업정보 검색은 참가자들이 자주 애용하는 스마트폰 또는 PC와 챗봇 방식을 번갈아 가며 검색속도 비교 테스트와 답변의 만족도, 편의성에 대한 테스트를 수행하였다.

주어진 진로(직업) 정보 검색 (11문항)



[그림 4-32] 검색 시간, 결과 만족도, 편의성 테스트 방식

테스트 결과 기존 PC나 스마트폰을 활용한 정보 검색보다 챗봇을 활용하는 방안이 훨씬 빠른 검색 속도를 보여주는 것으로 확인되었다. 하지만, 진로 조언을 위한 챗봇은 해당 기능을 수행하기에 최적화된 데이터 구조와 봇-프로그램이 설계되었기 때문에 이러한 속도 비교는 유의미한 결과로는 부족하지만, 특정된 기능 수행을 위해 개발된 정보 검색에서는 기존의 방식보다 빠른 정보를 찾을 수 있는 것으로 확인되었다.

[표 4-37]과 [표 4-38] 같이 직업정보 답변과 채용정보 답변도 기존 방식보다 높게 나왔으며, 이 역시 특정 기능을 수행하기 위해 개발된 챗봇의 만족도가 높았다.

[표 4-39] 의 정보 검색 편의성은 챗봇(85%), 기존검색(65.59%)으로 챗봇이 기존의 스마트폰이나 PC보다 편의성 부분에서는 더 높게 응답한 바와 같이 직업정보와 같이 특화된 분야의 정보 검색에서는 챗봇의 형태의 인터페이스의 편의성이 높은 것으로 확인되었다.

[표 4-37] 직업정보 검색 속도 평균

질문	기존 방식 평균	챗봇 방식 평균
[\$직업이름]은 뭐야	9.6초	1.5초
[\$직업이름]은 어떤 일을 해?	22초	1.8초
[\$직업이름]의 전망은 어때?	22.6초	2.1초
[\$직업이름]의 연봉은 얼마야?	17.6초	1.8초
[\$직업이름]의 만족도는 어때?	33초	2초
평균	21초	1.84초

[표 4-38] 채용정보 검색 속도 평균

질문	기존 방식 평균	챗봇 방식 평균
[\$직업이름]은 어떤 자격증이 있어야 해?	19.3초	2.3초
java기술을 가진 사람은 어떤 직업이 좋아?	105.3초	2.5초
[\$직업이름]은 어떤 성향을 가진 사람을 원해?	88.8초	2.75초
진취적인 성격은 어느 직업에서 원해?	82.4초	2.56초
[\$직업이름]은 어떤 사람을 우대해줘?	49.2초	2.31초
중국어 잘하는 사람을 우대해주는 직업은?	77초	2.75초
평균	70.3초	2.52초

[표 4-39] 직업정보 답변의 만족도와 정보 검색 편의성

질문	기존 방식 평균	챗봇 방식 평균
직업정보 답변의 만족도	72.82%	87.82%
채용정보 답변의 만족도	72.86%	82%
정보검색 편의성	65.59%	85%

5) 연구의 성과

본 연구를 통해 대학생 진로 문제 해결을 위한 진로 조언 플랫폼을 설계하고 이를 직접 구축하여 실제 이용자인 대학생들을 대상으로 직접 이용 테스트와 평가를 통해 진로 조언 플랫폼의 기능 만족도 85.7%로 매우 높았으며, 채용정보 분석과 직업정보 수집을 통해 구축된 챗봇 서비스의 답변 만족도는 70~85%대로 그래프 데이터베이스를 활용한 학생이 찾고자 하는 직업

의 관계를 및 역량을 기준으로 적합한 직업을 제안해 주는 진로 조언 방식의 선호도가 높았다. 챗봇을 활용한 멘토링 방식의 경우 47.6%의 응답자가 보통으로 답변한 것처럼 멘토링 부분에서의 정보기술 활용 연구가 더 필요하다.

e-포트폴리오를 통해 역량을 통한 직업 추천과 역량 관리가 매우 도움이 된다는 의견이 77.2%로 조사되었으며, e-포트폴리오의 역량 개발 기능이 더 추가된다면 활용하겠다(38.1%), 적극 활용하겠다(61.9%)고 높은 기대를 보여주고 있는 것처럼 단순한 정보의 제공이 아닌, 자신의 현재 역량을 기반으로 한 맞춤형 직업 추천과 진로에 대한 역량 개발에 도움 및 활용하기 위한 서비스를 원하는 것으로 확인되었다. 직업정보 및 채용정보 기능에 대해서는 92%의 응답자로부터 대부분 잘 구현되었다 이상으로 평가되었으며, 실제 기존의 검색 방식과 비교하여 검색 테스트를 시행한 결과 검색되는 속도와 답변의 만족도도 월등히 높았다. 하지만, 챗봇의 경우 특정 기능에 최적화된 지식 베이스 설계와 인터페이스가 설계되었기 때문에 직접적인 속도 비교는 무의미할 수 있으나, 진로 조언과 같은 특정 기능을 수행에는 인터페이스로는 효과적인 것으로 확인되었으며 정보 검색 편의성 역시 챗봇 방식이 높았다. 그러나 아직 수집한 지식과 관계의 정보가 많지 않고 특정 문장 중심으로 테스트를 수행했기 때문에 지식 베이스의 확장과 자연어처리 영역에서의 추가적인 연구가 더 필요하다.

제 5 장 결 론

청년 실업의 증가와 직무 능력이 있는 경력 위주의 채용이 증가하는 노동 시장 환경의 변화로 국내 대학생들의 가장 큰 고민은 졸업 후 진로·취업 문제이며, 대학생들의 진로에 대한 고민과 원인에 관한 많은 연구가 진행되고 있지만, 정보기술을 활용한 해결방안을 찾는 연구는 부족했다.

4차 산업혁명과 빠른 정보기술의 발전, 그리고 세대의 변화에 따라 기업이 원하는 역량과 기술 역시 빠르게 변화하고 있지만, 실제 이러한 직업 역량 정보들이 제대로 조사되지 않고 있으며, 실제 대학생들에게 제공되지 못하고 있다. 빠르게 변화하는 사회와 채용 환경의 변화에 따라 학생들이 진로를 결정하고 개발하기 위해서는 직업 관련 정보가 학생들에게 제공되어야 하며, 이러한 정보를 통해 진로 결정과 역량을 개발하기 위한 동기 부여가 될 수 있다. 또한, 기업이 요구하는 역량과 기술의 습득은 학생들의 진로에 대한 고민 감소와 성공적인 취업을 위한 지원할 수 있을 것이며, 기업 역시 요구하는 역량을 가진 직원 채용으로 교육비용 감소와 채용의 증가, 일자리 미스매치 문제를 줄일 수 있는 부가적인 효과를 기대할 수 있다.

본 연구는 대학생 진로 조언 플랫폼 연구의 성과로, 첫째, 진로 및 취·창업 준비를 위해 대학에서 지원받고 싶은 사항에 대한 응답 결과를 기준으로 학생들이 지원받기를 원하는 항목과 파슨스의 특성-요인 이론을 기반으로 정보기술을 활용한 진로 조언 모델을 도출하였다.

둘째, 실제 학생들에게 필요한 직업 역량 데이터를 수집하기 위해 채용정보 포털에서 기업의 채용정보를 수집·분석하여 기업이 특정 직업 채용에서의 기업이 요구하는 역량 데이터 수집·분석 모델을 제시하였다.

셋째, 직업과 직업, 직업과 역량 간의 관계 분석을 위해 그래프 데이터베이스를 활용하여 진로 조언을 위한 지식 그래프를 설계하고 실제 구축하기 위한 그래프 데이터베이스를 활용한 지식 그래프 설계·구축 방법론을 제시하였으며, 실제 구현과 테스트를 통해 효과성을 확인하였다.

넷째, 두 개의 서로 다른 진로 관련 지식의 관계 연결 방안 제시와 구현을 통해 다양한 지식의 관계를 통한 지식 그래프 확장 방안을 제시하였다.

다섯째, 대학생들에게 가장 친근한 챗봇 인터페이스를 활용하여 학생의 자연어 질문을 자연어처리기를 활용하여 발화 의도와 실체(엔티티)를 분류하고, 이를 지식 그래프에 매칭하여 답변을 찾아 이를 분석하여 조언해 줄 수 있는 모델 제시하였다. 여섯째, 학생들의 역량에 따른 맞춤형 직업 추천과 추천된 직업을 대상으로 역량을 검색하고 이렇게 검색된 역량을 학습하고 키워나가기 위한 모바일 e-포트폴리오 모델을 제시하였다.

구축된 모델을 평가하기 위해 구현된 진로 조언 플랫폼의 목표 대상인 정보기술 관련 대학생들을 평가단으로 선정하여 직접 시연과 테스트를 수행하고, 설문문을 통해 활용성, 만족도, 편의성을 평가하였으며, 정보기술을 활용한 진로 조언의 방향성을 도출하였다. 또한, 기존의 모바일 또는 웹검색을 활용한 정보 검색 방식과 챗봇을 활용한 방식과의 정보 검색 시간 테스트를 통해 챗봇을 활용한 방식의 검색 속도와 만족도, 편의성이 모두 높은 것으로 평가된 바와 같이 특정 기능을 수행하는 인터페이스로는 매우 효과적인 것으로 평가되었다.

본 연구는 그래프 데이터베이스를 활용한 진로 조언 플랫폼 구축에 관한 연구이나, 실제 대학 내부의 진로정보나, e-포트폴리오 및 다양한 학사 지원 시스템과 연계는 활용할 수 있는 데이터가 부족하였고, 법률적인 과정이 추가로 필요했기 때문에 연구의 한계성이 있었다.

대학 내부에는 다양한 진로나 학사 관련 정형·비정형 데이터가 저장되어 있다. 일부 대학들의 경우 맞춤형 진로·취업 서비스와 챗봇을 활용한 진로 컨설팅을 위한 고도화 계획을 수립하고 있는 바와 같이 본 연구를 활용하여 진로·취업 데이터 간의 관계를 분석하고 지식 그래프를 구축하여 학생들의 진로 지도에 활용할 수 있다면 실제 대학생들의 진로 결정과 진로역량 개발에 도움이 될 것으로 기대한다.

향후, 본 연구를 기반으로 대학 내 진로 관련 정보에 대한 지식 그래프가 구축되고, 대학 내 진로 관련 자료들이 지식 그래프에 통합되어 실제 학생들의 진로·취업 문제 해결에 직접적인 지원을 위한 추가 연구와 실제 적용 사례에 관한 연구를 바라며, 정보기술을 활용한 대학 내 다양한 문제 해결을 위한 후속 연구를 기대한다.

참 고 문 헌

1. 국내문헌

- 강민정. (2018). 사용목적에 따라 선호하는 챗봇의 성격에 관한 연구. 『한국 콘텐츠학회논문지』, 18(5), 319-329.
- 김봉환. (2000). 청소년 진로상담 정책의 현황과 과제. 『청소년상담연구』, 8, 146-169.
- 김종율, 노광현. (2019). 공공 데이터를 활용한 대학생 진로 조언을 위한 그래프데이터베이스 기반 지식베이스 구축. 『전자공학학회논문지』, 56(10), 41-48.
- 김종율, 노광현. (2019). AI 를 활용한 대학생 진로 조언 시스템 모델 및 데이터 수집과 융합에 대한 연구. 『디지털융복합연구』, 17(2), 177-185.
- 박동아. (2017). 인공지능 기반 대화형 공공 행정 챗봇 서비스에 관한 연구. 『멀티미디어학회논문지』, 20(8), 1347-1356.
- 박은선, 김광현. (2016). 학업중단요인 및 대학생활적응요인이 대학 신입생의 중도탈락의도에 미치는 영향. 『청소년학연구』, 23(10), 361-384.
- 박희석. (2018). 서울시 비영리스타트업 실태와 청년일자리. 『정책리포트』, (250), 1-24.
- 손정은, 박재한, 김성혜, 장원섭. (2019). 초기 경력자의 자발적 이직 경험에 관한 현상학적 연구 : 학습과 성장을 중심으로. 『교육문화연구』, 25.
- 이대용, 류동희, 김인환, 고태용. (2011). 한국대학생의 진로의식에 관한 연구. 『취업진로연구』, 1(2), 41-68.
- 이성희, 유동훈. (2019). 청년추가고용장려금사업 운영현황 및 정책과제. 『노동리뷰』, 167.
- 이지연. (2017). 모든 국민의 생애 전반을 지원하는 진로교육. 『서울교육』, 231.

- 이흥주. (2013). 제조업 분야의 창조적 융합. 『Korea Creative Content Agency』, 78-79.
- 윤상오, 현지우. (2019). 공공데이터 개방정책의 실태분석 및 개선방안에 관한 연구: 공공데이터 포털의 국가중점데이터 개방 사례를 중심으로. 『한국공공관리학보』, 33(1), 219-247.
- 윤수경, 한유경, 임소현, 김화영. (2015). 대학 및 전공 선택에 영향을 미치는 요인 분석. 『한국교육』, 42(2), 87-107.
- 장후은, 허선영, 이종호. (2017). 대학의 현장실습 운영 실태 및 정책 과제. 『한국산학기술학회 논문지』, 18(2), 493-500.
- 홍기석. (2018). 청년실업의 결정요인 연구. 『한국경제의 분석』, 24(2), 91-152.
- 김지경, 이상호, 이성용. (2016). 『대학생 졸업유예 실태 및 지원 방안 연구』. 세종: 한국청소년정책연구원 연구보고서.
- 김진숙. (2017). 『학습자 맞춤형 학습 실현을 위한 지능정보기술』. 한국교육신문.
- 김영애. (2010). 『입학사정관계의 효율적 운영을 위한 e-포트폴리오 설계 연구』. 대구: 한국교육학술정보원.
- 김은석. (2015). 『대학생과 인사담당자의 직업 역량별 비교』. 충북: 한국고용정보원.
- 리상섭. (2017). 『4차 산업혁명시대의 최신 직업진로설계』. 경기: 양서원.
- 류동희. (2010). 『대학에서 전문직업인 육성을 위한 생애교육』. 서울: 서경문화사.
- 문용식. (2018). 『인공지능 기반의 챗봇 서비스의 국내외 동향분석 및 발전 전망』. 대구: 한국정보화진흥원.
- 박인우. (2008). 『e-포트폴리오의 교육적 활용 방안 연구 : 초등학교급을 중심으로』. 대구: 한국교육학술정보원.
- 박홍성, 이상현, 전시홍. (2014). 『청년 채용박람회 운영실태』. 서울: 청년위원회.
- 서순식. (2016). 『로봇 활용 SW교육 효과성 검증 연구』. 대구: 한국교육학

술정보원.

어수봉. (2017). 『4차 산업혁명과 고용의 미래』. 서울: 경제사회발전노사정위원회.

오지은. (2011). 『성공적인 상담을 위한 커뮤니케이션의 첫 걸음』. 서울: 서울대학교 교수학습개발센터.

이요행, 강옥희. (2010). 『대학교 직업진로지도 현황과 과제』. 충북: Employment Issues.

장현진. (2017). 『대학진로교육 현황조사(2017) 기초통계표』. 세종: 교육부.

장현진. (2015). 『진로교육의 현황과 과제』. 세종: 한국직업능력개발원.

정보통신기술진흥센터. (2018). 『4차 산업혁명 시대에 대응한 주요국의 공공 데이터 개방 현황 분석』. 대전: Global ICT R&D Policy Trends.

정상현. (2017). 『‘대학 4.0’과 학습분석…맞춤형 고등교육 앞당긴다』. 서울: 한국대학신문.

전창욱, 최태균, 조중현. (2019). 『텐서플로와 머신러닝으로 시작하는 자연어 처리』. 경기: 위키북스.

조만업, 정은선. (2014). 『직업 상담 실무』. 서울: 정훈사.

천영민. (2015). 『채용 경향 변화 분석 및 이를 활용한 취업진로지도 방안 연구』. 충북: 한국고용정보원.

테크M. (2018). 『AI 거물 앤드류 응이 개발한 우울증 완화 챗봇』. 서울: MIT테크놀로지리뷰.

J. G. Maree. (2019). 『진로구성상담의 원리와 실제』. 서울: 학지사.

KT경제연구소. (2017). 『한국형 4차 산업혁명의 미래: KT경제경영연구소가 찾아낸 미래 한국의 7가지 성장전략』. 서울: 한스미디어.

관계부처 합동. (2019). 『2019년도 공공데이터 제공 및 이용 활성화 시행계획』.

관계부처 합동. (2018). 『데이터 산업 활성화 전략』.

관계부처 합동. (2017). 『제4차 산업혁명에 대응한 지능정보사회 중장기 종합대책』.

중소기업중앙회. (2014). 『2014 위상 지표』.

중앙대. (2019). 『인공지능 기술 기반의 e-Advisor 구축을 위한 BPR/ISP 영역 및 대화형 시스템 구축 사업 계획』.

한국고용노동부. (2016). 『대학경영진을 위한 진로·취업지원 전략가이드』.

한국경영자총연합회. (2012). 『기업과 세대 간 일자리 갈등에 관한 인식 조사』.

AWS. <https://aws.amazon.com/ko/neptune>

가트너. <https://www.gartner.com/>

오픈데이터헌장. <https://opendatacharter.net/>

월드와이드웹 재단. <https://webfoundation.org/>

한국고용정보원. <http://www.keis.or.kr/>



2. 국외문헌

- Bandura, A., & Simon, K. M. (1977). The role of proximal intentions in self-regulation of refractory behavior. *Cognitive therapy and research*, 1(3), 177-193.
- Crosling, G., Heagney, M., & Thomas, L. (2009). Improving student retention in higher education: Improving teaching and learning. *Australian Universities' Review*, The, 51(2), 9.
- Deroma, V. M., Leach, J. B., & Leverett, J. P. (2009). The relationship between depression and college academic performance. *College Student Journal*, 43(2), 325-335.
- Hsu, S. H., Wang, Y. C., Cheng, C. J., & Chen, Y. F. (2016). Developing a decomposed alumni satisfaction model for higher education institutions. *Total Quality Management & Business Excellence*, 27(9-10), 979-996.
- Phillips, E. D. (2013). Improving advising using technology and data analytics. *Change: The Magazine of Higher Learning*, 45(1), 48-55.
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. *Technological forecasting and social change*, 114, 254-280.
- Katz, M. (1966). A Model of Guidance for Career Decision-Making. *Vocational Guidance Quarterly*, 15(1), 2-10.
- Page, L. C., & Gehlbach, H. (2017). How an artificially intelligent virtual assistant helps students navigate the road to college. *AERA Open*, 3(4).
- Williamson, B. (2018). The hidden architecture of higher education: building a big data infrastructure for the 'smarter university'. *International Journal of Educational Technology in Higher*

- Education, 15(1), 12.
- Wolf, K., & Dietz, M. (1998). Teaching portfolios: Purposes and possibilities. *Teacher education quarterly*, 9–22.
- Zytowski, D. G. (2001). Frank Parsons and the progressive movement. *The career development quarterly*, 50(1), 57.
- Blumauer, A. (2018). Knowledge Graphs – Connecting the Dots in an Increasingly Complex World. Semantic Web Company.
- Baton, J., & Van Bruggen, R. (2017). Learning Neo4j 3. x: Effective data modeling, performance tuning and data visualization techniques in Neo4j. Packt Publishing Ltd.
- Carnevale, A. P., Jayasundera, T., & Repnikov, D. (2014). Understanding online job ads data: a technical report. Georgetown University, McCourt School on Public Policy, Center on Education and the Workforce.
- Dimock, M. (2019). Defining generations: Where Millennials end and Generation Z begins. Pew Research Center.
- Hunger, M., Boyd, R., & Lyon, W. (2016). The Definitive Guide to Graph Databases for the RDBMS Developer. Neo Technology.
- Iftene, A., & Vanderdonckt, J. (2016). Moocbuddy: a chatbot for personalized learning with moocs. In *RoCHI – International Conference on Human–Computer Interaction*.
- Page, L., & Gehlbach, H. (2018). How Georgia State University used an algorithm to help students navigate the road to college. *Harvard Business Review*.
- Martinez, A., Mora, R., Alvarado, D., Lopez, G., & Quiros, S. (2016). A Comparison between a Relational Database and a Graph Database in the context of a Personalized Cancer Treatment Application. In *Proceedings of the 10th Alberto Mendelzon International Workshop on Foundations of Data Management*.

- McIlveen, P., Brooks, S., Lichtenberg, A., Smith, M., Torjul, P., & Tyler, J. (2011). Career development learning frameworks for work-integrated learning. In *Developing Learning Professionals*. Springer.
- Cortez, M. B. (2018). Universities Deploy Chatbots to Aid Students in the Admissions Process and Beyond. *EdTech*.
- Shury, J., & IFF Research Ltd. (2017). Planning for Success: Graduates' Career Planning and Its Effect on Graduate Outcomes. Department for Education.
- Skjuve, M., & Brandtzæg, P. B. (2018). Chatbots as a new user interface for providing health information to young people. *Youth and news in a digital media environment – Nordic-Baltic perspectives*.
- Sullivan, R. (2018). *Exploring Emerging Technologies for Lifelong Learning and Success*. The State University of New York.
- Sullivan, D. (2015). *NoSQL for mere mortals*. Addison-Wesley Professional.
- Vicknair, C., Macias, M., Zhao, Z., Nan, X., Chen, Y., & Wilkins, D. (2010, April). A comparison of a graph database and a relational database: a data provenance perspective. In *Proceedings of the 48th annual Southeast regional conference*, 42.

부 록

진로 조언서비스 이용 편의성/만족도 설문지

안녕하십니까?

본 조사는 IT 관련 학과 전공 대학생을 대상으로 챗봇 인터페이스 진로 조언 서비스의 이용 편의성을 조사하기 위한 설문입니다.

주어진, 이용 안내서를 따라서 개발된 진로 조언 서비스를 이용해 보신 후 이용 편의성, 만족도에 대해 솔직하게 표기를 해주시기 바랍니다.

응답해 주신 모든 분의 개인정보는 통계법 제 13조에 의해 비밀이 보호됩니다.

■ 응답자 정보 및 진로 관련 내용

문1. 귀하의 성별은?

1) 남 2) 여

문2. 귀하의 전공은?

1) IT융합공학 2) 컴퓨터공학 3) 기타 ()

문3. 귀하의 학년은?

1) 1학년 2) 2학년 3) 3학년 4) 4학년

문4. 귀하의 나이는?

()

문5. 현재 귀하의 고민은

1) 졸업 후 진로 2) 학업 3) 기타

문6. 귀하는 이미 진로를 결정했는지요? (막연한 상태의 진로 결정이 아닌 확정적)

1) 예 2) 아니오 3) 아직 고민 중

문7. 진로 결정시 상담 또는 조언을 구하는 대상은? (복수 응답 가능)

1) 교수님 2) 친구 3) 선배 4) 부모님 5) 온라인 커뮤니티 6) 웹 검색
7) 기타 ()

문8. 귀하의 졸업 후 진로는?

1) 취업 2) 진학 3) 창업 4) 생각해 보지 않았다. 5) 기타 ()

문9. 진로 조언 서비스가 만들어진다면 어떤 형태를 원하시는지요?

- 1) 챗봇 2) 모바일앱 3) 학교홈페이지 웹서비스 4) PC 설치용 패키지

문10. 진로 취·창업 준비를 위해 대학에서 지원받고 싶은 사항은

()

■ 진로 조언 플랫폼 기능 구현/만족도 평가

문11. 테스트하신 진로 조언 서비스를 각 질문의 답변에 대한 만족도를 평가해 주시기 바랍니다. (5점 척도)

구 분	매우 도움이 된다 5	도움이 된다 4	보통 3	도움이 되지 않는다 2	전혀 도움이 되지 않는다 1
1. 기업의 채용정보 분석 관련	5	4	3	2	1
2. 직업의 종류와 세계에 대한 관련	5	4	3	2	1
3. 진로를 개발하기 위한 학습정보 관련	5	4	3	2	1
4. 취업 준비 과목 관련	5	4	3	2	1
5. 챗봇을 활용한 전문가 멘토링 관련	5	4	3	2	1
6. 챗봇을 활용한 진로·직업 관련 멘토링 관련	5	4	3	2	1
7. e-포트폴리오 관련 기능 (현재)	5	4	3	2	1
8. e-포트폴리오 관련 기능 (구축 완료 후)	5	4	3	2	1
9. e-포트폴리오의 적극 활용 여부 (구축 완료 후)	5	4	3	2	1

문12. 테스트하신 진로 조언 서비스를 활용성을 평가해 주시기 바랍니다.
5점 척도로 평가해 주세요

구 분	매우 도움이 된다 5	도움이 된다 4	보통 3	도움이 되지 않는다 2	전혀 도움이 되지 않는다 1
1. 챗봇을 진로 조언 인터페이스에 만족하십니까?	5	4	3	2	1
2. 진로 조언 플랫폼 조언 결과에 대해서 얼마만큼 만족하십니까?	5	4	3	2	1
3. 진로를 설계할 때 진로 조언 서비스가 도움이 될 것 같은가요?	5	4	3	2	1
4. 기업 채용정보 분석에 대한 만족도	5	4	3	2	1
5. 직업정보에 대한 만족도	5	4	3	2	1
6. 무료 진로 조언 서비스가 완성된다면 적극적으로 활용할 생각이 있으신지요?	5	4	3	2	1

문13. 진로 조언 서비스를 활용한 진로설계 시 불편한 사항/추가 필요한 사항/문제 되는
사항을 기재하여 주시기 바랍니다.
()

문14. 기능요구사항과 비교하여 구현된 정도를 평가해 주시기 바랍니다.

구 분	매우 잘 구현되었 다 5	잘 구현되었 다 4	구현되었 다 3	기능 구현이 부족하다 2	기능이 구현되지 않았다. 1
1. 직업정보 제공기능	5	4	3	2	1
2. 기업 채용정보 분석	5	4	3	2	1
3. 진로/직업 상담	5	4	3	2	1
4. 나의 정보 분석	5	4	3	2	1
5. 목표 설정 및 관리	5	4	3	2	1

문15. 아래의 정보를 기존 방식(스마트폰, PC) 와 진로 조언 챗봇을 활용하여 검색 후 정보 검색 시간을 입력해 주시기 바랍니다.

(1) 직업정보

- [\$직업이름]은 뭐야? (기존:)(챗봇:)
- [\$직업이름]는 어떤 일을 해?(기존:)(챗봇:)
- [\$직업이름]의 전망은 어때? (기존:)(챗봇:)
- [\$직업이름]의 연봉은 얼마야? (기존:)(챗봇:)
- [\$직업이름]의 만족도는 어때? (기존:)(챗봇:)

답변의 만족도 (기존:)(챗봇:)

※ 백분율로 표시해 주세요

(2) 채용정보

- [\$직업이름]은 어떤 자격증이 있어야 해? (기존:)(챗봇:)
- [\$직업이름]은 어떤 기술을 가져야 해? (기존:)(챗봇:)
- java 기술을 가진 사람은 어떤 직업이 좋아? (기존:)(챗봇:)
- [\$직업이름]은 어떤 성향이 있는 사람을 원해? (기존:)(챗봇:)
- 진취적인 성격은 어느 직업에서 원해? (기존:)(챗봇:)
- [\$직업이름]은 어떤 사람을 우대해줘? (기존:)(챗봇:)
- 중국어 잘하는 사람을 우대해주는 직업은? (기존:)(챗봇:)

답변의 만족도 (기존:)(챗봇:)

※ 백분율로 표시해 주세요

전체 편의성 (기존:)(챗봇:)

※ 백분율로 표시해 주세요

ABSTRACT

A Study on Career Advice Platform for College Students Using Graph Database

Kim, Jong-Yul

Major in Smart Convergence Product

Dept. of Smart Convergence Consulting

The Graduate School

Hansung University

According to the Ministry of Education's "2017 Career Education Survey", it was confirmed that 85% of university students worry about their careers, including career-related studies.

While many studies are being conducted on the causes and concerns of university students' career paths, suggestions on how to solve these problems are limited in scopes, such as career counseling and career education. This paper is a study on the solution using information technology on the career and employment problems of Korean university students.

The requirements of university students for research were defined by using answers to "What they want to receive from universities to prepare for employment and start-up" from the Ministry of Education's "Career Education survey in 2017" and answers from the "Why they are having difficulties deciding their careers?" in the 2017 career-conscious survey at

Chung–Ang University.

The career counseling model defines a counseling model by linking actual counseling cases with personal analysis, job analysis, and scientific advice based on Parsons' 'Trait–Factor Theory'. Based on the theory, the paper presents the career counseling model using information technology and the available technology elements in steps of information analysis, job information provision·recommendation, and counseling.

For career advice, job information and practical details were collected through the Worknet API of the Ministry of Employment and Labor, and because of the lack of practical information required by students for employment and preparation for start-ups, 89 job postings for IT-related jobs were collected from the private employment information portal, classified by capacity, and weighted by calculating their importance.

In the design process of using the collected data to provide customized advice to students, an analysis of relationships, such as each of the competency factors and occupations, and relationships with each competency, was introduced as an important technical issue. To solve these relationship analysis problems, knowledge graphs using the graph database were selected as knowledge bases for career advice through the relationship between relational databases and graph databases and comparative experiments and evaluations of query efficiency.

In response to natural language questions from college students for consultation and mentoring function design, the e-portfolio app was designed to receive customized jobs and support competency development based on chatbot-based career advice system and entered individual capabilities using natural language processing learning.

Through this study, we presented the knowledge base design and construction methodology using the graph database. Using the presented knowledge graph design methodology and the construction methodology,

we collected job and occupation data from the Ministry of Employment and Labor's API and job offer information from the private recruitment portal. We collected and implemented the knowledge graph for career and employment advice and proved the effectiveness of the methodology by comparing the construction speed with the method that did not use the methodology.

The career advice platform using the designed graph database was implemented to test the career advice scenario of chatbot and mobile app environments using knowledge graphs and selected college students related to information technology as evaluation teams to conduct actual demonstrations and tests.

Based on the survey and results of the evaluation team, the satisfaction of the career counseling and capacity development model is high as suggested in the results: the satisfaction of the company's recruitment information analysis and advice (4.08), interface satisfaction (4.24), whether to assist in career design(4.24), e-portfolio function and utilization (4.52) and career advice platform actively utilized (4.71). Also, it was evaluated that the search efficiency and satisfaction using the chatbot was high for certain functions such as career fields through the information retrieval time test with the existing information retrieval method using the mobile or web search and the chatbot method.

This study proposes a solution using information technology for career problems, the biggest concern of Korean university students, and newly proposed a model for constructing and using a knowledge graph based on graph database in knowledge where relationship analysis is important.

I hope that the results of this study will help university students to analyze and provide knowledge related to the construction of knowledge graphs using various occupational data, and to support students' career decisions and development of capabilities and that it will be a guide to

establish personalized support services for each student in conjunction with the academic support system, career path and job integration information system at school. Especially, the graph database can be explained as a structural perspective and there is not much data other than the data for model design. Therefore, further research is expected using this research as a model for building and utilizing various types of intelligent systems.

【Key word】 career advice, e-Advisor, chatbot, graph database, public data, edutech

