

연구보고서 2016-016

빅데이터 분석 기반

SW산업 직종·직무 분류체계 개발 및 동향 연구

Classifications of Software Occupation and Labor Force
Trends : Big Data Analytics Perspective

이호/공영일/이동현/임춘성

2017.04.

이 보고서는 2016년도 미래창조과학부 정보통신진흥기금을
지원받아 수행한 연구결과로 보고서 내용은 연구자의 견해이며,
미래창조과학부의 공식입장과 다를 수 있습니다.

목 차

제1장 서 론	1
제1절 연구 배경	1
제2절 연구 목적	5
제2장 빅데이터 분석을 통한 SW 분야 직무·직종 체계 개발	6
제1절 데이터 분석 개요	6
1. 분석 자료의 수집 대상 및 적합성	6
제2절 분석 자료의 전처리 과정	7
1. 범주형 및 숫자형 데이터의 전처리 사항	8
2. 문자형 데이터 전처리	14
제3절 SW산업 관련 기술통계 분석	17
1. 전체 구인 기업에 대한 현황	17
2. 워크넷 직업 분류 기반 채용공고 기술통계	27
3. 전체 구인 기업 및 워크넷 분류에 따른 SW분야의 특성	42
제4절 텍스트마이닝 분석	44
1. 채용공고 데이터	52
2. NCS 직종분류 데이터	68
제5절 분류체계 간의 기술역량 일치도 검토	86
1. 워크넷 직업 분류와 NCS 직무 분류 체계의 일치성 비교	86

2. 분류체계 비교를 위한 데이터 선택 및 분석 과정	87
3. 계산방법	90
4. 결과비교	92
5. 워크넷 분류와 NCS분류 체계 비교에 따른 시사점	95
제6절 빅데이터 분석 기반의 직무·직종 분류체계 개발	98
1. 핵심단어 주제 분류를 위한 분석 알고리즘 적용	98
2. LDA 분석 기반의 SW 직종직무 분류체계	106
3. 빅데이터 분석 기반의 직무·직종 분류 체계의 시사점	235
제3장 시사점 및 제언	237
제1절 시사점 및 향후 연구 방향	237
제2절 제언	239

표 목 차

〈표 1-1〉 주요국 정보통신 분야의 인력부족규모 추정	1
〈표 1-2〉 IT산업 인력의 고용추이	2
〈표 2-1〉 회사명 전처리의 예시	9
〈표 2-2〉 근로자수 전처리 예시	9
〈표 2-3〉 지역 전처리의 예시	10
〈표 2-4〉 자본금 전처리의 예시	10
〈표 2-5〉 고용형태 전처리의 예시	11
〈표 2-6〉 모집인원 전처리	11
〈표 2-7〉 임금조건 전처리의 예시	12
〈표 2-8〉 경력조건 전처리 규칙	12
〈표 2-9〉 학력조건 전처리 규칙	13
〈표 2-10〉 외국어 능력 전처리 규칙	13
〈표 2-11〉 원본 데이터와 가공 데이터 전처리 방식의 예시	15
〈표 2-12〉 워크넷 중·소분류	27
〈표 2-13〉 워크넷 소분류 201에 대한 세분류	28
〈표 2-14〉 워크넷 소분류 202에 대한 세분류	29
〈표 2-15〉 워크넷 소분류 203에 대한 세분류	29
〈표 2-16〉 워크넷 소분류 204에 대한 세분류	30
〈표 2-17〉 워크넷 소분류 205에 대한 세분류	30
〈표 2-18〉 워크넷 소분류 206에 대한 세분류	31

〈표 2-19〉 소분류별 채용공고	32
〈표 2-20〉 소분류별 자본금 기준 채용공고	34
〈표 2-21〉 모집직종과 직종코드 예시	52
〈표 2-22〉 정보통신 세부 분류	72
〈표 2-23〉 워크넷 직종코드와 일치도가 높은 NCS 세분류	74
〈표 2-24〉 소분류별 상위 추출 단어	87
〈표 2-25〉 소분류별 단어수와 빈도수	88
〈표 2-26〉 NCS 분류별 단어수와 빈도수	89
〈표 2-27〉 두 집합 간 유사성 측정식	91
〈표 2-28〉 일치도 결과 비교	92
〈표 2-29〉 201 분류와 20020315분류의 단어와 빈도수	92
〈표 2-30〉 202분류와 20020316분류의 단어와 빈도수	93
〈표 2-31〉 203분류와 20020306분류의 단어와 빈도수	93
〈표 2-32〉 204분류와 20010104분류의 단어와 빈도수	94
〈표 2-33〉 205분류와 20020316분류의 단어와 빈도수	99
〈표 2-34〉 206분류와 20020102분류의 단어와 빈도수	99
〈표 2-35〉 직무·직종 K토픽 모델링 예시	105
〈표 2-36〉 토픽모델링 기반의 직종직무 분류체계	110
〈표 2-37〉 직무 1의 핵심단어 목록	115
〈표 2-38〉 웹 관리자 직무 관련 국제직무분류표준	118
〈표 2-39〉 직무 2의 핵심단어 목록	122
〈표 2-40〉 직무군집 2의 국제직무분류표준	125

〈표 2-41〉 직무 3의 핵심단어 목록	128
〈표 2-42〉 웹 콘텐츠 디자이너 직무 관련 국제직무분류표준	131
〈표 2-43〉 직무 4의 핵심단어 목록	134
〈표 2-44〉 그래픽 디자이너 직무 관련 국제직무분류표준	137
〈표 2-45〉 직무 5의 핵심단어 목록	140
〈표 2-46〉 컴퓨터 네트워크 관리자 관련 국제직무분류표준	144
〈표 2-47〉 직무 6의 핵심단어 목록	148
〈표 2-48〉 소프트웨어 개발자 직무 관련 국제표준	151
〈표 2-49〉 직무 7의 핵심단어 목록	155
〈표 2-50〉 시스템 개발자 직무 관련 국제직무분류표준	158
〈표 2-51〉 직무 8의 핵심단어 목록	161
〈표 2-52〉 SW 분석가 직무 관련 국제표준	164
〈표 2-53〉 직무 9의 핵심단어 목록	168
〈표 2-54〉 시스템 설계분석가 직무 관련 국제표준	171
〈표 2-55〉 직무 10의 핵심단어 목록	175
〈표 2-56〉 미디어 콘텐츠 창작자 직무 관련 국제직무분류표준	178
〈표 2-57〉 직무 11의 핵심단어 목록	182
〈표 2-58〉 소셜 미디어 관리자 직무 관련 국제표준	185
〈표 2-59〉 직무 12의 핵심단어 목록	189
〈표 2-60〉 데이터베이스 관리자 직무 관련 국제표준	192
〈표 2-61〉 직무 13의 핵심단어 목록	195
〈표 2-62〉 HCI 컨설턴트 직종 관련 국제표준	198

〈표 2-63〉 직무 14의 핵심단어 목록	202
〈표 2-64〉 애플리케이션 개발자 직무 관련 국제표준	205
〈표 2-65〉 직무 15의 핵심단어 목록	208
〈표 2-66〉 모바일 개발자 직무 관련 국제표준	211
〈표 2-67〉 웹기획 관리자 직종 관련 국제표준	213
〈표 2-68〉 디지털 콘텐츠 디자이너 직종 관련 국제표준	215
〈표 2-69〉 컴퓨터 네트워크 관리자 직종 관련 국제표준	216
〈표 2-70〉 시스템 SW 개발자 직종 관련 국제직무분류표준	218
〈표 2-71〉 시스템 설계분석가 직종 관련 국제직무분류표준	220
〈표 2-72〉 콘텐츠 크리에이터의 역할	221
〈표 2-73〉 데이터베이스 관리자 직종 관련 국제직무분류표준	222
〈표 2-74〉 HCI 컨설턴트 직종 관련 국제표준	224
〈표 2-75〉 응용소프트웨어 개발자 직종 관련 국제직무분류표준	226
〈표 2-76〉 데이터 분석 기반 SW 산업 직무·직종 분류체계	228
〈표 2-77〉 주별 채용공고 등록 건수	230

그 립 목 차

[그림 2-1] 데이터 분석 절차 및 방법	8
[그림 2-2] 구인 기업 규모별 채용공고	17
[그림 2-3] 구인 기업 지역별 채용공고	18
[그림 2-4] 구인 기업 자본금별 채용공고	19
[그림 2-5] 구인 기업 채용인원별 채용공고	20
[그림 2-6] 구인 기업 고용형태별 채용공고	21
[그림 2-7] 구인 기업 임금조건별 채용공고	22
[그림 2-8] 구인 기업 고용형태에 따른 임금조건별 채용공고	23
[그림 2-9] 구인 기업 경력조건별 채용공고	24
[그림 2-10] 구인 기업 고용형태에 따른 경력조건별 채용공고	25
[그림 2-11] 구인 기업 학력조건별 채용공고	25
[그림 2-12] 구인 기업 외국어 능력별 채용공고	26
[그림 2-13] 소분류별 근로자수 기준 채용공고	33
[그림 2-14] 소분류별 지역 기준 채용공고	35
[그림 2-15] 소분류별 고용형태 기준 채용공고	36
[그림 2-16] 소분류별 채용인원 기준 채용공고	37
[그림 2-17] 소분류별 임금조건 기준 채용공고	38
[그림 2-18] 소분류별 경력조건 기준 채용공고	39
[그림 2-19] 소분류별 학력조건 기준 채용공고	40
[그림 2-20] 소분류별 외국어조건 기준 채용공고	41

[그림 2-21] 텍스트마이닝 프레임 워크	44
[그림 2-22] 텍스트마이닝 과정	45
[그림 2-23] 모집 직종 코드별 채용공고 수	53
[그림 2-24] 모집 직종 코드별 단어 수	54
[그림 2-25] 모집 직종 코드별 단어	54
[그림 2-26] 컴퓨터 및 통신공학 기술자 및 연구원(201번)	55
[그림 2-27] 컴퓨터 및 통신공학 기술자 및 연구원의 단어 공동 출현 네트워크	56
[그림 2-28] 컴퓨터시스템 설계 및 분석가(202번)	57
[그림 2-29] 컴퓨터시스템 설계 및 분석가의 단어 공동 출현 네트워크	58
[그림 2-30] SW개발전문가(203번)	59
[그림 2-31] SW개발전문가의 단어 공동 출현 네트워크	60
[그림 2-32] 웹전문가(204번)	61
[그림 2-33] 웹전문가의 단어 공동 출현 네트워크	62
[그림 2-34] 데이터베이스 설계·구축 및 관리기술자(205번)	63
[그림 2-35] 데이터베이스 설계·구축 및 관리기술의 단어 공동 출현 네트워크	64
[그림 2-36] 통신 및 방송 장비기사 설치 및 수리원(206번)	65
[그림 2-37] 통신 및 방송 장비기사 설치 및 수리원의 단어 공동 출현 네트워크 ..	66
[그림 2-38] 고용노동부 직업능력개발 혁신 3개년 실천계획	68
[그림 2-39] 능력중심사회 3개년 추진 경과	69
[그림 2-40] 국가직무능력표준 활용범위	70
[그림 2-41] NCS의 학습모듈의 개념	71
[그림 2-42] NCS 분야별 검색	71

[그림 2-43] NCS 세분류별 단어 개수	75
[그림 2-44] NCS 세분류별 단어	76
[그림 2-45] SW제품기획 (20010104)	76
[그림 2-46] 빅데이터 분석 (20010105)	78
[그림 2-47] 유선설비접속서비스 (20010301)	79
[그림 2-48] 구내통신구축 (20020102)	80
[그림 2-49] 네트워크구축 (20020103)	81
[그림 2-50] 전송시스템구축 (20020202)	82
[그림 2-51] 이동통신서비스 (20020306)	83
[그림 2-52] 부가통신응용중계서비스 (20020315)	84
[그림 2-53] 특수부가통신서비스 (20020316)	85
[그림 2-54] 소분류별 외국어조건 기준 채용공고	87
[그림 2-55] 문헌정보의 LDA 사례	99
[그림 2-56] 뉴스데이터의 LDA 사례	100
[그림 2-57] LDA 분석 모형	101
[그림 2-58] 채용공고 단어의 모형화	103
[그림 2-59] 단어 사전의 특정 단어 발생 수 도출과정	103
[그림 2-60] 워크넷 분류 체계 직무코드별 채용공고 수	111
[그림 2-61] 워크넷 분류 체계 직무코드별 핵심단어 수	111
[그림 2-62] 워크넷 분류 체계 직무 코드별 단어 Chord Chart	112
[그림 2-63] 직무 1의 특성	113
[그림 2-64] 직무 1의 단어 공동 출현 네트워크	114

[그림 2-65] 직무 1의 핵심단어 비율 그래프	116
[그림 2-66] 직무 1의 핵심단어 방사형 그래프	116
[그림 2-67] 직무 1의 핵심단어 워드클라우드	117
[그림 2-68] 직무 2의 특성	119
[그림 2-69] 직무 2의 공동 출현 단어 네트워크	121
[그림 2-70] 직무 2의 핵심단어 비율 그래프	122
[그림 2-71] 직무 2의 핵심단어 방사형 그래프	123
[그림 2-72] 직무 2의 핵심단어 워드클라우드	124
[그림 2-73] 직무 3의 특성	126
[그림 2-74] 직무 3의 단어 공동 출현 네트워크	127
[그림 2-75] 직무 3의 핵심단어 비율 그래프	129
[그림 2-76] 직무 3의 핵심단어 방사형 그래프	129
[그림 2-77] 직무 3의 핵심단어 워드클라우드	130
[그림 2-78] 직무 4의 특성	132
[그림 2-79] 직무 4의 단어 공동 출현 네트워크	133
[그림 2-80] 직무 4의 핵심단어 비율 그래프	135
[그림 2-81] 직무 4의 핵심단어 방사형 그래프	136
[그림 2-82] 직무 4의 핵심단어 워드클라우드	136
[그림 2-83] 직무 5의 특성	138
[그림 2-84] 직무 5의 단어 공동 출현 네트워크	139
[그림 2-85] 직무 5의 핵심단어 비율 그래프	141
[그림 2-86] 직무 5의 핵심단어 방사형 그래프	142

[그림 2-87] 직무 5의 핵심단어 워드클라우드	142
[그림 2-88] 직무 6의 특성	146
[그림 2-89] 직무 6의 단어 공동 출현 네트워크	147
[그림 2-90] 직무 6의 핵심단어 비율 그래프	148
[그림 2-91] 직무 6의 핵심단어 방사형 그래프	149
[그림 2-92] 직무 6의 핵심단어 워드클라우드	150
[그림 2-93] 직무 7의 특성	153
[그림 2-94] 직무 7의 단어 공동 출현 네트워크	154
[그림 2-95] 직무 7의 핵심단어 비율 그래프	156
[그림 2-96] 직무 7의 핵심단어 방사형 그래프	156
[그림 2-97] 직무 7의 핵심단어 워드클라우드	157
[그림 2-98] 직무 8의 특성	159
[그림 2-99] 직무 8의 단어 공동 출현 네트워크	160
[그림 2-100] 직무 8의 핵심단어 비율 그래프	162
[그림 2-101] 직무 8의 핵심단어 방사형 그래프	162
[그림 2-102] 직무 8의 핵심단어 워드클라우드	163
[그림 2-103] 직무 9의 특성	166
[그림 2-104] 직무 9의 단어 공동 출현 네트워크	167
[그림 2-105] 직무 9의 핵심단어 비율 그래프	169
[그림 2-106] 직무 9의 핵심단어 방사형 그래프	169
[그림 2-107] 직무 9의 핵심단어 워드클라우드	170
[그림 2-108] 직무 10의 특성	173

[그림 2-109] 직무 10의 단어 공동 출현 네트워크	174
[그림 2-110] 직무 10의 핵심단어 비율 그래프	176
[그림 2-111] 직무 10의 핵심단어 방사형 그래프	176
[그림 2-112] 직무 10의 핵심단어 워드클라우드	177
[그림 2-113] 직무 11의 특성	180
[그림 2-114] 직무 11의 단어 공동 출현 네트워크	181
[그림 2-115] 직무 11의 핵심단어 비율 그래프	183
[그림 2-116] 직무 11에 대한 핵심단어 방사형 그래프	183
[그림 2-117] 직무 11의 핵심단어 워드클라우드	184
[그림 2-118] 직무 12의 특성	187
[그림 2-119] 직무 12의 단어 공동 출현 네트워크	188
[그림 2-120] 직무 12의 핵심단어 비율 그래프	190
[그림 2-121] 직무 12의 핵심단어 방사형 그래프	190
[그림 2-122] 직무 12의 핵심단어 워드클라우드	191
[그림 2-123] 직무 13의 특성	193
[그림 2-124] 직무 13의 단어 공동 출현 네트워크	194
[그림 2-125] 직무 13의 핵심단어 비율 그래프	196
[그림 2-126] 직무 13의 핵심단어 방사형 그래프	196
[그림 2-127] 직무 13의 핵심단어 워드클라우드	197
[그림 2-128] 직무 14의 특성	200
[그림 2-129] 직무 14의 단어 공동 출현 네트워크	201
[그림 2-130] 직무 14의 핵심단어 비율 그래프	203

[그림 2-131] 직무 14의 핵심단어 방사형 그래프	203
[그림 2-132] 직무 14의 핵심단어 워드클라우드	204
[그림 2-133] 직무 15의 특징	206
[그림 2-134] 직무 15의 단어 공동 출현 네트워크	207
[그림 2-135] 직무 15의 핵심단어 비율 그래프	209
[그림 2-136] 직무 15의 핵심단어 방사형 그래프	209
[그림 2-137] 직무 15의 핵심단어 워드클라우드	210
[그림 2-138] 채용공고에 대한 직종 채용공고수 증감 그래프	231
[그림 2-139] 채용공고에 대한 직무 채용공고수 증감 그래프	232
[그림 2-140] 직무·직종 공고 기반 채용공고 변화에 대한 시계열 그래프	233

요 약 문

1. 연구 목적 및 필요성

산업 환경의 변화와 기술 기반의 4차 산업 혁명의 도래는 소프트웨어(이하 SW) 기반 산업 영역의 확산을 이끌어 오고 있다. 과거 기업의 생산성을 높여주는 지원 도구로서의 SW는 이제 기업의 핵심역량이 되어가고 있다. 이에 따라, 전 세계적으로 읽기, 쓰기, 계산하기의 기본 능력에서 추가적으로 갖추어야할 필수역량으로 SW역량이 대두되고 있다. 국내에서도 ‘초중등 SW교육 필수화’ 및 ‘SW중심대학’ 등을 통해 정부 주도로 국가 SW인재 양성을 강화하고 있다.

한편, 고질적인 학력위주 사회 타파 및 능력중심 사회 구현을 위해 국가직무능력표준(National Competency Standards, NCS)을 개발하였고, 국내 대학 교육 체계를 NCS와 일원화 시키려는 정책도 꾸준히 시행되고 있다. 그러나 기존 NCS 개발 과정에서 일부 전문가들 의견 중심의 Top-down 방식으로 SW분야의 체계가 개발 되면서, SW산업계의 수요를 면밀히 반영하지 못했다는 지적이 꾸준히 제기 되고 있다.

SW의 중요성 대두와 더불어, 대학 SW인재 양성의 근간이 되는 NCS의 SW분야 직무체계분류는 시급히 수정되어야 할 부분이다. 대학에서 배출되는 SW분야의 인력은 각 대학의 교육 체계 특성에 따라 다양한 방식으로 양성되어 오고 있으나, 여전히 SW전문인력 부족 현상은 심화 되고 있다. 이 현상은 SW전공 분야의 학문적 능력에 대해 기업에서 느끼는 활용도가 매우 낮음을 의미한다. 다시 말해, 산업 관점의 수요와 대학 관점의 교육 체계는 그 간극이 존재하고 있음을 알수 있다. 따라서 산업계 수요를 정확히 반영하는 전문 인력 양성 교육의 가이드라인이 필요한 시점이다.

이런 배경을 바탕으로 본 연구에서는 기존 소수의 의사결정자들에 의한 Top-down 방식에서 벗어나 SW분야 직무체계분류를 워크넷 SW분야 채용공고의 빅데이터 분석을 통해 분류하는 Bottom-up 방식의 접근을 진행하였다. SW 산업 분야의 실제 구인 수요를 바탕으로 SW직무·직종 분류체계를 개발 및 정의함으로써 NCS에 산업 수요를 반영하고, 산학간 수요·공급의 간극을 줄일 수 있는 단초를 제공할 수 있을 것이라고 기대된다.

2. 연구 범위 및 방법

본 연구 보고서의 구성은 다음과 같다. 먼저, 1장에서 SW 분야의 특성과 산업계 관점에서의 인력 수요에 대하여 구분하였다. 2장에서는 NCS와 워크넷의 SW분야 구분체계를 비교분석하였다. 이 후 워크넷의 SW인력 채용공고(총 20,408건)를 수집하여, 공고 내용의 핵심단어를 기반으로 직무·직종 분류체계를 정의하였다. 이를 위해, 텍스트마이닝(Text Mining)을 통해 채용공고 데이터에 대해 수집 및 단어별 전처리를 끝낸 후 잠재디리클레할당(Latent Dirichlet allocation, LDA) 알고리즘 기반의 토픽모델링(Topic Modeling) 분석을 수행하여 직무·직종의 분류체계를 신규 도출하였다. 이렇게 분류된 직무·직종에 대한 정의는 국제표준직업분류(International Standard Classification of Occupations, ISCO), 미국표준직업분류(Standard Occupational Classification, SOC), 직업정보네트워크(Occupational Information Network, O*NET)를 참고하여 작성되었고, 기존에 없었던 신규 직무·직종에 대한 정의는 여러 참고 문헌들을 기반으로 작성하였다. 3장에서는 전체 연구에서 확인된 결과의 해석 및 시사점을 도출하고 최종 정책적 제언을 제시하고자 하였다.

3. 연구 내용 및 결과

본 연구의 주요 연구 내용과 결과는 다음과 같다. 첫째, 본 연구는 NCS와 워크넷 기반의 직무 분류가 현재 어떤 형태로 구축되어 있는지 확인하였다. 분석 결과 워크넷의 직무 분류 형태와 NCS의 직무 분류 형태가 불일치되는 점을 확인하였고, 각 직종분류별로 세분화된 직무가 존재하지만 상호 매칭이 되지 않는 형태임을 확인하였다.

둘째, 워크넷과 NCS의 불일치 직무 체계를 보완하기 위하여, 워크넷의 SW관련 채용 공고 내용을 텍스트마이닝으로 분석하여 데이터를 수집하고 LDA 토픽모델링을 바탕으로 빅데이터 기반 직무·직종 분류체계를 도출하였다. 이 결과 총 9가지 직종별 15가지 직무를 구분해 내었으며, 이 중 신규 직종 3가지와 신규 직무 5가지를 도출해 내었다.

셋째, 도출된 직무·직종 분류체계의 각 직무 및 직종에 대해 3가지 국제직무표준을 참고하여 각 직종 및 직무를 정의하고 업무내용을 구분하였다. 기존 국제직무분류표준으로 활용되는 ISCO, SOC, O*Net의 분류체계로부터 유사 분류 정의를 비교하고 이를 바탕으로 9가지 직종, 15가지 직무에 대한 정의를 수립하였다.

4. 정책적 활용 내용

본 연구는 기존 불명확했던 워크넷 분류체계와 NCS 분류체계의 SW분야 산업계의 현실 반영이 미흡했다는 단점을 보완하기 위해 빅데이터 기반의 새로운 분류 체계를 개발하였는데 그 의의가 있다. 먼저, SW분야 기업들의 실제 채용공고를 전수 활용하여 직무·직종

분류체계를 반영하였다는 점에서 SW산업 분야의 인력 수요를 보다 실무적 관점에서 바라보고 SW인재 육성 정책 개발에 가이드라인을 제공할 수 있다. 둘째, 새로이 분류된 직무·직종 분류체계에서 각 직무·직종 별 채용공고상 핵심단어들을 도출하여, 보다 현실적인 직무·직종 별 필요역량을 파악할 수 있을 것이다. 본 연구 결과를 SW관련 정책 수립에 개선에 활용한다면, 더욱 효과적인 SW분야 전문 인력 양성이 가능할 것이다.

5. 기대효과

SW분야는 매우 빠르게 변화하고 있는 산업 군이다. 채용공고를 단어 단위로 분석한 빅데이터 기반의 분석방법은 향후 급변하는 SW분야 산업 직군의 특성을 파악하기 위한 최적화된 도구로써 판단할 수 있다. 본 보고서에서 도출한 직무·직종 분류체계의 경우, 분류된 체계를 기반으로 각 직무·직종 별 역량을 객관적으로 확인할 수 있고, 지속적으로 데이터를 수집한다면 시간 흐름에 따라 분야별 산업 수요의 변화를 확인할 수 있을 것이다. 이는 기존 NCS 체계의 보완뿐만 아니라 SW산업 영역의 수요 변화에 따른 시의적 정책 수립에 매우 유용할 것이다. 마찬가지로 SW분야 관련 산업계와 대학의 SW교육 체계의 간극을 줄이기 위해서 본 연구결과를 활용하여 교과과정을 개선한다면, 고질적인 SW업계의 산학간 수요·공급의 간극을 줄일 수 있을 뿐만 아니라, 대학별로 실제 산업계에서 요구하는 직무·직종 별로 특화된 다양한 교육과정이 개발 될 수 있을 것으로 기대된다.

SUMMARY

1. Research Purpose and necessity

Changes in the industrial environment and the advent of technology-based fourth industrial revolution have led to the spread of software-based industries. Software, as a support tool to enhance corporate productivity in the past, is now becoming a core competency of the enterprise. As a result, Software capabilities are emerging as essential competencies that must be additionally provided in the basic skills of reading, writing, and computing around the world. In Korea, it is strengthening the training of software specialist through ‘essential elementary Software education’ and ‘Software core university’.

On the other hand, National Competency Standards (NCS) have been developed for the achievement of solid academic-oriented society and ability-based society, and then The policy to unify domestic university education system with NCS is also being implemented steadily. However, it is pointed out that NCS has not been able to reflect the needs of the SW industry as a result of the development of the system of SW field as a top-down method based on opinions of some experts in the NCS development process.

In addition to the importance of SW, the classification of the SW field of the NCS, which is the basis of training the university SW human resources, should be urgently revised. The number of employees in the SW field of the university has been graduated in various ways according to the characteristics of the education system of each university. However, the demand of the SW company mainly requires 1-3 years of experience. This phenomenon means that companies’ demand for academic ability in the major field of university graduation is very low. In other words, there is a difference between the industry-oriented demand and the university-based education system. Therefore, it is necessary to establish guidelines for the training of professional manpower that accurately reflects SW industry demand.

Based on this background, this study conducted a bottom-up approach to classifying SW job classification through Big Data analysis of job announcements from Worknet SW category as against the top-down approach by a small number of existing decision-makers. It is expected that it will be possible to reflect the industrial demand of NCS and reduce the gap between demand and supply of industry by developing and defining SW job classification system based on actual job demand of SW industry.

2. Research Scope and Methodology

The contents of this research report are as follows. Based on our research purpose, section 1 described the characteristics of the software industry and the demands of the workforce from an industrial point of view. In Section 2, we compare and analyze the SW division of existing NCS and Worknet. After that, we collected SW job announcements (totally 20,408) and defined the job and occupation classification system based on the keyword of the job announcement.

To accomplish this, we collected the job announcement data through text mining and finished preprocessing by word, and then performed topic modeling analysis based on Latent Dirichlet allocation (LDA) algorithm to newly derive classification system of job and occupation. The developed job and occupation classifications were defined as International Standard Classification of Occupations(ISCO), Standard Occupational Classification(SOC), Occupational Information Network(O*NET). Definitions of new jobs and occupations that have not existed are based on several references. In Section 3, the interpretation and implications of the research results are presented and the final policy suggestions are presented.

3. Research Contents and Results

The main contents and results of this study are as follows. First, this study confirmed the current classifications of NCS and Worknet-oriented job classifications. In order to compensate the gap between two classifications, we developed a SW job classification system based on Big data Analysis. As a result of analysis, it was confirmed that Worknet 's Job Classification and NCS Job Classifications were inconsistent, and it is confirmed that there are subdivided jobs but not mutually matching between two classifications. Second, in order to improve the system of disagreement between the Worknet and NCS, this study analyzed the contents of the Worknet job announcement of the SW industry by text mining. We gathered the total job announcements from Worknet and derived a job classification system based on LDA topic modeling. As a result, 15 kinds of jobs were classified by 9 kinds of job families, and 3 of them were new job families and 5 were new ones. Third, we defined each occupation and divided their tasks by referring to three international job standards for each occupation classification(ISCO, SOC and O*Net).

4. Political Implications

This study has developed a classification system to improve the unclassified classification from Worknet classification system and NCS classification system.

First, it is possible to provide a guideline for the development of SW specialists training policy by the demand of manpower in the SW industry field from the practical point of view, since it reflects the job classification system by utilizing the actual all job posting of companies in the SW field. Second, it is possible to grasp keyword in job announcements by each job category in newly classified job classification system, and to identify more realized job competency.

5. Expected Results

The SW sector is one of the rapidly changing industries. Big data-based analysis method which analyzing the job announcements by word-level can be judged as an optimized tool to identify the characteristics of the rapidly changing SW industrial sector in the future. Through the job classification system that is developed from this report, it is possible to objectively confirm the competence of each job that is classified system, and identify the changes in industry demand according to the time if the data is continuously collected. This will be useful not only for complementing the existing NCS system but also for establishing a timely policy in response to changes in demand in the SW industry. Likewise, it is expected that not only the gap between demand and supply of industry-academy in SW industry can be reduced, but the results of this study be used to improve the curriculum to reduce gap between SW industry and university SW education system.

CONTENTS

Chapter 1. Intronduction	1
Section 1. Research Background	1
Section 2. Research Purpose	5
 Chapter 2. The Development of Occupation Classification for SW Industry based on Big Data Analysis	6
Section 1. Introduction of Data Analysis	6
Section 2. The Process of Data Cleansing	7
Section 3. Descriptive Statistics for SW Industry	17
Section 4. Text Mining Analysis	44
Section 5. The Matching Test of Technical Competence between previous Occupation Classifications	86
Section 6. The Development of Occupation Classification Based on Big Data Analysis	98
 Chapter 3. Implication and Critical Review	237
Section 1. Implication and Future Study	237
Section 2. Conclusions	239

제1장 서론

제1절 연구 배경

정보기술의 확산으로 현대 사회의 업무환경 및 사회 시스템은 컴퓨터 하드웨어 중심의 산업사회에서 소프트웨어(Software, 이하 SW) 기반의 사회로 변화하고 있다. 1960년대부터 급속도로 발전해오고 있는 컴퓨팅 기술의 발달과 1990년대 후반의 인터넷 기반 환경의 도래는 SW 중요성을 강화시키는 단초가 되었다. 이러한 산업 환경의 변화와 기술 기반의 4차 산업 혁명의 도래는 SW기반 산업 영역의 확산을 이끌어 오고 있다. 과거 기업의 생산성을 높여주는 지원 도구로서의 SW는 이제 기업의 핵심역량이 되어가고 있다. 특히 단순 업무 환경에서 활용되어 온 소프트웨어 환경 및 특정 전문가들의 고유한 영역이었던 SW분야는 그 적용영역이 초기 설계 및 네트워크 분야를 넘어 빅데이터 분석에 의한 기업전략 수립, 인공지능기술 기반 제품 생산 등 산업 전반으로 확산되고 있다. SW의 활용범위가 확장됨에 따라 대부분의 업무 환경에서 대다수의 인력이 갖추어야 할 필수역량의 하나로 자리매김하고 있다. 이에 따라, 전 세계적으로 읽기, 쓰기, 계산하기의 기본 능력에서 추가적으로 갖추어야 할 필수역량으로 SW역량이 대두되고 있다.

이러한 산업 환경 속에서 관련 분야의 SW 전문 인력에 대한 수요는 급속도로 증가하고 있지만 그에 따른 분야별 전문 인력의 부족 현상 또한 증가하고 있다.

〈표 1-1〉 주요국 정보통신 분야의 인력부족규모 추정

구분	추정 부족인원 (단위: 명)	추정기관
전 세계	600,000	유럽 정보기술조사국(EITO, 1998)
미국	346,000	미국 정보기술협회 (ITAA, 1998)
미국	450,000	Microsoft
캐나다	20,000~30,000	미국 기술정책국(OTP)
EU	320,000	IDC
독일	60,000	유럽 정보기술조사국
영국	20,000	유럽 정보기술조사국

출처: 정보통신정책연구원 (2014)

<표 1-1>과 같이 대다수 국가들의 정보통신 분야 인력 부족이 심화되는 현상이 나타나고 있으며, SW역량이 범용적으로 갖추어야 할 역량으로 여겨지면서 다수의 국가들은 SW인력에 대한 산업계 수요에 의한 인력 양성과 관련 분야의 SW 교육 체계화에 집중하고 있는 실정이다(김갑수, 2016).

특히 SW 관련 활용 산업은 단순 제조분야에서 콘텐츠, 정보서비스 등 다양한 분야로 그 직무가 확장되어 오고 있으며 대학에서도 SW 관련 과목과 교육훈련이 진행되고 있다. 그에 따라 다양한 SW융합 분야의 전공을 개설하여 SW전문 인력을 배양하고 있다(황성수, 2016). <표 1-2>과 같이 통신, 제조 등의 전반적인 인력 수요는 줄어드는 반면 SW관련 인력 수요는 급증하는 것을 알 수 있다. 따라서 SW산업계 인력 수요의 급증은 대학의 SW관련 교육의 수요를 급증시키며 관련 전공자 수를 양성하고 있지만 여전히 산업계의 수요에 맞는 실무능력을 보유한 인력을 양성하는 측면에서는 인력부족 문제가 지속적으로 논의되고 있다.

<표 1-2> IT산업 인력의 고용추이

(단위: 명)	2000	2005	2010	2000-2005	2005-2010
IT제조업	433,705	496,576	474,948	62,871	-21,628
통신서비스	79,856	66,762	52,102	-13,094	-14,660
SW 및 정보서비스업	124,984	161,083	248,104	36,099	87,021
콘텐츠	111,532	109,407	115,335	-2,125	5,928
IT 임대 및 도소매업	101,642	98,196	116,387	-3,446	18,191
IT 산업	851,719	932,024	1,006,876	80,305	74,852
전 산업 대비 비중	4.00%	4.10%	4.20%	4.70%	7.70%

출처: 통계청 (2012)

이는 ‘삼성의 SW기술 반성’¹⁾이라는 기사와 같이, 기업에서는 국내 SW인력에 대한 교육은 실무 기술 보유 관점에서 산업계 수요를 따라가지 못하고 있기 때문이다. 이 기사에 따르면, 전 세계 3억 명 이상이 사용하고 있는 소셜네트워킹 서비스 ‘인스타그램’ 애플리케이션을 개발자 4명이 단 6주를 할애하여 만들었던 것에 반해 유사 서비스를 삼성이 개발한다면 수 백 명이 붙어 1년 정도가 소요될 것이라고 한다. 기사 내용이 다소 과장 됐다고 하더라도 우리나라 SW전문 인력의 현주소가 어디에 있는지를 가늠하게 한다.

1) 삼성의 반성..."SW인력 절반이 기초수준 실력", 김현석 기자, 한국경제, 2016. 06. 22

실무에 바로 투입 가능한 SW 역량을 보유한 전문가 인력 양성을 기준으로 본다면, 최근 대학 졸업 인력들에 대한 산업계와 교육계의 관점은 상당한 차이가 존재한다(홍장의, 2012). 산업계에서는 실무능력을 보유한 인력을 요구하는 반면, 대학에서는 학문 기반의 인력을 배출하면서, 이러한 간극은 채용 시장의 경력직 선호 현상의 원인이 되고 있다.

정부는 학력위주사회 문제 개선 및 능력위주사회 안착, 현장중심의 교육개선을 목표로 2013년 5월 21일 국가직무능력표준(NCS, National Competency Standard) 개발 계획을 수립, 2014년 833개 직종을 개발하여 제시하였다. 그러나 소수 전문가들의 의견을 기반으로 단기간에 작성되어 세부 SW 분야별 특수성에 대한 고려가 부족하고, 분류체계의 직무 구분이 현업에서 사용되는 구분과 일치하지 않는 부분이 있어 지속적인 보완이 필요하다. 더욱이 SW역량이 기업의 핵심역량으로 변해가면서, 산업계에서 활용하는 SW인력 범위가 SW개발자, 데이터베이스 관리자 등 SW전문분야에 국한되지 않고 소셜미디어관리자와 같이 기타 연관 직무들로 확장되고 있기 때문에 기존 SW전문 인력 위주로 고려된 직종 및 직무 분류체계의 개선은 지속적으로 이루어져야 한다.

최근 미래부에서는 4차 산업 혁명 도래에 대한 준비하는 차원에서 SW산업 현황을 조사한 결과 매출, 영업이익, R&D 투자, 일자리 모든 분야에서 지속적인 증가추세를 보이고 있다는 조사 결과를 발표하였다.²⁾ 청년실업의 심화와 경기둔화 속에서 SW산업의 약진은 SW인력 양성의 중요성을 여실히 나타내고 있다. 보다 효과적인 SW인력양성 정책 수립을 위해서는 명확한 시의적인 SW인력 변화 파악이 선행되어야 한다.

그러나 대부분의 관련 연구들은 주로 SW인력 수급 실태 조사 등 설문조사를 기반으로 하는 연구들이 대부분이다. 따라서 현재까지의 연구 결과들은 급변하는 SW산업 현황을 명확히 반영하지 못할 뿐만 아니라, 6개월에서 1년간의 장기간 연구기간 소요로 인한 시의성 있는 결과의 도출에 대한 한계가 존재한다. 더욱이 대부분의 과거 연구들은 횡단연구(Cross-sectional Study)³⁾ 방식으로 진행되어 시간 변화에 따른 파악이 불가능하다는 한계가 있다. 이를 해결하기 위해서는 실시간으로 수집이 가능한 산업 현장의 직무 관련 정보들에 대한 데이터 분석을 기반으로

2) 미래부 “제 15차 민관합동 SW TF 회의”, SW 중심사회, 2015. 05. 28

3) 횡단연구(Cross-sectional Study): 일정시기에 자료를 수집하여 분석하는 연구

SW산업계의 기술 및 인력 수요 변화를 확인하는 것이 필수적인 상황이다. 그러므로 현실적인 SW인재양성을 위한 분야별 특성을 고려한 직무 및 직종의 가이드라인을 제공하는 것은 향후 SW관련 산업의 발전 가능성에 중요한 의의를 가진다.

따라서 본 연구과제는 SW 산업 현장에서 요구하는 직무 관련 지식 및 직종에 대한 산업 수요 체계를 SW분야 채용공고 데이터를 전수 조사하여 종단연구(Longitudinal Study)⁴⁾방법으로 분석하여 SW분야의 실수요에 적합한 직무·직종 분류 체계 개발 및 인력 수요를 확인하고자 하였다.

연구 목적에 따라 본 보고서에서는 SW 산업의 직무 및 직종 분류 환경이 어떤 상황인지 진단하는 단계에서부터 시작하여 최근 고용노동부에서 진행하고 있는 능력중심사회의 한 부분인 ‘국가직무능력표준(NCS)’의 SW산업에 대한 적합성 실태를 분석하였다. 또한 실무 관점에서 해당 분류체계가 SW산업계에 적용 가능한 영역을 확인하고자 빅데이터 분석 기반의 조사를 진행하였다. 특히 현재까지 SW 직무직종 분류체계에 대한 사회전반의 관심이 증가하고 있음에도 불구하고 관련 연구들이 소수의 전문가 인터뷰 중심으로 분류체계가 정해져왔다는 한계를 보완하기 위해서 본 연구에서는 실제 산업계 수요를 반영한 분류체계 수립을 시도하였다. 이를 위해서 SW인력 채용공고에서 추출된 핵심단어(Keyword) 기반의 빅데이터 분석을 바탕으로 SW산업계가 요구하는 직무·직종 체계를 개발하였다.

따라서 본 연구는 빅데이터 기반의 SW분야 채용공고에 대한 추출된 핵심단어(Keyword) 기반의 텍스트마이닝 분석을 바탕으로 산업계의 실수요를 반영한 직무·직종 체계 개발 하고자 함에 그 목적이 있다. 이에 따라 2절에서 본 보고서의 연구 목적을 체계적으로 정리하였다.

4) 종단연구(Longitudinal Study): 일정기간에 걸쳐 반복적으로 자료를 수집하여 분석하는 연구

제2절 연구 목적

1절에서 제시한 SW관련 환경에 대한 내용에 따라 본 보고서의 목적은 아래와 같이 설정하였다. 특히 SW분야의 직무·직종 체계에 대한 채용공고 분석을 통한 산업계의 현실적 수요를 반영한 직무·직종 분류체계를 개발하고 SW인력의 효과적인 양성을 위한 산업현장이 요구하는 직종 및 직무 역량에 대한 조사를 진행함이 필요하다.

1. 국가직무능력표준 (NCS) 및 워크넷의 현황 파악 및 문제점 도출
2. 빅데이터 분석 기반 新SW 직무·직종 분류체계 개발
3. 新직무·직종 분류체계의 명칭 및 세부 정의 개발을 위하여 구인 정보 분석 및 국제직무분류표준 방식과의 비교를 통한 직무별 특성 확인

위의 연구 목적에 따라 이어지는 2장에서는 빅데이터 분석 기법을 적용한 직무·직종 체계를 도출하기 위해서 사용된 워크넷 데이터에 관한 설명, 워크넷을 이용한 채용공고 현황 분석 및 공고 기업에 대한 기술통계 분석 결과를 서술하였다. 이후, 본 연구에서 사용된 빅데이터 분석 방법에 대한 정의 및 특성을 설명하고 이를 이용하여 각 직무·직종의 정의 및 핵심단어를 도출하고, 직무·직종 체계를 개발하였다.

제2장 빅데이터 분석을 통한 SW 분야 직무·직종 체계 개발

제1절 데이터 분석 개요

1. 분석 자료의 수집 대상 및 적합성

본 연구과제는 실무 현장에서 요구하는 직무·직종 분류와 핵심 요구사항을 확인하기 위하여 워크넷(www.work.go.kr)의 SW분야 채용공고 데이터를 수집하여 분석하고자 하였다. 워크넷은 1998년 서비스를 시작으로 고용노동부에서 제공하는 채용정보 제공 사이트의 일종이다. 워크넷은 구직·구인정보와 진로 정보 등을 확인 할 수 있는 정부 운영 사이트로 일자리 창출 및 구인 효율 증가를 목적으로 운영되고 있다. 정부 3.0 출범을 바탕으로 현재 워크넷은 Open API 제공 및 지역별 실시간 일자리 현황 등의 서비스가 활용되고 있으며 사람인이나 잡코리아 등과 같이 민간에서 운영되고 있는 채용정보 서비스 보다 훨씬 방대한 양의 데이터를 보유하고 있기 때문에 본 연구에 가장 적합한 분석 대상으로 판단되었다.

특히 워크넷은 SW산업과 관련하여 다양한 정보 및 최근 업계 동향을 반영하고 있어서 실제 업계가 요구하는 전문 역량 및 SW산업의 현재 동향 정보를 포함하고 있다. 따라서 워크넷 자료를 통해 SW분야의 실무적 특성을 분석하고, 산업계가 실제 요구하는 구인 수요와 주로 인력이 집중되고 있는 직무·직종 역량체계를 확인할 수 있을 것으로 기대된다. 본 보고서는 전술한 연구 목적에 따라 워크넷의 SW분야 채용공고 자료를 수집 및 전처리한 후 빅데이터 분석을 통한 직무·직종 분류체계를 도출함으로써 SW 산업의 실제 수요에 대응할 수 있는 SW 인력 양성 정책 수립에 가이드라인을 제공할 수 있다.

본 장에서는 분석대상으로서 워크넷 채용공고 자료의 적절성을 확인하였다. 본격적인 직무·직종 분류체계를 구성하기 위하여 2절에서는 워크넷에서 획득한 채용공고 데이터에 대한 분석데이터로 변환하기 위한 전처리 과정을 기술하였다. 이후 3절에서는 전처리된 분석용 데이터를 바탕으로 SW산업에 대한 기초 분석과 워크넷 내 직무 분류와 NCS의 분류에 따른 채용공고 데이터를 분석하고 두 직무 분류 형태의 불일치성 분석을 진행하였다. 이후 4절에서는 데이터 기반의 실수요

지원을 위한 직무 및 직종 분류와 특성 파악을 위하여 워크넷 데이터에서 획득된 문자형 데이터를 바탕으로 분석을 수행하기 위한 텍스트마이닝 절차를 기술하였다. 5절에서는 전처리된 텍스트 데이터를 바탕으로 직무·직종 군집 및 핵심단어를 도출하기 위하여 활용된 빅데이터 분석 알고리즘의 일종인 잠재디리클레할당(Latent Dirichlet allocation, LDA)기반 토픽모델링 분석 알고리즘의 개념과 분석 결과를 기반으로 각 직무·직종 군집별 핵심단어 기반의 직무·직종 분류 군집들을 도출하였다. 또한 각 직무·직종 군집에 대하여 국제직무분류표준의 직무·직종 정의와 설명을 바탕으로 각 직무와 직종 군집의 명칭 및 정의 등을 도출하였다. 이 결과를 바탕으로 도출된 핵심단어 기반 직무·직종 체계에 따른 직종별 채용 공고의 내용을 파악하였다. 결론에서는 특정 직무·직종 종사자에게 산업계에서 실제 요구하는 핵심 역량을 발견하고 SW 관련 직종의 분류 체계를 구성하였다.

제2절 분석 자료의 전처리 과정

수집된 채용공고 자료를 바탕으로 획득된 데이터는 각 채용공고 별로 공고명, 기업정보, 지원 자격, 모집요강 등의 자료들로 구성되었다. 해당 구인 정보들을 분석 전 처리 과정을 통해 분석에 활용 가능한 수준으로 변환하는 전처리 과정을 거쳐 분석에 이용하였다. 또한 한국고용직업분류표⁵⁾에 따른 워크넷의 모집직종 단어가 수집되었다. 이를 바탕으로 워크넷과 NCS 기반의 한국고용직업분류표에 따른 직종별 기초 채용공고 경향을 확인하였으며 워크넷과 NCS의 구인 정보에 대한 직무 분류 형태의 일치성 정도를 확인하였다. 이후 수집된 자료들을 단어별로 전처리 한 후 텍스트마이닝을 통해 신규 직무·직종 분류체계를 도출하는 과정을 거쳤다. 이를 바탕으로 최종적으로 실무 수요 중심의 직무·직종 분류체계와 각 직무별 핵심단어들을 도출함으로써 직종에 따른 산업계가 중요하게 생각하는 역량을 직무별로 나타내고자 하였다. [그림 2-1]은 전체 자료 수집 단계에서부터 전처리 과정을 거쳐 분석을 진행하는 SW 관련 채용공고 데이터를 분석하는 과정을 묘사한 것이다.

5) 통계청, 고용직업분류, KECO, https://kssc.kostat.go.kr:8443/ksscNew_web/index.jsp#

[그림 2-1] 데이터 분석 절차 및 방법

분석 단계	데이터 수집	데이터 정제	데이터 탐색	분류 체계 도출
분석 내용	• 워크넷 사이트 데이터 수집 • 구인공고 내역 • 모집직종, 키워드	• 단위 변환 • 불용어 처리 • 단어 추출 • Term-Document 매트릭스 생성	• 기술 통계 분석 • 직종별 워드 클라우드 • 직종간 내용 일치도	• 구인공고 토픽 생성 • 유사도 기반 구인공고 클러스터 생성 • 직종 분류 체계 도출
적용 방안	• Web crawling	• NLP • Text mining	• Text mining • Word cloud • Jaccard similarity	• LDA

1. 범주형 및 숫자형 데이터의 전처리 사항

분석에 앞서 수집된 채용공고 데이터들은 연구목적에 맞도록 각각의 분석에 적합한 형태로 변환하기 위한 정제 과정을 거쳤다. 워크넷에서 수집된 채용공고들은 공통된 구인 정보 양식을 활용하지 않고 자유롭게 입력되어 작성된 자료들이 대부분이었다. 따라서 분석을 위한 항목의 값을 입력하는 방식도 구인정보를 등록한 회사에 따라 모두 상이한 형태로 작성되어 있었다. 그러므로 자료 전처리 과정을 통해 워크넷에서 수집된 자료를 동일한 기준에서 분석이 용이한 형태로 변환하는 것은 필수적인 작업이다. 이는 각 채용공고들을 동일한 의미로 분석 및 해석할 수 있음을 의미한다. 마찬가지로 같은 값을 뜻하면서도 상이하게 작성된 항목들도 발견되어 모두 동일한 형태의 자료 값으로 변환하였다. 예를 들어 ‘연봉’의 경우 ‘2천만 원’이라고 표시된 공고와 ‘20,000천원’이라고 표시된 공고들은 같은 값을 뜻하지만 다른 형태의 수치로 입력되어 있는 값으로 해당 형태의 값들은 모두 단일한 형태인 ‘20,000,000원’으로 변환하였다. 또한 ‘연봉’ 대신 ‘월급’으로 표기된 공고도 ‘연봉’으로 환산하여 자료 값을 변환하였다. 추가적으로 분석의 목적에 맞도록 얻어진 수치 값 중에서 연속형 변수를 ‘구간’ 변수로 구분하여 계급값을 만들었으며, 계급값에 따른 빈도를 구하였다. 범주형 변수와 숫자형 변수의 처리 방안과 문자형 변수의 전처리 과정 또한 변환하여 해당 사항을 기술하였다. 다음의 각 세부 내용은 기초 분석을 위한 변수 별 전처리 내역이다.

1) 회사명

‘회사명’과 같은 변수는 각 회사별로 단일한 회사명을 보유하고 있는 경우가 대다수이기 때문에 이 변수항목의 경우 범주화하여 빈도를 구하거나 숫자로 변환하여 통계치를 작성할 필요가 없다. 따라서 <표 2-1>과 같이 ‘회사명’ 변수의 값은 데이터 수집기로 얻어낸 원본 데이터 (워크넷에서 수집된 원 데이터를 의미함)와 전처리 가공을 끝낸 데이터가 동일하다.

<표 2-1> 회사명 전처리의 예시

이름	원본 데이터	가공 데이터
회사명	WELCO	WELCO
	(주)카이스스	(주)카이스스
	보엠컴퍼니	보엠컴퍼니
	sc엔지니어스	sc엔지니어스

2) 근로자 수

‘근로자 수’는 기업별로 상호 비교가 가능하도록 원본 데이터에서 수량과 단위를 분리하여 <표 2-2>와 같이 숫자로 표기하였다. 이후 근로자 수를 10인 미만에서부터 10인 이상과 50인 미만 그리고 50인 이상과 100인 미만 그리고 100인 이상, 미등록으로 범주화하였으며 각 범주별로 어느 정도의 기업 근로자 규모를 보유하게 되는 지 빈도를 계산 하였다.

<표 2-2> 근로자수 전처리 예시

이름	원본 데이터	가공 데이터
근로자수	7명	10인 미만
	40명	10인 이상 50인 미만
	19명	10인 이상 50인 미만
	200명	100인 이상

3) 지역

수집된 워크넷 원본 데이터에서 수집한 회사 주소의 경우 <표 2-3>과 같이 행정 구역을 구분하는 특별시, 광역시, 특별자치시, 도, 특별자치도 단위로 구분하였다. 하지만 일부 행정 구역 단위를 표시하지 않은 채용공고의 경우 해당 회사명을 직접 검색하여 검색된 주소 지역으로 할당했다.

<표 2-3> 지역 전처리의 예시

이름	원본 데이터	가공 데이터
지역	(42168) 대구광역시 수성구 들안로 1, 임시변경 (상동)	대구
	(336030) 충청남도 아산시 충무로 100 (권곡동)	충남
	(22314) 인천광역시 중구 제물량로 214 (해안동2가)	인천
	(150714) 서울 영등포구 여의도동 44-1 대영빌딩3층	서울

4) 자본금

‘자본금’은 <표 2-4>처럼 한글과 숫자가 혼용되어 있기 때문에 동일한 분석단위로서 공고내용을 숫자로 변환하였다. 숫자로 변환된 자본금은 금액에 따라 12개 수준의 단위 (1억 원 이하, 1억 원 초과~2억 원 이하, 2억 원 초과~3억 원 이하, 3억 원 초과~5억 원 이하, 5억 원 초과~10억 원 이하, 10억 원 초과~20억 원 이하, 20억 원 초과~30억 원 이하, 30억 원 초과~50억 원 이하, 50억 원 초과~100억 원 이하, 100억 원 초과~200억 원 이하, 200억 원 초과, 미등록)로 구분해 범주별 빈도를 구하는 방식을 적용하였다.

<표 2-4> 자본금 전처리의 예시

이름	원본 데이터	가공 데이터
자본금	오억	3억 원 초과 5억 원 이하
	1,300 백만 원	10억 원 초과 20억 원 이하
	140,000,000 원	1억 원 초과 2억 원 이하
	1천5백3십만 원	1억 원 이하

5) 고용형태

〈표 2-5〉 고용형태 전처리의 예시

이름	원본 데이터	가공 데이터
고용형태	기간이 정함이 있는 근로계약 12개월 / 계약기간 만료 후 상용직전환검토	계약직
	정규직	정규직
	계약직 후 정규직 전환 검토·계약직	계약직
	인턴 후 정규직 전환 검토	인턴

수집된 원본데이터의 채용광고에서 ‘고용형태’는 일정한 형식으로 작성되지 않았다. 또한 채용공고를 위해 등록된 기업들마다 다양한 유형의 고용형태가 제시되어 있다. 예를 들어서 〈표 2-5〉의 첫 줄과 세 번째 줄의 경우 모두 계약직 고용형태에 해당하지만 다르게 표현되고 있으며, 다양한 조건들이 고용형태에 기술되고 있었다. 따라서 전처리 과정을 거쳐 ‘고용형태’는 정규직, 계약직, 프리랜서, 인턴, 파 시간제, 파견직, 병역특례, 전임, 위촉직, 별정직, 미등록으로 구분하였다. 계약직이나 시간제의 경우 1년, 2년 등 기간이 명시된 경우가 있지만 본 분석의 전처리 과정에서는 별도의 분류 값으로 구분하지 않고 기본적인 고용형태인 계약직이나 시간제로 분류했다.

6) 모집인원

모집인원은 근로자 수를 처리하는 방법과 동일한 방식으로 인원 수 만을 추출하여 〈표 2-6〉과 같이 모집인원을 1명, 2명, 3명, 4명, 5명, 5명 초과 10명 이하, 10명 초과, 미등록으로 구분하였다. 구분된 범주를 바탕으로 각각의 빈도를 계산하여 분석 자료로서 활용하였다.

〈표 2-6〉 모집인원 전처리

이름	원본 데이터	가공 데이터
모집인원	‘2명 한 지원자가 알선과 이메일 입사지원을 모두 한 경우 총인원수에서는 한명으로 카운트됩니다.’	2
	20명	10명 초과
	1명	1
	‘4명 한 지원자가 알선과 이메일 입사지원을 모두 한 경우 총인원수에서는 한명으로 카운트됩니다. ‘	4

7) 임금조건

‘임금조건’은 월급과 연봉으로 상이하게 명시된 공고내용을 연단위로 환산하여 전처리 과정을 진행하였다. 예를 들어 <표 2-7>과 같이 월급 160만원으로 표기되어있는 공고는 19,200,000원으로 환산되었다. 또한 2,400만원 ~ 2,600만원과 같이 연봉이 범위로 나와 있는 경우 범위의 평균값을 연봉으로 환산 적용하였다.

<표 2-7> 임금조건 전처리의 예시

이름	원본 데이터	가공 데이터
임금조건 (연봉)	연봉 (2,400만원 ~ 2,600만원)	25,000,000
	월급 160만원	19,200,000
	연봉 3600만원	36,000,000
	연봉 2,000만원	20,000,000

8) 경력조건

‘경력조건’은 채용공고에서 요구하는 경력을 경력, 신입, 무관, 미등록으로 구분하였다. <표 2-8>과 같이 기업이 요구하는 경력 기간을 1년, 3년, 5년, 10년 단위로 구분하였으며, 경력과 신입을 모두 모집하는 경우는 경력의 ‘년’ 단위를 구분하지 않고 ‘신입, 경력’으로 구분하여 자료화 하였다.

<표 2-8> 경력조건 전처리 규칙

이름	원본 데이터	가공 데이터
경력조건	경력 (최소 1년 이상)	경력 1년
	무관	무관
	신입 , 경력 (최소 2년 이상)	신입, 경력
	경력 (최소 3년 이상)	경력 3년

9) 학력조건

‘학력조건’은 학력무관, 고졸, 대졸(2~3년), 대졸(4년), 석/박사, 미등록으로 구분했다. <표 2-9>와 같이 학력조건을 6개로 구분한 형태에 해당되지 않는 공고 형태 또는 더 넓은 범위의 학력조건으로 채용공고가 등록된 경우 가장 낮은 학력으로 구분했다. 예를 들어 채용공고에서 요구하는 학력조건이 ‘고졸~박사’라면 최하 기준인 ‘고졸’로 구분했다.

<표 2-9> 학력조건 전처리 규칙

이름	원본 데이터	가공 데이터
학력조건	학력무관	학력무관
	고졸 ~ 박사	고졸
	고졸	고졸
	대졸(2~3년) ~ 석사	대졸(2~3년)

10) 외국어 능력

‘외국어 능력’의 경우 원본데이터에서는 영어, 중국어, 일본어, 베트남어, 독일어가 각각 상·중·하로 구분되어 있었다. 그러나 본 보고서의 목적상 ‘외국어 능력’을 요구하는 것이 직무에 필요한 것을 의미하고 그 수준을 중요시하는 것은 연구 범위에서 벗어나기 때문에 <표 2-10>과 같이 ‘등급’을 구분하지 않고 ‘언어’로만 구분하여 데이터를 가공하였다. 추가적으로 하나의 언어가 아니고 두 개 이상의 언어가 외국어 능력에서 요구되는 경우에는 두 개 이상의 언어를 모두 구분하여 처리하였다.

<표 2-10> 외국어 능력 전처리 규칙

이름	원본 데이터	가공 데이터
외국어능력	일어(하)	일어
	중국어(상)	중국어
	영어(상), 중국어(상)	영어, 중국어
	영어(중), 중국어(상)	영어, 중국어

2. 문자형 데이터 전처리

채용공고에서 수집된 범주형 및 수치형 데이터와 같이 기술통계 기반의 현황 분석에 이용되는 자료들과 다르게 문자형 자료의 경우 더욱 복잡한 전처리 과정을 거치게 된다(강성경, 2016). 예를 들어 형태소 구분에 의한 품사에 따라 조사와 같은 분석에 의미가 없는 자료들을 구별하고 제외한 후 명사, 형용사 등 의미를 내포하는 자료들만을 구분해 내는 것이 필요하다. 따라서 본 연구과제는 문자형 데이터에 대하여 형태소 분석을 수행하여 얻어진 단어들을 바탕으로 유사 의미를 가진 단어들을 하나의 단어로 통일된 기준의 단어 목록을 규정하였으며 단위를 통일하는 과정을 거쳤다.

따라서 워크넷 채용공고에서 추출된 데이터를 바탕으로 직종별 채용공고에 많이 등장하는 단어를 활용한 분석을 수행하기 위하여 직종별 단어에 기반을 두어 NCS 직종의 기술, 태도, 지식에 작성된 내용과 비교하여 유사도를 계산하였다. 특히, 범주 값으로 변환하기 어렵거나 변환이 불가능한 문자형 데이터는 문장에서 필요한 단어를 추출한 후 해당 단어의 빈도를 구하였다. 단어를 추출하기 위한 전처리 작업은 다음의 세 단계를 기반으로 진행되었다.

첫 번째 단계는 단어를 치환하는 과정이다. 이 과정은 채용공고 입력자에 따라서 영어와 한글로 다르게 입력된 것을 동일한 것으로 판단하기 위한 과정이다. 예를 들어 ‘Java’ 와 ‘자바’ 라고 표현된 것이 같은 단어임을 인지하고 두 표현을 공통적인 ‘JAVA’ 로 변환하였다. 두 번째 단계는 숫자, 특수기호, 불용어 제거 과정이다. 두 번째 단계는 채용공고에 등장하는 불필요한 특수 문자들이나 숫자들을 제거하기 위하여 필수적으로 수행해야 하는 단계이다. 세 번째 과정은 희소 단어 제거 과정이다. 이 과정에서는 많은 채용공고에서 매우 희소하게 등장하는 단어를 제거하는 것을 주목적으로 하고 있다. 단어의 등장빈도가 희소한 것은 실제로 잘 쓰이지 않는 표현일 수도 있고, 오타로 인해 등장빈도가 낮은 단어일 가능성도 있다. 특히 매우 희소한 단어들은 채용공고를 구분하는 데 많은 영향을 미치지 못하므로 특정 수준을 정해 희소 단어를 제거하였다.

〈표 2-11〉 원본 데이터와 가공 데이터 전처리 방식의 예시

이름	원본 데이터	가공 데이터
제목	웹디자이너 모집(고용촉진 지원금 대상자)	웹디자인
	java 웹개발자(경력2년 이상 초급) 채용	JAVA 웹개발
단어	웹디자이너, 웹 UI디자이너, 포토샵 디자이너,쇼핑몰관리,웹디자인,제품촬영,	웹디자인 웹UI디자인 포토샵디자인 쇼핑몰관리
	웹디자이너, 포토샵 디자이너	웹디자인 포토샵디자인
직무 내용	1. 국내 및 해외수입물품의 마켓 판매용 제작 2. 소셜 및 마켓 상품 등록 및 수정 업무 3. 오픈 마켓	국내 해외 수입물품 마켓 판매 소셜 오픈 상품 등록 수정
	신용카드 단말기 설치 및 포스 설치에 관련된 업무	신용카드 단말기 설치 포스 관련

상기 제시한 문자형 데이터의 3단계 전처리 과정을 통해서 채용공고 데이터의 ‘제목’, ‘핵심어’, ‘직무내용’을 전처리 과정을 진행한 결과는 〈표 2-11〉과 같다. 전처리 과정 이후 각 채용공고에서 추출한 단어에 기반을 두어 공고에 주로 등장하는 단어들의 관계를 발견하였다. 군집별 해당 채용공고들을 유사한 군집으로 묶어서 직종 분류 체계 생성을 위한 기반 정보로 활용하였다. 또한, 단어에 기반을 두어 공고들을 유사한 군집으로 직접 묶는 LDA 토픽 모델링 방법 등 다양한 기법들을 적용하여 분류 체계 도출을 위한 기반 정보를 생성하였다. 이와 같은 과정은 본 보고서의 목적인 데이터 분석 기반의 SW 산업 직무·직종 분류체계의 개발에 적합한 분석을 수행하기 위한 것이다. 수집된 자료의 전처리 과정을 수행하는 이유는 문자형 자료의 분석 가능성을 높일 수 있다.

전처리가 완료된 이후 처리 과정에서 정제된 데이터의 오류가 없었는지 확인하기 위하여 텍스트마이닝 분야의 전문가인 교수 및 데이터 분석 분야의 기업 분석 전문가 5인에게 처리 과정에 대한 타당성 및 신뢰성을 확인하였다. 특히 원본

데이터와 가공 데이터의 변환이 필요한 사안이 발생한 경우 변환 방법에 문제가 없는지 또는 개선 방안이 존재하는 것은 없는지에 대한 타당성 검토에 주안을 두어 자료의 도출 과정 및 타당성 분석을 실시하였다. 그에 따라 수집된 자료들의 전처리 과정에서 오류가 없음을 확인하였다. 이 후 최종 안정화된 정제된 데이터를 바탕으로 채용공고를 게시한 기업들의 특성 분석 및 워크넷 직무 분류 형태와 NCS 직무 분류 형태의 일치성에 대한 기술통계 및 데이터 기반 직무·직종 분류체계를 위한 텍스트마이닝과 토픽모델링 분석을 수행하였다. 이를 바탕으로 이후 3절~5절에서는 정제된 데이터를 바탕으로 수행된 기술통계 및 텍스트마이닝의 절차 및 분석 결과에 대하여 기술하였다.

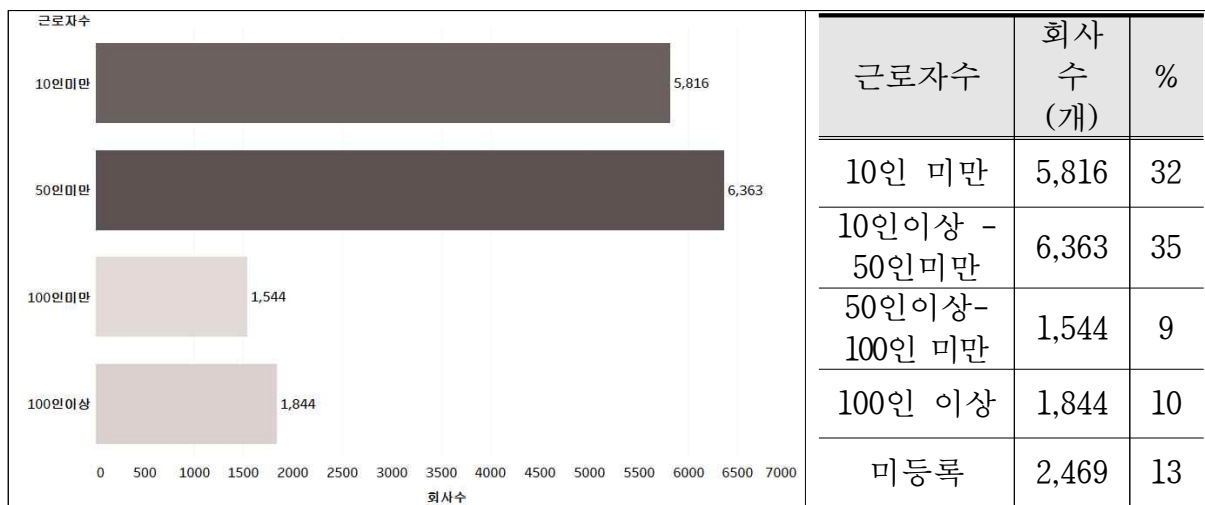
제3절 SW산업 관련 기술통계 분석

1. 전체 구인 기업에 대한 현황

앞 절에서 설명한 워크넷에서 수집된 채용공고 데이터를 대상으로 구인 내용을 등록한 기업들의 특성을 파악하기 위하여 기술통계 분석을 수행하였다. 2절에서 제시한 것과 같이 채용공고 내 데이터의 전처리 과정을 거쳐 타당성을 확보한 데이터를 바탕으로 분석을 수행하였다. 분석을 위해 시각화 프로그램인 Microsoft Power BI, R 그리고 Tableau를 활용하여 자료 분류와 시각화를 수행하였다.

1) 구인 기업의 근로자수

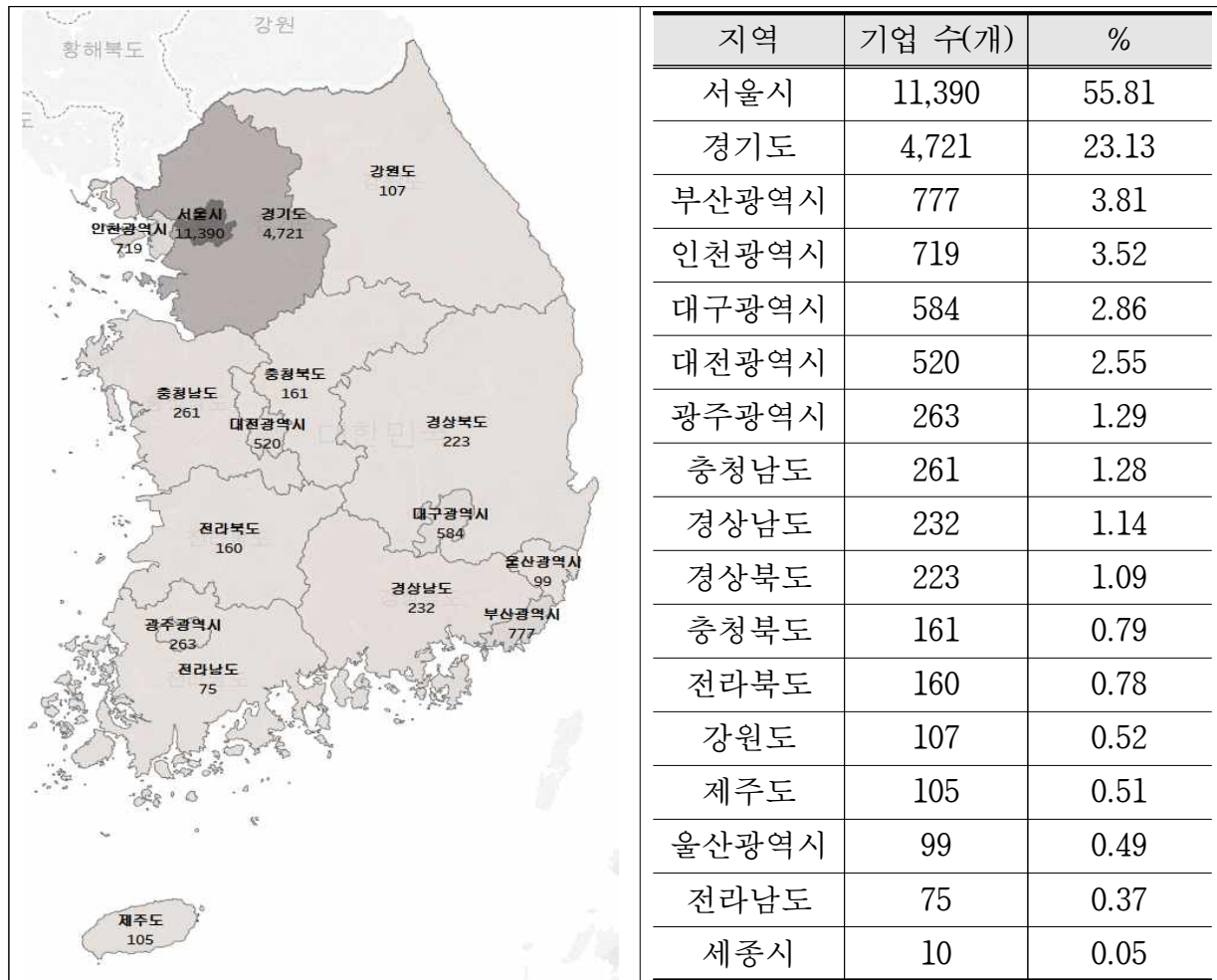
[그림 2-2] 구인 기업 규모별 채용공고



[그림 2-2]와 같이 채용공고를 게시한 기업 중 10인 미만의 근로자가 근무하는 기업의 수는 5,816개로 전체의 32%에 해당하며, 10인 이상 50인 미만은 6,363개의 회사로 35%에 해당하였다. 50인 이상 100인 미만은 1,544개 회사로 9%, 100인 이상 회사는 10%를 차지하였으며 근로자를 표기하지 않은 회사는 미등록으로 분류하였다. 미등록에 해당하는 회사는 2,469개의 회사로 전체의 13%로 나타났다. 이 결과 SW관련 채용공고를 게시한 기업의 대부분은 50인 미만의 회사가 주도하고 있으며 SW산업분야의 기업들이 대부분 다수의 인원을 보유하지 않은 작은 크기의 기업들이 많음을 알 수 있다. 이는 워크넷이 대기업의 공채정보를 포함하지 않고 있기 때문에 발생하는 결과라고 추측된다.

2) 지역별 채용공고

[그림 2-3] 구인 기업 지역별 채용공고

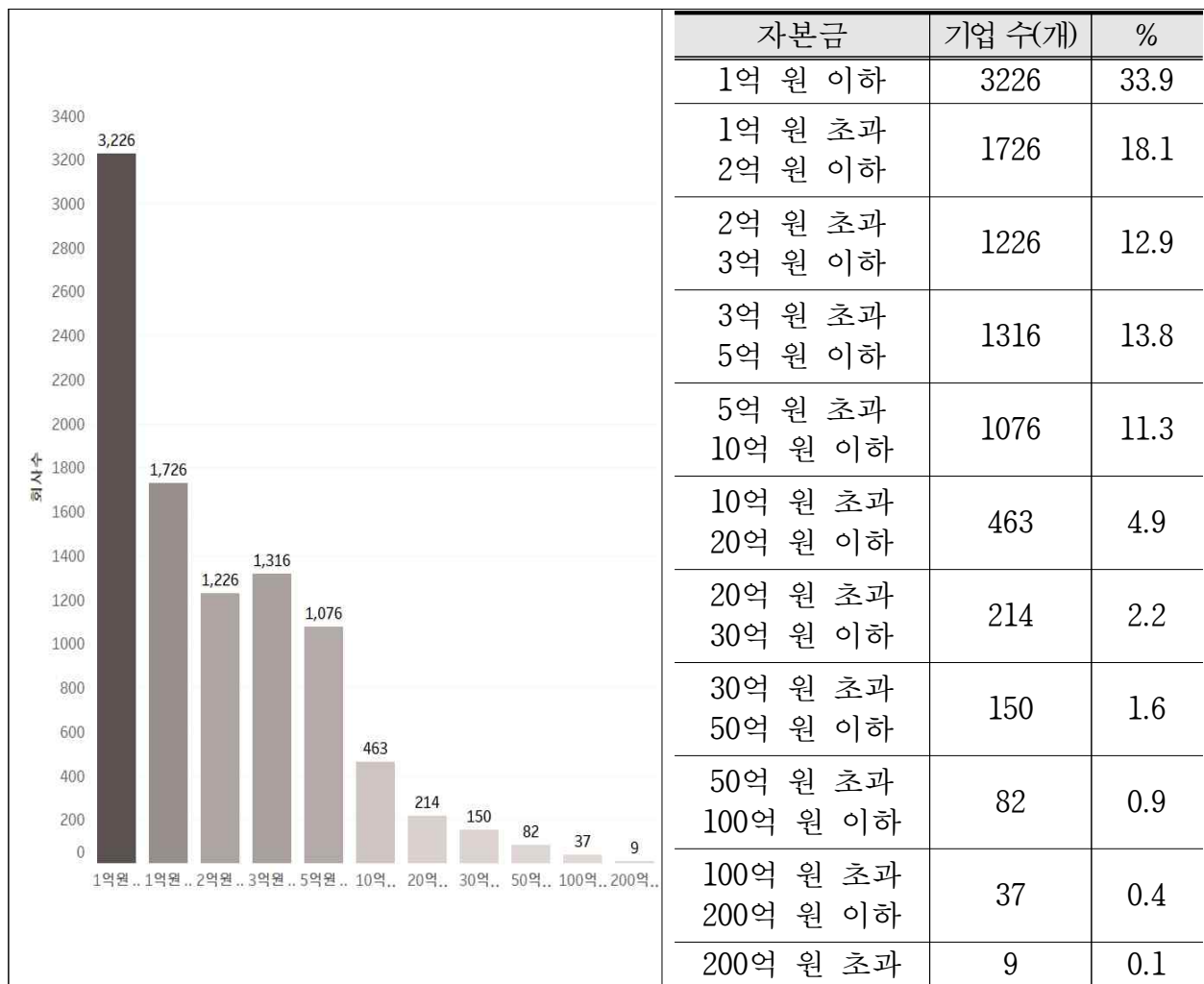


[그림 2-3]과 같이 지역별 채용공고는 서울시 지역이 11,390개 전체의 55.81%로 가장 많은 채용공고가 있는 것으로 파악되었으며, 세종시가 0.05%로 채용공고가 가장 낮은 것으로 파악되었다. 경기도가 4,721개(23.13%), 인천광역시가 719개(3.52%)를 차지하였다. 그 결과 채용공고의 약 83%가 수도권에 집중되는 현상을 보였다. 그 외 지역으로는 부산·울산·경남 권은 약 5.5%, 대구·경북 권은 약 4%, 충청권은 약 2.5%, 전라도 권역은 약 1.7%의 각 채용공고를 차지하는 것으로 나타났다.

3) 자본금

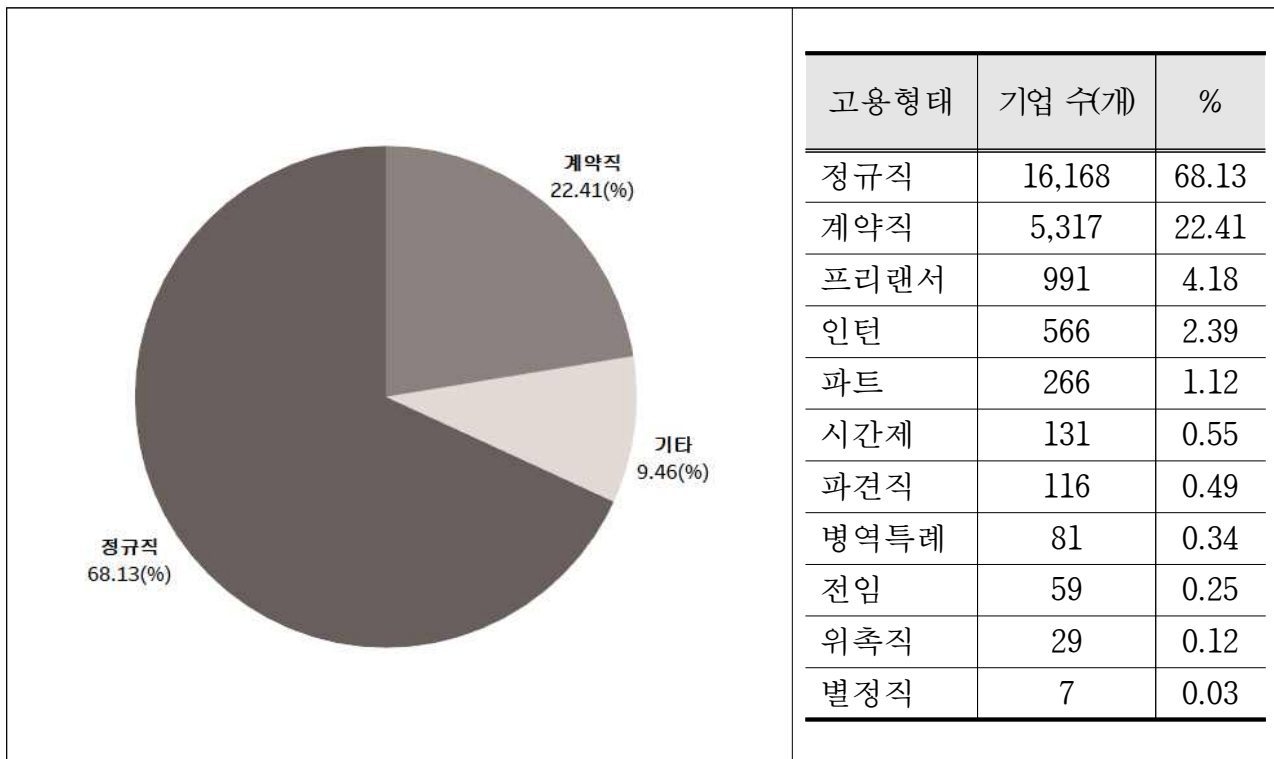
자본금과 관련하여 분석을 수행한 결과 [그림 2-4]와 같이 채용공고를 낸 회사들의 자본금 분포는 1억 원 이하의 회사가 3,226개(33.9%)로 가장 많았으며, 200억 원 초과하는 회사는 9개(0.1%)로 가장 적었다. 2억 원 이하의 자본금을 지닌 회사가 전체 채용공고의 약 50%를 차지하였다. 하지만, 이러한 결과는 전술한 바와 같이 워크넷에서는 대기업의 채용공고를 포함하고 있지 않기 때문에 모든 SW 기업의 채용 특성을 대변하는 것은 아님을 주지하여야 한다.

[그림 2-4] 구인 기업 자본금별 채용공고



4) 고용형태

[그림 2-5] 구인 기업 고용형태별 채용공고

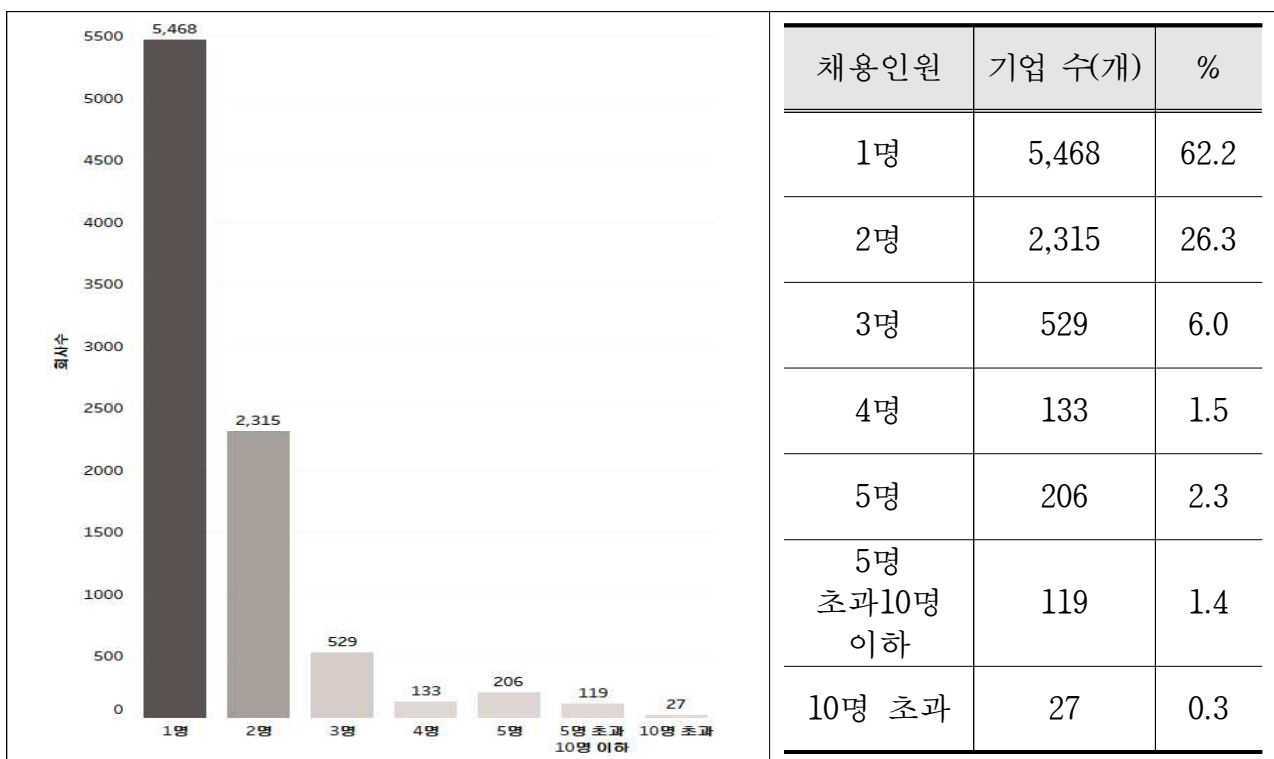


고용형태는 [그림 2-5]와 같이 정규직, 계약직, 기타의 순으로 나타났다. 정규직을 채용하는 회사는 16,168개로 전체의 68.13%, 계약직은 5,317개(22.41%)에 해당하였다. 기타는 인턴, 파트타임, 시간제, 파견직 등으로 구성되어 있으며 전체의 9.46%에 해당하였다.

5) 채용인원

[그림 2-6]과 같이 채용인원을 1명만 모집하는 회사는 전체의 약 62%(5,468개)를 차지하였다. 2명을 채용하는 회사는 26.3%로 1명 또는 2명을 채용하는 회사가 대다수인 것으로 확인되었다. 10명을 초과하여 채용하는 회사는 적었지만, 이들 중에는 100명단위의 인원을 채용하는 경우도 확인되었다.

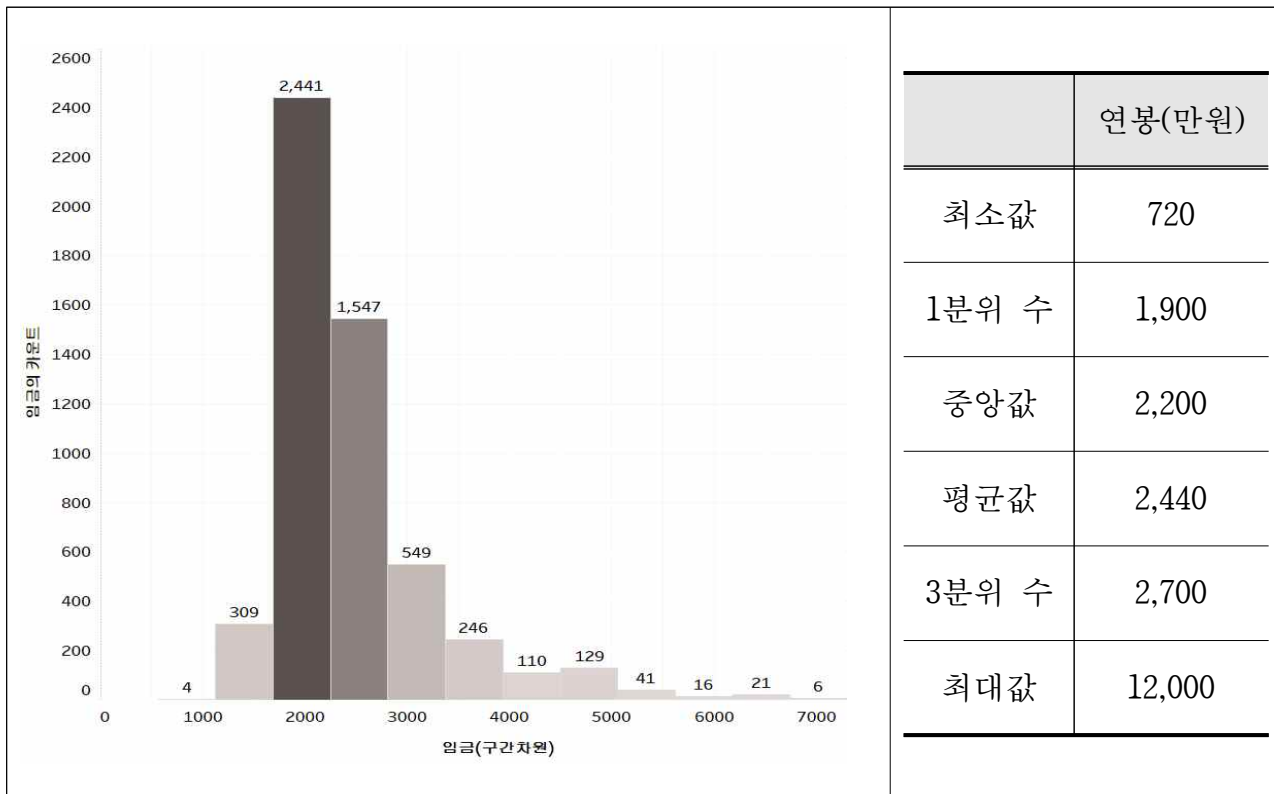
[그림 2-6] 구인 기업 채용인원별 채용공고



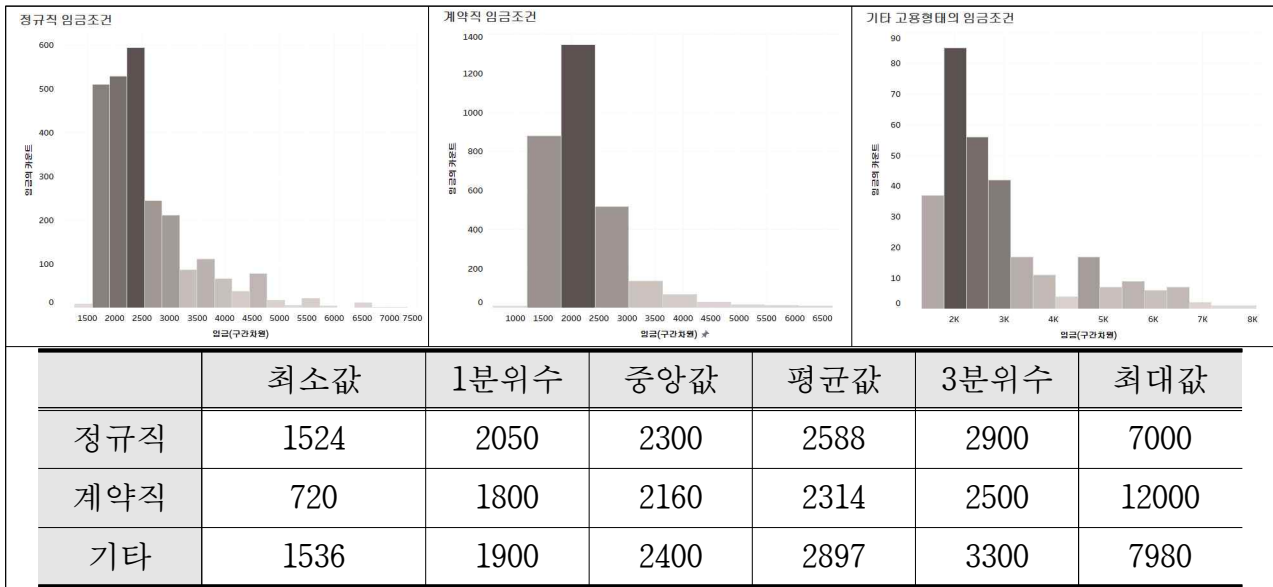
6) 임금조건(연봉)

[그림 2-7]과 같이 연봉조건은 평균 2,440만원으로 확인되었다. 중앙값은 2,200만원으로 나타났으며, 최소 연봉은 720만원, 최대 연봉은 12,000만원으로 나타났다. 대부분의 연봉조건은 2,000만 원대를 형성하였다. 분석 결과와 같이 워크넷에서 구인되는 SW분야 인력의 급여 조건은 큰 격차를 보이는 것으로 나타났다. 그러나 전반적으로는 낮은 조건의 급여 조건을 보이고 있다.

[그림 2-7] 구인 기업 임금조건별 채용공고



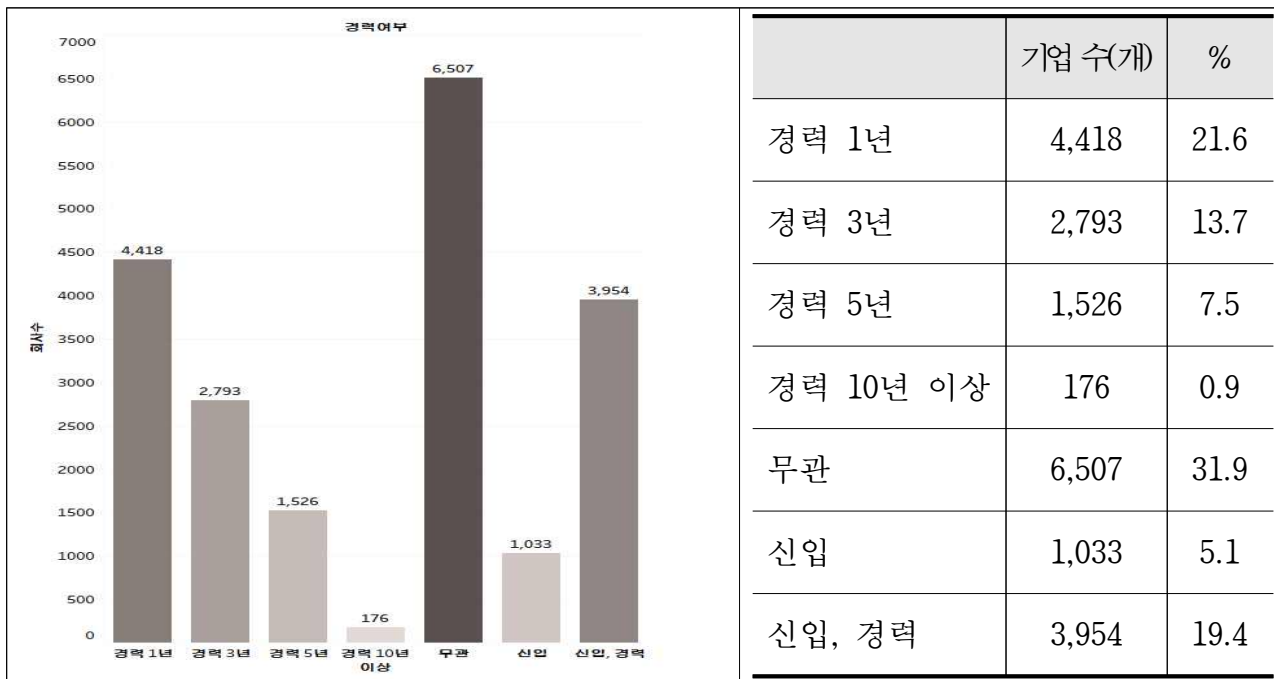
[그림 2-8] 구인 기업 고용형태에 따른 임금조건별 채용공고



고용형태에 따른 임금조건을 보면 상기 [그림 2-8]과 같은 결과를 보이고 있다. 분석 결과 정규직의 임금조건이 계약직에 비해 더 나은 조건으로 나타나는 것을 볼 수 있다. 정규직의 최소 연봉조건은 1,524만원이며, 중앙값은 2,300만원, 평균값은 2,588만원이다. 반면 계약직의 경우, 최소 연봉조건은 720만원, 중앙값 2,160만원, 평균값 2,314만원으로 나타났다. 계약직의 경우, 시간제근무 근무 조건을 포함하고 있기 때문에, 최소 연봉조건이 풀타임을 기준으로 한 것은 아니다. 또한, 정규직이 68%의 분포를 보이는 것으로 나타나 전반적으로는 정규직의 활용에 의해 운영이 되고 있는 것을 알 수 있다.

7) 경력 조건

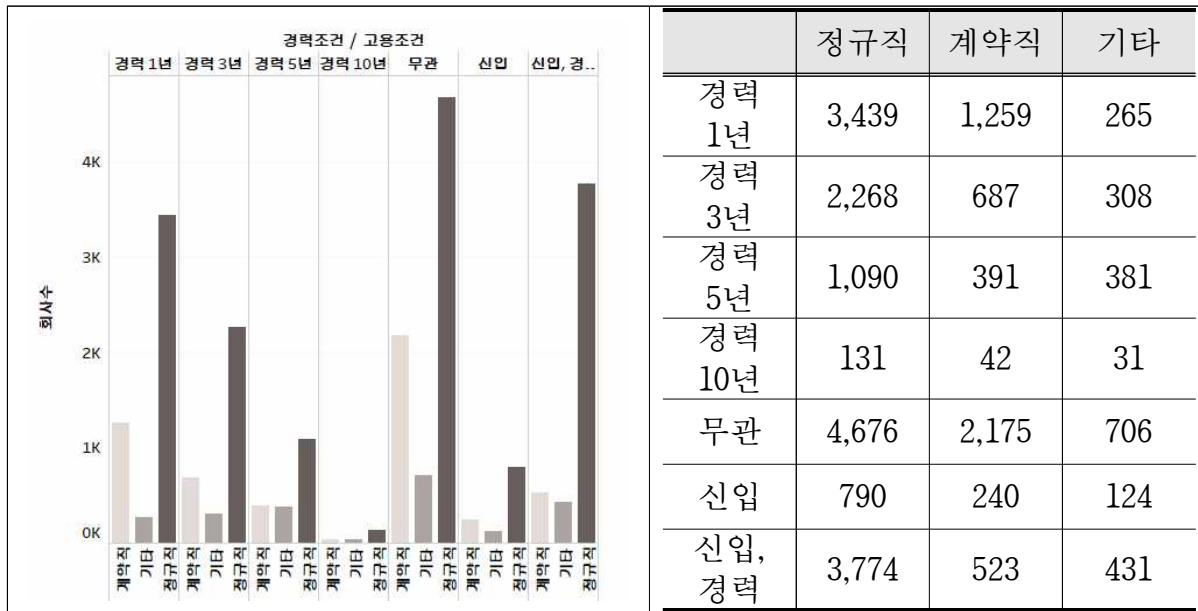
[그림 2-9] 구인 기업 경력조건별 채용공고



[그림 2-9]와 같이 채용공고에서의 경력조건은 무관인 경우가 약 32%로 가장 많이 나타났고, 경력 1년에서 경력 5년 사이를 채용하는 경우가 약 40%에 해당하였다. 신입만 채용하는 경우는 5.1%이었고, 신입과 경력을 동시에 모집하는 경우는 19.4%였다. 경력여부는 경력 1년~5년의 인력을 모집하는 경우가 약 40%로서 대부분의 채용이 경력자를 선호하는 것으로 보인다. 이는 기업에서 실무 경험 어느 정도 있는 인력을 요구하고 있으며, 산업계가 요구하는 능력과 신규 인력의 격차가 존재한다는 것을 알 수 있다.

[그림 2-10]과 같이 고용형태에 따른 경력조건을 살펴보아도 상기 [그림 2-9]와 같은 분포를 보이는 것을 확인할 수 있었다. 경력 무관인 경우와 경력 1년이 대부분을 차지하였다. 다만 정규직은 경력 3년을 원하는 경우가 계약직이나 기타에 비해 많음을 확인 할 수 있다. 역시 마찬가지로 분석 결과는 신규 졸업자보다는 경력직을 채용하고자 하는 경향을 보이고 있다. 이는 대학 졸업을 한 신규 인력들에 대한 능력이 산업계가 원하는 수준과는 격차가 존재한다는 점을 유추해볼 수 있다.

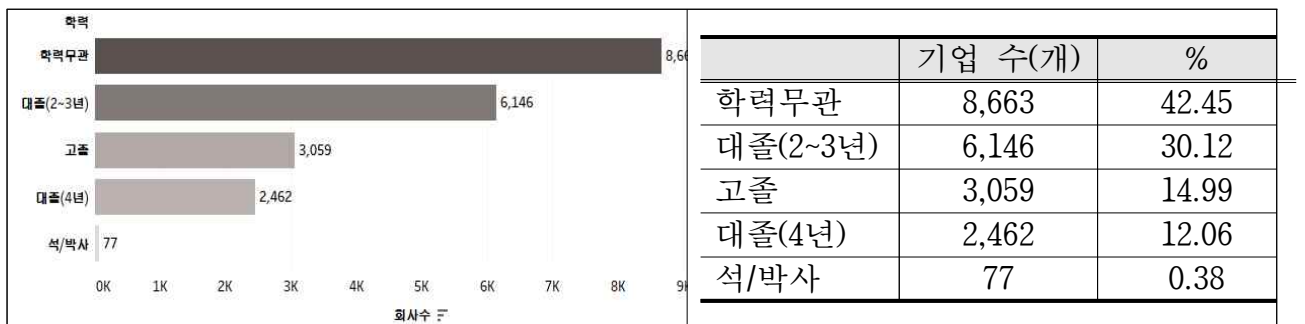
[그림 2-10] 구인 기업 고용형태에 따른 경력조건별 채용공고



8) 학력 조건

[그림 2-11]과 같이 채용공고에서 학력조건은 무관인 경우가 42.45%로 가장 많이 나타났다. 2에서 3년제 대학의 졸업자와 4년제 대학의 졸업자를 채용하는 회사는 전체의 약 42%로 나타났으며, 고졸 이상의 경우는 약 15%에 해당하는 것으로 나타났다. 일반적으로는 고학력의 인력보다는 대졸 인력 수준의 인력을 요구하는 것으로 보이고 있으나 전반적으로는 학력 무관 공고가 42%정도를 보이고 있다. 이는 실제 산업계에서는 실무적 능력을 보유하고 있다면 학력과는 무관하게 채용이 이루어지고 있다는 점을 시사하고 있다.

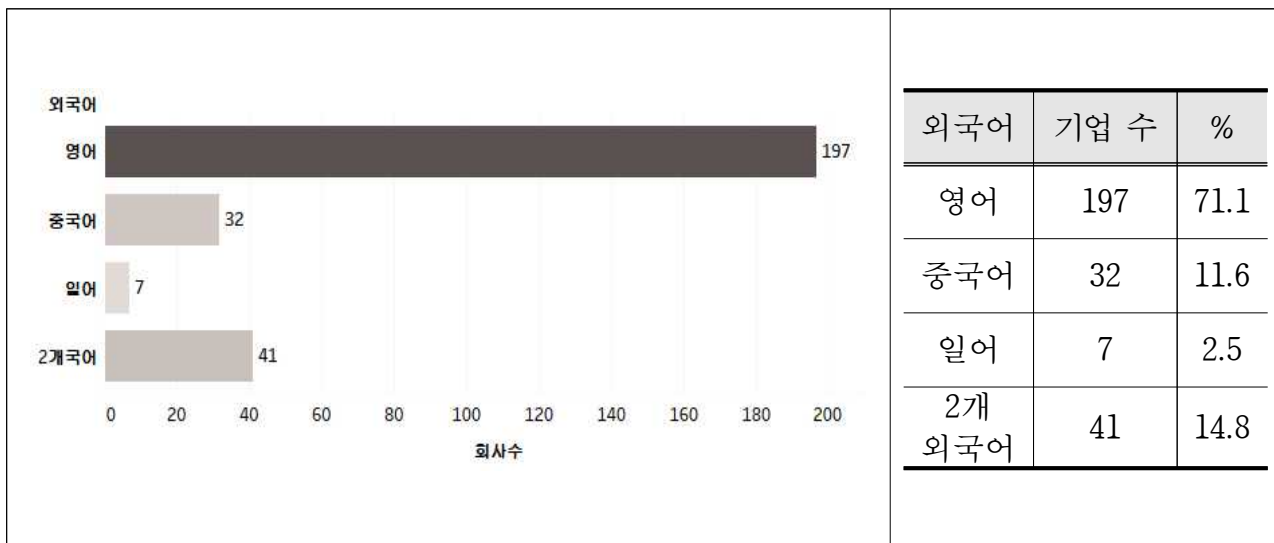
[그림 2-11] 구인 기업 학력조건별 채용공고



9) 외국어 능력

[그림 2-12]와 같이 채용공고에서 외국어 조건의 경우 영어를 조건으로 하는 회사가 86.8%로 나타났다. 중국어(11.6%), 일어(2.5%)는 다른 제 2 외국어와 다르게 단일 언어로 채용조건에 포함되어 있었다. 채용공고에서 2개 외국어를 요구하는 경우는 영어를 기본으로 독일어 또는 베트남어, 중국어 및 일어를 포함하는 형태로 채용조건에 나타났다. 마찬가지로 대다수의 기업은 영어 능력을 매우 집중하고 있다. 이는 대부분의 SW 프로그램이 영어를 기반하고 있다는 점도 반영되고 있는 것으로 보인다.

[그림 2-12] 구인 기업 외국어 능력별 채용공고



2. 워크넷 직업 분류 기반 채용공고 기술통계

1) 워크넷 직업 분류

워크넷 직업 분류는 한국고용직업분류표⁶⁾를 기초로 하여 분석되었다. 한국고용직업분류표는 <표 2-12>와 같은 기준으로 나뉘어져 있다. 본 보고서의 목적이 되는 SW 관련 직군은 중분류 번호 20 정보통신 관련 직으로 분류된다. <표 2-12>을 따라 워크넷 채용 공고에서 모집직종이 20으로 시작되는 자료를 가지고 6개의 소분류로 구분했다.

<표 2-12> 워크넷 중·소분류

중분류 코드	중분류 이름	소분류 코드	소분류 이름
20	정보통신관련직	201	컴퓨터 및 통신공학 기술자 및 연구원
		202	컴퓨터시스템 설계 및 분석가
		203	SW개발전문가
		204	웹전문가
		205	데이터베이스 설계·구축 및 관리기술자
		206	통신 및 방송 장비기사 설치 및 수리원

소분류 201(컴퓨터 및 통신공학 기술자 및 연구원)은 <표 2-12>와 같이 2011과 2012로 나뉘지는데 이는 컴퓨터 하드웨어 기술자 직군과 통신공학 기술자 직군으로 분류되어 있다. 이들 소분류는 아래 <표 2-13>과 같이 세분류와 세세분류로 각각 세분화 되어 있다.

6) 통계청, 고용직업분류, KECO, https://kssc.kostat.go.kr:8443/ksscNew_web/index.jsp#

〈표 2-13〉 워크넷 소분류 201에 대한 세분류

세분류 코드	세분류 이름	세세분류 코드	세세분류 이름
2011	컴퓨터 하드웨어 기술자 및 연구원	201100	컴퓨터하드웨어 기술자 및 연구원
		201101	컴퓨터하드웨어 기술자 및 연구원
		201102	컴퓨터 부품 및 주변기기 개발기술자 및 연구원
2012	통신공학 기술자 및 연구원	201200	통신공학 기술자 및 연구원
		201201	휴대폰·스마트폰 개발자 및 연구원
		201202	통신기기 및 장비 개발자 및 연구원 (휴대폰 등 제외)
		201203	통신기술개발자 및 통신망 운영기술자
		201204	통신감리기술자

소분류 202(컴퓨터시스템 설계 및 분석가)는 세분류 2021, 2022, 2023으로 나뉜다. 컴퓨터 시스템, 네트워크 시스템의 설계 분석과 이와 관련한 보안 전문가 직군이 이에 해당한다. 이들 세분류는 <표 2-14>과 같이 각 세세분류로 세분화 되어 있다.

<표 2-14> 워크넷 소분류 202에 대한 세분류

세분류 코드	세분류 이름	세세분류 코드	세세분류 이름
2021	컴퓨터시스템 설계 및 분석가	202100	컴퓨터시스템설계 및 분석가
		202101	IT 컨설턴트
		202102	컴퓨터 부품 및 주변기기 개발기술자 및 연구원
		202103	컴퓨터시스템관리기술자
2022	네트워크시스템 분석가 및 개발자	202201	네트워크시스템분석가 및 개발자(운영자제외)
2023	컴퓨터시스템 및 네트워크 보안전문가	202301	컴퓨터시스템 및 네트워크 보안전문가

소분류 203(SW개발전문가)은 세분류 2031, 2032로 나누어지며 시스템SW, 응용SW 관련 개발자 직군이 이에 해당한다. 이들 세분류는 <표 2-15>와 같이 각 세세분류로 세분화된다.

<표 2-15> 워크넷 소분류 203에 대한 세분류

세분류 코드	세분류 이름	세세분류 코드	세세분류 이름
2031	시스템SW개발자	20310	시스템SW개발자
		20311	시스템SW 및 운영체제 프로그래머
		20312	임베디드 및 펌웨어 프로그래머
2032	응용SW개발자 (응용프로그래머)	20320	응용SW개발자(응용프로그래머)
		20321	응용SW프로그래머
		20322	네트워크프로그래머
		20323	컴퓨터 및 모바일 게임프로그래머

소분류 204(웹전문가)는 세분류 2041, 2042로 나뉘게 되며 웹개발자와 웹 및 멀티미디어 기획자에 관련한 직군이 이에 해당한다. 이들 세분류는 <표 2-16>와 같이 각 세세분류로 세분화된다.

<표 2-16> 워크넷 소분류 204에 대한 세분류

세분류 코드	세분류 이름	세세분류 코드	세세분류 이름
2041	웹마스터 및 웹개발자 (웹프로그래머)	20410	웹마스터 및 웹개발자(웹프로그래머)
		20411	웹마스터 및 웹PD
		20412	웹개발자(웹프로그래머 및 웹엔지니어)
2042	웹 및 멀티미디어 기획자	20420	웹 및 멀티미디어 기획자
		20421	웹기획자
		20422	멀티미디어기획자
		20423	컴퓨터 및 모바일 게임기획자

소분류 205(데이터베이스 및 정보시스템 운영전문가)는 세분류 2051, 2502로 나뉜다. 데이터베이스 설계·구축 또는 정보시스템 운영자 관련 직군이 이에 해당한다. 이들 세분류는 <표 2-17>과 같이 각 세세분류로 세분화한다.

<표 2-17> 워크넷 소분류 205에 대한 세분류

세분류 코드	세분류 이름	세세분류 코드	세세분류 이름
2051	데이터베이스 설계·구축 및 관리 기술자	20511	데이터베이스 설계·구축 및 관리 기술자
2052	정보시스템운영자	20520	정보시스템운영자
		20521	서버운영관리자
		20522	네트워크운영관리자
		20523	IT기술지원전문가

소분류 206(통신 및 방송 장비기사 설치 및 수리원)은 2061, 2062, 2063, 2064의 세분류 코드로 나뉜다. 통신 또는 방송, 인터넷 장비에 관련한 직군이 이에 해당한다. 이들 세분류는 <표 2-18>과 같이 각 세세분류로 세분화 되어 있다.

<표 2-18> 워크넷 소분류 206에 대한 세분류

세분류 코드	세분류 이름	세세분류 코드	세세분류 이름
2061	통신 및 방송 송출 장비기사	20610	통신 및 방송 송출 장비기사
		20611	통신송출기술자(통신송출기사)
		20612	방송송출기술자(방송송출기사)
2062	방송장비 설치 및 수리원	20620	방송장비 설치 및 수리원
		20621	방송장비 설치 및 수리원
		20622	위성방송안테나 설치 및 수리원
2063	통신 및 관련 장비 설치 및 수리원	20630	통신 및 관련 장비 설치 및 수리원
		20631	통신장비(설비) 설치 및 수리원
		20632	기지국 설치 및 수리원
2064	통신·방송 및 인터넷 케이블 설치 및 수리원	20640	통신·방송 및 인터넷 케이블 설치 및 수리원
		20641	통신케이블 설치 및 수리원

2) 워크넷 직업 분류 기반 채용공고 현황 분석

워크넷에서 수집된 전체 데이터 20,408개 중 직업분류 코드가 표시된 데이터는 1,165개 이다. 직업분류 코드가 표시된 채용공고 1,165개를 소분류 201에서 206로 재분류를 하면 <표 2-19>와 같이 나뉜다. 소분류 203 SW개발전문가 직군이 379개로 채용공고가 가장 많으며, 컴퓨터 및 통신공학 기술자 및 연구원 직군이 가장 적은 71개로 나타났다. 응용SW를 다루거나 웹개발·기획에 많은 공고가 있으며, 통신 및 방송장비를 다루는 직군 또한 채용공고가 많은 것으로 나타났다.

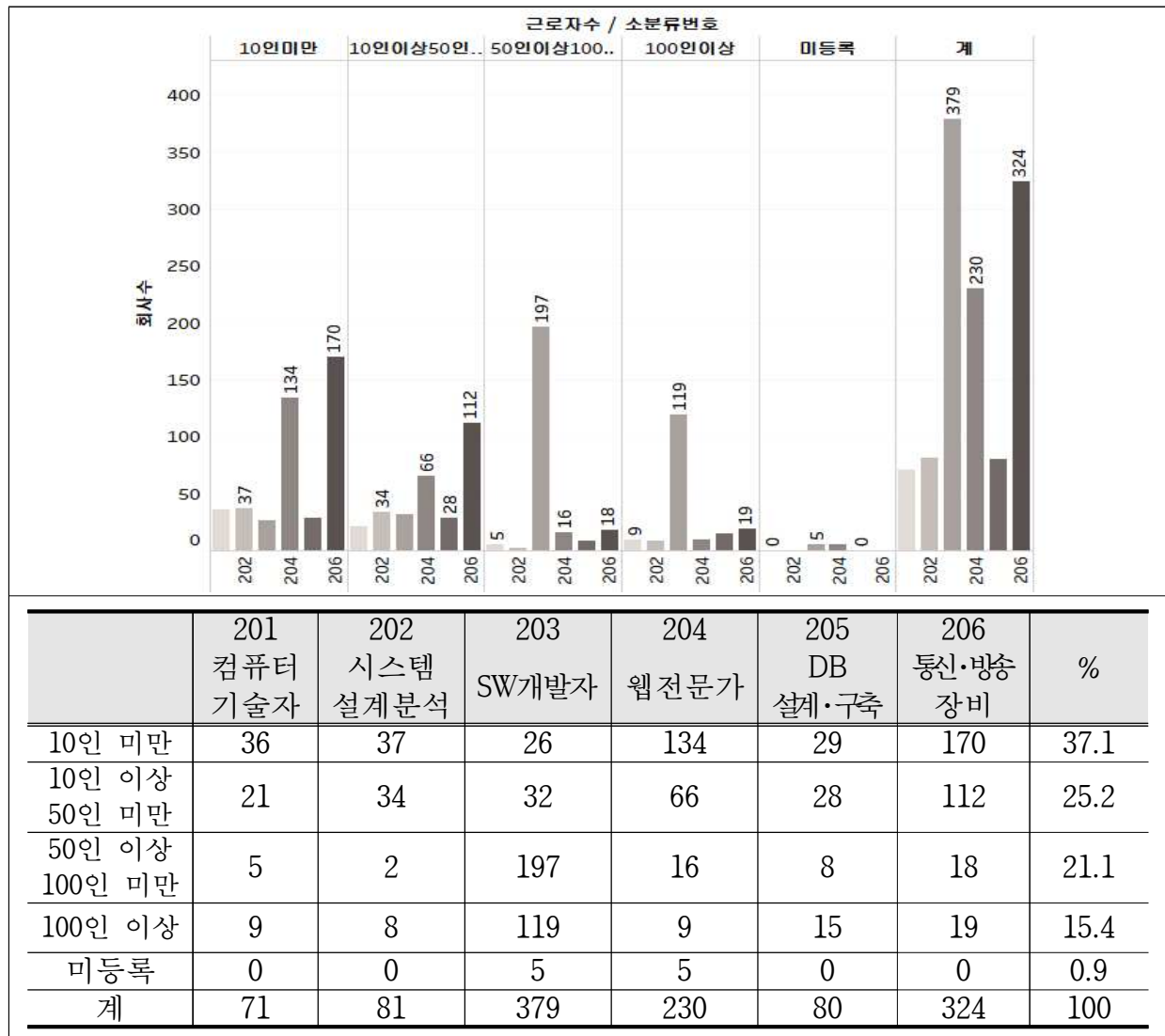
<표 2-19> 소분류별 채용공고

소분류 코드	소분류 이름	채용공고 수(개)
201	컴퓨터 및 통신공학 기술자 및 연구원	71
202	컴퓨터시스템 설계 및 분석가	81
203	SW개발전문가	379
204	웹전문가	230
205	데이터베이스 설계·구축 및 관리기술자	80
206	통신 및 방송 장비기사 설치 및 수리원	324

3) 소분류별 근로자 수에 따른 채용회사의 수

소분류별 근로자 수에 따른 채용회사의 수는 [그림 2-13]과 같다. [그림 2-13]과 같이 채용공고를 낸 회사의 근로자수는 SW 직군 전체로 봤을 때 50인 미만 회사가 약 50%를 차지하는 것을 볼 수 있다. 다만, 소분류 단위로 보면 구성이 달라진다.

[그림 2-13] 소분류별 근로자수 기준 채용광고



201 컴퓨터 기술자 직군과 202 시스템 설계 분석 직군, 204 웹전문가 직군, 205 DB설계·구축 직군, 206 통신·방송장비 관력 직군은 모두 50인 미만의 회사가 대부분을 차지하였다. 하지만, 203 SW 개발자 직군은 50인 이상의 회사가 약 83%를 차지하는 것을 확인할 수 있다.

4) 자본금 기준 채용공고 현황

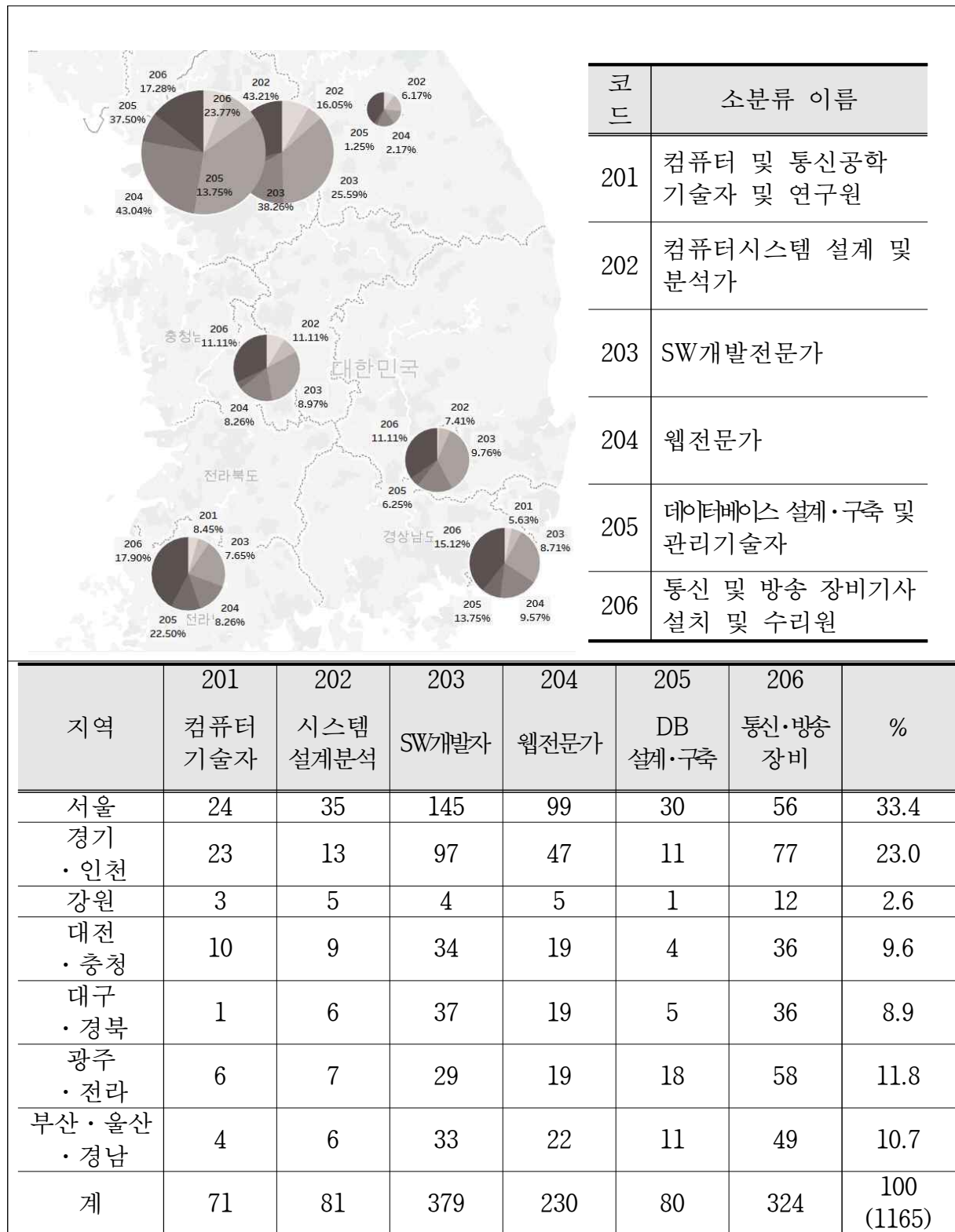
〈표 2-20〉과 같이 채용공고를 낸 회사가 자본금을 공개한 경우는 전체 1,165개 회사 중에 759개이다. 공개한 회사의 자본금은 1억 원 이하인 경우가 약 38.7%로 가장 많은 것으로 나타났으며, 2억 원 이하의 자본금을 보유한 회사가 전체의 약 64%를 차지하는 것으로 나타났다.

〈표 2-20〉 소분류별 자본금 기준 채용공고

자본금	컴퓨터 기술자	시스템 설계분석	SW개발자	웹전문가	DB 설계·구축	통신·방송 장비	%
1억 원 이하	18	18	112	77	25	44	38.7
1억 원 초과 2억 원 이하	7	16	56	21	14	77	25.2
2억 원 초과 3억 원 이하	4	7	22	6	4	23	8.7
3억 원 초과 5억 원 이하	6	3	25	11	6	21	9.5
5억 원 초과 10억 원 이하	5	1	25	11	4	14	7.9
10억 원 초과 20억 이하	1	2	7	2	3	4	2.5
20억 원 초과 30억 원 이하	1	0	3	2	2	1	1.2
30억 원 초과 50억 원 이하	0	1	7	3	1	7	2.5
50억 원 초과 100억 원 이하	2	1	1	0	0	3	0.9
100억 원 초과 200억 원 이하	2	1	2	0	0	0	0.7
200억 원 초과	2	0	3	5	2	5	2.2
미등록	23	31	116	92	19	125	-
계	71	81	379	230	80	324	-

5) 지역별 채용공고

[그림 2-14] 소분류별 지역 기준 채용공고

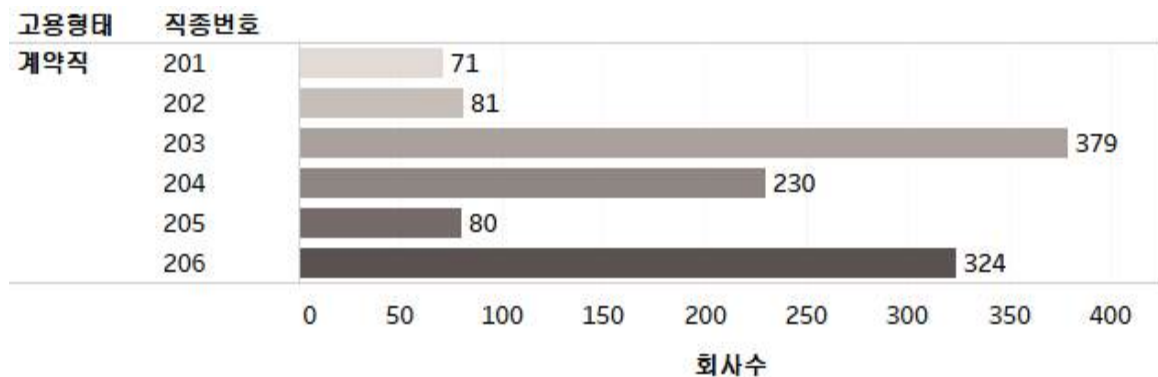


[그림 2-14]와 같이 지역별 채용공고를 보면 서울이 33.4%, 경기·인천이 23%로 전체 채용공고의 약 56%가 수도권에서 진행됨을 알 수 있다. 다만, 소분류번호 206 통신 및 방송 장비기사 설치 및 수리원 직군의 경우는 다른 직군에 비해 채용공고가 전국에 나타나 있는 것을 확인할 수 있다.

6) 고용형태 기준 채용공고

고용형태는 소분류 201에서 206에 해당하는 모든 직군 중 1,165개의 채용공고가 계약직으로 나타났다. 그 중 203, 204, 206 직군이 두드러지게 채용공고의 수가 높게 나타났다.

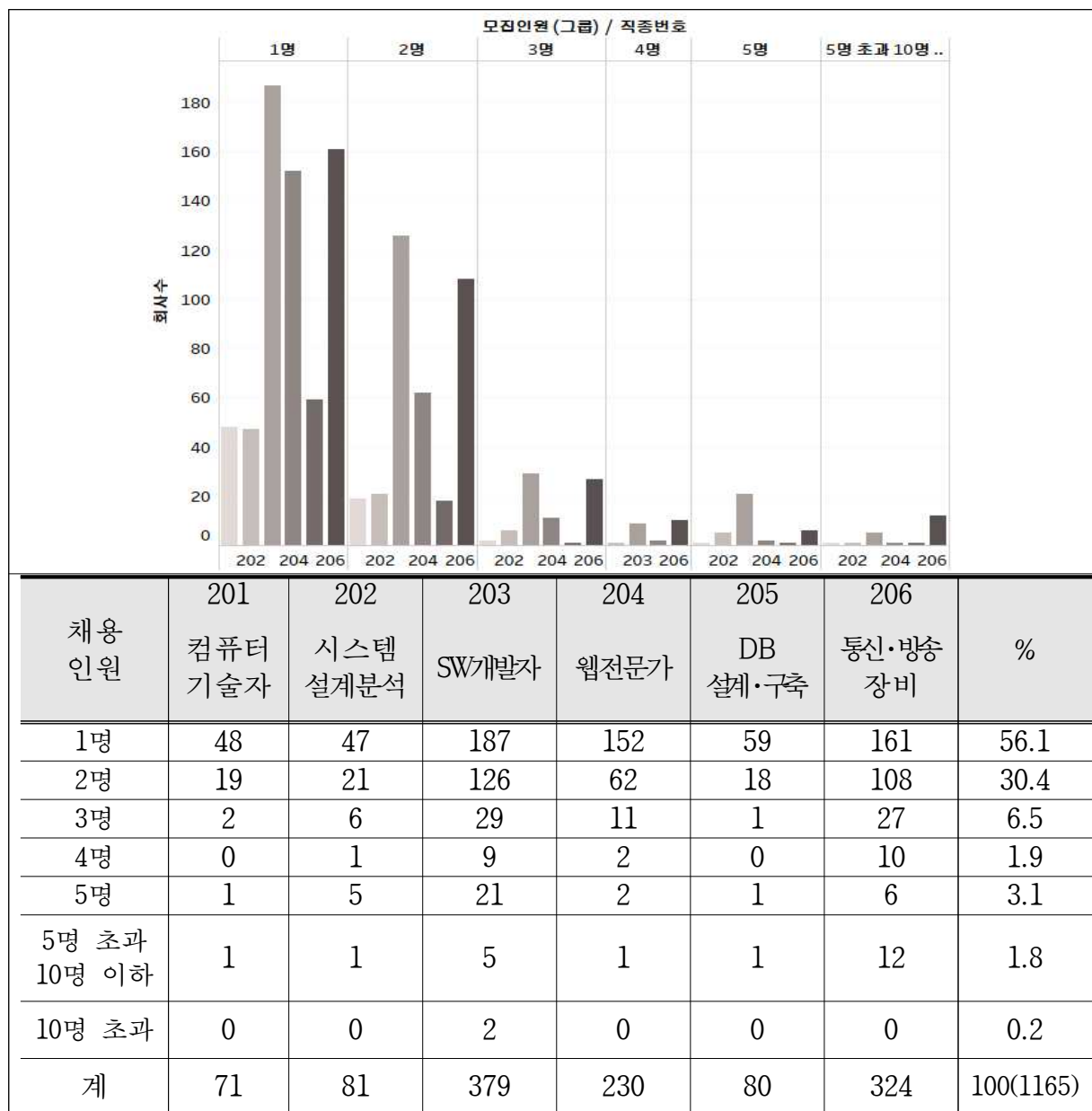
[그림 2-15] 소분류별 고용형태 기준 채용공고



7) 채용인원

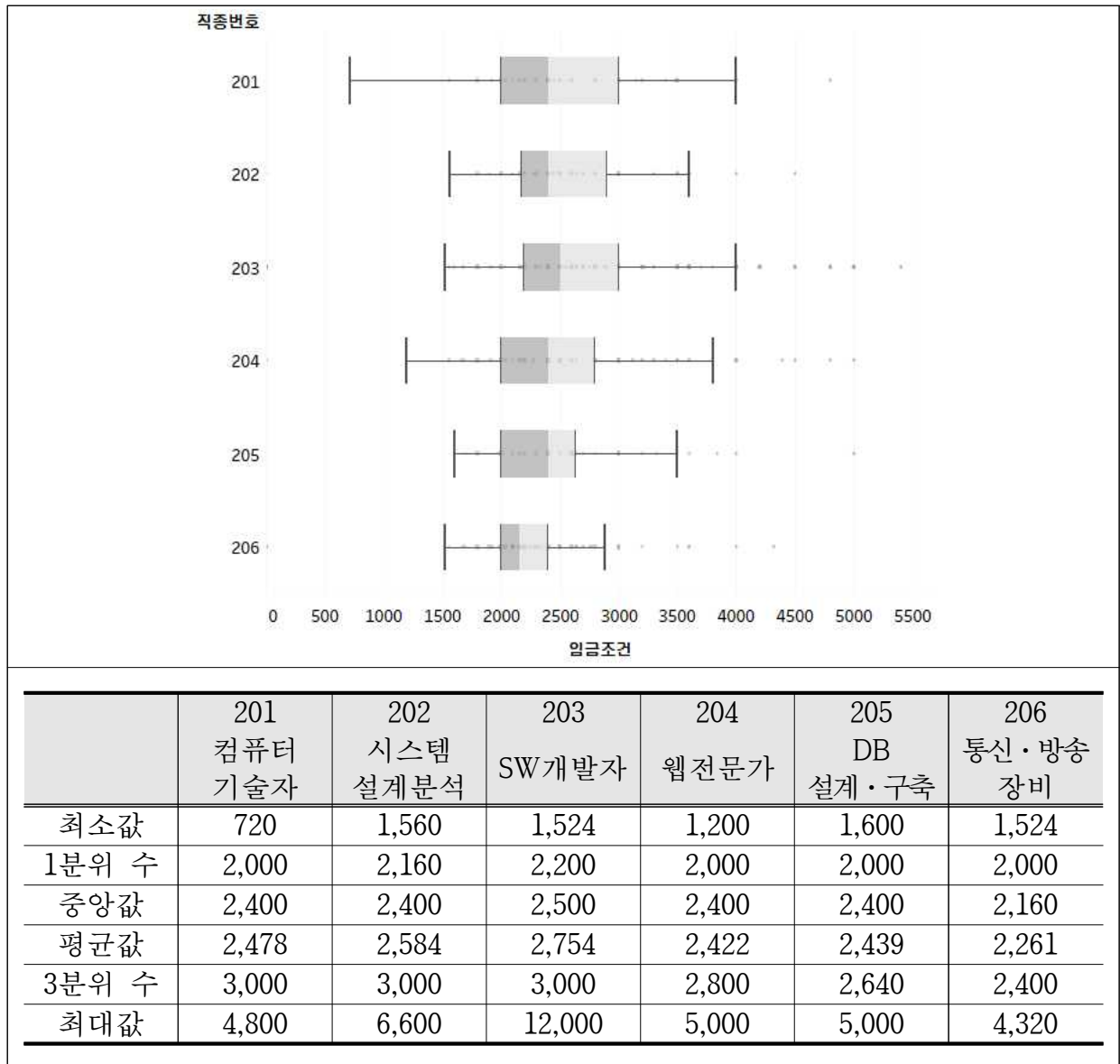
[그림 2-16]과 같이 채용형태는 1명만 채용하는 회사가 56.1%, 2명이 30.4%로 전체의 약 86%를 차지하였다. 직종에 관계없이 1에서 2명만 채용하는 회사가 대부분임을 확인할 수 있다.

[그림 2-16] 소분류별 채용인원 기준 채용광고



8) 임금조건

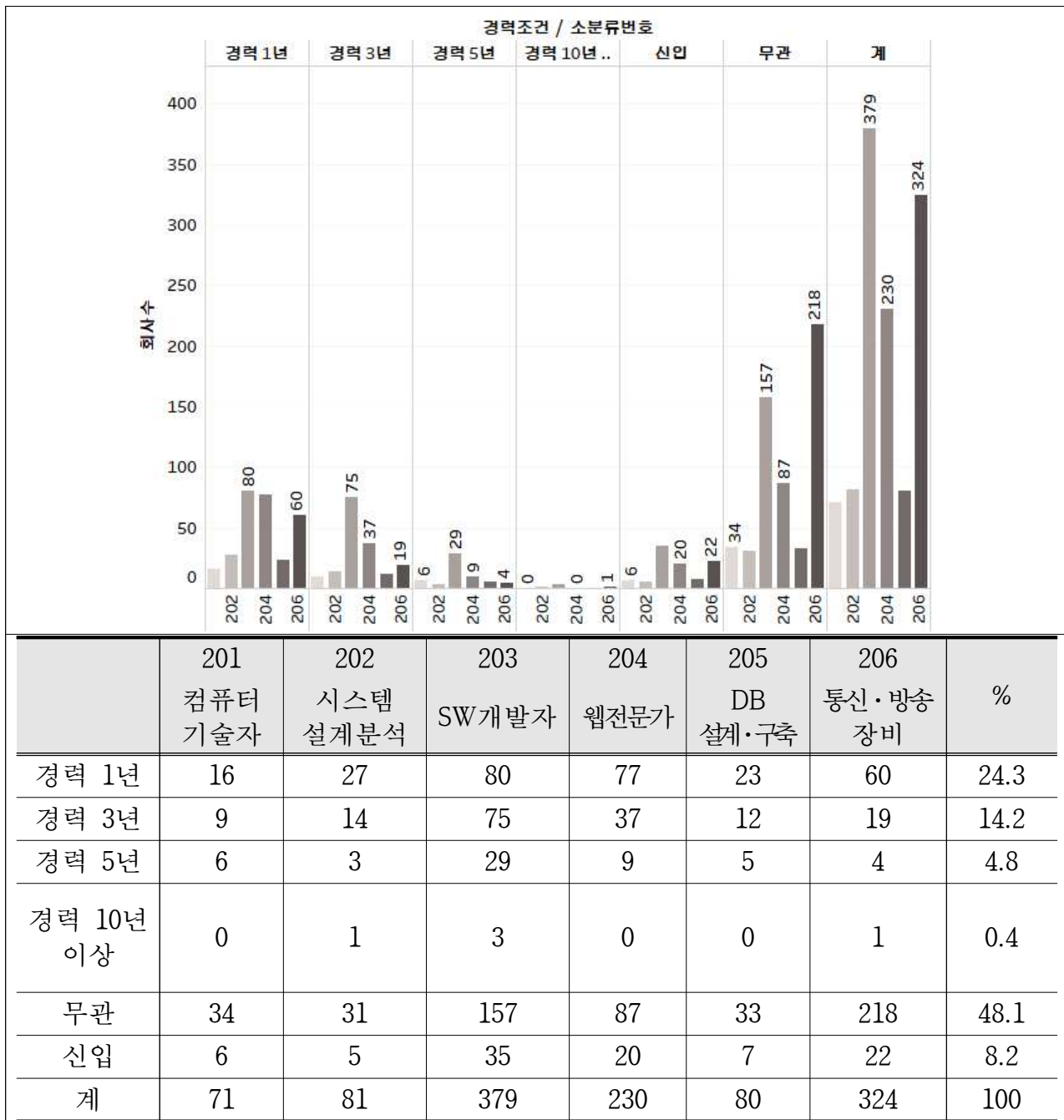
[그림 2-17] 소분류별 임금조건 기준 채용공고



[그림 2-17]과 같이 임금조건은 평균값을 기준으로, 소분류 번호 203 SW개발전문가 직군이 2,754만원으로 평균임금이 타 직군에 비해 높고 소분류번호 206 통신 및 방송 장비기사 설치 및 수리원이 2,261만원으로 타 직군에 비해 낮게 나타났다. 최대임금 또한 206 통신 및 방송 장비기사 설치 및 수리원 직군이 다른 직군에 비해 가장 낮게 형성돼 있음을 확인할 수 있다.

9) 경력조건

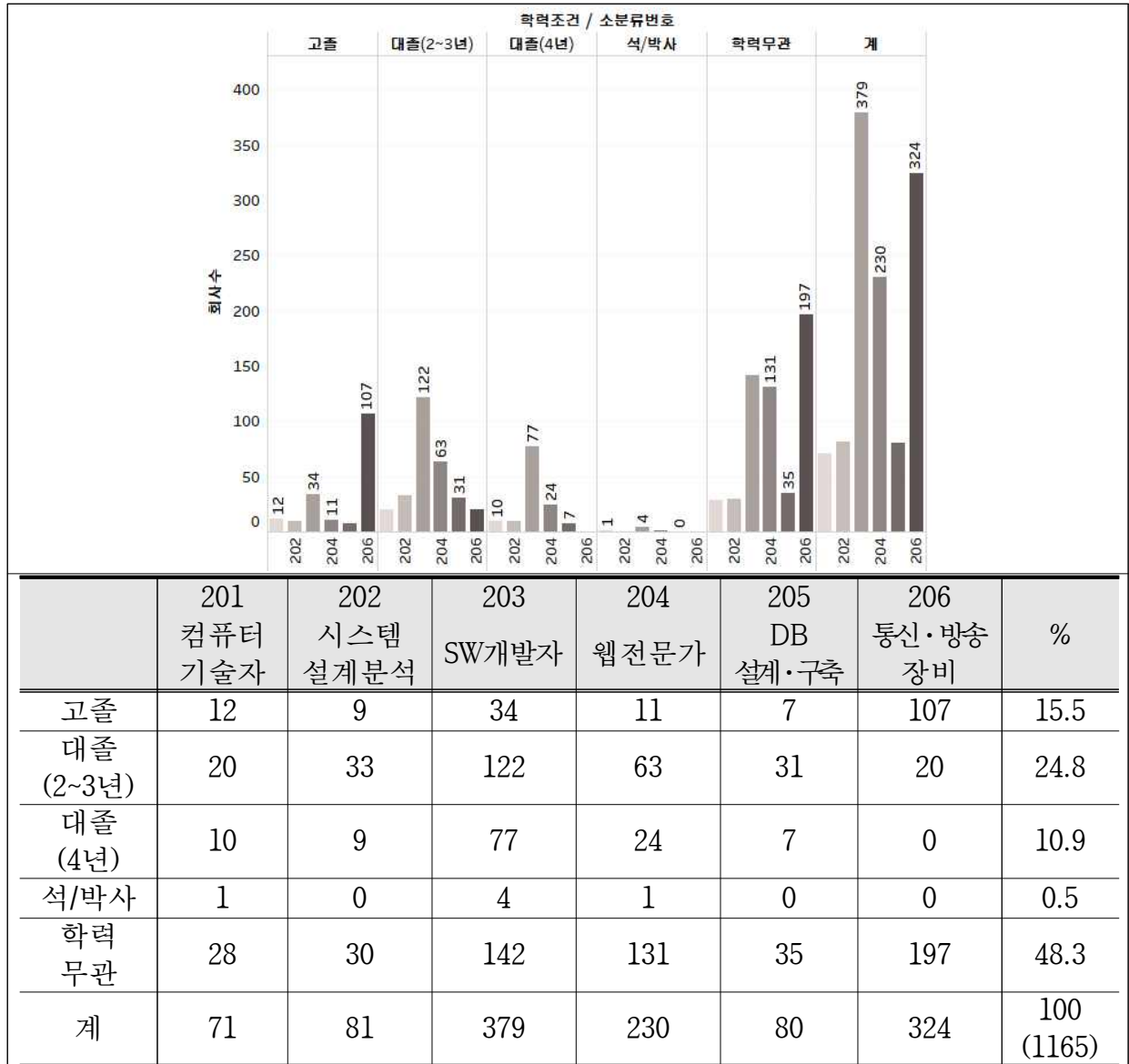
[그림 2-18] 소분류별 경력조건 기준 채용공고



[그림 2-18]과 같이 채용공고에서 경력여부는 무관이 48.1%로 가장 많았다. 신입의 경우는 8.2%이었고, 경력직은 전체의 약 43.%에 해당하였다. 경력직의 경우 경력 1년의 채용공고가 24.2%로 나타나 경력 1년의 구직자의 수요가 높은 것으로 판단된다.

10) 학력조건

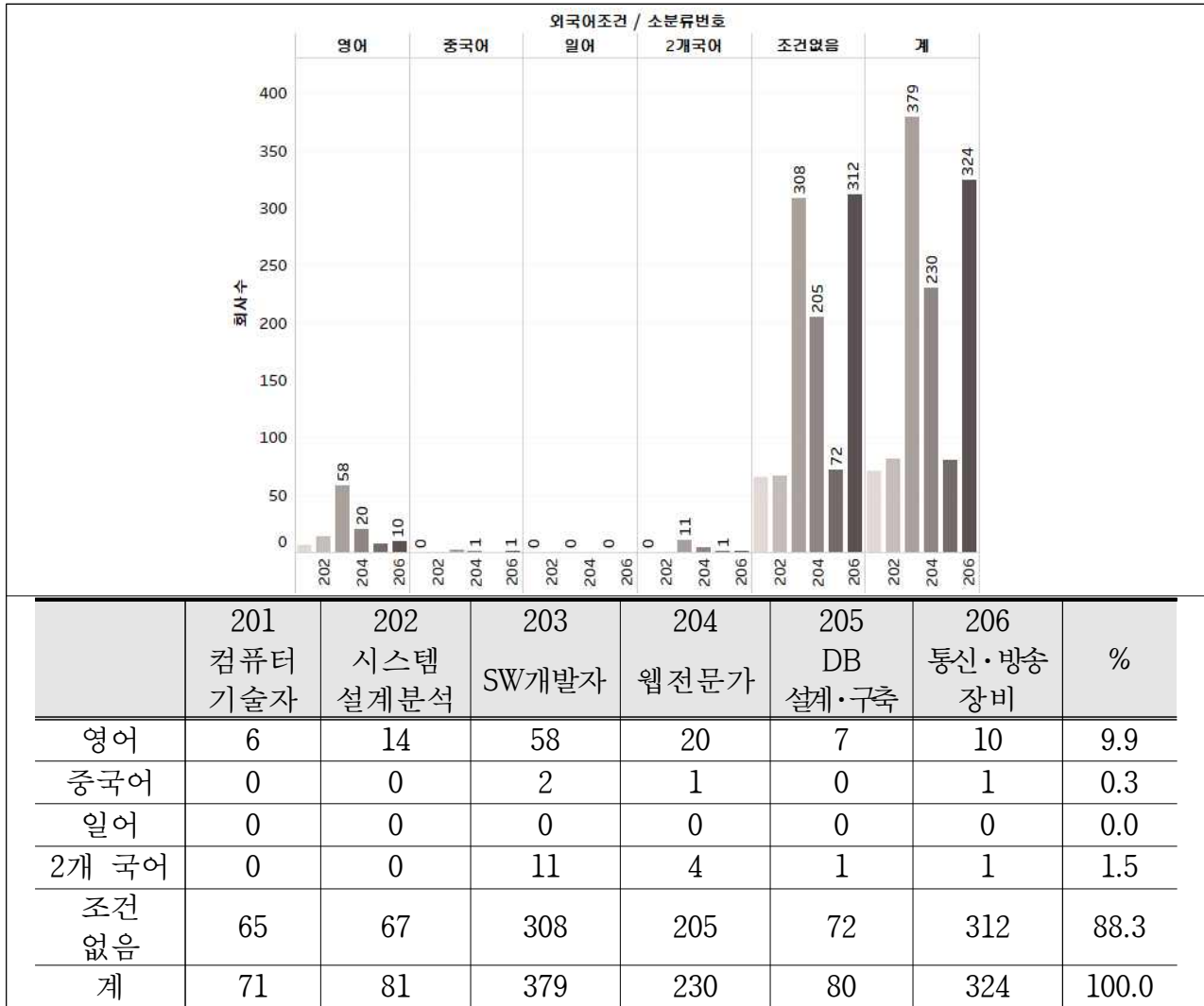
[그림 2-19] 소분류별 학력조건 기준 채용공고



[그림 2-19]와 같이 채용공고에서 학력 조건은 학력무관이 48.3%, 고졸 15.5%, 대졸(2~4년) 35.7%, 석/박사 0.5% 순으로 나타났다. 201 컴퓨터 기술자 직군과 202 시스템 설계분석, 203 SW개발자, 204 웹전문가, 205 DB설계·구축 관련 직들은 모두 채용조건이 학력무관이나 대졸이상을 선호하는 것으로 나타난다. 하지만 206 직군의 경우 고졸 또는 학력무관을 더 선호하는 것을 볼 수 있다.

11) 외국어조건

[그림 2-20] 소분류별 외국어조건 기준 채용공고



[그림 2-20]과 같이 채용조건에서 외국어 능력은 특별한 조건으로 달고 있지 않은 회사가 88.3%로 대부분을 차지함을 볼 수 있다. 또한 소분류 번호 201에서 206에 해당하는 직군 모두 외국어 조건은 크게 중요하지 않음을 알 수 있다. 다만 그중에서 가장 많이 나온 외국어 능력조건은 영어였다. 특히 203 SW 개발자의 경우 약 19%가 영어를 외국어 능력 조건으로 삼아 타 직군에 비해 영어 능력을 중시하는 것을 볼 수 있다.

3. 전체 구인 기업 및 워크넷 분류에 따른 SW분야의 특성

워크넷 직업 분류를 바탕으로 SW분야의 구인 기업에 대한 분석을 진행한 결과 대부분의 기업들은 수도권에 위치한 기업들이 대다수인 것으로 나타났다. 워크넷의 특성상 경력요건을 입력하지 않으면 경력무관으로 표기되는 것을 고려한다면, SW 분야 기업들의 경우 예측되었던 것과 같이 채용공고 상에서 막 대학을 졸업한 인력보다는 1년~5년 사이의 전문 능력을 가진 인력을 선호하는 것으로 판단된다. 이는 기본적인 대학 졸업자들의 학업 수준에 대한 측면 보다는 실제 보유하고 있는 능력에 더 집중하는 것을 의미한다.

또한 대부분 기업들의 경우 급여 체계가 2000만 원대로 나타났으며 대다수의 경우는 60:40의 비율로 정규직과 비정규직 비율을 유지하고 있다. 이는 업계 특성상 SI 중심의 인력 구성을 하고 있기 때문으로 판단되어진다. 또한 급여 수준의 경우 대다수는 높게 나타나지 않아 단기적 인력을 상당수 채용하고 있는 것으로 나타났다. 대다수의 SW 관련 기업들이 대학 및 고학력 졸업자와 관련 없이 보유하고 있는 실무 경력에 따라 채용 진행하고 있다는 점에서 대학 교육의 현 체계와는 기업의 인력 수요 측면에 대한 관점의 차이가 존재하는 것을 알 수 있다.

또한 구인에 대한 수요가 높은 분야의 경우 SW 개발자, DB설계, 통신 및 방송 장비 등과 관련된 채용공고들이 나타나고 있다. 그러나 대부분의 경우 직무 및 직종과 관련하여 구인 기업의 자유입력에 의해 중복 선택되기 때문에, 단순 워크넷의 분류체계 기준으로는 정확한 분석이 어렵다. 더욱이 SW 개발자의 직무의 경우에도 SW 개발자가 단순 SW가 아니라 시스템, SW, 모바일 등 다양한 분류가 존재할 수 있으나 이에 대한 구분이 명확하지 않아 현재 워크넷에서 제시하고 있는 분류체계 만으로는 각 직무의 중요 요구사항을 확인하기 어렵다고 판단된다.

결과적으로 좀 더 명확한 분석을 위해서는 SW 관련 기업들이 워크넷에 제시한 채용공고 내용을 취합하여 SW 관련 직무 분류체계를 실제 데이터를 기반으로 Bottom-up 방식의 빅데이터 분석방법으로 정립하여 확인할 필요가 있다. 이를 통해, 각 직무 및 직종의 실제 기업의 수요에 맞는 분류를 확인 할 수 있을 뿐만 아니라 각 직무 및 직종에서 요구하는 역량을 확일 할 수 있을 것이다.

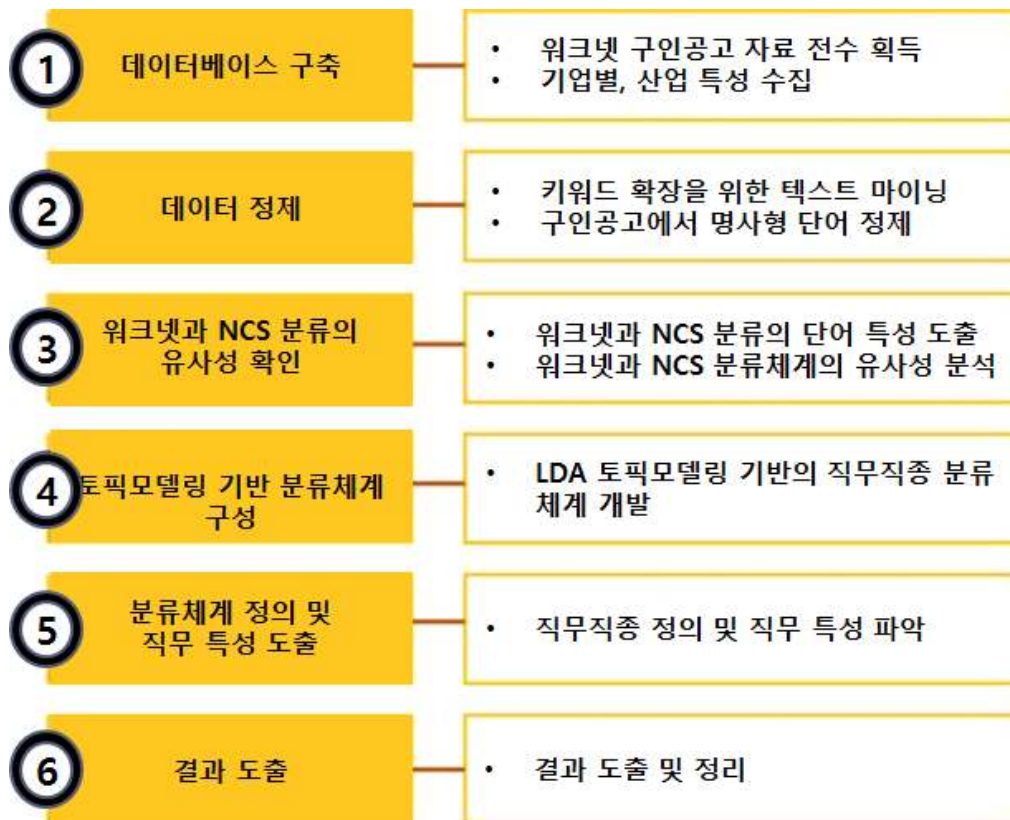
이와 같은 분석 결과에 따라 본 연구과제는 4절에서 채용공고 내용을 분류하기

위한 텍스트마이닝 분석을 제시하고 핵심단어들을 바탕으로 군집을 도출하는 직무 및 직종 체계를 제시하고자 하였다. 이를 위해 먼저 전처리된 채용공고 단어 데이터를 바탕으로 워크넷의 직업 체계와 NCS 직무 분류를 바탕으로 두 분류의 직종들에 대해 워크넷에서 획득한 채용공고 단어 데이터를 매칭 하여 두 분류체계의 분류 체계 현황을 비교하였다. 이후 LDA 알고리즘 기반의 토픽모델링을 통한 신규 직무·직종 체계를 도출하고자 하였다. 이에 더하여 본 연구를 통해 도출된 각 직무·직종 체계를 바탕으로 SW 분야의 인력 수요 동향을 확인하였다.

제4절 텍스트마이닝 분석

본 보고서의 목적인 채용공고 기반의 데이터 분석을 바탕으로 직무·직종 분류체계를 개발하기 위해서 워크넷 채용공고 자료는 텍스트마이닝 기법을 이용하여 채용공고 내용의 형태소를 분석하고 단어별 데이터를 도출하여 직무 특성을 확인하고자 하였다. 텍스트마이닝 분석 기법은 다양한 분석 방법이 존재하는데 (Sullivan, 2001; Tan, 1999; 이순근, 2015) 연구 목적에 따라 본 보고서는 동시단어분석, 분류 분석, 연관성 분석 등의 기법을 이용하여 분석을 진행하였다 (Feldman and Sanger, 2007). 형태소 분석과 관련된 연구들의 경우에도 주로 동시단어분석을 이용한 분석을 이용한 프레임워크를 인용하여 분석에 차용한 연구들이 일반적이다 (Tseng et al., 2007; weiss et al., 1999; 박만희 2016; 안형준, 2013).

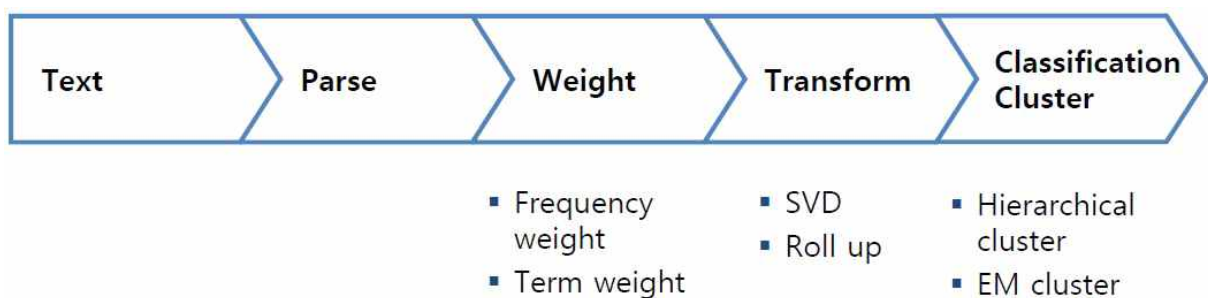
[그림 2-21] 텍스트마이닝 프레임 워크



본 연구는 텍스트마이닝을 통한 분석을 위하여 [그림 2-21]과 같은 연구 프레임워크를 구성함으로써 구조적 분석을 이행 할 수 있었다. 데이터베이스를 구축하고 데이터 정제를 통하여 유효 단어를 도출, 관련 단어 연관성 분석을 실시하고 도출된 단어들의 군집을 구성하여 해당 단어들의 분류를 찾아내고 관련 결과를 찾는 순서로 연구를 진행하였다(박만희 2016).

텍스트마이닝은 자연어 처리 기반을 통한 언어 분석으로 언어학, 통계학, 기계 학습 등을 기반으로 정형/비정형 텍스트 자료를 정형화한다. 그에 따라 내포된 특징을 추출하기 위한 기술과 추출된 결과로부터 의미 있는 정보를 발견할 수 있도록 하는 기술이다(노형남, 2014). 자연어 처리 기반 텍스트마이닝의 주요 기술로는 자연어 처리(파싱, 형태소 분석, 품사 태깅, 관계 추출, 의미 추출), 언어모델링(언어 감지, 규칙기반 개체명 · 상용어 인식), 기계 학습 알고리즘(반복훈련을 통하여 습득한 정보 사용능력을 개선), 마이닝 기술(각종 통계적 기법을 활용한 정보 분류 및 분석 기술 · 기법) 등이 있다 (Manning and Schütze, 1999). 텍스트마이닝 과정을 개략적으로 나타내면 [그림 2-22]와 같다. 특히 자연어 처리 기반 텍스트마이닝 기술을 구현하기 위해서는 자연어 텍스트를 컴퓨터가 학습할 수 있는 요소로 데이터를 구조화하여 문장 수준의 텍스트에서 의미를 추출해야 한다(Hutchins, 2005; Kishorjit et al., 2012; Duan and Cruz, 2011).

[그림 2-22] 텍스트마이닝 과정



출처: Manning and Schütze (1999) 내용 재구성

텍스트마이닝 과정에서 가장 먼저 해결해야 할 부분은 문서에 대한 탐색이다. 이를 바탕으로 텍스트 분석은 문장을 형태소, 즉 명사, 조사, 부사, 동사원형 등 문법상 의미를 분해할 수 없는 최소의 단위로 분리하여 주어, 목적어 등의 주요

언어학적 구조를 나누고, 문장의 의미를 파악하는 것이 중요하다. 파싱은 단순히 단어를 일치시키는 것이 아닌 문장의 본 의미를 파악하는 것이 비구조화 데이터를 분석하는데 매우 중요하다(장재영, 2013; 장필성, 2016).

비구조화된 자료를 분석하기에 적합한 비구조화 자료를 구조화된 자료로 재구성하는 방식을 바탕으로 채용공고 내에 있는 텍스트를 추출하기 위하여 파이썬(python) 언어를 기반으로 문서 내 텍스트 들을 자연어 분석 처리 하였다. 그에 따라 명사형 단어를 파싱(Parsing)하여 필터링하고 중복되는 명사들을 수량화하여 가장 많이 나온 순으로부터 상위 50개를 추려 데이터로서 구성하였다. 이에 추가적으로 파싱된 명사 중에서 컴퓨터 및 SW에 관련된 용어사전을 개발하여 데이터 내의 불필요한 오류가 될 단어들을 모두 구분하여 배제하였다.

본 연구과제는 채용공고 문서가 관련 단어의 집합으로 이루어져 있다는 개념을 바탕으로 행렬 데이터로 도출하였다. 이에 따라 각 조건의 문서와 단어 간의 상호관계로 이루어진 단어-문서 행렬(term-document matrix)을 통해 각 셀 안에 문서에 포함된 단어의 빈도를 도출하는 과정은 다음과 같다.

문서의 분류 자동화를 위해 각 문서의 내용을 대표하는 단어와 해당 단어의 가중치를 결정하는 방식을 활용하여 가중치의 계산에 단어의 빈도를 사용하는 ‘용어빈도 가중치’를 활용하였다. 용어 빈도 가중치는 지역적 가중치라고도 불리며 단순히 특정 용어가 전체 문서 중에 차지하는 비율을 도출한다. a_{ij} 는 문서 j에서 단어 i의 가중된 도수를 의미하며 이를 계산하기 위해서는 G_i 는 용어 i의 가중치와 L_{ij} 는 문서 j와 용어 i의 빈도 가중치를 활용해야 한다.

$$a_{ij} = G_i L_{ij}$$

문서와 단어에서 추출된 빈도 가중치와 용어 가중치를 결합함으로써 채용공고 문서 내, 그리고 채용공고 문서 간 관련 정보를 획득할 수 있다. 행렬 안에 들어가는 값들은 빈도 가중치와 용어 가중치의 곱으로 계산된다. 비구조화된 텍스트 문서를 일련의 숫자 배열로 구성된 구조로 변환하면 대부분이 0으로 값이 도출되며(sparse data) 실제 자료 크기가 매우 크다는 특징이 존재한다. 따라서 행렬 매트릭스 자료의 데이터가 없는 부분의 차원을 줄이기 위해서 변환 과정을 수행하였다.

그러나 용어 빈도 가중치는 특정 문서가 전체 문서 중에서 차지하는 정보를 반영하지 못한다. 즉, 분석에 이용된 문서의 크기가 클수록 비교적 작은 문서보다 훨씬 높은 빈도(frequency)를 도출해낼 가능성이 증가하게 되므로 이것을 지역적 가중치(local weight)라고도 한다. 이에 대한 해결책으로 단어 가중치를 동시에 고려하는 전역적 가중치를 고려하게 되며 기본적인 빈도 가중과 방법은 다음과 같다.

$$\text{Binary : } L_{ij} = \begin{cases} 1 & \text{if term } i \text{ is } \in \text{document } j \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Log: } L_{ij} = \log_2(a_{ij} + 1)$$

$$\text{None: } L_{ij} = a_{ij}$$

a_{ij} = 문서 j에서 용어 I가 나타낸 횟수

L_{ij} = 문서 j와 용어 I의 빈도 가중치

일반적으로 빈도 가중치 함수는 단순 빈도를 활용하거나 또는 단순 빈도를 로그 변환하거나 이항 변환하는 형태로 활용한다. 빈도 가중치 함수에서 문서 j에 포함되어 있는 단어 i의 가중치를 표현하며 $L_{(i,j)}$ 로 나타낸다. a_{ij} 는 분석한 문서 j에서 단어 i가 나타나는 빈도를 의미하며 도출된 빈도 수 자체를 빈도 가중치의 함수로 사용하기도 한다. 로그 값으로 변환한 용어 빈도 가중치 함수는 a_{ij} 의 값이 클 경우 이를 감소시켜주는 역할을 이행한다. 이항 변수 개념을 활용한 지역적 가중치에 대한 함수 값은 문서 j에 단어 i가 한 번이라도 나타나는 경우를 1 값으로 적용하고, 한 번도 나타나지 않는 경우를 0의 값으로 할당하는 방법을 사용하여 단어의 수가 많지 않은 경우 보편적으로 사용한다.

용어 빈도 가중치와 관련된 가중치 함수는 다양한 종류가 존재하며 정보이론⁷⁾의 정보 엔트로피⁸⁾ 개념을 이용하여 생겨난 1) 전역적 가중치 함수(Entropy), 2) 전체 문서에서 단어의 빈도수와 단어가 나타난 문서 수의 비율을 비교 분석한 전역적

7) 정보이론: 정보 데이터를 상호 비교 분석이 가능하도록 정량화하는 응용수학의 한 분야

8) 정보 엔트로피(information entropy): 확률 분포에 담긴 불확실성의 척도. 불확실성이 높아질수록 정보의 양이 늘어나며 엔트로피 또한 증가한다.

가중치 함수(Global Frequency Times Inverse Document Frequency, GfIdf), 3) 단어가 갖고 있는 정보를 양적으로 표현한 전역적 가중치 함수(Inverse Document Frequency, Idf) 또는 4) 표준화를 활용하여 계산하는 전역적 가중치 함수 등이 있다. 본 보고서에서는 일반적으로 특정 문자의 문제 해결이 필요하지 않기 때문에 일반적으로 활용되는 GfIdF를 사용하였다.

p_{ij} 는 문서 j 에서 단어 i 가 나타난 빈도를 i 가 나타난 문서의 개수로 나누어 계산한다. 단어 d_j 는 단어 i 가 나타난 문서의 수, g_i 는 전체 문서 모음에서 단어 i 가 나타난 빈도, N 은 전체 문서의 수를 의미한다. 엔트로피 개념을 활용한 지역적 가중치는 전체 문서 안에서 단어의 빈도가 드물게 도출될 때 높은 값을 가진다. 전체 문서에서 특정 단어가 정확하게 한 번씩 도출된다면 해당 단어는 각 문서를 구분해 주는 특징이 없다. 상기의 경우, 엔트로피 개념을 활용한 지역적 가중치는 0의 값을 갖게 되고 문서 전체 중 단 하나의 문서에서만 특정한 단어가 나타나게 된다면 1의 값을 갖는다. 문서들의 모음 안에서 한정된 소수의 문서에 존재하는 단어만을 강조하고 있다는 특징을 갖으며 GfIdf도 엔트로피 개념을 이용한 지역적 가중치와 똑같은 특징을 갖게 된다. 표준화 개념을 이용하고 있는 전역적 가중치 함수는 문서들의 모음 안에서 존재하고 있는 단어의 비율을 표현한다. 다음은 GfIdf에 대한 구현을 위한 수식이다.

$$\text{Entropy: } G_i = 1 + \sum_{j=1}^{d_j} \frac{p_{ij} \log_2(p_{ij})}{\log_2(n)}, 0 \leq G_i \leq 1$$

$$\text{Inverse Document Frequency: } G_i = 1 + \log_2\left(\frac{n}{d_i}\right), 1 \leq G_i \leq \infty$$

$$\text{Mutual Information: } G_i = \max_2\left(\log\left(\frac{P(x_i, C_k)}{p(x_i)P(C_k)}\right)\right)$$

$$G_i = 1$$

Global Frequency Times Inverse Document Frequency: $G_i = \frac{g_i}{d_i}$

$$\text{Normal: } G_i = \frac{1}{\sqrt{\sum_j f_{ij}^2}}$$

Information Gain:

$$G_i = - \sum_k P(C_k) \log P(C_k) + P(t_i) \sum_k P(C_k|t_i) \log P(C_k|t_i) + P(\bar{t}_i) \sum_k P(C_k|\bar{t}_i) \log P(C_k|\bar{t}_i)$$

a_{ij} : 문서 j에서 용어 i가 나타난 횟수

f_i : 전체 문서에서 용어 I가 나타난 빈도

n: 전체 문서의 수

d_i : 용어 i가 나타난 문서의 수

$$p_{ij} = a_{ij}/f_i$$

G_i : 용어 i의 용어 가중치

문서 변환은 단어와 문서 간의 행렬이 차원이 크고 값이 비어 있기 때문에 0값이 들어가 있는 행렬 구조가 대부분인 희소 행렬이다. 따라서 문서에 포함되어 있는 단어는 항상 동일하지 않기 때문에 단어와 문서 간 행렬에서 단어의 개수는 문서의 개수 보다 훨씬 크다. 문서의 차원이 크면 단어 계산을 하는데 할애되는 시간이 많이 소요될 뿐만 아니라 해당 두 문서 사이의 거리 및 밀도를 정의하기가 어렵기 때문에 좋은 군집 결과를 얻지 못한다. 차원을 축소하는 방법은 연속형 자료를 활용하는 경우 고차원에서 저차원으로 투영시키는 선형대수학 기법을 많이 활용한다. 본 분석에서는 대표적인 차원 축소 방법인 단어를 추출하는 알고리즘을 개선한 방식을 다음과 같이 Tsenf et al. (2007)의 연구를 참고하여 분석에 활용하였다.

```

1. Convert the input text into a LIST of words.
2. Do Loop
2.1 Set MergeList to empty.
2.2 Put a separator to the end of LIST as a sentinel and set the occurring frequency of the separator to 0.
2.3 For I from 1 to NumOf(LIST) - 1 step 1, do
2.3.1 If LIST[ I ] is the separator, Go to Label 2.3.
2.3.2 If Freq(LIST[ I ]) > threshold and Freq(LIST[ I+1]) > threshold, then
Merge LIST[ I ] and LIST[ I +1] into Z.
Put Z to the end of MergeList.
Else
If Freq(LIST[ I ]) > threshold and LIST[ I ] did not merge with LIST[ I - 1], then
Save LIST[ I ] in FinalList.
If the last element of MergeList is not the separator, then
Put the separator to the end of MergeList.
End of For loop
2.4 Set LIST to MergeList.
Until NumOf(LIST) < 2.
3. Filter terms in FinalList based on some criteria.

```

알고리즘을 바탕으로 연관된 알고리즘에 의한 추출사례는 다음과 같다⁹⁾.

```

Example: Given an input string: BACDXAYCDBACD.
Let threshold=1, separator=x.
Step 1: Create a list of single tokens:
LIST = (B:2, A:3, C:3, D:3, X:1, A:3, Y:1, C:3, D:3, B:2, A:3, C:3, D:3, x)
Step 2:
After 1st iteration :
MergeList = (BA:2, AC:2, CD:3, x, CD:3, DB:1, BA:2, AC:2, CD:3, x)
FinalList = (A:3)
After 2nd iteration :
MergeList = (BAC:2, ACD:2, x, BAC:2, ACD:2, x)
FinalList = (A:3, CD:3)
After 3rd iteration :
MergeList = (BACD:2, x, BACD:2, x)
FinalList = (A:3, CD:3)
After 4th iteration :
MergeList = (x)
FinalList = (A:3, CD:3, BACD:2)

```

이를 바탕으로 단어 중심의 워드클라우드 알고리즘을 작성하고 시각화 정보를 도출하였다. 워드알고리즘에서 각 단어별 유사도를 위하여 코사인 유사도(Cosine Similarity), 자카드 유사도(Jaccard Similarity), 어휘적 유사성(Lexical Similarity) 중(김성근, 2016)에서 코사인 유사도 방식을 사용하였다(Amit, 2001). 또한 이후 워크넷 직업체계와 NCS 직업체계를 비교하는 방식에서는 두 집반의 비교가 가능한 자카드 유사도를 사용하였으며 자카드 유사도의 계산 방식은 5절의 2. 분류체계 비교를 위한 데이터 선택 및 분석 과정과 3. 계산방법에서 자세히 설명하였다.

9) Tsenf et al. (2007)의 사례 재구성

워드클라우드 알고리즘을 작성하고 시각화 정보를 도출하기 위한 코사인 유사도는 코사인 거리를 바탕으로 단어 간 유사도를 측정하는 방식으로 아래와 같이 두 개의 영벡터를 바탕으로 유클리디안 방정식을 도출하여 작성할 수 있다 (Spruill, 2007).

$$a \cdot b = \|a\| \|b\| \cos\theta$$

두 벡터 a와 b에 대하여 유사도를 계산하면 다음과 같다.

$$\text{Similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

도출된 유사도를 바탕으로 N개 차원 벡터 a와 b를 통해 코사인 유사도 (Gregori et al., 2014)를 계산하는 알고리즘은 다음과 같다.

$$\text{soft_cosine}(a,b) = \frac{\sum_{i,j}^N s_{ij} a_i b_j}{\sqrt{\sum_{i,j}^N s_{ij} a_i a_j} \sqrt{\sum_{i,j}^N s_{ij} b_i b_j}}$$

where $s_{ij} = \text{similarity}((feature_i, feature_j))$

1. 채용공고 데이터

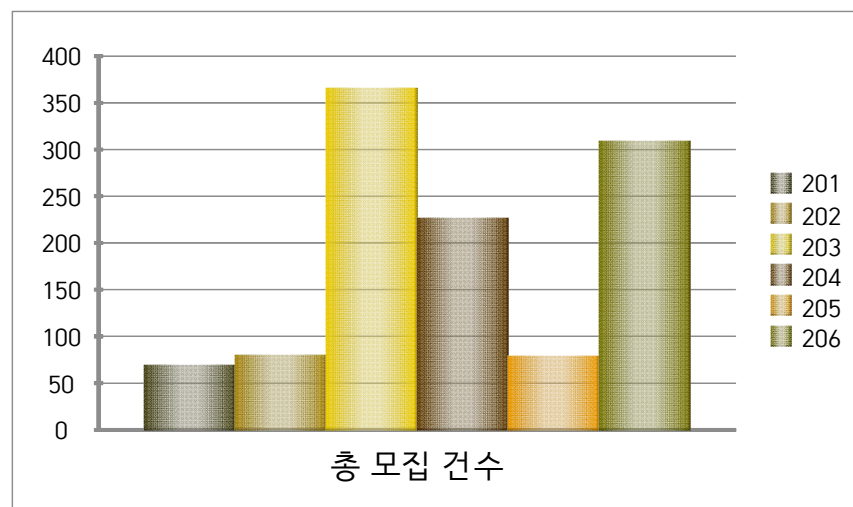
본 연구에서는 워크넷에서 수집한 전체 채용공고 데이터 중에서 모집직종 코드가 붙어있는 채용공고의 데이터만을 추출하여 분석을 진행하였다. 분석에 사용한 채용공고에 제시되고 있는 제목, 단어, 직무내용에 해당하는 텍스트 데이터를 분석했다. 각 텍스트에 등장하는 단어를 추출하여 각 단어의 빈도수가 5를 초과하는 단어들만 선별하여 원형 차트와 워드클라우드를 제작하였다. 또한 단어들 간의 관계를 알아보기 위해서 단어 공동 출현 네트워크를 제작하였다. 원형 차트에서 각 단어가 차지하는 비중은 단어의 빈도수를 나타낸다. 즉, 비중이 크면 해당 단어가 다른 단어들과 비교했을 때 상대적으로 더 많이 등장한다는 것을 의미한다. 워드클라우드에서는 단어의 크기가 해당 단어의 상대적 빈도를 나타내며, 상대적 빈도가 높을수록 크게 등장한다. 단어 공동 출현 네트워크에서 단어들이 연결되어 있는 선의 굵기는 단어들 간의 관계를 의미한다. 선의 굵기가 굵을수록 단어들이 동시에 발생하는 것이고 굵기가 얇을수록 단어들이 동시에 발생하지 않는 것을 의미한다.

〈표 2-21〉 모집직종과 직종코드 예시

모집직종	직종코드
홈페이지자료관리원(정보검색 포함)(029203)	X
통신장비(설비) 설치 및 수리원(206301)	206
통신 및 관련 장비 설치 및 수리원(206300)	206
JAVA프로그래밍언어 전문가(203202)	203
감시카메라 및 보안장치 설치 및 수리원(193901)	X
총무 및 일반 사무원(024301)	X
응용SW개발자(응용프로그래머)(203200)	203
웹개발자(웹프로그래머 및 웹엔지니어)(204102)	204
서버운영관리자(205201)	205
웹개발자(웹프로그래머 및 웹엔지니어)(204102)	204

<표 2-21>은 워크넷에 게시된 채용공고의 모집직종에 해당하는 내용과 이를 처리하여 구한 직종코드이다. 워크넷에서 활용한 직종코드는 총 6자리로 구성되며 한국고용직업분류표¹⁰⁾의 기준에 따라서 제작되었다. 본 절의 텍스트마이닝 분석에서는 6개의 번호에서 앞 2개의 번호가 20인 정보통신 관련 직종에 해당하는 채용공고를 선정하여 분석하였다. 또한 세 번째 자리 번호를 기준으로 201, 202, 203, 204, 205, 206 여섯 개의 범주에 따라서 채용공고를 분류하고 분석을 진행하였다.

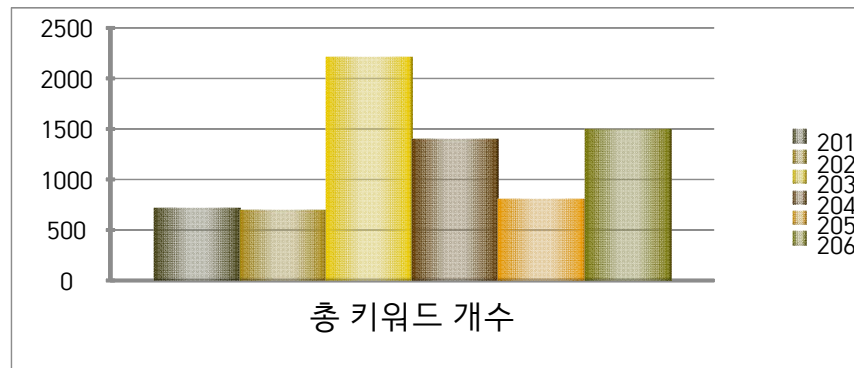
[그림 2-23] 모집 직종 코드별 채용공고 수



[그림 2-23]은 모집 직종 코드별 채용공고 수이다. 모집 건수는 직종 코드별로 50건 이상이었으며, 워크넷 직종코드가 203(SW개발전문가)에 해당하는 모집 건수가 366건으로 가장 많았다. 직종코드 206(통신 및 방송 장비기사 설치 및 수리원)번과 204번(웹전문가)이 다음으로 가장 많았으며 200건을 초과하였다. 나머지 직종코드들은 모집 건수가 50 ~ 100건이었다.

10) 통계청, 고용직업분류, KECO, https://kssc.kostat.go.kr:8443/ksscNew_web/index.jsp#

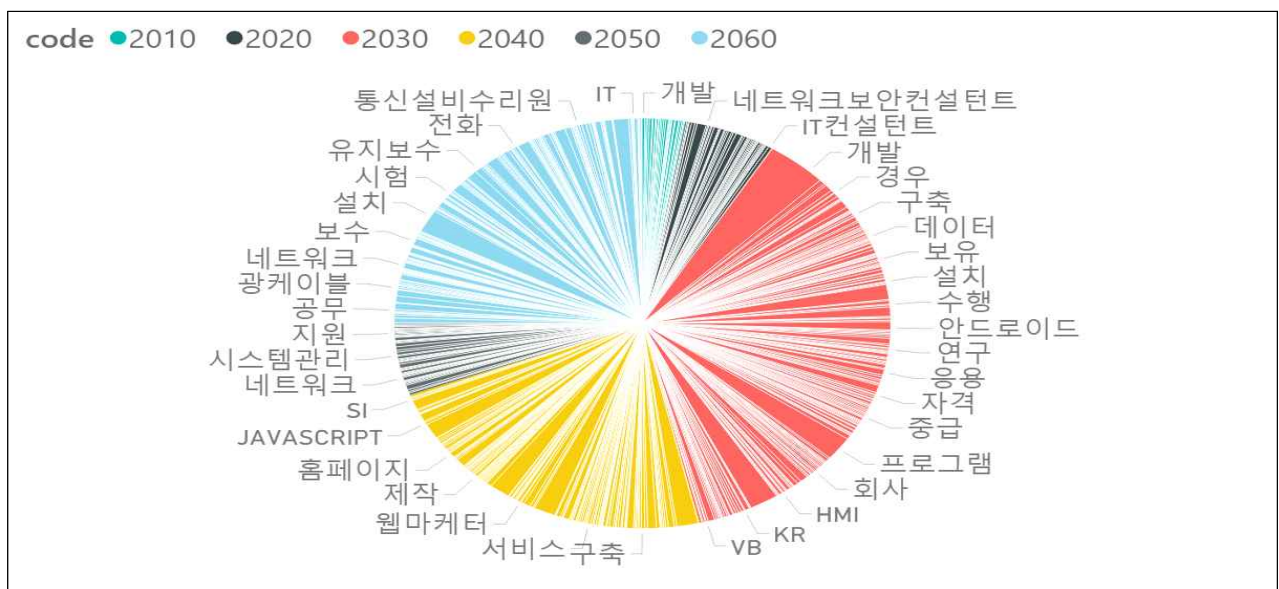
[그림 2-24] 모집 직종 코드별 단어 수



[그림 2-24]는 모집 직종 코드별 단어 수이다. 분석에 사용한 단어는 채용공고에 등장하는 제목, 단어, 직무내용에 해당하는 단어를 추출한 것이다. 직종 코드 203에 해당하는 단어 수가 2,216개로 가장 많으며 206번과 204번이 다음으로 많았다. 6개의 코드번호가 모두 총 500개의 단어 수를 초과하였으며 평균 1,222개의 단어가 추출되었다.

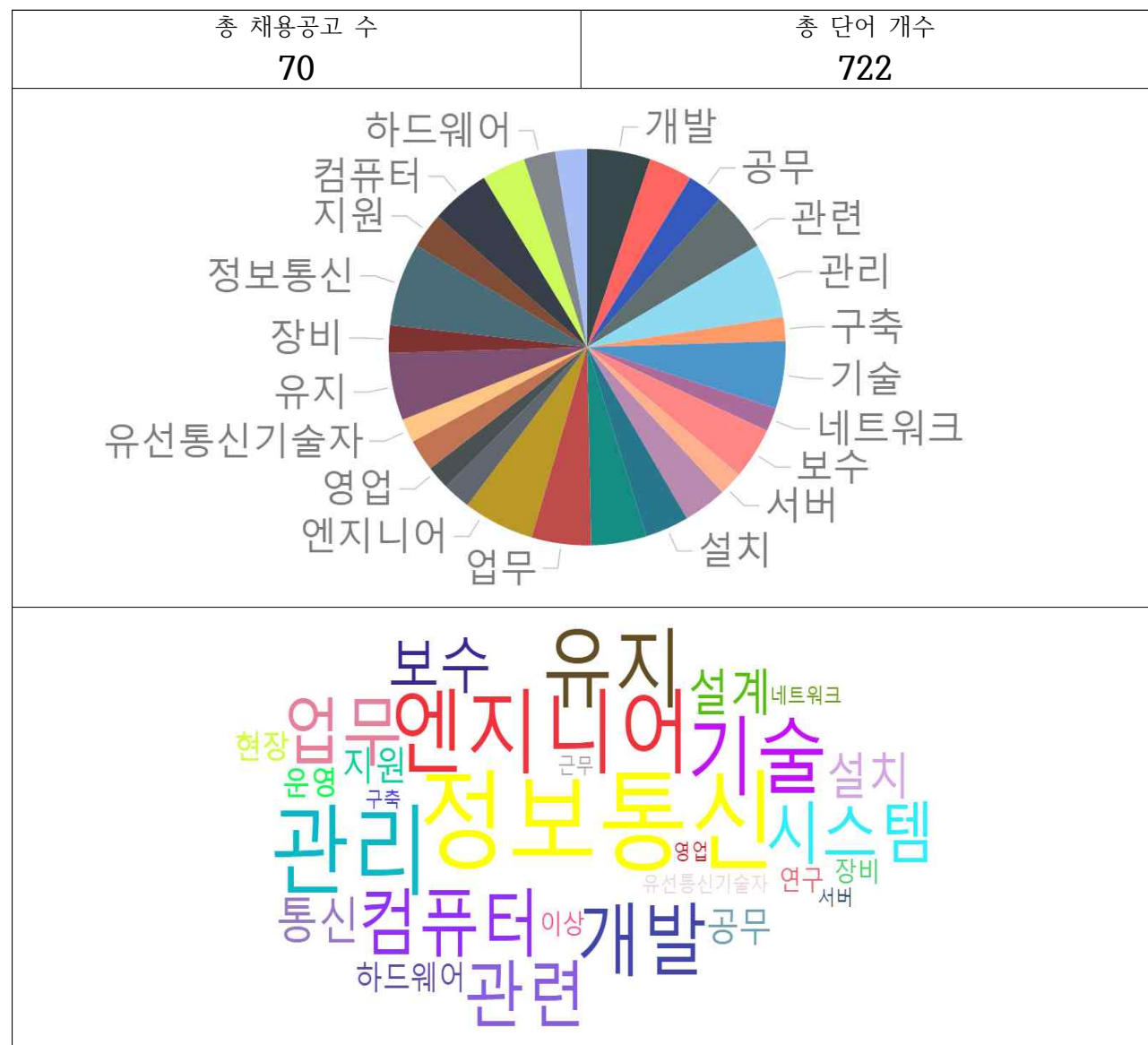
[그림 2-25]는 모집 직종 코드별 단어로 코드 번호에 따라 색깔을 달리하여 코드별 주요 단어들을 보여주고 있다. 시각화 결과는 각 코드별로 다양한 단어들이 할당됨을 알 수 있으나 단어의 수가 매우 많기 때문에 한눈에 어떤 단어들이 어떤 직업에 할당되는 지 확인하기 어렵다. 따라서 각 코드 별 분석을 별도로 수행이 필요하다고 생각되어 각 직업별로 분석을 추가 수행하였다.

[그림 2-25] 모집 직종 코드별 단어



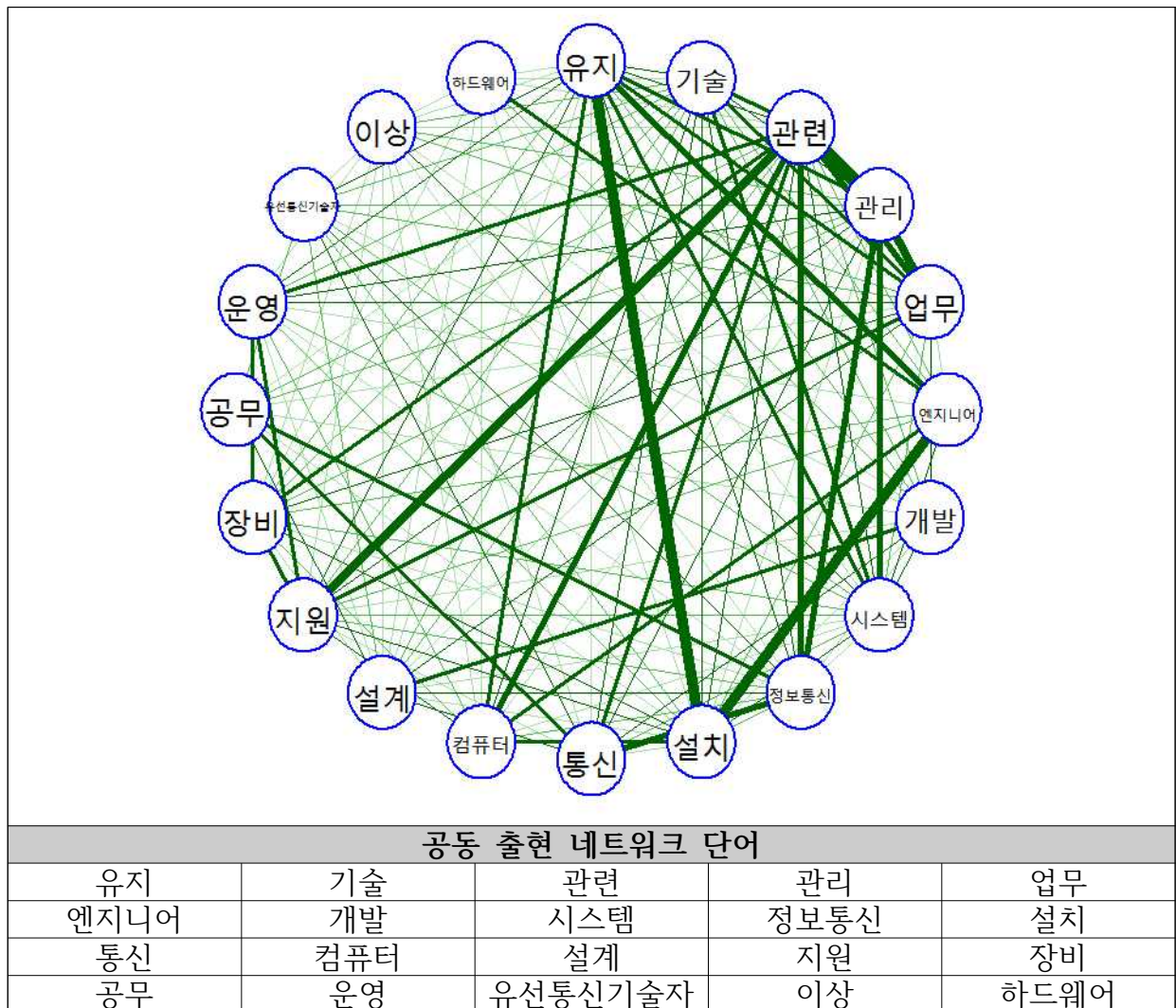
1) 컴퓨터 및 통신공학 기술자 및 연구원(201번)

[그림 2-26] 컴퓨터 및 통신공학 기술자 및 연구원(201번)



직종코드 201번에 해당하는 채용공고는 [그림 2-26]에서 보이는 것과 같이 70건으로 722개의 단어가 등장한다. 가장 많이 등장하는 단어는 ‘정보통신’으로 ‘엔지니어’, ‘기술’, ‘시스템’이라는 단어들이 그 뒤로 나타난다. 해당 공고는 정보통신과 관련된 기술 직무로 보인다. 또한 워크클라우드 결과는 ‘컴퓨터’와 ‘네트워크’와 같은 단어들이 등장하기 때문에 하드웨어 설치와 더불어 소프트웨어 관련 역량을 요구하는 것으로 나타났다.

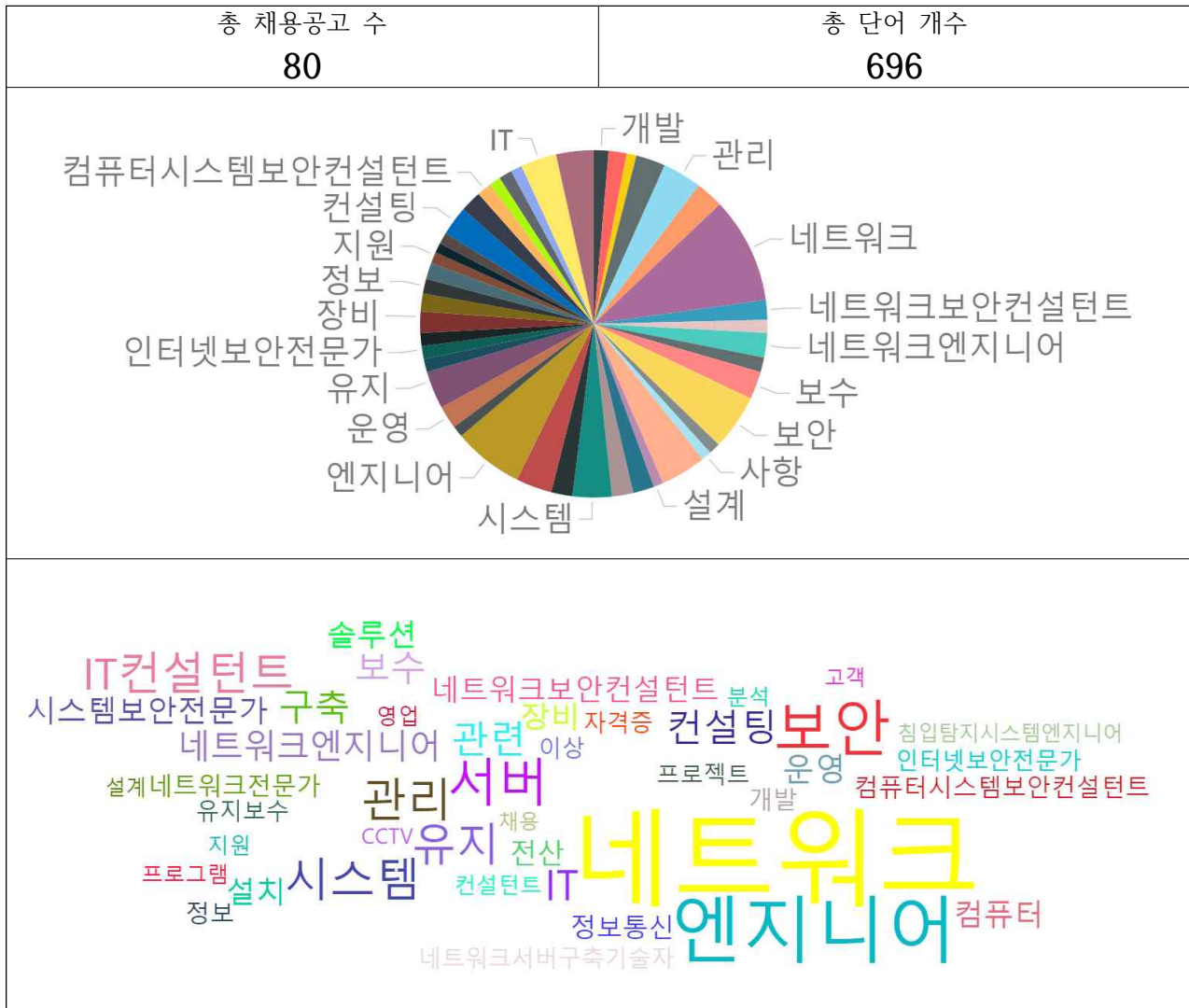
[그림 2-27] 컴퓨터 및 통신공학 기술자 및 연구원의 단어 공동 출현 네트워크



[그림 2-27]은 소셜네트워크 분석을 통한 컴퓨터 및 통신공학 기술자 및 연구원의 단어 공동 출현 네트워크이다. 분석 결과 {유지 - 설치 - 엔지니어}, {유지 - 컴퓨터 - 엔지니어}의 관계를 통해서 컴퓨터와 관련한 설치, 유지의 역량을 보유한 엔지니어를 요구하는 것으로 보인다. {통신 - 설치 - 정보통신}의 관계는 통신과 관련되어 설치를 하는 인력을 필요로 하는 것으로 보인다. 또한 {개발 - 설계}의 관계를 통해 설치, 유지와 다르게 연구를 하는 인력도 필요한 것으로 분석된다. 즉, 컴퓨터 및 통신공학 기술자 및 연구원은 컴퓨터의 유지 및 설치와 관련된 역량, 통신 설치와 관련된 역량 개발 연구와 관련된 역량을 요구 한다.

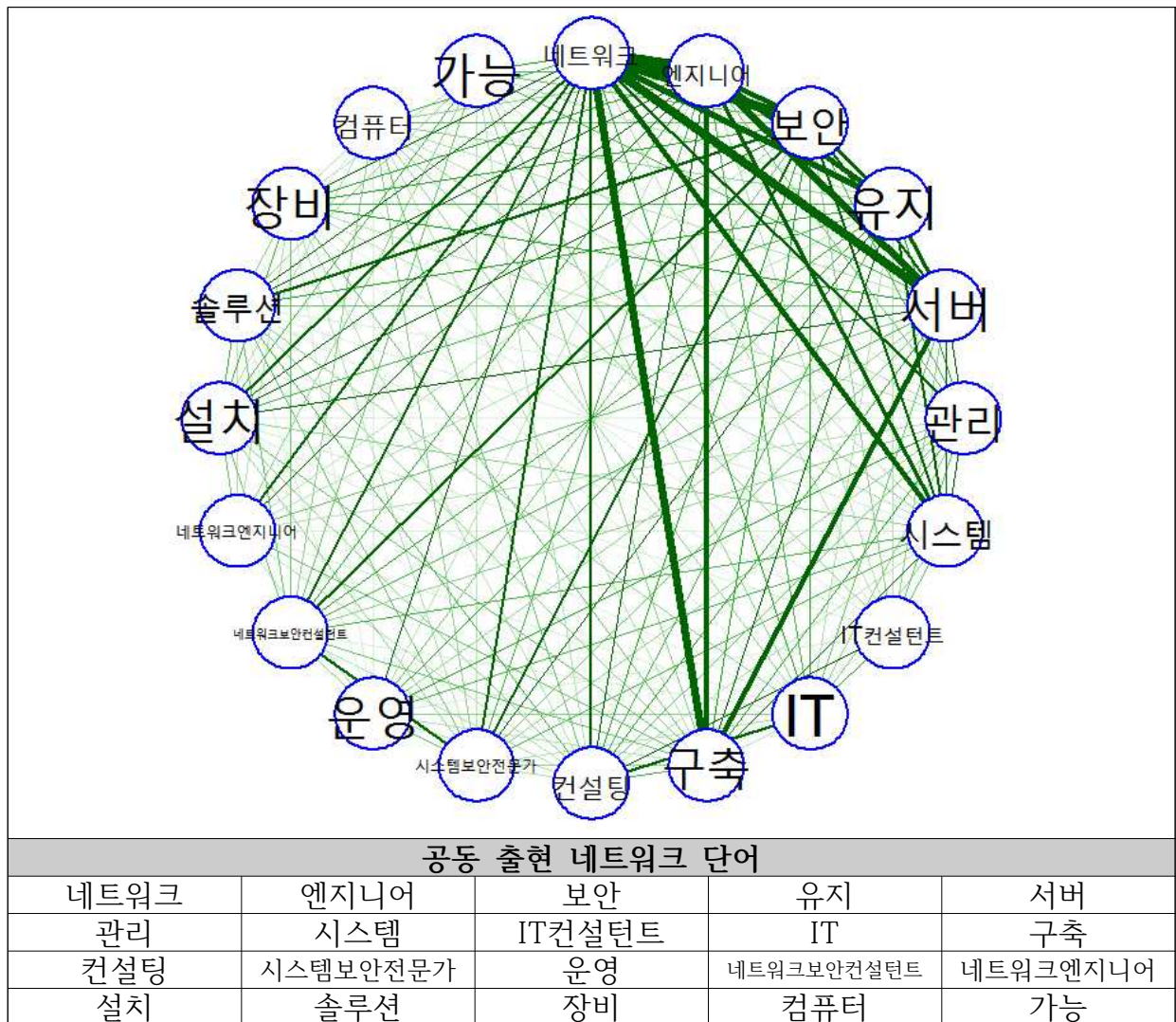
2) 컴퓨터시스템 설계 및 분석가(202번)

[그림 2-28] 컴퓨터시스템 설계 및 분석가(202번)



[그림 2-28]과 같이 직종코드 202번에 해당하는 채용공고는 80건으로 696개의 단어가 등장한다. 가장 많이 등장하는 단어는 ‘네트워크’로 ‘엔지니어’, ‘보안’, ‘시스템’이라는 단어들이 그 뒤로 나타난다. 해당 공고는 네트워크 보안과 관련된 기술 직무로 보인다. 또한 위드클라우드 분석 결과 ‘보수’와 ‘유지’와 같은 단어들이 등장하기 때문에 네트워크와 서버를 설치하는 역량을 요구하는 것을 알 수 있다. 또한 ‘IT컨설턴트’, ‘컴퓨터시스템보안컨설턴트’ 등 컨설팅과 관련된 역량을 필요로 하는 직무로 판단된다.

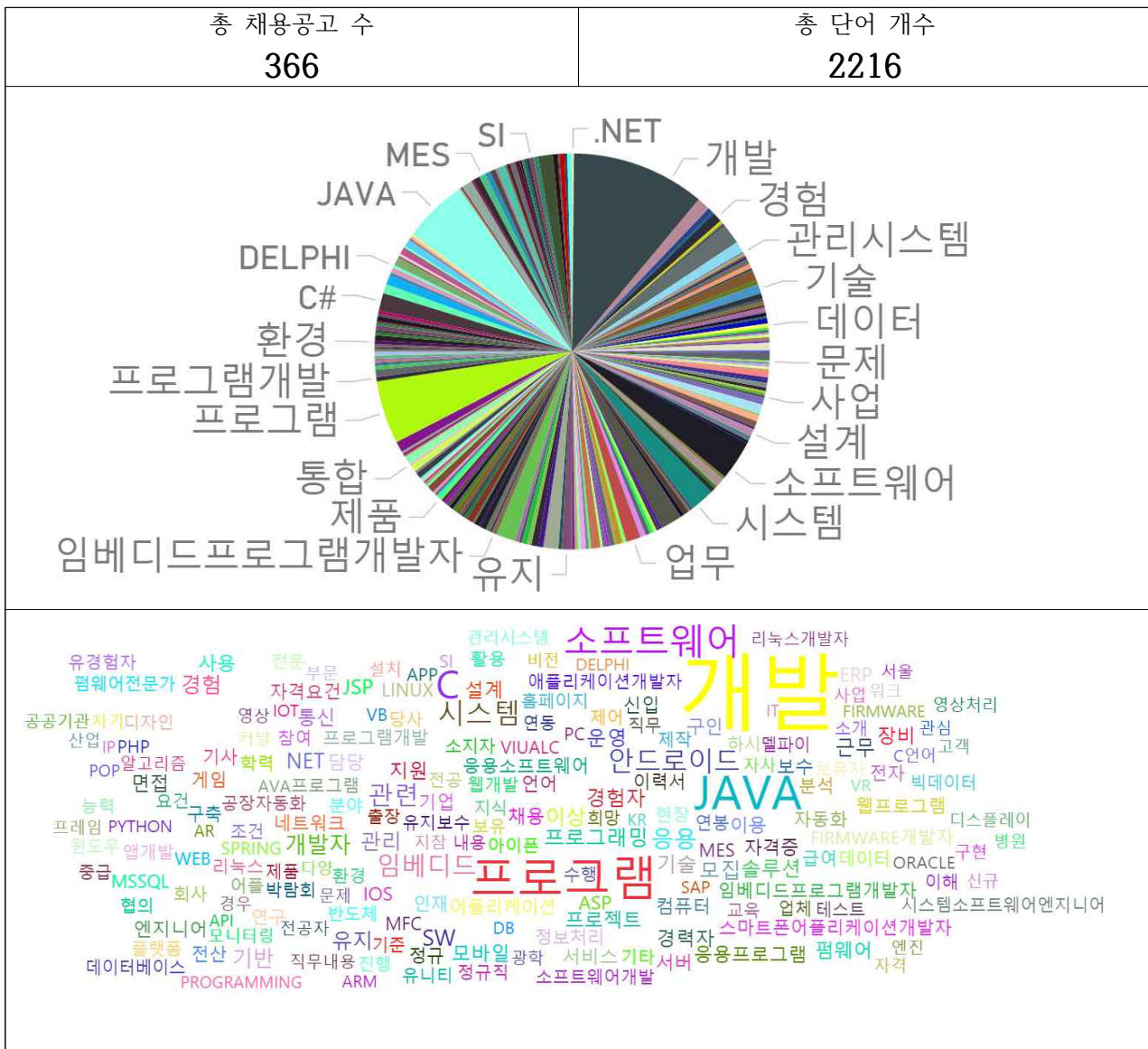
[그림 2-29] 컴퓨터시스템 설계 및 분석가의 단어 공동 출현 네트워크



[그림 2-29]는 소셜네트워크 분석을 통한 컴퓨터시스템 설계 및 분석가의 단어 공동 출현 네트워크이다. {네트워크-엔지니어-서버}, {네트워크, 서버, 구축}의 관계를 통해서 서버 구축과 관련한 엔지니어를 모집하는 것을 확인 할 수 있다. {네트워크-보안-네트워크엔지니어}의 관계는 보안과 관련되어 인력을 필요로 하는 것으로 보인다. 또한 {솔루션-보안}, {컨설팅-IT}의 관계가 나타난다. 즉, 컴퓨터시스템 설계 및 분석가는 서버 구축과 관련된 역량, 보안과 관련된 역량을 요구하며, 보안 솔루션이나 IT 컨설팅과 관련된 지식과 관계있는 것으로 보인다.

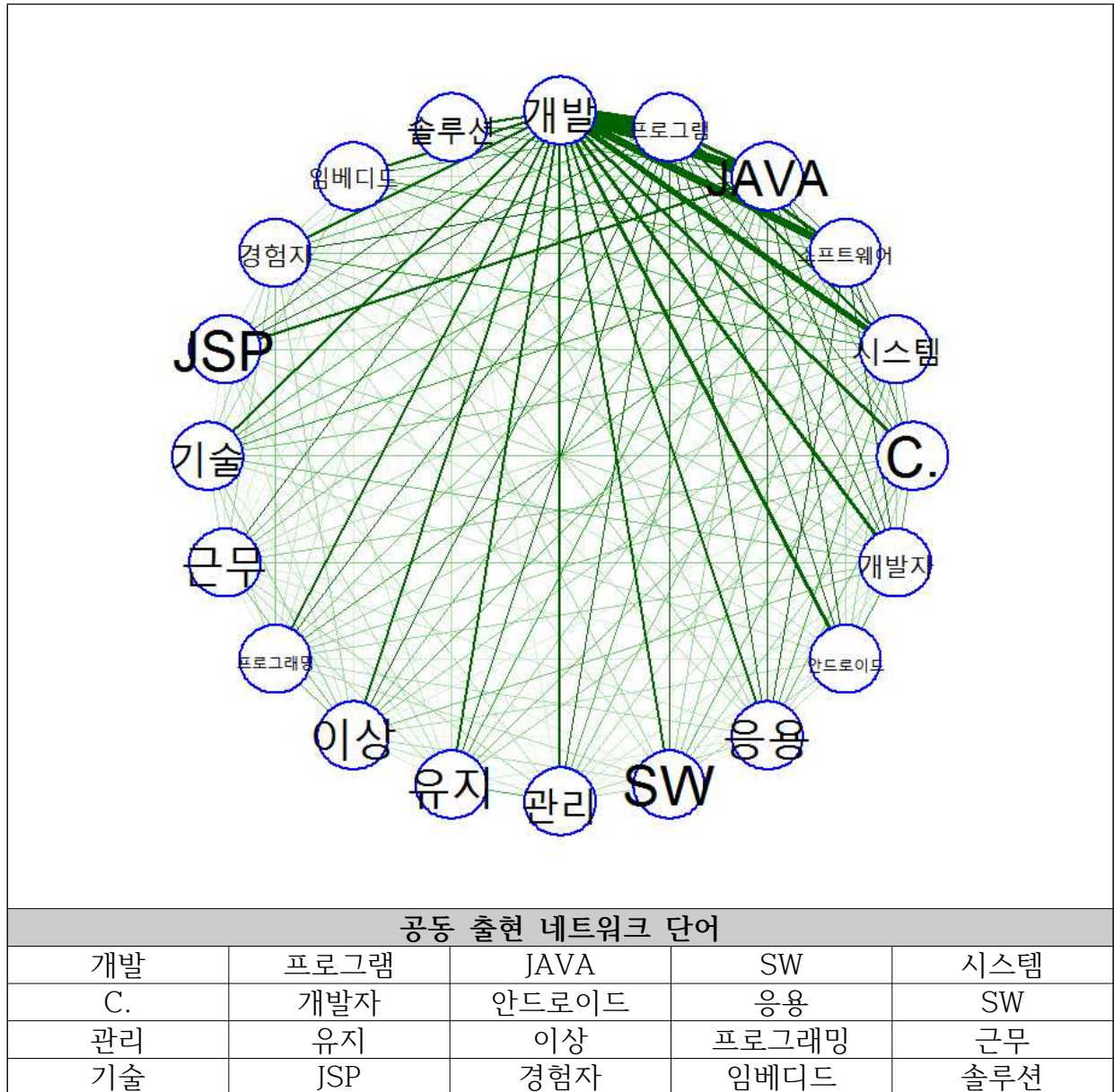
3) SW개발전문가(203번)

[그림 2-30] SW개발전문가(203번)



[그림 2-30]과 같이 직종코드 203번에 해당하는 채용공고는 366건으로 2216개의 단어가 등장한다. 가장 많이 등장하는 단어는 ‘개발’로 ‘프로그램’, ‘SW’, ‘JAVA’라는 단어들이 그 뒤로 나타난다. 또한 워드클라우드 결과와 같이 프로그래밍 언어인 C, JAVA, 임베디드프로그램 등 SW와 관련된 역량을 요구하는 것으로 해당 직무는 SW 개발자와 같은 직무로 분석된다.

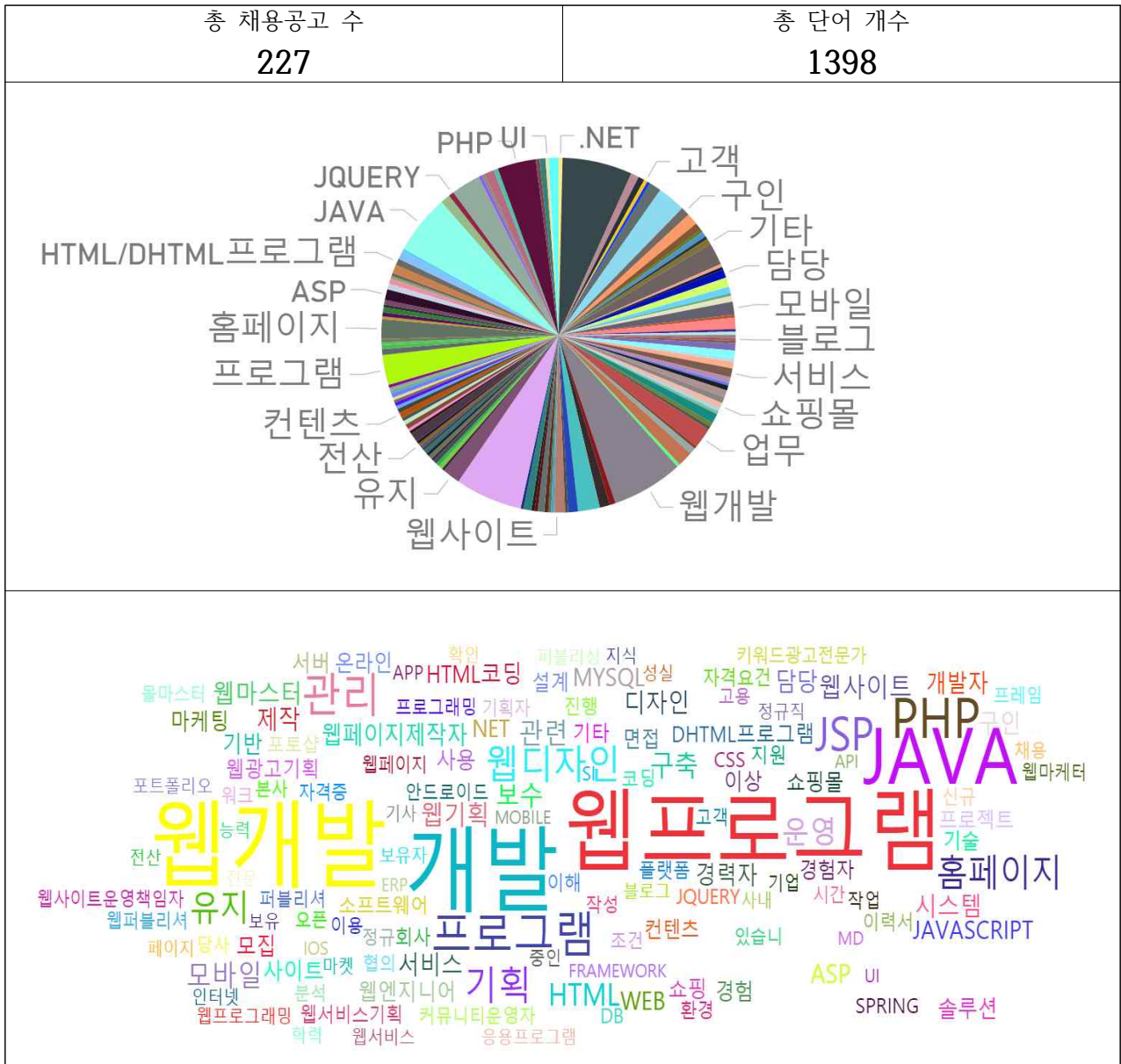
[그림 2-31] SW개발전문가의 단어 공동 출현 네트워크



[그림 2-31]은 소셜네트워크 분석을 통한 SW개발전문가의 단어 공동 출현 네트워크이다. {개발-프로그램-SW}의 관계를 통해서 SW 프로그램을 개발하는 인력을 모집하는 것으로 보인다. ‘개발’의 경우 {개발-C.}, {개발-안드로이드}, {개발-SW}, {개발-JAVA}, {개발-JSP}, {개발-임베디드}, {개발-솔루션}의 관계가 나타난다. 즉, SW개발전문가는 각종 프로그래밍 언어를 통해 개발할 수 있는 역량을 요구하며, 솔루션, 임베디드 등 특정 영역의 개발자의 역량을 요구하고 있다.

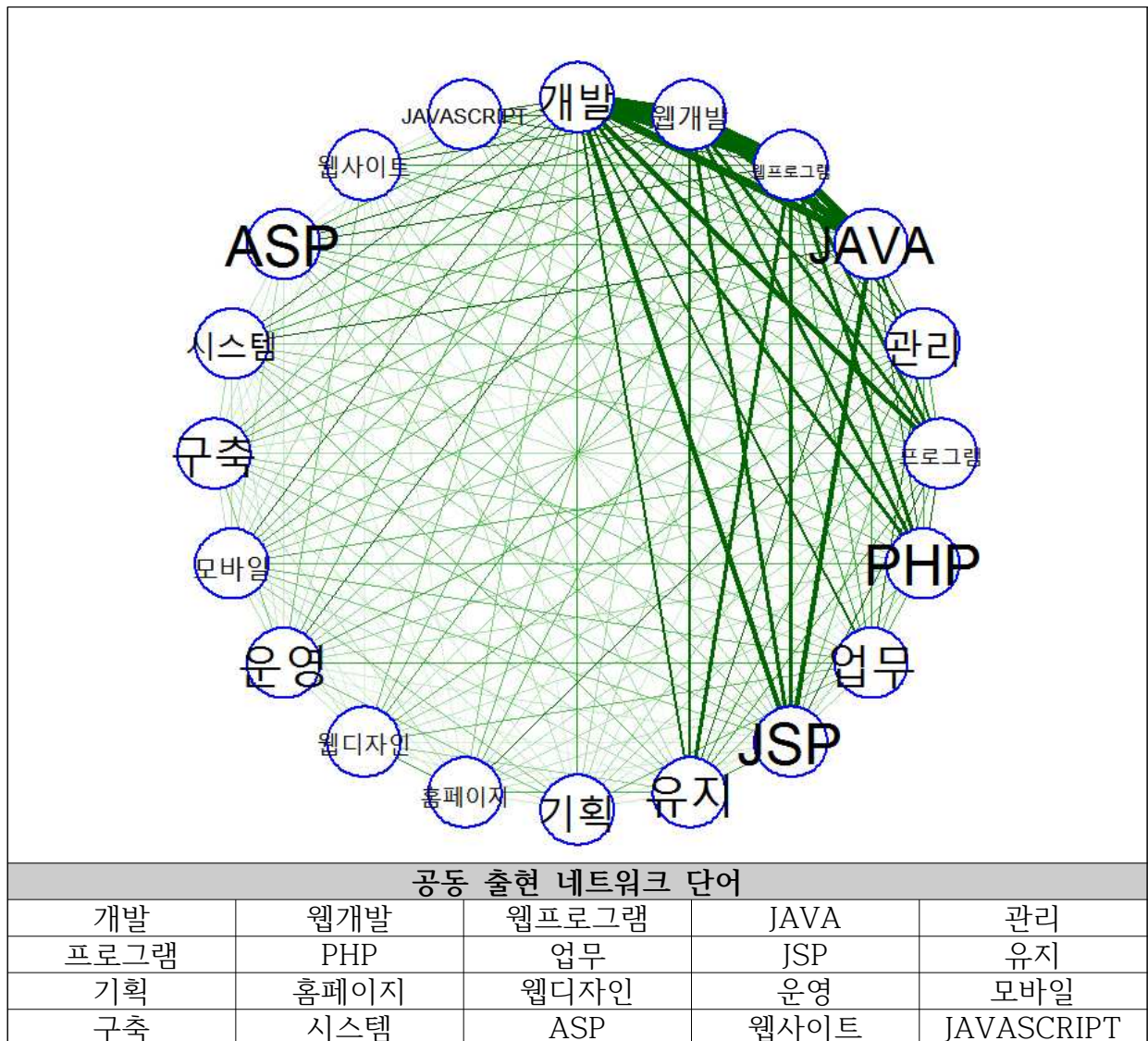
4) 웹전문가(204번)

[그림 2-32] 웹전문가(204번)



직종코드 204번에 해당하는 채용공고는 [그림 2-32]와 같이 227건으로 1398개의 단어가 등장한다. 빈번하게 등장하는 단어들은 ‘웹개발’, ‘웹프로그램’, ‘개발’, ‘JAVA’, ‘PHP’, ‘JSP’로 웹 관련된 단어들 혹은 프로그래밍 언어들이다. 해당 직무는 웹과 관련된 역량을 필요로 하는 것으로 분석된다.

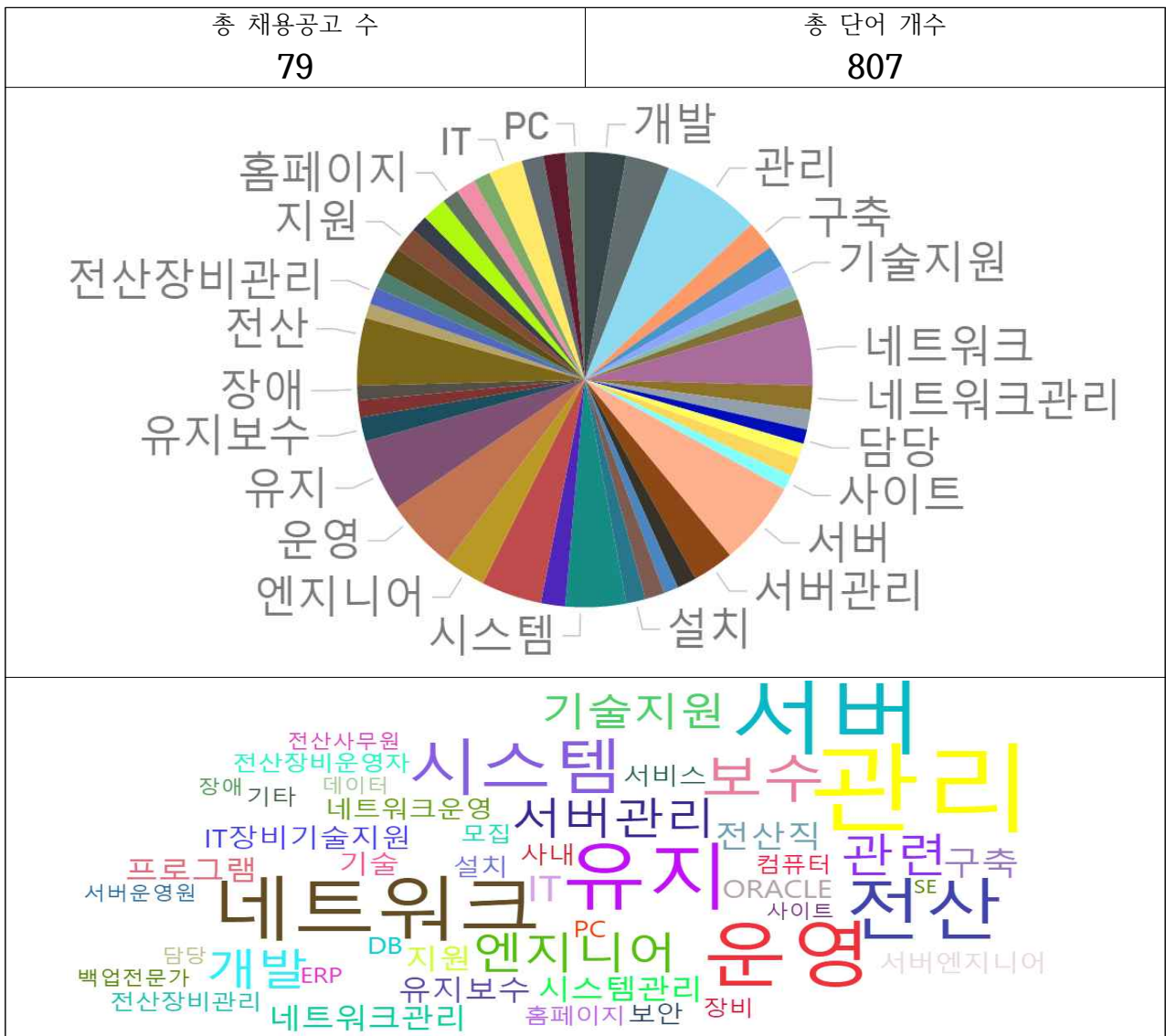
[그림 2-33] 웹전문가의 단어 공동 출현 네트워크



[그림 2-33]은 소셜네트워크 분석을 통한 웹전문가 단어 공동 출현 네트워크이다. {웹개발-웹프로그래밍-프로그램-JAVA-JSP}의 관계를 통해서 웹과 관련한 프로그래밍 역량을 갖춘 인력을 모집하는 것으로 보인다. 온라인상에서 서버와 통신하기에 적합한 프로그래밍 언어인 PHP와 JSP가 포함된 {PHP-웹프로그램-개발}, {JSP-JAVA-개발}의 관계를 통해서 서버와 관련된 인력을 필요로 하는 것을 알 수 있다. 즉, 웹전문가는 웹과 관련된 역량을 요구하며, PHP, JSP, JAVA와 같은 웹프로그래밍 언어를 사용할 수 있는 역량을 요구하는 것으로 보인다.

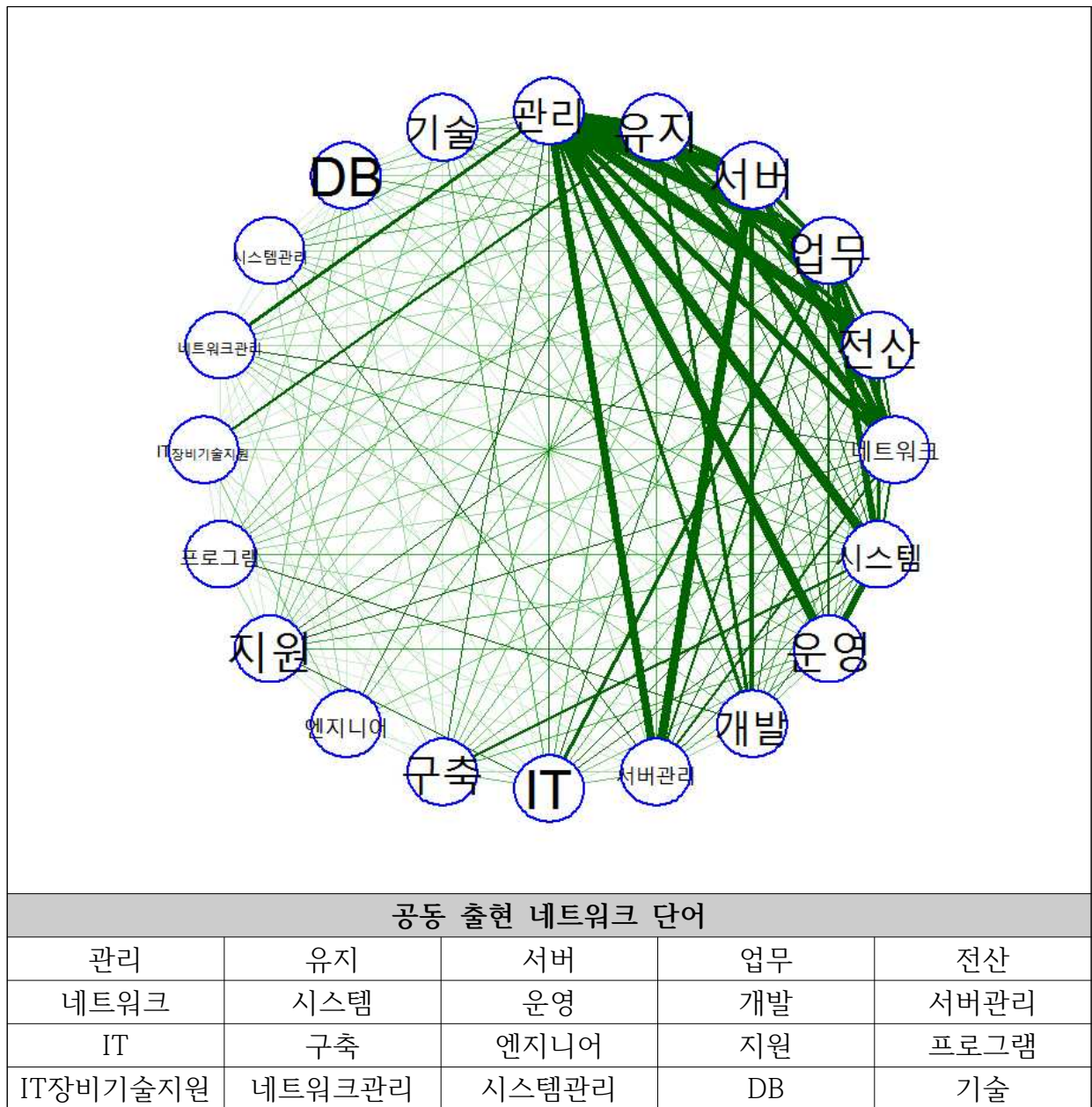
5) 데이터베이스 설계·구축 및 관리기술자(205번)

[그림 2-34] 데이터베이스 설계·구축 및 관리기술자(205번)



워크넷 번호 205번에 해당하는 채용공고는 [그림 2-34]와 같이 79건으로 807개의 단어가 등장한다. 빈번하게 등장하는 단어들은 ‘네트워크’, ‘서버’, ‘관리’, ‘시스템’ 등이 있다. 해당 단어들은 워크넷 번호 202번과 유사하다. 워드클라우드 결과에서 보이는 것과 같이 하지만 컨설팅과 관련한 단어는 등장하지 않고 ‘유지’, ‘보수’, ‘운영’ 단어들이 등장하는 것으로 네트워크와 서버의 관리적 측면의 역량을 요구하는 것으로 분석된다.

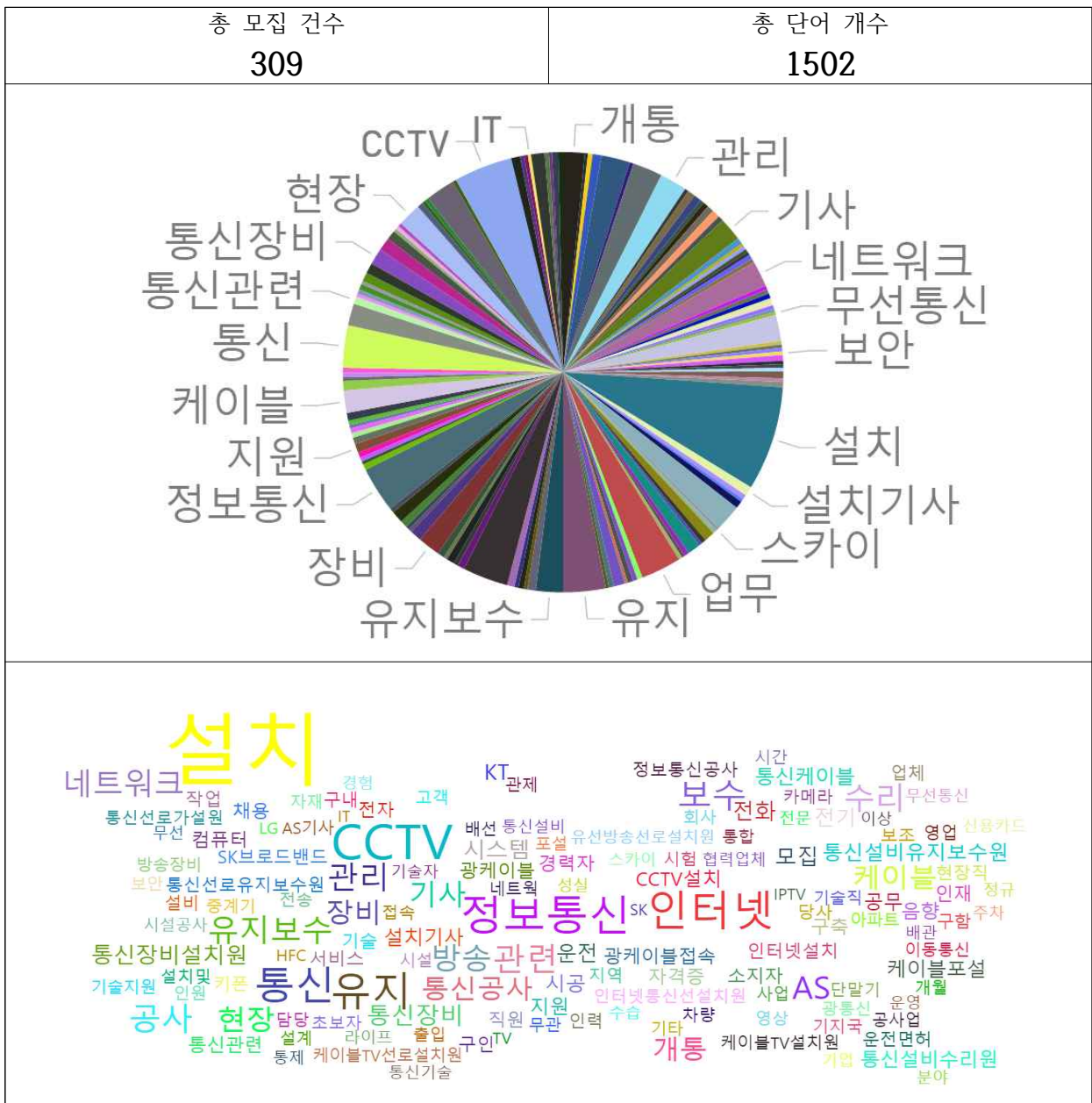
[그림 2-35] 데이터베이스 설계·구축 및 관리기술의 단어 공동 출현 네트워크



[그림 2-35]는 소셜네트워크 분석을 통한 데이터베이스 설계·구축 및 관리기술의 단어 공동 출현 네트워크이다. {관리-유지-서버-전산-네트워크}, {관리-서버-서버관리}의 관계를 통해서 서버를 사용할 수 있는 역량을 갖춘 인력을 모집하는 것으로 보인다. 또한 {서버-관리-개발-유지}의 관계는 서버를 개발부터 유지 관리까지 하는 관계를 나타낸다. 즉, 데이터베이스 설계·구축 및 관리기술은 서버와 네트워크를 사용할 수 있는 역량을 요구하며, 서버 개발의 영역도 관계가 있는 것으로 보인다.

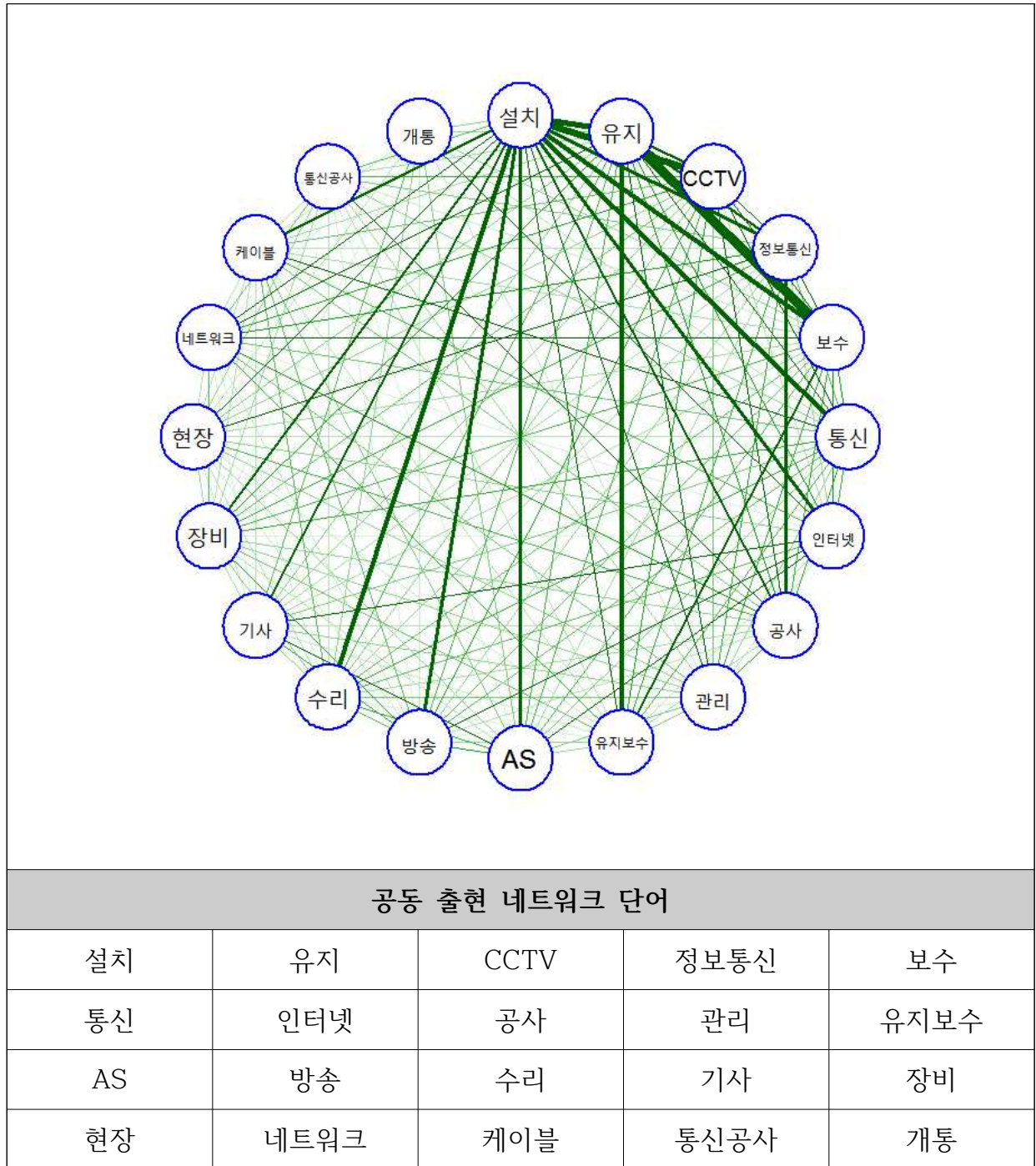
6) 통신 및 방송 장비기사 설치 및 수리원(206번)

[그림 2-36] 통신 및 방송 장비기사 설치 및 수리원(206번)



직종코드 206번에 해당하는 채용공고는 309건으로 1502개의 단어가 등장한다. [그림 2-36]과 같이 가장 많이 등장하는 단어는 ‘설치’로 ‘정보통신’, ‘인터넷’, ‘CCTV’이라는 단어들이 그 뒤로 나타난다. 또한 워드클라우드 결과 해당 공고에서 ‘설치기사’, ‘케이블’, ‘개통’이라는 단어들이 동시에 등장하는 것으로 분석하면, 케이블, 네트워크의 설치와 관련된 직무로 보인다.

[그림 2-37] 통신 및 방송 장비기사 설치 및 수리원의 단어 공동 출현 네트워크



[그림 2-37]은 소셜네트워크 분석을 통한 통신 및 방송 장비기사 설치 및 수리원의 단어 공동 출현 네트워크이다. ‘설치’와 관련하여 {설치-통신}, {설치-케이블}, {설치-방송}, {설치-CCTV}, {설치-인터넷}의 관계는 각종 설치와 관련된 관계있는 것으로 분석된다. 또한 {설치-유지-보수}의 관계를 통해서 설치부터 유지 보수 단계까지의 역량을 갖춘 인력을 필요로 한다. {정보통신-공사-설치}의 관계를 통해서 정보통신의 경우 설치와 공사의 역량이 요구된다. 하지만 {인터넷-AS}, {인터넷-설치-개통}의 관계에서 인터넷의 경우 공사가 아닌 수리와 개통의 역량을 요구한다. 즉, 통신 및 방송 장비기사 설치 및 수리원은 각종 설치와 관련된 역량을 요구하며, 수리, 공사와 관련된 역량도 관계가 있는 것으로 보인다.

이와 같이 워크넷 직업 분류를 바탕으로 매칭 되는 직업 별로 채용공고 내용의 단어들을 활용하여 직종별 주요 단어들을 도출해 볼 수 있었다. 따라서 워크넷 체계가 최근 제시되고 있는 NCS 분류체계와 일치되는 직무 내용을 제시하고 있는지 알아보기 위하여 NCS 분류체계를 통하여 각 직무별 중요 단어들을 도출해 보았다.

2. NCS 직종분류 데이터

상기 워크넷 내용을 텍스트마이닝을 실시한 후 직무에 관련한 능력표준을 정해놓은 NCS와 비교 분석하기 위하여 자료 수집을 실시하였다. 일반적으로 제시되고 있는 NCS의 정의 및 활용은 다음과 같다.

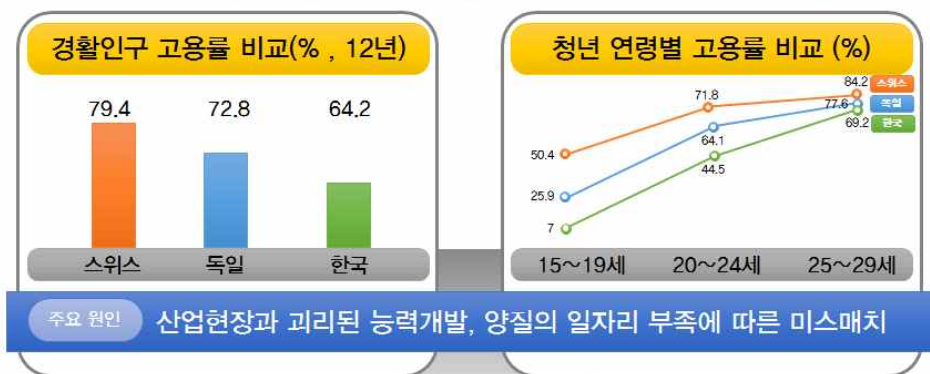
1) 국가직무능력표준, NCS (National Competency Standards)

국가직무능력표준(이하 NCS)은 현장에서 직무를 수행하기 위한 필요로 되는 기술·지식·태도 등의 내용을 체계적으로 기술하여 정립해 놓은 것으로 [그림 2-38]과 같이 고용노동부에서 진행되는 ‘능력중심사회’ 기반의 직업능력개발 3개년 실천계획의 일환으로 산업계의 현장 능력을 갖추기 위한 기본적인 소양 및 기준을 정하여 그 기준에 맞게 직무 평가를 할 수 있는 제도를 만든 것이다.

[그림 2-38] 직업능력개발 혁신 3개년 실천계획

추진 필요성

우리의 청년 고용률은 독일과 스위스 대비 매우 낮음

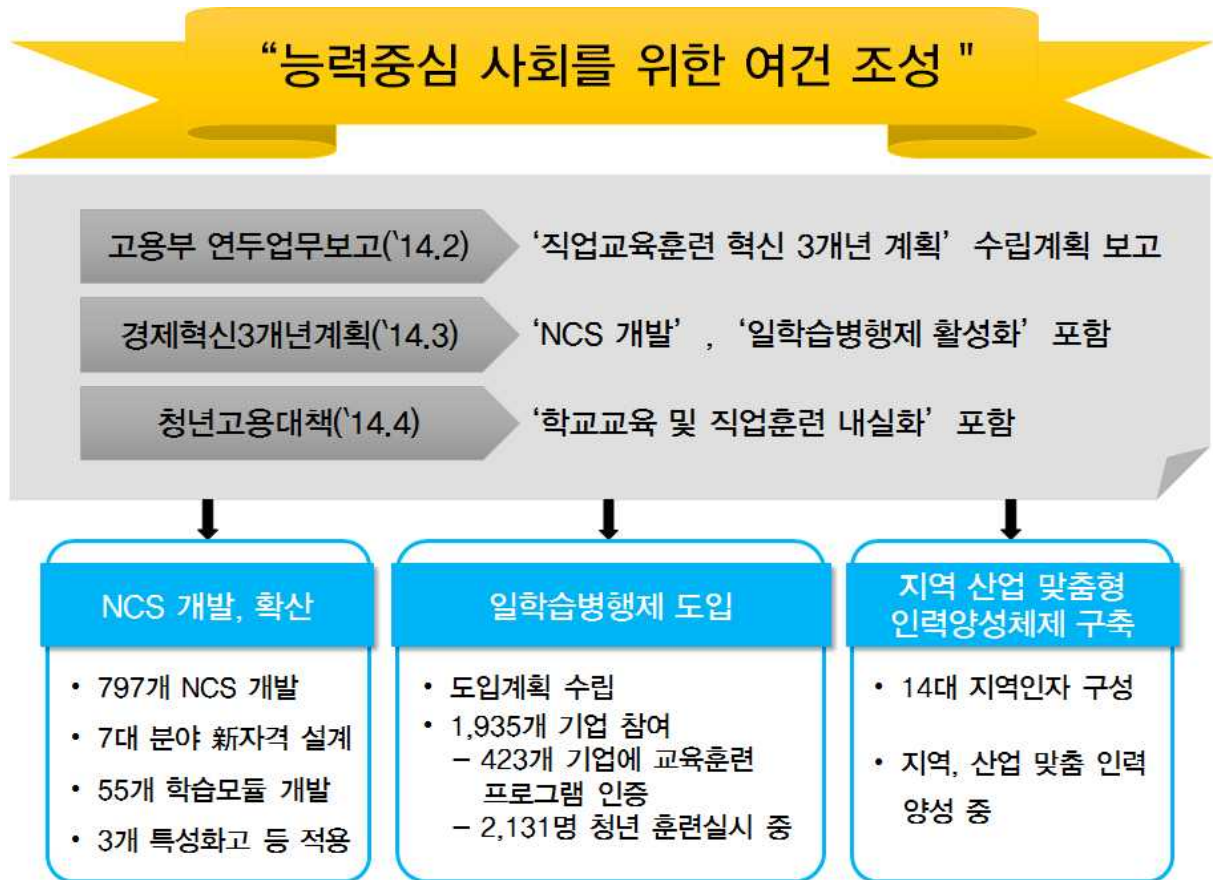


직업능력개발 혁신을 통해
청년 고용률을 높이고, 능력중심사회의 기반 마련

출처: 고용노동부 '14.12.18 발표자료¹¹⁾

11) 고용노동부, '능력중심사회를 위한 직업능력개발 혁신 3개년 실천계획', 관계부처합동, 보도자료 2014.09.25

[그림 2-39] 능력중심사회 3개년 추진 경과



출처: 한국소프트웨어기술진흥협회 (2014)¹²⁾

능력중심사회 3개년 추진 경과에 따라서 NCS 개발 강화 및 일학습병행제, 지역 산업 맞춤형 인력양성체제를 구축하고 있으며 특히 NCS는 능력중심의 인재를 산학이 연계성 있게 연계시키기 위한 표준을 제공함으로써 산업계와 교육의 징검다리 위치에서 지속적으로 개발 되고 있다. NCS는 다음과 같이 3가지의 활용범위가 존재한다.

첫째, 기업체에서 활용 할 수 있는 부분은 현장에 대한 목소리를 담은 표준 모듈로 인하여 취업 또는 진로를 준비하는 이들에게 가이드로서 제공될 수 있다는 점이다. 둘째, 교육훈련기관에서는 직업 교육 훈련 과정에 있어 표준을 갖고 진행한다는 점을 내세우고 있다. 셋째, 자격시험의 경우 자격종목을 개설하여 원하는 산업 인력에 대한 인재 역량 측정에 도움을 주는 역할을 하도록 구성하고 있다. 그러나 SW 분야에 적합한 직무와 관련하여 광의의 직무 개념을 제공하고 있다는 점에서는 보완이 필요한 실정이다.

12) 한국소프트웨어기술진흥협회, 고용노동부 『능력중심을 위한 직업능력개발 혁신 3개년 실천계획』, 2014, http://www.kosta.or.kr/mail/2014/download/05-SW_NCS_conference_141217_KimYoungJu.pdf

[그림 2-40] 국가직무능력표준 활용범위

● 국가직무 능력표준 활용범위



출처: 한국산업인력공단, 국가직무능력표준 웹페이지¹³⁾

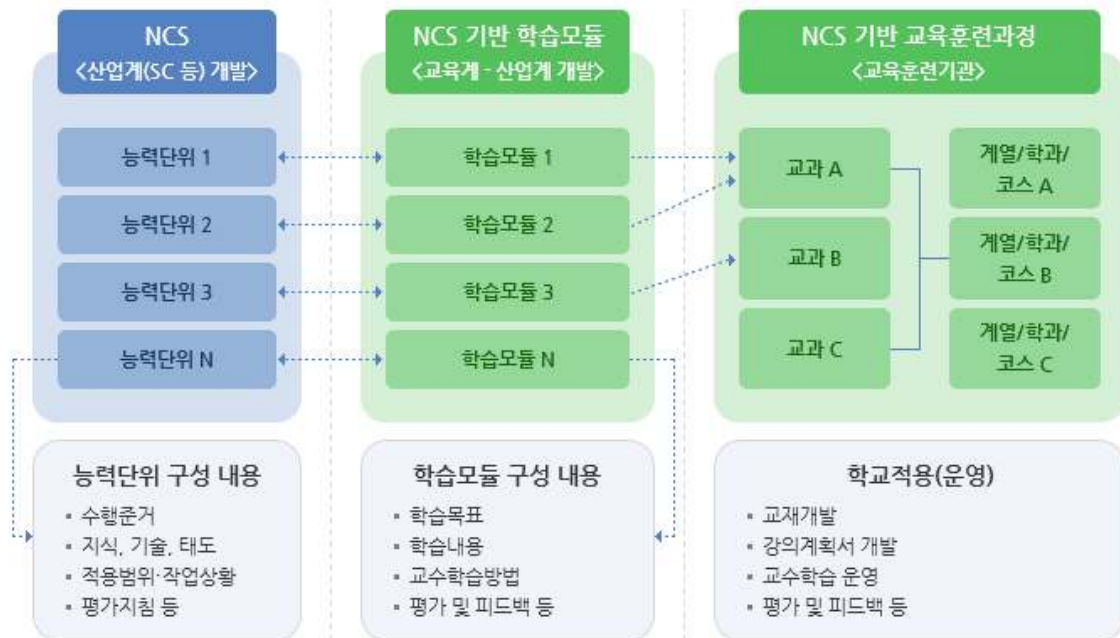
2) NCS 학습 모듈의 개념

NCS 자체가 직무 기술서의 역할을 한다고 하면 학습 모듈은 이를 기반으로 직무에 대하여 어떤 방법 및 어떠한 도구로 교수 및 학습을 해야 하는 것인지를 가이드 해주는 일종의 가이드라인을 제공한다. 특히, 산업계에서 요구하는 직무능력을 모듈화 하여 학습함으로써 산업현장에서 바로 사용 가능한 준비된 인재를 만드는 것을 목표로 하여 이에 대한 각각의 성취 목표를 설정하고 학습 방향을 제시하는 모듈을 구성했다.

NCS 학습모듈을 활용할 수 있는 기관은 크게 특성화고, 마이스터고, 전문대학, 4년제 대학 등의 교육 훈련 기관을 대상으로 할 수 있으며, 직장 교육 기관 등에서도 관련 직무의 전문성 제고와 기초 자질에 대한 평가 등을 활용 할 수 있다고 언급하고 있다. 그러나 직무체계 분류의 측면에서는 구체적인 직무의 개념이 명확하지 않기 때문에 NCS의 활용도에 대한 지속적인 관심은 필요한 실정이며 어떠한 방향성이 직무에 대한 접근인지에 대하여 연구할 필요성을 도출할 수 있다.

13) 국가직무능력표준, 국가직무 능력표준 활용범위, http://www.ncs.go.kr/ncs/page.do?sk=P1A1_PG01_001

[그림 2-41] NCS의 학습모듈의 개념



출처: 국가직무능력표준 홈페이지, ‘NCS 학습모듈의 개념’, 2013¹⁴⁾

[그림 2-42] NCS 분야별 검색

01 사업관리	02 경영·회계·사무	03 금융/보험	04 교육/자연	05 법률/경찰	06 보건/의료	07 사회복지·종교	08 문화예술·디자인
09 운전·운송	10 영업판매	11 경비·청소	12 숙박·여행·오락	13 음식서비스	14 건설	15 기계	16 재료
17 화학	18 섬유·의복	19 전기전자	20 정보통신	21 식품가공	22 인쇄·복제·가구·공예	23 환경·에너지	24 농림어업

출처: 국가직무능력표준 홈페이지, ‘NCS 및 학습모듈 검색’, 2013¹⁵⁾

NCS는 학습 모듈을 기반으로 총 24개의 직군을 분류하고 있으며 분류 내용은 [그림 2-42]와 같다. 24개의 분류 중 SW 인력과 관련이 있는 분류는 ‘20. 정보통신’에 관련 되어 있기 때문에 본 보고서에는 NCS 직무와 관련하여 ‘20. 정보통신’ 분류만을 선택하여 분석 하였다.

14) https://www.ncs.go.kr/ncs/page.do?sk=P1A1_PG01_006

15) https://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_001#none

3) 정보통신 세분류

NCS 분류상 정보통신 분야는 3가지 중분류 정보기술, 통신기술, 방송기술 분야로 나뉘어 구분되어 있으며 이를 표로 정리한 내용은 <표 2-22>와 같이 소분류는 직무에 맞게 다양하게 배치하였다. 소분류는 각 중분류에 따라서 많게는 6개의 분류로 나뉘지고 있으며 해당 분류에 대한 단어를 도출하여 분석에 사용하였다.

<표 2-22> 정보통신 세부 분류

중분류		소분류		세분류		
01	정보기술	01	정보기술전략·계획	01	정보기술전략	
				02	정보기술컨설팅	
				03	정보기술기획	
				04	SW제품기획	
				05	빅데이터분석	
				06	IoT 응용서비스기획	(신규)
		02	정보기술개발	01	SW아키텍처	
				02	응용SW엔지니어링	
				03	임베디드SW엔지니어링	
				04	DB엔지니어링	
				05	NW엔지니어링	
				06	보안엔지니어링	
				07	UI/UX엔지니어링	
				08	시스템SW엔지니어링	
		03	정보기술운영	01	IT시스템관리	
				02	IT기술교육	
				03	IT기술지원	
		04	정보기술관리	01	IT프로젝트관리	
				02	IT품질보증	
				03	IT테스트	
				04	IT감리	
		05	정보기술영업	01	IT기술영업	
				02	IT마케팅	
		06	정보보호	01	정보보호관리·운영	(신규)
				02	보안사고분석·진단	(신규)
				03	정보침해사고대응	(신규)
02	통신기술	01	유선통신구축	01	교환시스템구축	
				02	구내통신구축	
				03	네트워크구축	

중분류		소분류		세분류		
		02	무선통신구축(이동통신포함)	01	무선통신시스템구축	
				02	전송시스템구축	
				03	무선통신망구축	
				04	위성통신망구축	
		03	통신서비스	01	유선설비접속서비스	
				02	전용회선서비스	
				03	초고속망서비스	
				04	부가네트워크서비스	
				05	전신서비스	
				06	이동통신서비스	
				07	콘텐츠사용자서비스	
				08	콘텐츠네트워크서비스	
				09	무선초고속인터넷서비스	
				10	주파수공용통신	
				11	무선호출메시징서비스	
				12	위성통신서비스	
				13	특수이동통신서비스	
				14	인터넷지원서비스	
				15	부가통신응용중계서비스	
				16	특수부가통신서비스	
				17	무선데이터통신서비스	
				18	디지털비즈니스지원서비스	(신규)
		04	실감형콘텐츠제작	01	실감형콘텐츠제작	(신규)
03	방송기술	01	방송제작기술	01	중계방송	
				02	방송품질관리	
		02	방송플랫폼기술	01	라디오방송	
				02	지상파TV방송	
				03	지상파 DMB 방송	
				04	케이블 방송	
				05	인터넷멀티미디어방송	
		03	방송서비스	01	유무선통합서비스	
				02	방송시스템운영	
				03	정보시스템운영	
				04	방송기술지원서비스	
				05	방송장비설치유지보수	

4) NCS 직종분류 데이터 비교

워크넷에서 수집된 채용광고들을 내용을 직종코드별로 탐색해 보았다. 워크넷 직종코드가 부여된 채용광고들과 내용적으로 가장 유사한 직종을 NCS 직업분류 체계에서 파악하였다 (유사한 직종을 파악하는 방안은 제5절 참조).

워크넷에서 수집된 채용광고 중에서 직업코드가 부여된 광고가 주로 포함하고 있는 단어들과 NCS에서 세분류 직군에 대해 정의하고 있는 기술, 지식, 태도에 대하여 워크넷 채용광고 데이터를 분석했다. 각 텍스트에 등장하는 단어를 추출하여 각 단어의 빈도수가 5를 초과하는 단어들만 선별하여 원형 차트와 워드클라우드를 제작하였다.

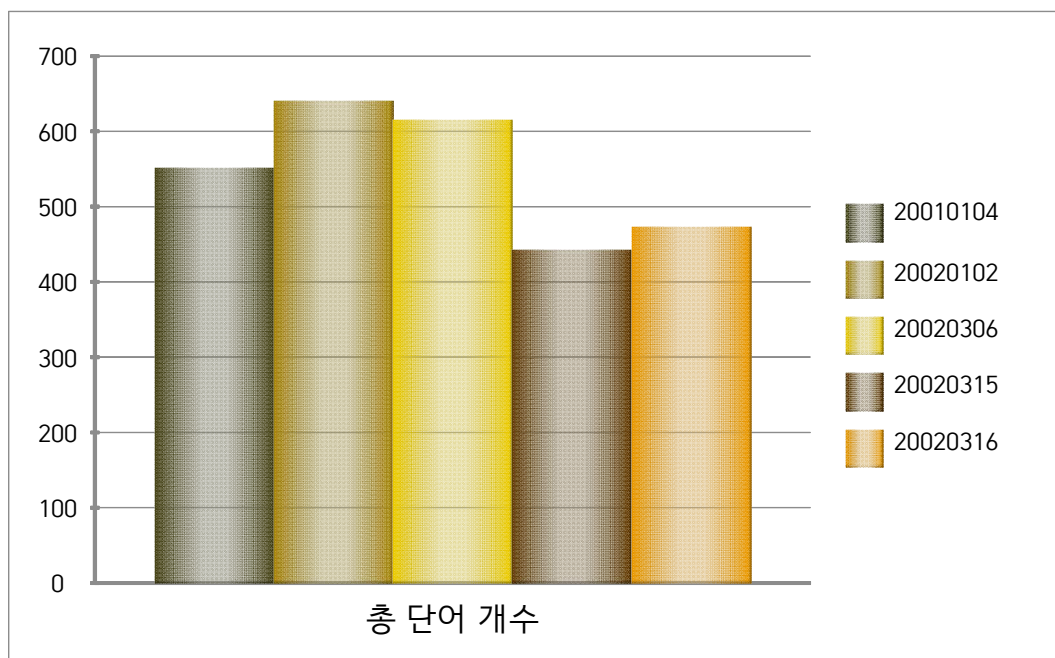
〈표 2-23〉 워크넷 직종코드와 일치도가 높은 NCS 세분류

워크넷 직종코드	NCS 세분류		
	1순위	2순위	3순위
201 (컴퓨터 및 통신공학 기술자 및 연구원)	20020315 (부가통신응용중계 서비스)	20020316 (특수부가통신 서비스)	20020103 (네트워크구축)
202 (컴퓨터시스템 설계 및 분석가)	20020316 (특수부가통신 서비스)	20010301 (유선설비접속 서비스)	20020315 (부가통신응용계 서비스)
203 (SW개발전문가)	20020306 (이동통신서비스)	20010301 (유선설비접속 서비스)	20020316 (특수부가통신 서비스)
204 (웹전문가)	20010104 (SW제품기획)	20020316 (특수부가통신 서비스)	20010105 (빅데이터 분석)
205 (데이터베이스 설계·구축 및 관리기술자)	20020316 (특수부가통신 서비스)	20020306 (이동통신서비스)	20010301 (유선설비접속 서비스)
206 (통신 및 방송 장비기사 설치 및 수리원)	20020102 (구내통신구축)	20020202 (전송시스템구축)	20020316 (특수부가통신 서비스)

<표 2-23>는 워크넷의 직종코드에 해당 하는 채용공고의 내용과 가장 유사한 NCS 세분류이다. 워크넷 직종코드와 유사한 1순위, 2순위, 3순위의 NCS 세분류 번호이다. 가장 유사한 1순위의 경우 워크넷 직종코드 202번과 205번이 NCS 세분류 20020316(특수부가통신서비스¹⁶⁾)으로 동일하고 나머지 번호들은 각각 1개 워크넷 직종코드에 1개 NCS 세분류가 유사한 것으로 나타났다.

[그림 2-43]은 NCS 세분류별 단어 수이다. 분석에 사용한 단어는 NCS의 기술, 지식 태도에 해당하는 단어를 추출한 것이다. 세분류 코드 20020102에 해당하는 단어 수가 604개로 가장 많으며 20020306과 20010104가 다음으로 많았다. 6개의 번호가 모두 총 500개의 단어 수를 초과하며 평균 544개의 단어가 추출되었다. 또한 [그림 2-44]는 NCS 세분류별 단어로 NCS 번호에 따른 단어들을 보여주고 있다.

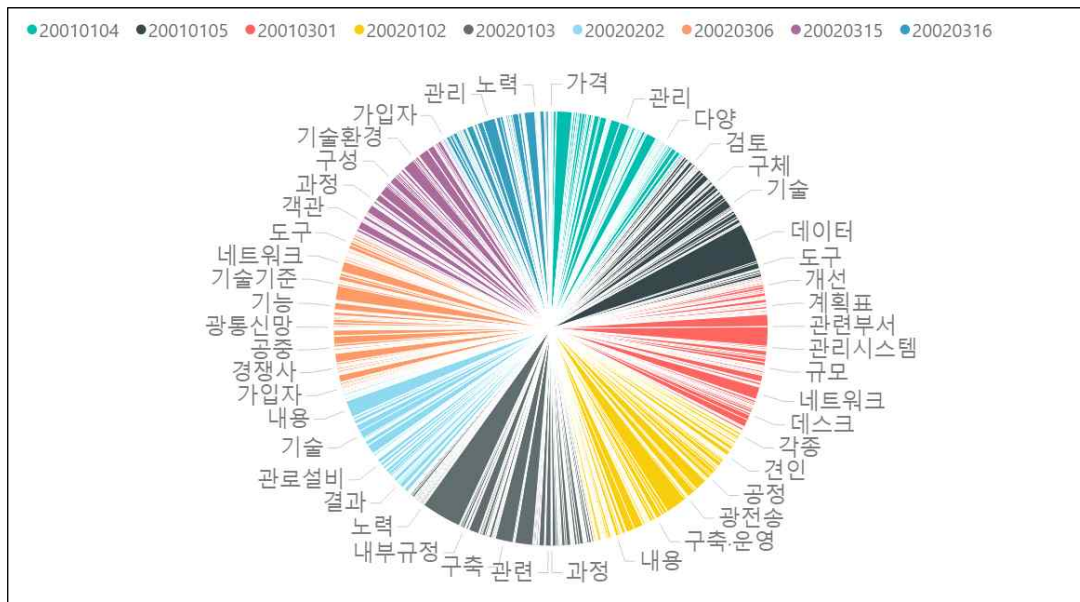
[그림 2-43] NCS 세분류별 단어 개수



16) 특수부가통신서비스 NCS 직무정의,

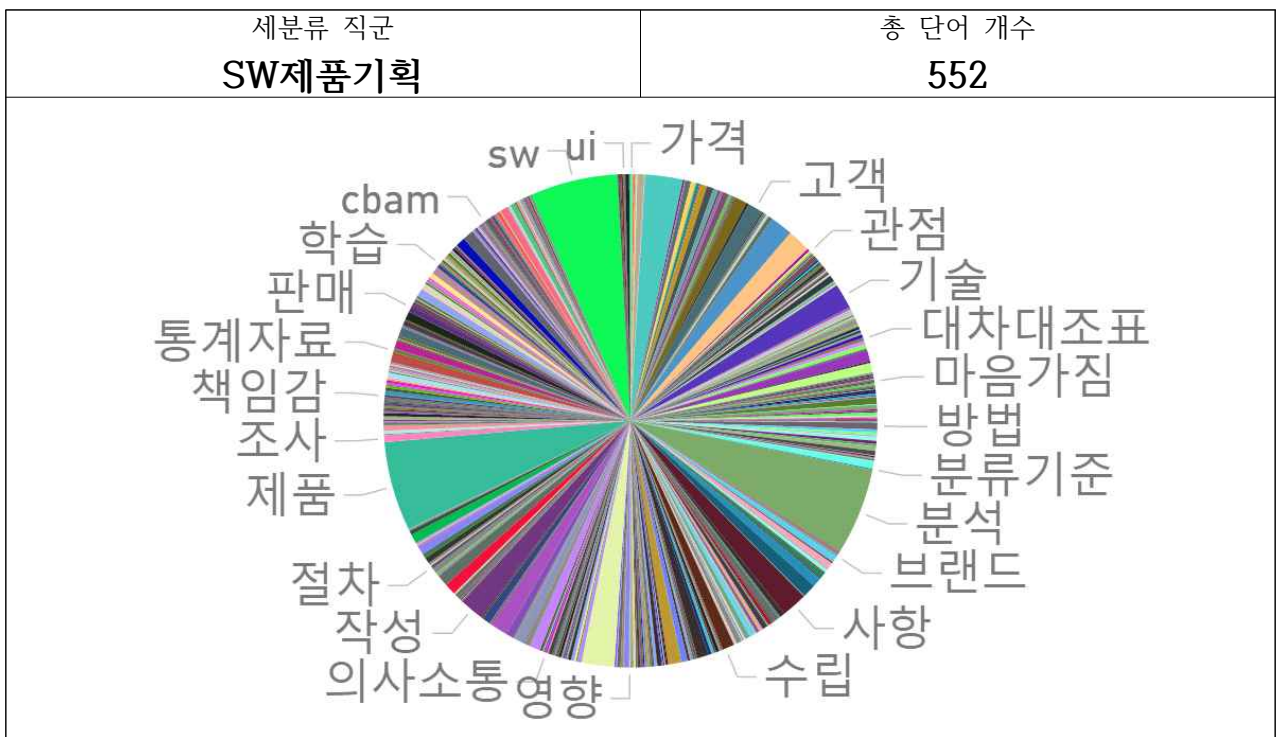
http://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_002&mk=MU000000099&uk=MU000000067

[그림 2-44] NCS 세분류별 단어



1) SW제품기획 (20010104)

[그림 2-45] SW제품기획 (20010104)

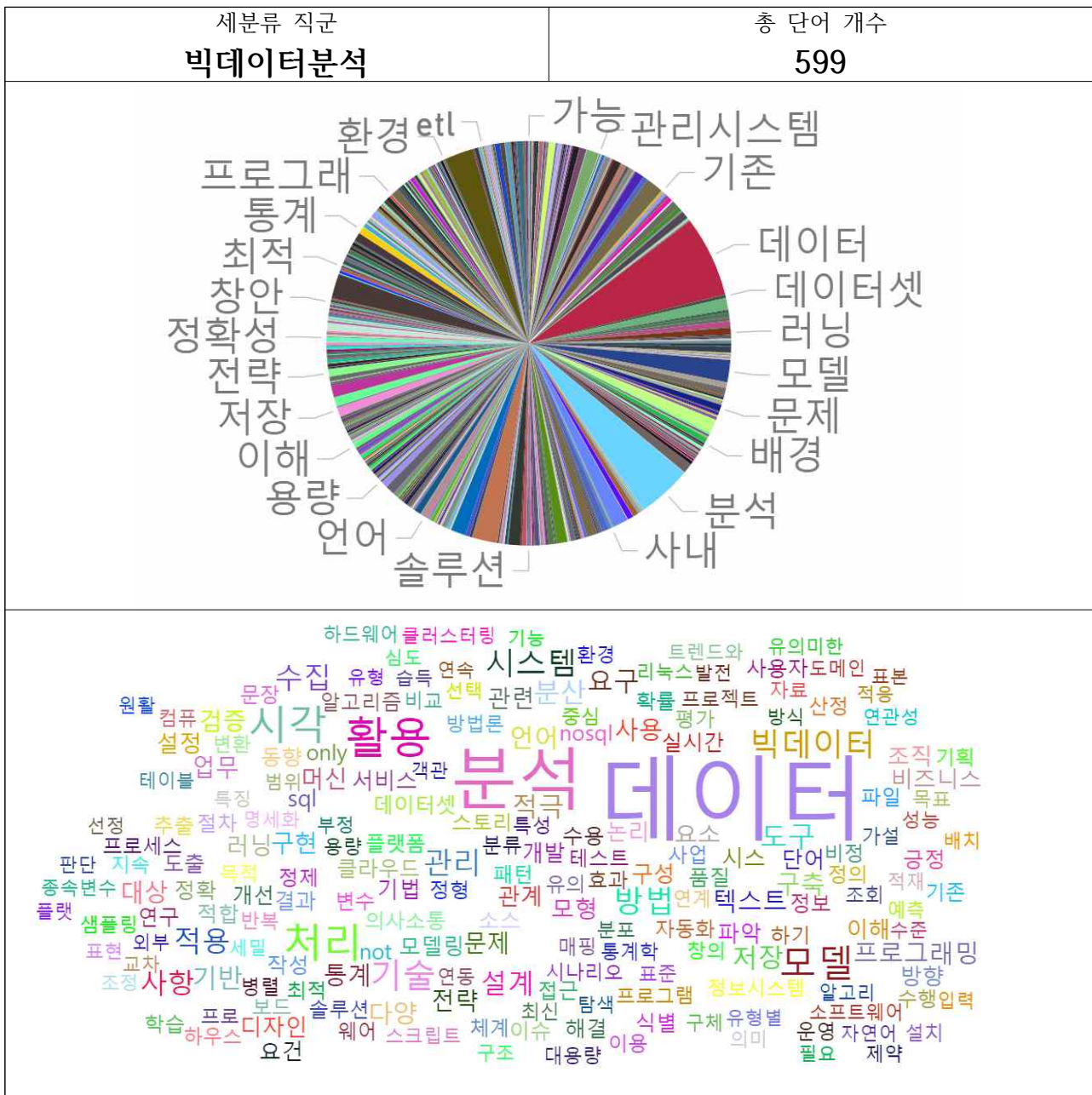




NCS 번호 20010104번에 해당하는 직군은 [그림 2-45]의 SW제품기획 직군이다. 총 552개의 단어가 등장하는데, 가장 많이 등장하는 단어는 ‘제품’이다. 그 뒤로 ‘분석’, ‘SW’, ‘개발’이 빈번하게 등장한다. 워드클라우드 결과와 같이 제품기획과 관련하여 ‘절차’, ‘의사소통’, ‘브랜드’ 등의 단어가 등장하고 SW의 특성이 반영되어 ‘통계자료’라는 단어들이 추가적으로 등장한다.

2) 빅데이터 분석 (20010105)

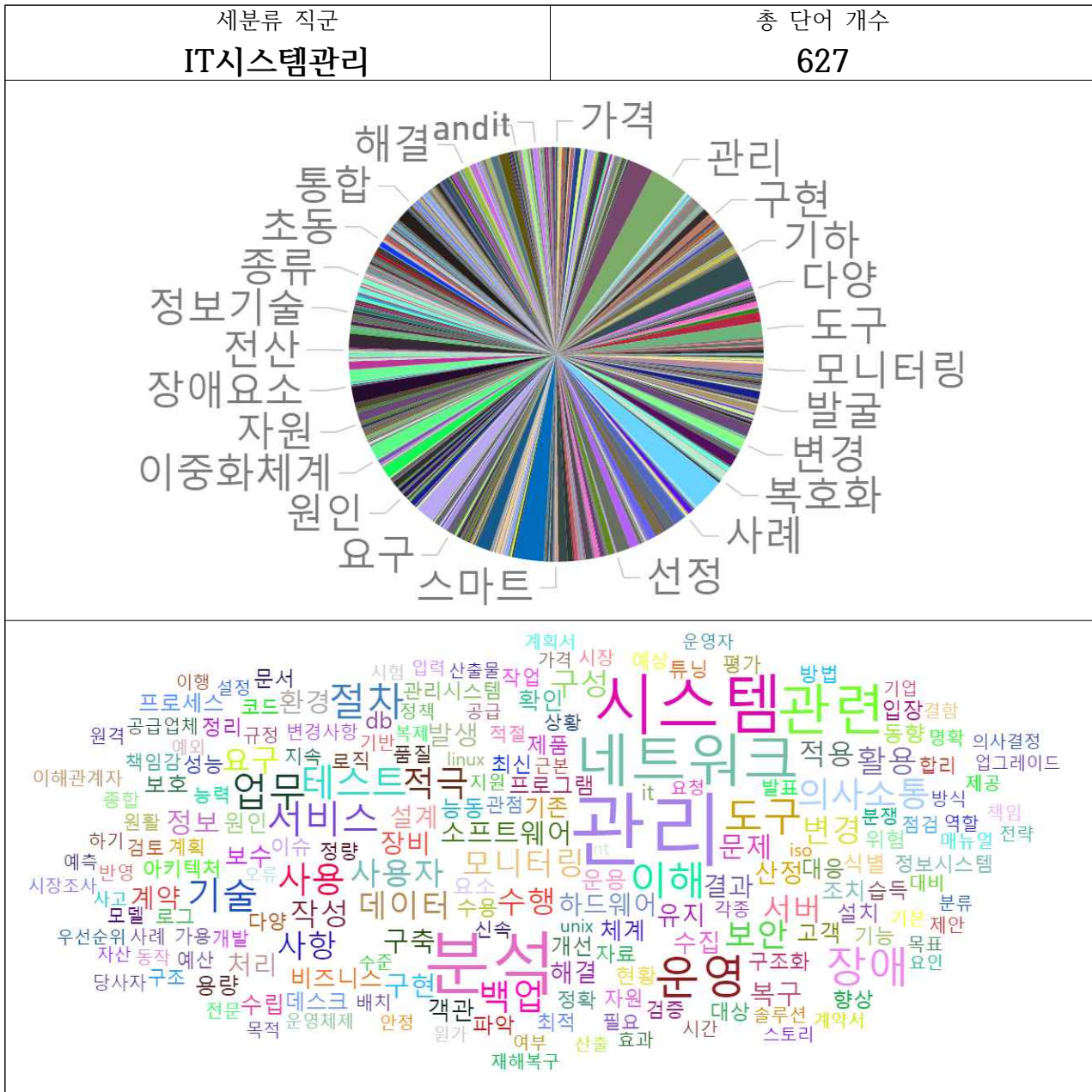
[그림 2-46] 빅데이터 분석 (20010105)



NCS 번호 20010105번에 해당하는 직군은 [그림 2-46]의 빅데이터분석 직군이다. 총 599개의 단어가 등장하는데 가장 많이 등장하는 단어는 ‘데이터’이다. 또한 워드클라우드 결과와 같이 ‘분석’, ‘활용’, ‘모델’이 빈번하게 등장하며 ‘통계’, ‘프로그래밍’ 등 매우 다양한 단어가 등장하는 특성을 보인다.

3) 유선설비접속서비스 (20010301)

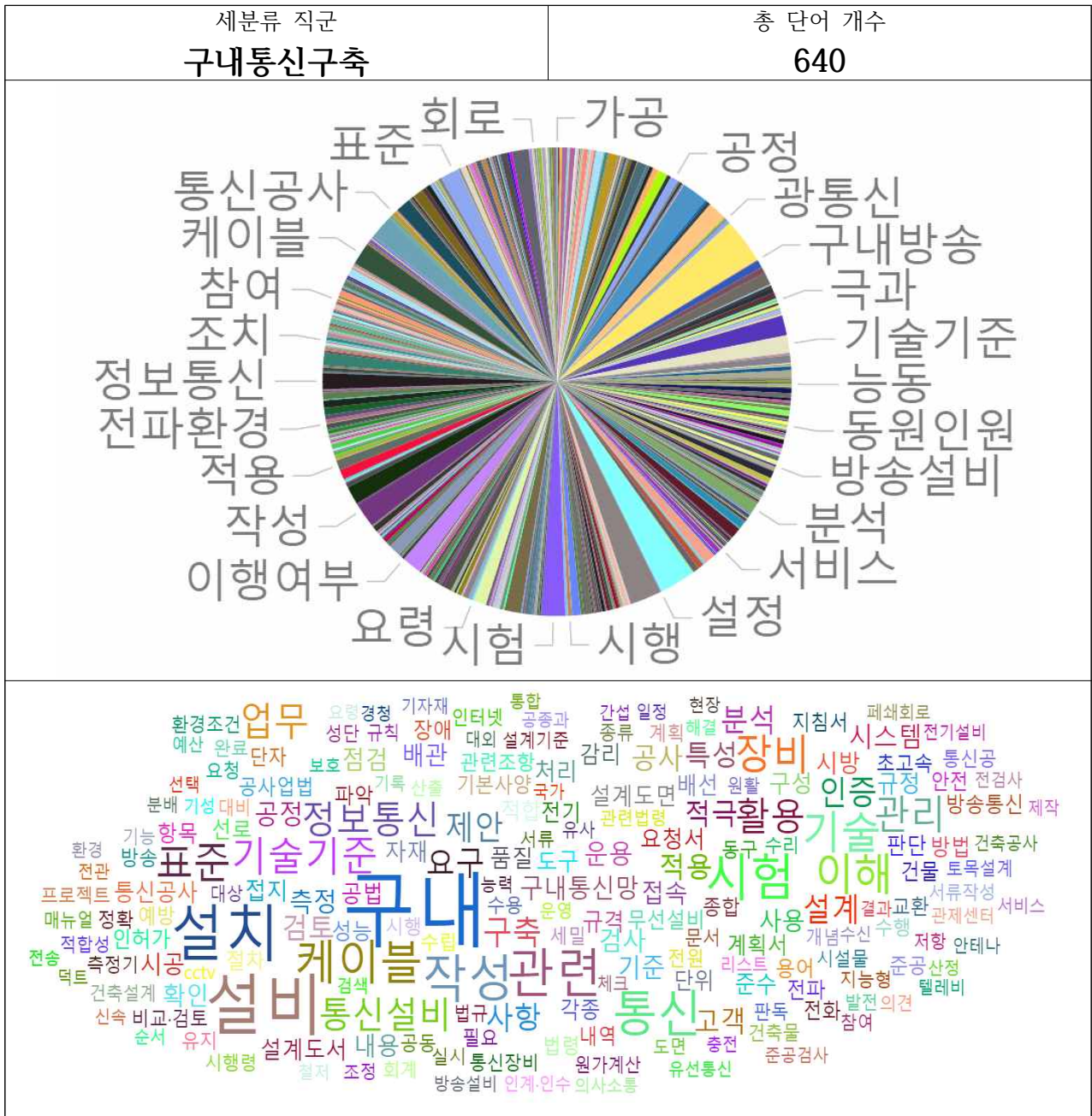
[그림 2-47] 유선설비접속서비스 (20010301)



[그림 2-47]은 NCS 번호 20010301번에 해당하는 IT시스템관리 직군이다. 총 627개의단어가 등장하는데 워드클라우드 결과와 같이 가장 많이 등장하는 단어는 ‘관리’이며 그 외 ‘시스템’, ‘네트워크’, ‘분석’이 빈번히 등장한다.

4) 구내통신구축 (20020102)

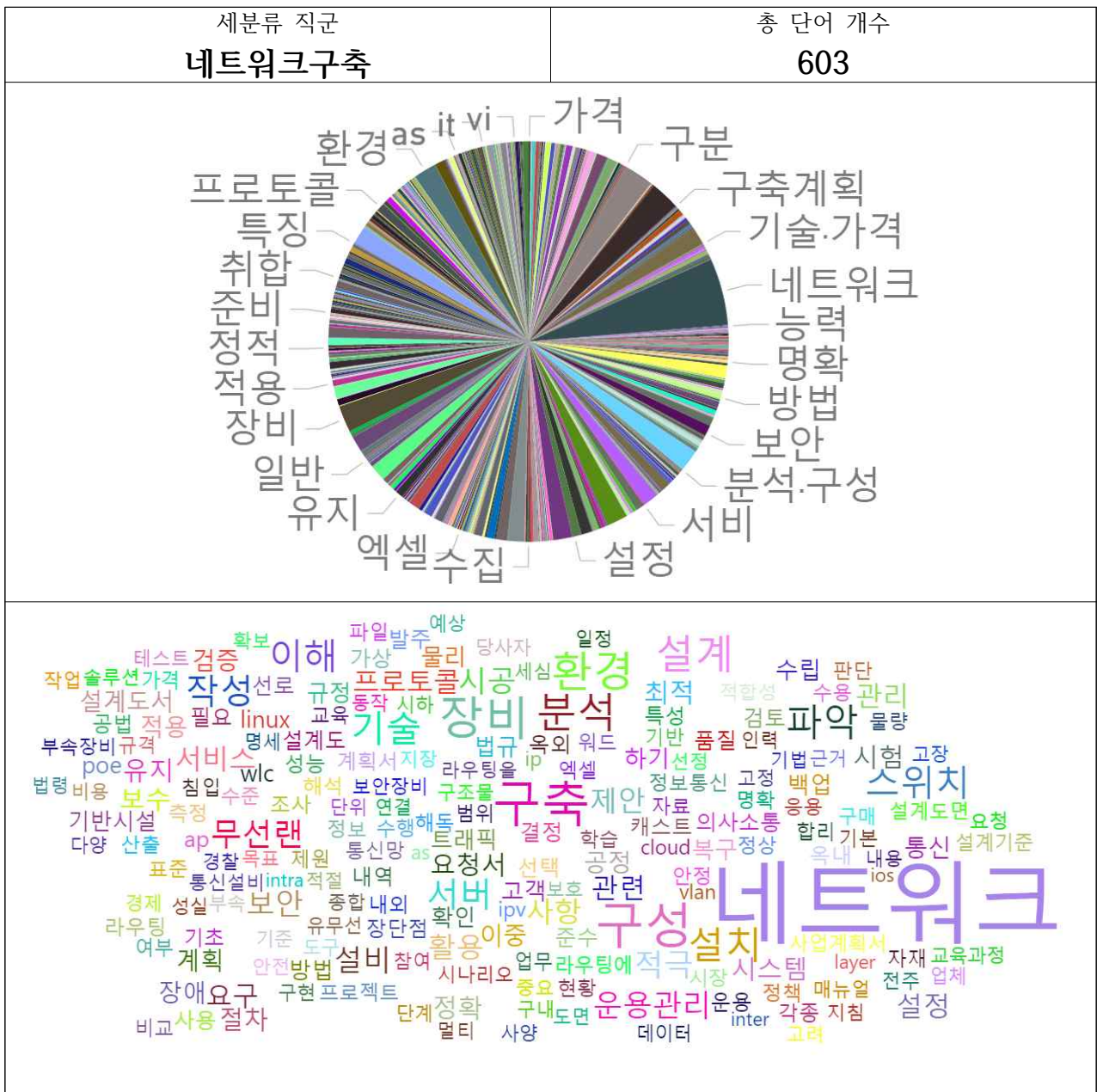
[그림 2-48] 구내통신구축 (20020102)



[그림 2-48]은 NCS 번호 20020102번에 해당하는 구내통신구축 직군이다. 총 640개의 단어가 등장하며 가장 많이 등장하는 단어는 워드클라우드 결과 ‘구내’, ‘설비’, ‘설치’, ‘케이블’ 이 등장하여 통신구축과 관련하여 ‘정보통신’, ‘전파환경’, ‘통신공사’, ‘광통신’ 등의 단어가 등장한다.

5) 네트워크구축 (20020103)

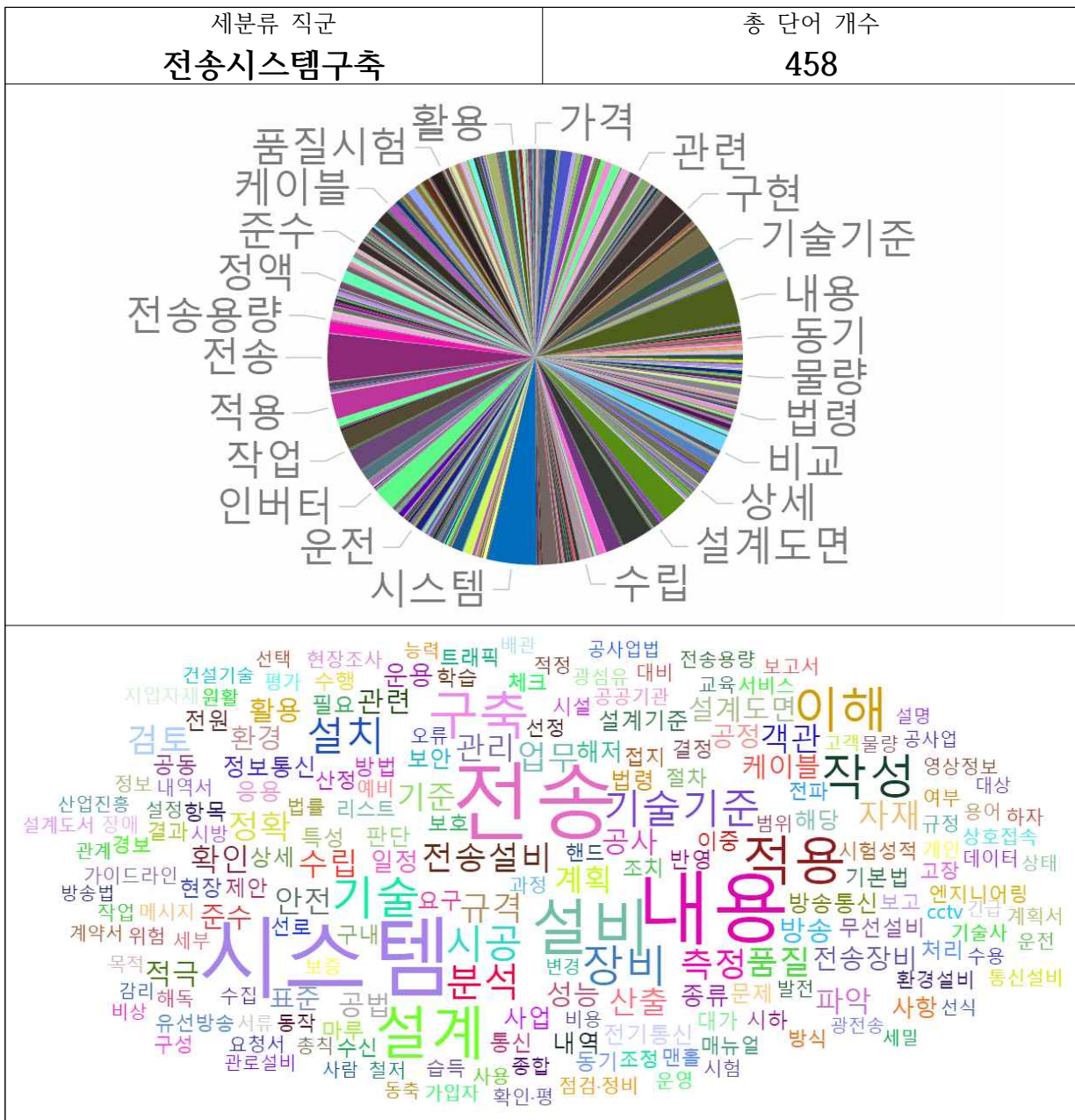
[그림 2-49] 네트워크구축 (20020103)



[그림 2-49]는 NCS 번호 20020103번에 해당하는 네트워크구축 직군으로 총 603개의단어가 등장한다. 위드클라우드 결과 가장 많이 등장하는 단어는 ‘네트워크’이며 ‘구축’, ‘장비’, ‘분석’ 이 차례로 많이 등장한다. 네트워크와 관련하여 ‘서버’, ‘보완’, ‘프로토콜’ 등의 단어가 등장한다.

6) 전송시스템구축 (20020202)

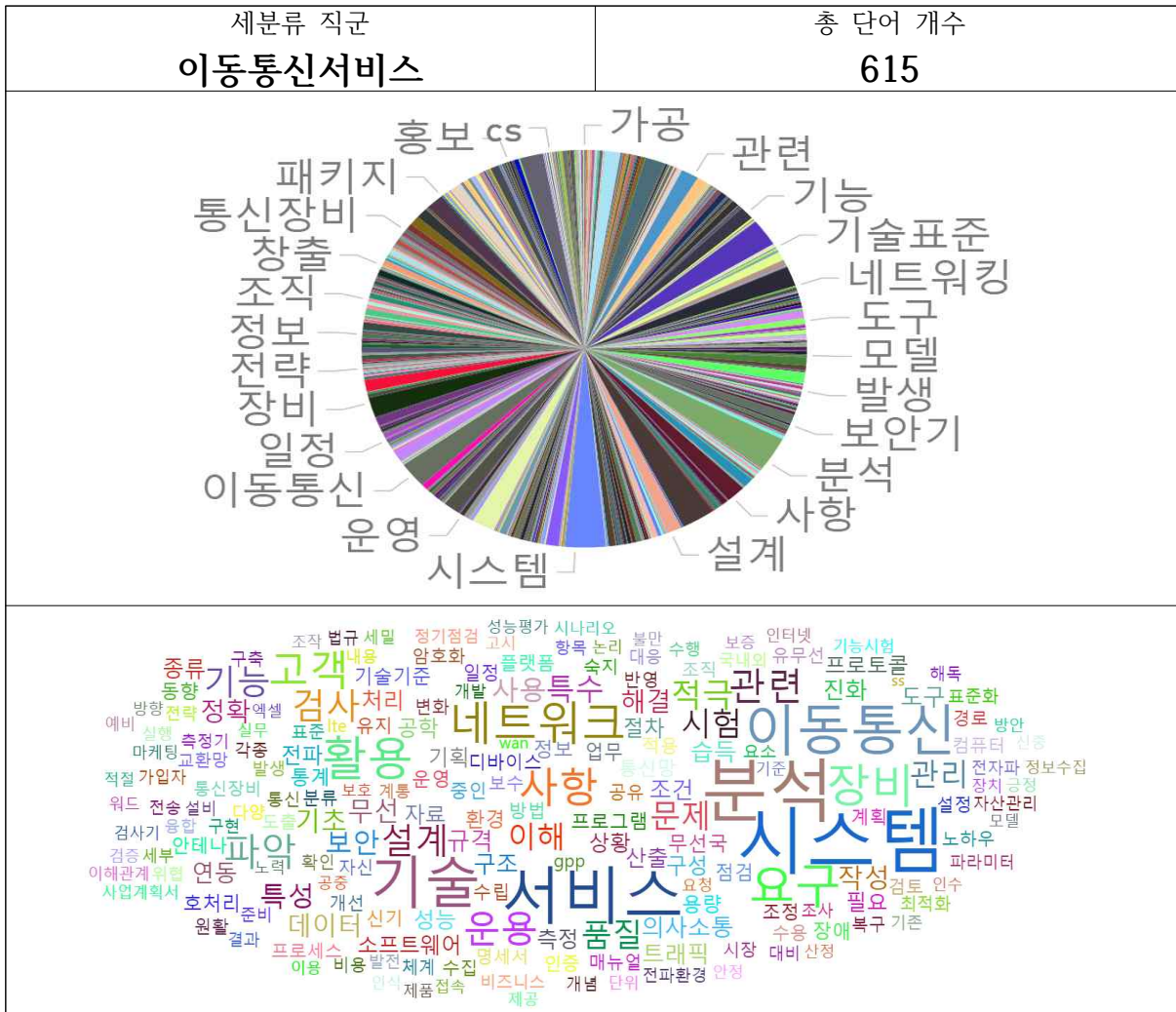
[그림 2-50] 전송시스템구축 (20020202)



[그림 2-50]은 NCS 번호 20020202번에 해당하는 전송시스템구축 직군이다. 총 458개의 단어가 등장하는데, 가장 많이 등장하는 단어는 ‘시스템’이며 워드클라우드 결과와 같이 ‘전송’, ‘내용’, ‘적용’이 빈번하게 등장한다.

7) 이동통신서비스 (20020306)

[그림 2-51] 이동통신서비스 (20020306)

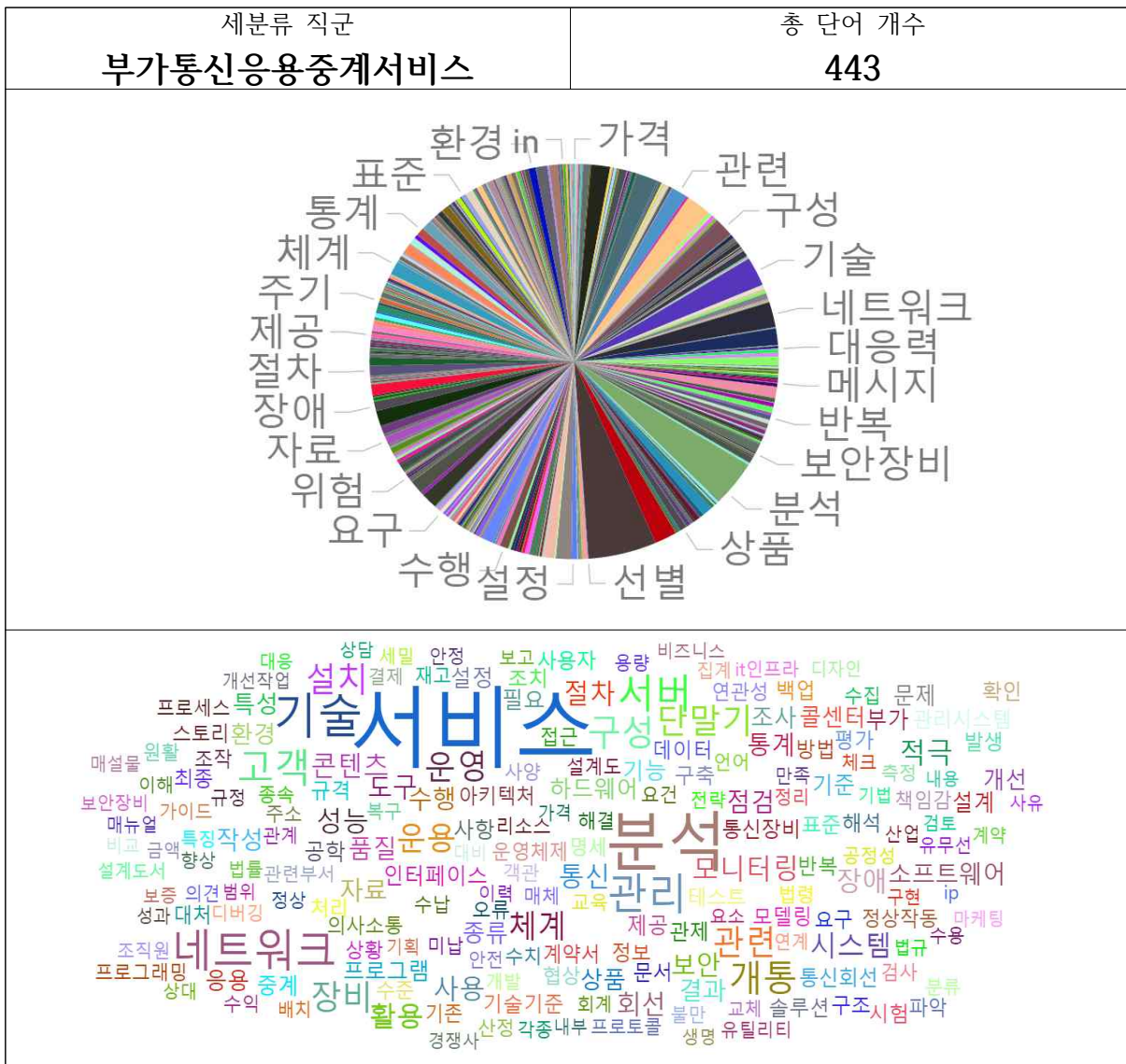


[그림 2-51]은 NCS 번호 20020306번에 해당하는 이동통신서비스 직군¹⁷⁾이다. 총 615개의 단어가 등장하는데, 가장 많이 등장하는 단어는 ‘시스템’이다. 그 뒤로 ‘분석’, ‘이동통신’, ‘서비스’, ‘기술’이 빈번하게 등장한다. 서비스와 관련하여 ‘홍보’, ‘CS’라는 단어들이 등장하고 이동통신과 관련하여 ‘네트워크’, ‘통신장비’라는 단어들이 등장한다.

17) 이동통신서비스 NCS직무정의,
http://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_002&mk=MU00000099&uk=MU00000067

8) 부가통신응용중계서비스 (20020315)

[그림 2-52] 부가통신응용중계서비스 (20020315)



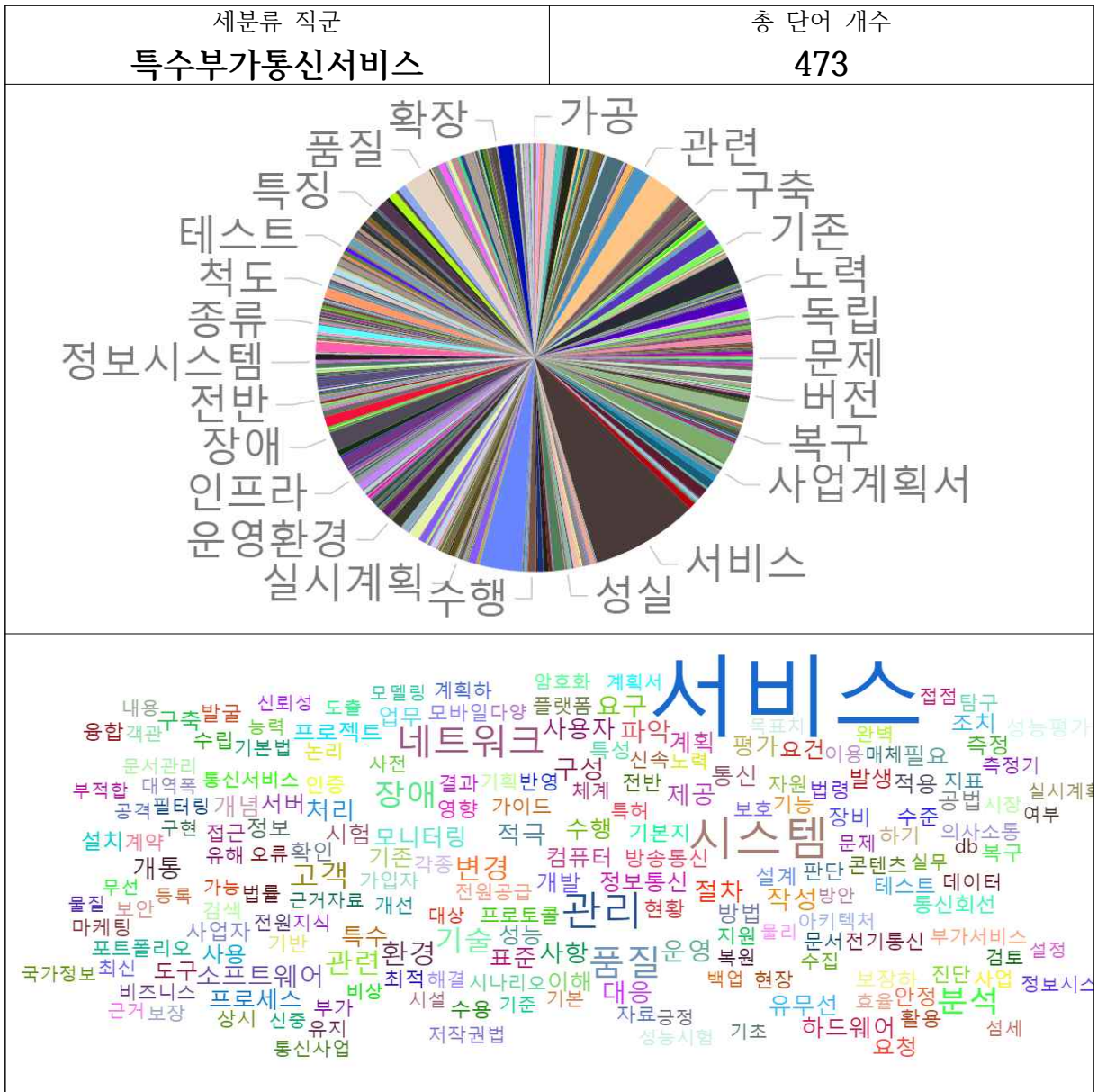
[그림 2-52]는 NCS 번호 20020315번에 해당하는 부가통신응용중계서비스 직군¹⁸⁾이다. 총 443개의 단어가 등장하는데, 가장 많이 등장하는 단어는 ‘서비스’이다. 그 뒤로 ‘분석’, ‘네트워크’, ‘운용’, ‘개통’이 빈번하게 등장한다.

18) 부가통신응용중계서비스 NCS 직무정의,

http://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_002&mk=MU00000099&uk=MU00000067

9) 특수부가통신서비스 (20020316)

[그림 2-53] 특수부가통신서비스 (20020316)



NCS 번호 20020316번에 해당하는 직군은 특수부가 통신 직군이다. 총 473개의 단어가 등장하는데, 가장 많이 등장하는 단어는 ‘서비스’이다. 그 뒤로 ‘시스템’, ‘네트워크’, ‘관리’가 빈번하게 등장한다. 통신서비스와 관련하여 ‘품질’, ‘인프라’라는 단어들이 등장한다.

제5절 분류체계 간의 기술역량 일치도 검토

1. 워크넷 직업 분류와 NCS 직무 분류 체계의 일치성 비교

기술역량 일치성 비교는 기업들이 원하는 인력이 가져야 하는 기술과 역량과 NCS 세부분류에 표기되어 있는 기술과 역량이 얼마나 일치하는 지를 확인하는 분석이다. 워크넷에서 수집된 채용공고로부터 각 기업들이 필요로 하는 인력이 가져야 하는 기술과 역량을 파악하고, 워크넷 직업분류와 NCS의 세부분류에 대한 기술, 태도, 지식에 대한 기술과 유사도를 비교한다.

NCS 세부분류와 워크넷 채용공고 모두 필요한 역량들을 문장으로 표현하고 있다. 문장의 구성성분은 단어들로 구성되어 있는 바, 단어들의 구성과 빈도의 차이를 확인하면 실제 현업에서 필요로 하는 역량과 NCS에서 정의하고 있는 역량 사이에 차이가 있는지를 확인할 수 있다.

본 절의 목표는 NCS 분류체계와 워크넷 채용공고간의 단어 구성 차이를 확인하여 유사도를 측정하고자 한다. 기업이 필요로 하는 인력이 가져야 하는 역량과 NCS에서 직종별로 가져야 하는 역량에 대한 기술이 차이가 없다면 현행 NCS 분류체계가 적절하게 구성되어 있음을 확인하는 분석이 될 것이다.

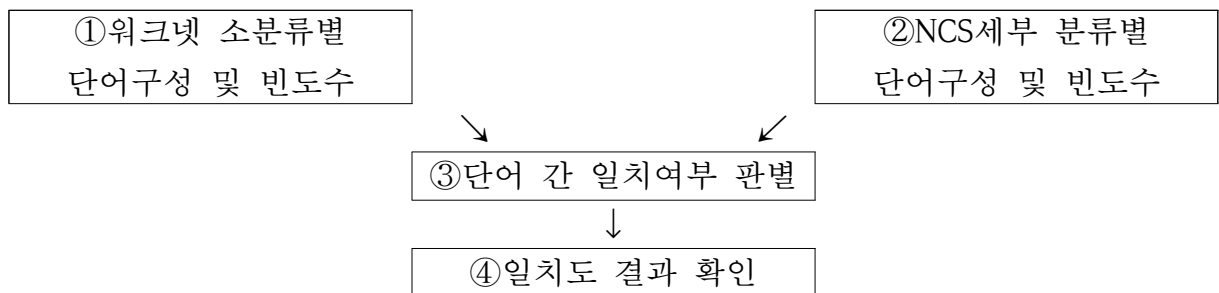
하지만 두 분류체계의 차이가 있을 경우 워크넷에서의 분류체계와 NCS 분류체계에 대한 SW 산업의 적합성을 확인하기 어렵다는 점을 판단할 수 있다. 따라서 분석 결과에 따라 새로운 정의와 분류체계 재구성의 필요성을 확인하는 방법론과 개선점이 요구된다.

연구 목적에 따라 데이터 분석을 위한 전처리과정은 앞서 3절에 기술하였다. 기술된 전처리 과정을 거친 후에 채용공고 단어의 구성과 빈도수를 확인하였다. 이 과정을 통하여 NCS 분류체계의 기술·역량·태도에 표현된 문장의 단어 구성과 빈도수를 확인하여 채용공고에서 파악된 단어의 구성과 빈도와 비교하여 일치도를 계산하였다. 채용공고에는 직종코드가 포함되어 있어 어떤 분야의 채용공고인지를 파악할 수 있으며, 가장 유사한 NCS 세부분류를 파악하여 채용공고의 직종과 유사한지를 판단하게 된다.

2. 분류체계 비교를 위한 데이터 선택 및 분석 과정

워크넷 직업분류별 모집내용과 NCS 분류기준에 따른 세분류와 비교를 위해 수행한 과정은 [그림 2-54]와 같이 진행 되었다. 먼저 워크넷과 NCS의 분류 및 단어 구성을 파악한 후 유사 직업 별 단어 일치여부를 확인하여 일치도를 비교 하였다.

[그림 2-54] 소분류별 외국어조건 기준 채용공고



1) 워크넷 소분류별 단어구성 및 빈도수

<표 2-24> 소분류별 상위 추출 단어

번호	소분류 이름	상위 6개 추출 단어					
201	컴퓨터 및 통신공학 기술자 및 연구원	정보통신	관리	엔지니어	기술	유지	개발
202	컴퓨터시스템 설계 및 분석가	네트워크	엔지니어	보안	서버	관리	시스템
203	SW개발전문가	개발	JAVA	프로그램	SW	시스템	안드로이드
204	웹전문가	웹개발	개발	웹 프로그램	JAVA	PHP	프로그램
205	데이터베이스 설계·구축 및 관리기술자	관리	서버	운영	유지	네트워크	전산
206	통신 및 방송 장비기사 설치 및 수리원	설치	CCTV	인터넷	정보통신	유지	통신

워크넷 모집직종의 분류인 한국고용직업분류표¹⁹⁾를 기준으로 따라 소분류 201에서 206 에 해당하는 것을 각각 DB화 하였다. 이후 소분류 번호에 해당하는 워크넷 직종별 채용공고에서 채용공고 제목과 채용공고의 직무내용, 채용공고에 포함된 단어 등에 대해 단어를 추출하였고, 해당 단어의 빈도수를 확인하였다.

그에 따라 <표 2-24>는 각 소분류 별 단어 카운트 자료의 상위 6개 단어를 추출한 것이다. 소분류 201에서는 총 722개가 단어가 추출되었고, 소분류 202는 696개, 소분류 203은 2,216개의 단어가 추출되었다. 소분류 204는 1,398개의 단어가 추출되었고, 소분류 205는 807개의 단어가, 소분류 206은 1,502개의 단어가 추출되었다. 각 소분류별 단어 수와 단어의 총 단어 빈도수는 아래의 <표 2-25>과 같다.

<표 2-25> 소분류별 단어수와 빈도수

소분류	소분류이름	단어수	빈도수
201	컴퓨터 및 통신공학 기술자 및 연구원	722	1,270
202	컴퓨터시스템 설계 및 분석가	696	1,647
203	SW개발전문가	2,216	7,445
204	웹전문가	1,398	4,611
205	데이터베이스 설계·구축 및 관리기술자	807	1,736
206	통신 및 방송 장비기사 설치 및 수리원	1,502	4,856

19) 통계청, 고용직업분류, KECO, https://kssc.kostat.go.kr:8443/ksscNew_web/index.jsp#

2) NCS 세부분류별 단어구성 및 빈도수

NCS 직종별 분류는 대분류번호 20, 대분류이름 정보통신부분에 해당하는 NCS 각 세부분류 58개의 직종별 설명이 기술된 지식, 기술, 태도로 부터 단어를 추출하였고, 추출된 각 단어들에 대한 빈도수를 확인하였다. NCS 분류의 지식, 기술, 태도를 이용한 이유는 각 세부분류 별로 DB화가 되어있으며, 워크넷 모집공고에서의 직종별 필요역량이 NCS 분류의 지식, 기술, 태도와 가장 유사하다고 판단하였기 때문이다. NCS 세부 분류의 지식, 기술, 태도의 워드카운트 결과의 일부는 아래 <표 2-26>과 같다.

<표 2-26> NCS 분류별 단어수와 빈도수

NCS 분류번호	NCS 세분류 이름	단어수	빈도수
20010101	정보기술전략	409	986
20010102	정보기술컨설팅	442	1,814
20010103	정보기술기획	489	1,641
20010104	SW제품기획	552	2,433
20010105	빅데이터 분석	599	2,388
20010201	SW아키텍처	448	2,051
20010202	응용SW엔지니어링	463	1,878
20010203	임베디드SW엔지니어링	275	754
20010204	DB엔지니어링	376	1,465
20010205	NW엔지니어링	494	1,857

3) 단어 간 일치여부 판별

상기 분류별 단어수와 빈도수를 이용하여, 워크넷 소분류와 NCS 세분류의 비교를 진행 한다. 주요 측정 방법은 자카드 유사성(Jaccard Similarity) 측정을 이용한 방법으로 유사성을 측정하였고, 자세한 내용은 3. 계산방법에 기술되어있다.

4) 일치도 결과 확인

단어 간 일치여부를 판별하였으면 가장 높은 값을 나타내는 것을 하나씩 추출하여, 워크넷 소분류 명칭과 채용공고 등의 내용을 확인하고, NCS 세분류의 직무정의와 비교하여 두 집단이 정말 유사한지를 확인한다. 자세한 내용은 4. 결과비교에 기술되어 있다.

3. 계산방법

자카드 유사성 측정²⁰⁾은 그림 5-2와 같이 두 집단 A, B의 교집합 원소 수를 분자로 하고 두 집단의 합집합의 원소 수를 분모로 하여 두 집합 간 유사성을 수치화하는 방법이다. 자카드 유사성 측정 공식은 아래와 같다. 값의 범위는 0에서 1사이의 값을 가진다. 교집합이 없을 때 0, 두 집합이 동일할 때 1의 값을 가진다. 즉, 두 집단이 유사할수록 1에 가까운 값을 나타내게 되며, 0에 가까울수록 두 집단은 유사하지 않다고 볼 수 있다.

$$J(A,B)=\frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A|+|B|-|A \cap B|}$$

20) Jaccard index, https://en.wikipedia.org/wiki/Jaccard_index

아래의 <표 2-27>의 식은 자카드 유사성 측정을 응용하여 이번 보고서에 적용한 수식이다.

<표 2-27> 두 집합 간 유사성 측정식

$$C_{ij} = \frac{E_{ij}}{W_i + N_j - E_{ij}} \times 100$$

C_{ij} : 워크넷 소분류 i와 NCS 세부분류 j간의 일치도

E_{ij} : 워크넷의 소분류와 NCS 분류의 동일한 단어에 대한 빈도수
(빈도수는 워크넷 소분류와 NCS 분류 중 큰 것으로 결정한다.)

W_i : 워크넷 소분류에 해당하는 단어의 총 빈도수.

i는 소분류 번호에 해당한다.

N_j : NCS 분류에 해당하는 단어의 총 빈도수

j는 NCS 분류 번호에 해당한다.

특정 워크넷 소분류 i (예: 201 컴퓨터 및 통신공학 기술자 및 연구원)의 채용 공고에 나타난 단어의 빈도수를 모두 더한 값을 W_i 로 하고, 이와 비교하는 NCS 세부 분류의 기술, 지식, 태도에 나타난 단어의 빈도수를 모두 더한 값을 N_j 로 한다. E_{ij} 는 W_i 와 N_j 의 단어를 비교하여 동일한 단어를 분류하고, 확인된 동일한 단어의 최대 빈도수의 합을 E_{ij} 로 한다. 최대 빈도수는 W_i 나 N_j 에서의 최대 빈도수를 의미한다. 분모($W_i + N_j - E_{ij}$)는 워크넷 소분류와 NCS 세부분류에 등장하는 단어의 합집합을 의미하며, 분자(E_{ij})는 워크넷 소분류와 NCS 세부분류에 등장하는 단어의 교집합을 의미한다.

본 보고서에서 단어의 빈도수를 고려한 이유는 중요 기술일수록 그리고 현업에서 많이 요구하는 기술일수록 공고에서의 언급 빈도가 높기 때문이다. 이를 반영하기 위해 빈도수를 적용하였다. <표 2-27>에 표기된 식은 0에서 100의 값을 가지며, 100에 가까울수록 유사한 집단으로 판단하게 된다.

4. 결과비교

〈표 2-28〉 일치도 결과 비교

워크넷 번호	분류명	NCS 번호	세분류명	일치도 (%)
201	컴퓨터 및 통신공학 기술자 및 연구원	20020315	부가통신응용중계서비스	49.8
202	컴퓨터시스템 설계 및 분석가	20020316	특수부가통신서비스	49.3
203	SW개발전문가	20020306	이동통신서비스	45.7
204	웹전문가	20010104	SW제품기획	41.3
205	데이터베이스 설계·구축 및 관리기술자	20020316	특수부가통신서비스	44.0
206	통신 및 방송 장비기사 설치 및 수리원	20020102	구내통신구축	48.4

워크넷 소분류와 NCS 비교 결과를 정리하여 일치도 결과를 판단한 결과는 다음과 같다. 〈표 2-28〉과 같이 워크넷 소분류 201 컴퓨터 및 통신공학 기술자 및 연구원 직군과 가장 일치하는 NCS 분류체계는 20020315는 부가통신응용중계서비스로 나타났으며, 일치도는 약49.8%에 해당하였고 NCS와 워크넷의 분류체계는 일치도가 낮은 것으로 판단된다.

그에 따라 분석 결과에 대한 워크넷과 NCS의 일치하는 단어의 빈도수 상위 10개 단어는 〈표2-29〉와 같다. 이를 통해 확인해보면 서비스가 가장 높은 빈도로 제시되고 있으나 워크넷 201 분류와는 다른 결과임을 확인할 수 있다.

〈표 2-29〉 201 분류와 20020315분류의 단어와 빈도수

단어	서비스	분석	기술	네트워크	관리	서버	고객	개통	구성	단말기
빈도수	105	71	46	40	38	34	33	29	28	26

워크넷 소분류 202 (컴퓨터시스템설계 및 분석가) 직군과 가장 일치하는 NCS 분류체계는 20020316 (특수부가통신서비스)로 나타났으며, 일치도는 약 49.3%였다. 이 결과 역시 50%를 넘지 못하여 두 체계의 매우 유사성이 낮은 것으로 나타났다. 그에 따라 두 체계의 일치하는 단어의 빈도수 상위 10개의 단어는 아래 <표 2-30>과 같다. 이 결과는 네트워크 및 시스템이 주요 일치 단어로 제시되고 있으나 202 분류의 정의를 고려할 때 시스템 설계 분석가에 대하여 큰 상관성은 나타나지 않는다.

<표 2-30> 202분류와 20020316분류의 단어와 빈도수

단어	네트워크	보안	서버	관리	시스템	유지	업무	관련	구축	보수
빈도수	65	34	28	48	61	23	22	21	17	17

워크넷 소분류 203 (SW개발전문자) 직군과 가장 일치하는 NCS분류체계는 20020306 (이동통신서비스)로 나타났으며, 일치도는 약 45.7%에 해당하였다. 마찬가지로 두 체계의 일치하는 단어의 빈도수 상위 10개의 단어는 아래 <표 2-31>와 같다. 분석 결과에 주요 단어는 ‘개발’ 과 ‘프로그램’, ‘SW’ 로 나타나고 있어 203 분류의 정의에 적합한 것으로 보이지만 NCS 20020306의 ‘이동통신서비스’ 와 관련하여 매우 광범위한 정의가 보임에 따라 역시 명확한 직무 체계를 파악하기 어렵다.

<표 2-31> 203분류와 20020306분류의 단어와 빈도수

단어	개발	프로그램	SW	시스템	관련	임베디드	가능	응용	업무	기반	관리
빈도수	439	207	131	80	66	64	61	50	49	37	36

워크넷 소분류 204(웹전문가) 직군과 가장 일치하는 NCS분류체계는 20010104 (SW제품기획)로 나타났으며, 일치도는 약 41.3%에 해당하였다. 일치하는 단어의 빈도수 상위 10개의 단어는 아래 <표 2-32>와 같다. 분석 결과 웹전문가에 대하여 웹개발 및 개발과 같은 역량이 나타나고 있지만 웹전문가와 SW 제품기획이라는 분류 체계는 NCS분류가 더 광의의 개념을 가지고 있고 개념상으로는 명확히 같은 분류로 판단하기 어렵다.

〈표 2-32〉 204분류와 20010104분류의 단어와 빈도수

단어	웹개발	개발	프로그램	가능	관리	업무	유지	운영	관련	보수	시스템
빈도수	160	157	67	63	63	45	40	34	36	25	22

워크넷 소분류 205 (데이터베이스 설계·구축 및 관리기술자) 직군과 가장 일치하는 NCS분류체계는 20020316 (특수부가통신서비스)로 나타났으며, 일치도는 약 44%에 해당하였다. 두 체계의 일치하는 단어의 빈도수 상위 10개의 단어는 아래 〈표 2-33〉과 같다. 그러나 개념적으로 데이터베이스 설계와 특수부가통신서비스는 정의상 일치하지 못하고 NCS의 경우 좀 더 광의의 개념으로 나타나고 있다.

〈표 2-33〉 205분류와 20020316분류의 단어와 빈도수

단어	관리	서버	운영	유지	네트워크	시스템	업무	보수	관련	개발	구축
빈도수	48	35	30	30	40	61	25	21	21	17	12

워크넷 소분류 206 (통신 및 방송 장비기사 설치 및 수리원) 직군과 일치하는 NCS분류체계는 20020102 (구내통신구축)로 나타났으며, 일치도는 약 48.4%에 해당하였다. 두 직무상의 일치하는 단어의 빈도수 상위 10개의 단어는 아래 〈표 2-34〉과 같다.

〈표 2-34〉 206분류와 20020102분류의 단어와 빈도수

단어	설치	인터넷	정보통신	유지	통신	업무	보수	관련	공사	관리	수리
빈도수	200	87	87	81	81	77	62	57	55	52	51

5. 워크넷 분류와 NCS분류 체계 비교에 따른 시사점

워크넷 분류와 NCS 직무 분류 체계의 일치도 결과를 워크넷 소분류와 NCS 분류체계 세분류를 단어를 기준으로 유사도를 계산하였을 때, 하드웨어와 같은 장비를 다루는 직종(201, 206)은 워크넷 채용공고에 기술된 내용과 NCS 세부분류사이를 포함하여 전반적으로는 두 직무체계의 유사도가 높지 않다. 특히 SW 직종(202, 203, 204, 205)에 대해서는 명확히 일치하지 않는 점을 확인할 수 있었다. 이 결과는 현재의 워크넷 및 NCS 분류체계가 생각보다 관련 산업에 대한 직무에 대한 정확한 기술이 어렵다는 점을 보여준다. 특히 두 체계별로 직무에 대한 정의나 의미가 약간씩 다른 점이 있기 때문에 SW 분야에 대한 직무 체계를 정확히 구분하기에는 보완할 측면이 존재한다.

특히 워크넷 직종코드 202, 203, 204, 205에 해당하는 SW 관련 직종이 NCS 직종에 대한 기술내용과 상이한 이유는 다음과 같다. 첫째, 워크넷 소분류 202 컴퓨터시스템설계와 NCS 20020316 특수부가통신서비스의 경우는 단어 간 일치도가 높게 나온 이유는 컴퓨터시스템설계(202) 직종에 네트워크 설계 및 보안 관련 직종이 포함되어 있기 때문이다. 특수부가통신서비스는 스마트 IT 환경에서 기간 통신사업자의 회선설비 및 서비스를 근간으로 특수한 서비스를 생성 및 파생 또는 융·복합하여 고객의 다양한 요구를 수행하는 형태의 일이다. 특수부가통신서비스의 대표적인 예는 이러닝(e-Learning) 서비스이다. 특수부가통신서비스를 위해서는 네트워크 관련 지식과 기술이 필요하지만, 소분류 202와 같이 컴퓨터시스템 간 네트워크 및 보안에 더 전문화된 지식 기술과는 상이하므로 연관성이 없다고 판단할 수 있다.

둘째, 워크넷 소분류 203 SW개발전문가와 NCS 20020306 이동통신서비스는 워드 카운트된 단어 간 일치도는 높으나 실제 각 직군 간 연관성에 큰 차이가 있음을 볼 수 있다. SW개발전문가(203)의 경우 시스템SW와 응용SW개발자와 같은 직종이 주를 이루고 있으나, NCS 20020306 이동통신서비스의 경우 이동통신 주파수 대역의 전파를 전송 매체로 이용하여 음성, 문자, 데이터, 영상 등 모든 종류의 정보를 송·수신하는 서비스를 제공하기 위하여 기획, 설계, 검사, 시험, 운용, 최적화, 유지보수, 관리를 수행하는 직종이다. 따라서 서로 연관성이 전혀 없는 독립된 직종이다.

셋째, 워크넷 소분류 204 웹전문가 NCS 20010104 SW제품기획의 경우는, 웹전문가(204)의 하부 분류에 웹 및 멀티미디어 직종과 기획자 직종이 있어 일치도가 높게 평가된 것으로 보인다. 워크넷 소분류 204는 웹, 멀티미디어 개발자 및 기획자가 해당하는 직군이다. NCS에서 SW제품기획자는 개발자 직군이기보다는 기업의 경제적 활동을 위하여 SW 기반의 기업 내/외부 환경과 기술, 시장성 등을 분석한다. 따라서 SW기반 제품의 사양 및 개발, 지원, 판매 계획을 수립하는 직무를 가지고 있으나 직무의 일부가 유사한 것과 달리 전반적으로는 차이가 있다.

넷째, 워크넷 소분류 205 데이터베이스 설계·구축 및 관리기술자와 NCS 20020316 특수부가통신서비스의 경우는, 단어 일치도 결과와 달리 워크넷과 NCS 직군이 일치하지 않음을 알 수 있다. 단어의 결과가 일치하게 나온 이유는, 두 직군 모두 네트워크 및 서비스의 기술과 지식이 필요하기 때문이다. 하지만 데이터베이스와 통신서비스에 대해 필요한 네트워크 기술과 지식이 상이하므로 두 직군은 서로 연관성이 없는 상이한 직군이라 볼 수 있다.

다음으로 워크넷의 채용공고 내용과 NCS 직종 기술이 서로 유사한 의미를 가지는 소분류 201, 206에 대해 확인한 결과는 다음과 같다. 첫째, 워크넷 소분류 201 컴퓨터 및 통신공학 기술자 및 연구원과 NCS 20020315 부가통신응용중계서비스의 경우, 단어의 일치도가 높지 않지만 다른 분류에 비해 약간 일치하는 편이다. 이는 연관된 지식 및 활용하는 기술이 유사한 것으로 나타났기 때문이다. 예를 들어 소분류 201에는 컴퓨터 하드웨어 기술자 및 연구원은 통신에 대한 설계와 인프라 관리, 유지보수 등을 수행한다. 이는 20020315 부가통신응용중계서비스 통신서비스가 컴퓨팅 기술을 결합하여 부가가치서비스를 제공하는데 필요한 기획, 설계, 구현, 영업 관리, 개통, 모니터링, 유지보수, 인프라관리, 고객서비스 및 교육하는 일을 행하는 것과 유사하다고 볼 수 있다.

둘째, 워크넷 소분류 206 통신 및 방송 장비기사 설치 및 수리원과 NCS 20020102 구내통신구축의 경우 유사한 직군으로 볼 수 있다. 구내통신구축은 기업 내에 설치하는 통신설비의 구축계획을 수립하고 설계, 시공, 감리와 운영관리 업무를 수행하는 직무이다. 워크넷 소분류 206 또한 통신설비를 설치하고 유지 관리하는 직무를 수행하기 때문에 서로 유사하다고 볼 수 있지만 상호 분류의 의미는 명확히 동일한 의미를 담고 있지 않다는 점에서 역시 개선의 측면이 필요하다.

분석 결과에 따라 6절에서는 산업계에서 직접 게시한 채용공고 자료를 바탕으로 직무 및 직종 군집을 도출하기 위한 빅데이터 분석을 수행하였다. 특히 추출된 단어들을 분류하는 분석 알고리즘을 적용하여 직무체계를 개발하고 각 직무별로 현재 산업계의 수요를 확인할 수 있는 핵심단어들을 도출하였다.

제6절 빅데이터 분석 기반의 직무·직종 분류체계 개발

1. 핵심단어 주제 분류를 위한 분석 알고리즘 적용

5절의 워크넷과 NCS 분류체계의 적합성 분석 결과를 바탕으로 SW분야의 산업 수요를 보다 정확하게 예측하기 위하여 SW분야 직무·직종 체계를 도출 및 개선하고 각 직무별 산업계 요구를 확인하기 위하여 채용광고상의 핵심단어들을 도출하고자 LDA 토픽모델링 기법을 통한 빅데이터 분석을 수행하였다.

1) 잠재 디리클레 할당 모델링 (Latent Dirichlet Allocation)

잠재 디리클레 할당 모델링 (Latent Dirichlet Allocation, 이하 LDA)은 텍스트마이닝 기법에서 사용되는 알고리즘으로 문서에서 숨겨져 있는 주제 (이하 토픽, Topic)들을 찾는 모델링 기법이다 (Blei et al., 2003). 토픽 모델링은 방대한 텍스트 데이터를 바탕으로 구분된 단어들을 분석함으로써 단어 활용에 의한 특정 주제들을 도출하고 해당들이 서로 어떻게 연결되어 있는지 또는 시간 흐름에 따라 어떻게 변화되는지 분석하는 기법이다.

LDA는 본 보고서의 목적인 워크넷 채용광고 자료를 바탕으로 현재까지의 직무 및 직종은 기존의 NCS와 같은 복잡하고 다양한 분류체계와 비교할 때 최근 동향을 반영한 직무 토픽 군집을 도출할 수 있다는 장점을 가지고 있다. 또한 기존 분류체계에서 구분하는 이론적 구분 방식에 비해 각 직무 별 산업계가 요구하는 핵심단어(직무 내 역량)의 중요도를 확인할 수 있다는 추가적인 장점을 가지고 있으며 이는 지속적인 SW인력 수요에 대한 시간에 따른 수요 변화 정도를 확인할 수 있게 한다.

최근 다양한 분석 영역에서 LDA 모델링을 통한 토픽의 탐색 및 도출을 활용하는 사례가 증가하고 있다. 문헌정보학에 분야에서 LDA를 통해 토픽을 추출하는 연구를 예를 들 수 있다 (박자현, 송민, 2013). 이 연구는 LDA를 사용하여 연구주제를 나타내는 단어들이 어떤 학문영역에 속하는지 판단하였다. 정보관리학회지, 한국문헌정보학회지, 한국도서관정보학회지, 한국비블리아학회지에 등재된 논문의 초록을 데이터로 사용하여 [그림 2-57]과 같이 총 3,834개의 문서(M)에서 20개의 토픽(k)을 추출했다.

문헌정보학에서 시행한 LDA 모델링은 토픽에 추출된 단어들(w_n)과 기존의 문헌정보학 주제 분류표(오세훈, 2005)를 비교할 수 있음을 확인하였다.

[그림 2-55] 문헌정보의 LDA 사례

토픽0	토픽1	토픽2	토픽3	토픽4
information service technology user resources environment development resource society	books dynasty king tripitaka published historical edition made history	social communication cultural development theory process society culture life	model quality user performance suggested developed develop process factor	library public service user community local facilities cultural people
토픽5	토픽6	토픽7	토픽8	토픽9
library university collection academic materials user service collections number	knowledge management structure system organization business marketing environment strategy	system information record retrieval user text time developed full	digital library electronic copyright access resources law legal open	school reading education library children program learning literacy instruction
토픽10	토픽11	토픽12	토픽13	토픽14
document retrieval term indexing performance subject index query clustering	metadata element ontology model web semantic thesaurus xml core	korea national management korean policy record system development government	library librarian education science subject professional korea organization educational	classification subject scheme ddc kdc system edition korean class
토픽15	토픽16	토픽17	토픽18	토픽19
cataloging catalog bibliographic rules record authority description control access	journal science citation korean articles medical field published korea	user web search service internet information reference site searching	factor satisfaction group students significant level variables groups differences	book books korean number publications japanese period materials literature

출처: 오세훈 (2005)

[그림 2-56] 뉴스데이터의 LDA 사례

“Arts”	“Budgets”	“Children”	“Education”
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

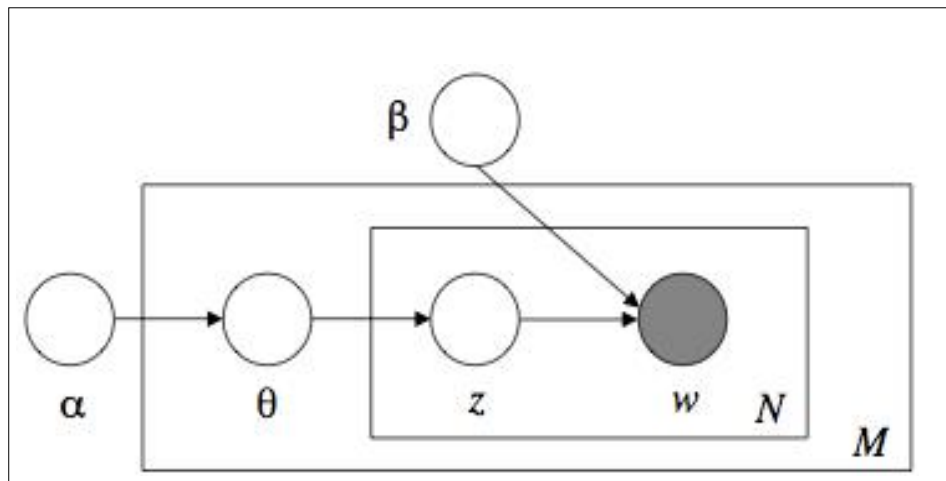
The William Randolph Hearst Foundation will give \$1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. “Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services,” Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center’s share will be \$200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive \$400,000 each. The Juilliard School, where music and the performing arts are taught, will get \$250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual \$100,000 donation, too.

출처 : Blei et al. (2013)

뉴스 데이터를 사용하여 토픽을 추출하는 연구도 있다 (Blei et al., 2013). 해당 연구는 AP 뉴스 16,333건에서 23,075개의 단어를 추출하여 토픽 모델링을 진행했으며 그 결과 도출된 토픽에서 판별된 단어들을 기반으로 ‘Arts’, ‘Budgets’, ‘Children’, ‘Education’ 이라는 중요 주제를 도출할 수 있었다.

마찬가지로 LDA는 텍스트와 같은 이산형 데이터에 적용되는 확률적 모형으로서 이미지, 소리 등의 비정형화된 영역에서도 사용이 가능하다 (Harman, 1992). 따라서 본 보고서의 목적과 비교할 때 LDA 방식은 데이터 기반의 SW분야 직무·직종 분류체계를 도출하고 각 직무별 핵심 역량 (핵심 단어의 중요도 정도를 중심으로)을 확인하는데 최적의 방법론 중 하나이다(이수연, 2016; 최해옥, 2015).

[그림 2-57] LDA 분석 모형



출처: Wikipedia, 잠재 디리클레 할당²¹⁾

본 보고서에서 적용한 LDA 분석 모형은 [그림 2-57]과 같은 절차로 수행된다. 활용한 LDA 모델링의 기본 가정으로는 첫째, 수집된 전체 채용공고 자료에는 k개의 주제와 M개의 문서가 존재하고 모든 채용공고 문서는 하나의 구인을 위한 토픽을 가지고 있다고 설정한다. 이를 바탕으로 둘째, 각 채용공고 문서는 N개 단어의 연속으로 나타나며, $W = (w_1, w_2, \dots, w_n)$ 으로 표현할 수 있다. 이를 바탕으로 토픽을 도출하기 위해서는 다음의 단계를 거치게 된다.

- 1) 푸아송 분포²²⁾를 사용하여 임의의 N을 선정한다.
- 2) 각 채용공고 문서가 어떤 토픽(z)의 비율로 구성될지 나타내는 θ (theta)을 디리클레(Dirichlet) 분포를 따라 선택한다.
- 3) N개의 단어에 대한 개별 단어들에 대해서
 - 3_1) θ 의 다항 분포에 따른 토픽들을 선택한다.
 - 3_2) 토픽들이 주어졌을 때 $p(w_n|z_n, \beta)$ 로부터 단어들을 구한다.

21) 위키피디아, LDA 도식,

https://ko.wikipedia.org/wiki/%EC%9E%A0%EC%9E%AC_%EB%94%94%EB%A6%AC%ED%81%B4%EB%A0%88_%ED%95%A0%EB%8B%B9

22) 푸아송 분포(Poisson distribution): 단위시간 안에 특정 사건의 발생 빈도를 나타내는 확률분포

즉, 전체 채용공고 문서에 대한 디리클레 분포를 따르는 특정 토픽인 θ 가 도출되고, 토픽의 벡터로부터 문서에서 분류된 단어들을 선택하는 방식으로 모델링을 진행한다.

LDA 알고리즘은 다양한 형태의 응용이 이루어져 오고 있다. 특히 상관 주제 모형은 디리클레 분포가 아닌 로지스틱 정규분포를 바탕으로 주제 간 상관관계를 분석하는 기법이다(Blei et al., 2006). 그러나 본 연구에서는 로지스틱을 이용하는 상관관계가 아닌 단어들에 의한 군집 배열을 확인하는 연구를 수행함에 따라 상관주제 모형은 활용하지 않는다.

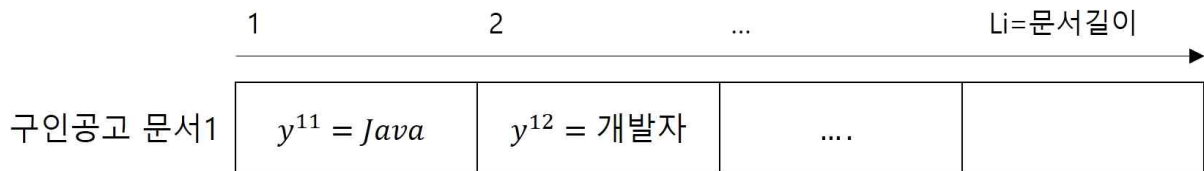
또한 추가적인 알고리즘으로는 계층적 잠재 디리클레 할당(hierarchical LDA, 이하 HLDA)이 있다(Griffiths and Tenenbaum, 2004). HLDA는 주제들 간의 계층 구조를 활용하지만 본 연구의 텍스트 데이터는 1단계 계층 구조가 사전에 정해져 있어야 한다는 HLDA 전제조건에 부합되지 않기 때문에 HLDA를 활용할 수 없다. 또한 본 연구에서와 같이 직무들의 군집인 직종을 정하는 계층적 군집을 정하는 경우, HLDA에서는 각 단어 간 명확한 상위와 하위 관계가 정해져 있어야 한다는 전제조건을 가지고 있다. 그러나 채용공고의 경우 단어의 상하 관계는 파악하기 어렵다는 점에 따라 HLDA는 활용이 불가능하다. 그에 따라 본 연구과제에서는 일반적으로 사용하는 잠재 디리클레 분포를 활용한 LDA 알고리즘을 적용하여 분석을 진행하였다. 그에 따라 우선적으로 1차적으로 1단계 LDA 토픽모델링을 통해 토픽 15가지를 구분하고 이 중에서 2단계 LDA 분석을 통해 유사 토픽끼리 2단계 군집화를 진행하여 직종을 구분하였다.

2) LDA를 활용한 직무·직종 토픽모델링 개발 방법

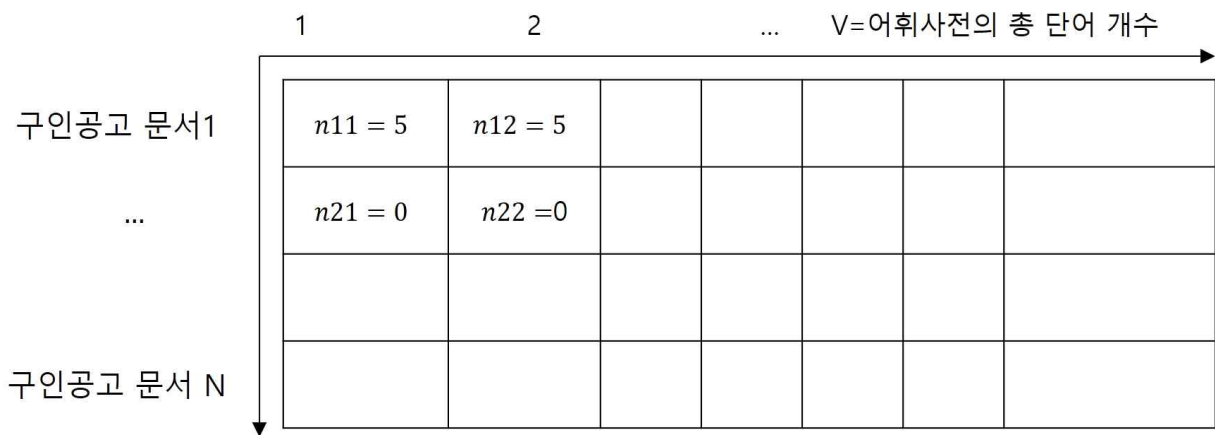
LDA는 문서의 단어 분포 값을 추정하여 잠재 변수인 “토픽”을 형성한다. 전체 공고데이터는 토픽들의 혼합재(Mixture of Topic)이며, 토픽에 따라 공고마다 단어 분포가 달라진다. 토픽이 같은 글에는 단어들이 나타나는 패턴이 유사함을 기본 전제로 할 때 본 연구에서는 LDA를 사용하여 토픽모델링 기반의 직무·직종 분류체계를 정하는 것이 적합하다고 판단하였다.

워크넷 공고(문서)를 단어의 나열로 모형화 하면, y_{il} 는 공고 i 에서 l 번째 단어를 표현한다. y_{il} 는 어휘 집합의 원소로 볼 수 있다. $i = 1:N$ 이며, N 은 공고의 개수다. $l = 1 : L_i$ 을 가정하고, L_i 는 공고문서 i 의 길이이다.

[그림 2-58] 채용공고 단어의 모형화



[그림 2-59] 단어 사전의 특정 단어 발생 수 도출과정



[그림 2-58]은 본 연구의 분석을 위한 채용공고에서 어휘사전(Word Dictionary)에 존재하는 특정 단어가 발생하는 개수를 표현한 그림이다. 이를 표현하면 [그림 2-59]와 같이 $N \times V$ 의 카운트 행렬로 표현할 수 있다. 본 보고서의 데이터 분석에서 사용한 단어집합, 즉 어휘사전에 구성된 단어수는 1,419개의 단어를 포함하고 있다.

어휘사전을 개발한 과정은 다음과 같다. 소프트웨어 분야의 용어사전에 나타나는 단어들을 어휘사전 구성에 활용하였으며 영문 용어들은 TechTerms의 IT관련 용어사전²³⁾에서 인터넷, HW, SW, 기술 관련 단어 범주에 정리된 단어들을 활용하였다. 또한 SW 관련 단어들이 주로 영문 약어를 주로 쓰고 있으나 영문을 쓰지 않는 국문 어휘는 한국저작권위원회의 소프트웨어 용어 사전²⁴⁾을 통해 공식 사용 용어로 어휘사전에 포함하였다.

워크넷 SW분야 채용 공고들에서 추출된 총 8,000여개의 단어들 중에서 등장 빈도수 50개 이상인 단어들을 대상으로 SW산업 관련 전문가 5인 (기업체: 3명, 학계: 2명, 이하 코더)이 SW 직업 및 직무와 관련 있는 모든 단어들을 독립적으로 선정하였다. 분석 결과의

23) <http://techterms.com/category/software>

24) <http://www.copyright.or.kr/information-materials/dictionary>

신뢰성 확보를 위해서, 5인의 코더에 대해 연구에 대한 이해와 단어 분류 기준의 일치성을 위해 약 1시간 동안 교육과정을 진행 하였다. 이 같은 교육과정은 코더 간 분석 결과상의 일치도를 확보하고 있는가를 추후 확인하기 위하여 상호 코더 간 신뢰성 조사 (Inter-coder Reliability)를 수행하기 위해서 분석 전에 수행해야하는 선행적 조치이다.

상호코더간신뢰성 검증은 정성적인 분석 결과에 대한 일치도를 확보하고 있는 가를 확인하기 위해 실시하는 것으로서, 구체적인 상호 코더 간 신뢰도는 Perreault & Laurence (1989)의 P/L인덱스를 활용하였다. 최종 분류결과 상호코더간신뢰성 조사에서 전문가 5인이 뽑은 단어목록과 개발된 어휘사전의 일치도가 79.4%의 신뢰성을 보임에 따라 개발된 어휘사전의 구성이 유의하다고 판명하였다. 결과적으로 총 1,419개의 단어를 분석을 위한 어휘사전으로 결정하였다.

이를 바탕으로 어휘사전에 포함된 모든 단어의 출현빈도에 대하여 행렬 내의 확률 값을 계산한 후 전체 확률 분포의 거리가 가까운 공고와 단어들이 군집을 구성함에 따라 각 군집별 특징을 가진 토픽 모델링을 수행하였다.

정의된 어휘사전을 바탕으로 <표 2-35>와 같이 LDA를 활용하여 K개의 토픽 모델링을 도출함에 따라 단어와 토픽 내 핵심단어의 확률분포²⁵⁾를 계산하였다. 토픽 내 각 핵심단어의 확률분포 값은 도출된 각 직무 군집 내의 단어가 군집 내에서 단어가 다른 단어들과 함께 등장할 가능성을 의미한다. 그러므로 각 토픽 내의 단어별 확률분포 값은 특정 토픽에서 특정 단어의 ‘동시등장 빈도율 (Co-occurrence)’ 로서 정의할 수 있다. ‘동시등장 빈도율’은 채용공고 문서에서 토픽 내 특정 단어가 다른 단어들과 동시에 등장하는 빈도의 비율을 의미한다. 따라서 특정 단어의 동시등장 빈도율이 커질수록 해당 토픽 내에서 많이 등장하는 ‘핵심단어’ 라고 정의할 수 있다.

25) 직무 내 핵심단어의 확률분포는 핵심단어가 해당 직종·직무 채용공고에서 등장하는 확률을 의미한다. 그에 따라 본 보고서에서는 동시등장 빈도율 (Word Co-occurrence)로 표기하였다.

〈표 2-35〉 직무·직종 K토픽 모델링 예시

토픽1	개발 (0.38)	Java (0.24)	운영 (0.002)	...	...	...
토픽2	디자인 (0.60)	UI (0.07)	UX (0.06)	...	...	...
토픽3	개발 (0.12)	앱 (0.10)	모바일 (0.10)	...	...	...
...	...					
토픽K	네트워크 (0.45)	시스템 (0.18)	운영 (0.12)	...	...	...

1차적으로 토픽 모델링이 완료된 후에는 신규 추가되는 채용공고는 특정 토픽으로 할당될 수 있다. 모든 공고 데이터는 LDA 확률 계산식에 따라서 공고상의 핵심 단어의 동시등장 빈도율을 구하고, 가장 유사한 직무 내 핵심단어의 동시등장 빈도율을 갖는 토픽 군집에 할당된다. 이에 따라서 도출된 토픽 결과들을 바탕으로 직무·직종 분류 체계를 구성하고 신규 생성되는 공고들을 적정 직무에 할당하면 시간 흐름에 따라 각 직무 및 직종의 변화 추이의 관찰이 가능하다. 특히 LDA는 신규 공고들의 단어들도 고려하기 때문에 특정 직무 내 핵심 단어들의 변화를 분석한다면, 특정 직무 특성 변화도 추가적으로 파악 할 수 있다는 장점이 있다.

2. LDA 분석 기반의 SW 직종직무 분류체계

1) LDA 토픽모델링 기반의 분류체계 개발 과정

본 연구에서는 워크넷에서 추출된 단어들을 기반으로 앞서 설명한 어휘사전을 구성하고, 어휘사전을 기반으로 워크넷 데이터를 분석하였다. 분석방법은 LDA 토픽 모델링 방법을 사용 하였으며, 이를 통해 총 15개 직무, 9개 직종 및 각 직무·직종에 해당하는 핵심 단어들을 도출하였다. LDA 토픽 모델링 기반 분류체계 개발 과정을 간략히 하면 다음과 같다.

① 최적의 군집(직무)수 결정을 위한 포커스 그룹 인터뷰 수행

→ 최적 군집 수 15개 결정

② 인터뷰 결과의 상호신뢰성 검증

→ 신뢰성 검증 통과로 인한 군집수 결과에 대한 유의성 확인

③ 최적 군집수에 의한 LDA기반 핵심단어(동시등장빈도율 기준)도출 및 상위 계층 군집(직종) 도출

→ 직무 15개, 직종 9개

④ 도출된 직무 군집별로 군집 내에 출현하는 전체 단어 빈도수 기반 원형차트 분석 및 관계네트워크 분석을 통한 군집별 특성 확인

→ 원형차트 기반의 직무 군집별로 가장 많이 출현하는 단어 확인 및 관계 네트워크 분석 기반의 군집별 동시 출현하는 연관 단어들을 통한 군집 특성을 파악

⑤ LDA기반 핵심단어(동시등장빈도율 기준)기반 분석

→ ‘핵심단어 방사형 그래프’, ‘핵심단어 비율 그래프’, ‘핵심단어 워드클라우드’ 등 핵심단어 기반 빅데이터 분석 결과의 시각화를 통한 분석

⑥ 기존 분석 결과들을 기반으로 주요 국제직무분류표준체계들을 참고하여 직무 명칭 및 정의 확정

⑦ 직무 명칭 및 정의 상호신뢰성 검증

→ 신뢰성 검증 통과로 인한 연구 결과에 대한 유의성 확인

⑧ 직무별 특성 개발 결과와 주요 국제직무분류표준체계들을 참고하여 직종 명칭 및 정의 확정

⑨ 직종 명칭 및 정의 상호신뢰성 검증

→ 신뢰성 검증 통과로 인한 연구 결과에 대한 유의성 확인

먼저 LDA 분석을 위해서는 분류될 토픽의 수 즉, 군집의 수를 정하는 과정이 선행되어야 하기 때문에 워크넷에서 추출된 동시등장빈도율을 기준으로 5개에서 30개로 분류된 군집의 상위 10개 핵심단어들을 대상으로 최적의 군집수를 결정하였다. 이를 위해 SW산업 관련 전문가 5인 (기업체: 3명, 학계: 2명)을 대상으로 각각의 결과에 대한 독립적인 인터뷰를 진행하였으며, 인터뷰 결과에 대한 신뢰성 검증을 위해 어휘사전을 작성했을 때와 유사하게 상호코더간신뢰성 조사를 실시하였다. 검증 결과 군집수에 대한 코더 상호간 신뢰도는 5인 평균값이 0.853, 즉 85.3%의 일치도를 획득하였다. 이러한 상호 코더간 신뢰도 결과는 LDA를 통해 도출한 직무·직종 토픽분류가 실제 활용될 수 있는 수준에 이르렀음을 의미한다. 그에 따라서 LDA기법을 기반으로 추출된 분류체계에서 k를 15로 선정하였을 때 가장 적합한 분류가 되었음을 확인하였다.

또한 다수의 직무들로 구성되는 상위 계층의 분류인 직종의 경우, 동일한 LDA를 적용하여 분석한 결과 9개의 직종으로 분류되었다. <표2-36>와 같이 LDA 토픽 모델링을 바탕으로 15개 직무, 9개의 직종으로 분류된 군집 내 특정 단어의 동시출현 빈도율을 기준으로 전체 단어 중에서 가장 중요한 핵심단어들을 도출하고 그에 따라 군집별 특성을 파악하였다. 각 단어들의 동시출현 빈도율을 기준으로 2%이하의 값을 가지는 단어들은 전체 군집에 미치는 영향력이 미미하다고 판단되어 추후 분석에서 제외 하였다.

분류된 군집별로 특성을 좀 더 명확히 파악하기 위하여 각 군집별로 등장하는 전체 단어를 추출하여 원형 차트와 워드클라우드를 제작하였다. 그에 따라 제시된 각 직무 군집별 원형 차트는 각 단어가 차지하는 비중을 보여주는 단어의 빈도수로 제시하였다. 다시 말해서 직무 군집 내 전체 단어들과 비교하여 해당 단어가 비중이 크면 해당 단어가 다른 단어들과 비교했을 때 상대적으로 더 많이 등장한다는 것을 의미한다. 마찬가지로 워드클라우드에서 보이는 단어의 크기는 다른 단어

대비 해당 단어의 상대적 빈도를 의미한다. 따라서 워드클라우드 결과에서 큰 단어로써 등장되는 단어들은 다른 단어와 비교할 때 상대적 빈도가 높음을 의미한다.

이후 군집 내 단어들 간의 관계를 알아보기 위해서 소셜네트워크 분석을 바탕으로 공동 출현 단어 네트워크를 제작하였다. 단어 공동 출현 네트워크는 직무 군집에서 동시에 발생하는 단어 간의 관계를 빈도로 표현한 것이다. 따라서 네트워크 그림에서 보이는 단어 간 연결선의 굵기는 단어들 간의 동시에 쓰이고 있는 관계를 의미한다. 따라서 네트워크 내의 연결 선 굵기가 굵을수록 단어들이 동시에 채용공고 내에서 발생하는 것이고 굵기가 얇을수록 단어들이 함께 쓰이지 않는 관계임을 의미한다.

LDA기반 핵심단어 데이터를 기반으로 분류된 각각의 직무·직종에 대해 ‘핵심단어 방사형 그래프’, ‘핵심단어 비율 그래프’, ‘핵심단어 워드클라우드’로 시각화 하고 각 군집들의 특성 분석에 활용하였다. ‘핵심단어 비율 그래프’는 직무·직종에 해당되는 채용공고의 동시등장 빈도를 상위 10순위에 해당하는 핵심단어들을 100%로 환산하여 각 단어들의 비중을 시각화한 그래프이다. ‘핵심단어 워드클라우드’는 직무·직종에 해당되는 채용공고의 동시등장 빈도율이 높은 단어들을 시각화한 결과이다. 또한 ‘핵심단어 방사형 그래프’는 직무·직종에 해당되는 채용공고의 동시등장 빈도율이 가장 높은 핵심단어를 100점으로 환산하여 나머지 주요 핵심단어들과 비교 분석이 가능하도록 시각화한 방사형 그래프(radar graph)이다.

전술한 다양한 LDA 분석 결과에 따라 <표 2-36>의 추출된 각각의 군집들의 직무 명칭과 정의를 구하기 위하여 국제직무분류표준 3가지를 참고하여 직무·직종 체계 개발에 활용하였다.

분류체계 도출을 위해 참고한 국제직무분류표준체계는 다음의 3가지를 고려하였다. 첫째, 국제노동기구(International Labour Organization, 이하 ILO)에서 구분한 국제표준직업분류체계²⁶⁾(International Standard Classification of Occupation, 이하 ISCO)와 둘째, 미국 노동부(U. S. Bureau of Labor Statistics)에서 2010년 개발한 미국표준직업분류체계(Standard Occupational Classification 2010, 이하 SOC)²⁷⁾와 2018년 개정을 위해 2016년 사전 정의된 Preliminary 2018 SOC (이하 SOC-2016)²⁸⁾, 마지막으로 미국 노동부 직업훈련청(Employment and Training Administration of the United States

26) <http://www.ilo.org/public/english/bureau/stat/isco/>

27) <https://www.bls.gov/soc/>, 2010년 재정 판이 최신 표준이며, 2018년도에 개정 예정임

28) <https://www.bls.gov/soc/>, 2018년 개정을 위한 사전 직무 정의 판으로 2016년 웹사이트에 제공 중

Department of Labor)이 1998년 개발하여 현재 운영 중인 직무역량 데이터베이스 (Occupational Information Net, 이하 O*Net)²⁹⁾을 참고하였다. 이 중 특히 가장 최신버전인 SOC의 예비 개정버전인 SOC-2016에 적합한 분류가 있을 경우 SOC-2016을 우선적으로 고려하였다.

이를 바탕으로 도출된 직무 군집별 명칭과 정의는 역시 SW분야의 전문가 5인 (SW관련 학과 교수 2인, SW관련 기업인 3인)에게 적합도 검사를 위해 상호코더 간신뢰성 분석을 수행하였다. 선택된 5명의 전문가들(이하 코더)들은 각각 독립적으로 LDA에 의해 구분된 군집에 속한 동시등장 빈도율 상위 10개의 핵심 단어들과 본 연구를 통해 정해진 명칭 및 정의를 서로 매칭하게 하여 그 결과들을 비교하여 일치성을 검사하였다. 그 결과 5명의 코더들에 의한 결과들과 본 연구의 개발결과에 대한 신뢰도가 5인 평균 84.7%로 나타나 본 연구 결과의 유의성을 확인하였다.

직무 군집별 명칭과 정의의 유효성을 확인한 이후 직무와 마찬가지로 LDA 분석 결과인 <표 2-36>의 추출된 각 직종 군집들의 명칭과 정의를 확정하기 위하여 직무 군집에서와 동일한 국제직무분류표준 3가지(ISCO, SOC, O*Net)를 참고하여 직종 체계 개발에 활용하였다.

도출된 9개 직종에 대한 명칭 및 정의에 대한 유효성을 확인하기 위하여 SW 분야 전문가 5인을 대상으로 상호코더간신뢰성 분석을 추가 수행하였다. 직무 군집별 개발과 마찬가지로 각 5명의 전문가들이 독립적으로 구분된 직종 군집들에 대해 정해진 명칭 및 정의를 각각 매칭한 후 해당 매칭 결과를 비교하여 전문가들의 결과에 대한 일치성을 판단하였다. 분석 결과 5명의 전문가 집단에 의한 결과와 본 연구에서 개발한 직종 체계의 결과에 대한 신뢰도가 5인 평균 82.2%로 나타나 개발된 직종 체계에 대한 유의성을 확인하였다.

그에 따라 도출된 각 직무 및 직종에 대한 세부 정의 및 분석은 <표 2-36>부터 각각의 직무별로 토픽모델링 결과와 시각화 자료들을 바탕으로 설명하였다.

29) <https://www.onetonline.org/>

〈표 2-36〉 토픽모델링 기반의 직종직무 분류체계

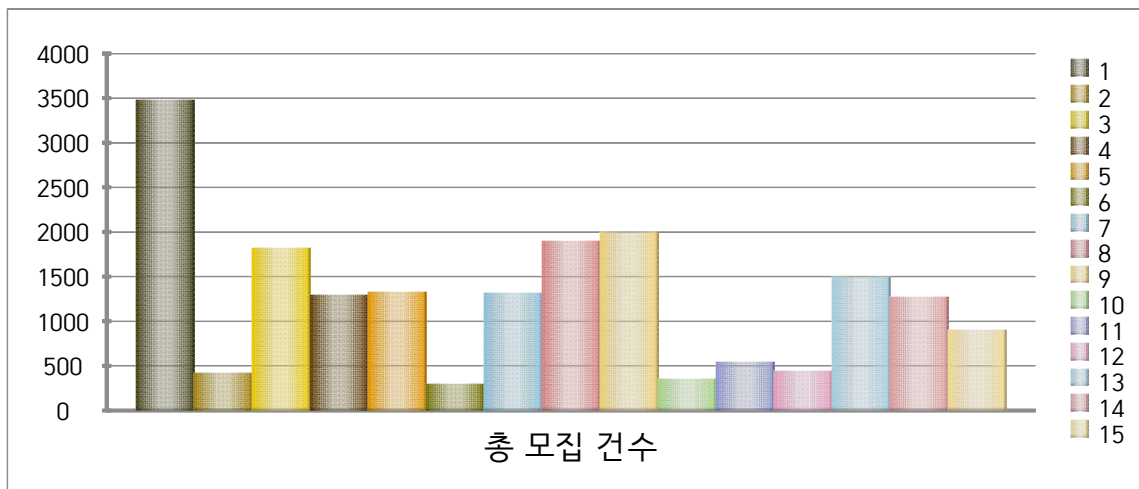
직종	직종1		직종2		직종3	직종4			직종5	직종6		직종7	직종8	직종9	
직무 군집	직무1	직무2	직무3	직무4	직무5	직무6	직무7	직무8	직무9	직무10	직무11	직무12	직무13	직무14	직무15
핵심 단어 (동시 등장 빈도율)	관리 (0.2406)	디자인 (0.1862)	디자인 (0.259)	디자인 (0.259)	엔지니어 (0.1336)	개발 (0.0192)	개발 (0.1327)	개발 (0.1302)	하드웨어 (0.1267)	웹 (0.109)	IT (0.1466)	개발 (0.1263)	마케팅 (0.1693)	개발 (0.2862)	개발 (0.3717)
	쇼핑몰 (0.0832)	쇼핑몰 (0.1256)	웹 (0.2222)	시스템 (0.0632)	네트워크 (0.1117)	웹 (0.0761)	프로그램 (0.1293)	운영 (0.1286)	설계 (0.1154)	편집 (0.0913)	개발 (0.0891)	웹 (0.0875)	온라인 (0.111)	java (0.1262)	웹 (0.1351)
	상품 (0.0675)	웹 (0.0701)	관리 (0.0564)	전산 (0.0627)	정보보안 (0.0899)	c (0.0703)	서버 (0.1225)	소프트웨어 (0.0938)	디자인 (0.0696)	기획 (0.0754)	프로젝트 (0.0759)	모바일 (0.0852)	디자인 (0.0869)	운영 (0.0483)	android (0.0709)
	홈페이지 (0.058)	오픈마켓 (0.0667)	포토샵 (0.048)	광고 (0.0577)	설치 (0.0854)	c++ (0.0454)	컴퓨터 (0.0837)	관리 (0.0609)	개발 (0.0687)	제작 (0.0726)	java (0.0699)	디자인 (0.0473)	기획 (0.0826)	시스템 (0.0374)	앱 (0.0687)
	온라인 (0.0556)	3d (0.0543)	쇼핑몰 (0.0398)	IT (0.0369)	시스템 (0.0557)	디자인 (0.0426)	시스템 (0.0452)	분석 (0.0445)	보조 (0.0473)	방송 (0.0515)	바이럴 (0.0515)	데이터 베이스 (0.0423)	cs (0.053)	jsp (0.0274)	ios (0.0499)
	디자인 (0.0505)	기획 (0.0505)	설치 (0.0286)	웹 (0.0344)	장비 (0.0506)	영상 (0.0395)	linux (0.0442)	기획 (0.0421)	제어 (0.036)	프로젝트 매니저 (0.0437)	웹 (0.0398)	기획 (0.0345)	UI (0.0462)	설계 (0.025)	php (0.041)
	제작 (0.042)	온라인 (0.0321)	인터넷 (0.0276)	생산 (0.03)	관리 (0.0497)	컨텐츠 (0.0321)	엔지니어 (0.0368)	서비스 (0.0356)	그래픽 (0.0355)	교육 (0.0404)	소셜 (0.0373)	관리 (0.0301)	UX (0.0422)	c (0.0217)	모바일 (0.0271)
	페이지 (0.0363)	운영 (0.0254)	홈페이지 (0.025)	cg (0.0294)	정보 (0.0474)	게임 (0.0266)	웹 (0.0333)	그래픽 (0.0287)	웹 (0.0313)	코딩 (0.0391)	마케팅 (0.037)	제작 (0.0232)	쇼핑몰 (0.0411)	연구 (0.0212)	java (0.0243)
	블로그 (0.0245)	컨텐츠 (0.022)	기획 (0.0237)	기획 (0.0241)	구축 (0.0265)	html5 (0.0263)	제안 (0.0216)	센터 (0.0278)	기계 (0.0301)	모바일 (0.0386)	오픈마켓 (0.0346)	운영 (0.0231)	디지털 (0.0307)	데이터 베이스 (0.02)	기획 (0.0177)
	기획 (0.0243)	제작 (0.0186)	운영 (0.0153)	편집 (0.016)	개발 (0.0227)	쇼핑몰 (0.0257)	네트워크 (0.0206)	프로그램 (0.0268)	전자 (0.0296)	판매 (0.0328)	PC (0.0338)	프로젝트 (0.0219)	웹 (0.0293)	웹 (0.0196)	운영 (0.0174)

2) LDA 토픽모델링 기반의 직종/직무 분류체계

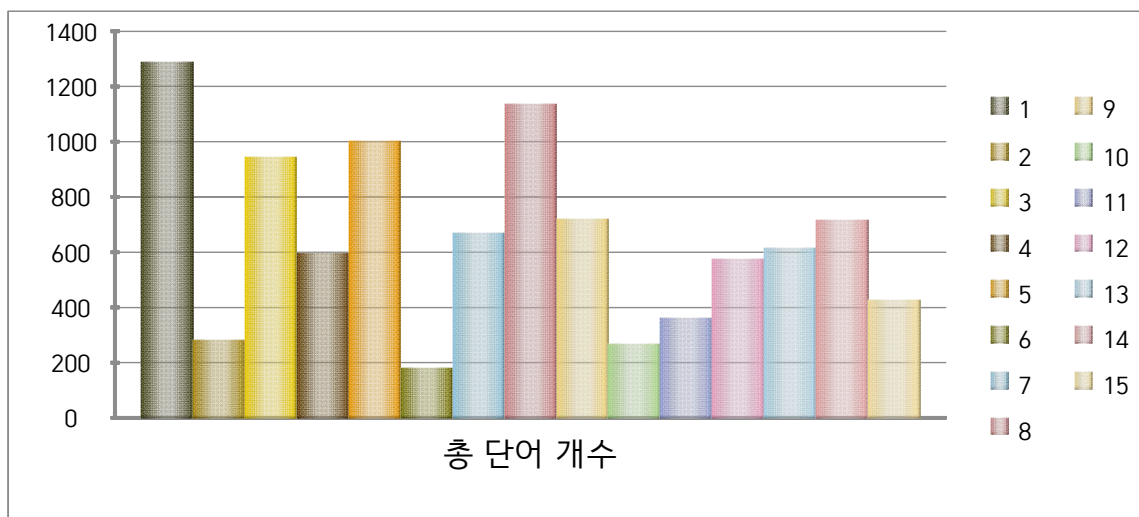
2.1) 총 수집 단어 기반 군집별 기본 현황

[그림 2-60]은 LDA 토픽모델링 기반 워크넷 자료 분석과 분류 체계의 직무코드별 채용공고 수이다. 모집 건수는 직무군집 1번에 해당하는 모집 건수가 3,486건으로 가장 많고 8번, 3번이 다음으로 많았다. 가장 적은 359건을 모집한 직무군집은 10번이며, 평균적으로 직무군집별로 1,259건의 모집을 하는 것으로 나타났다.

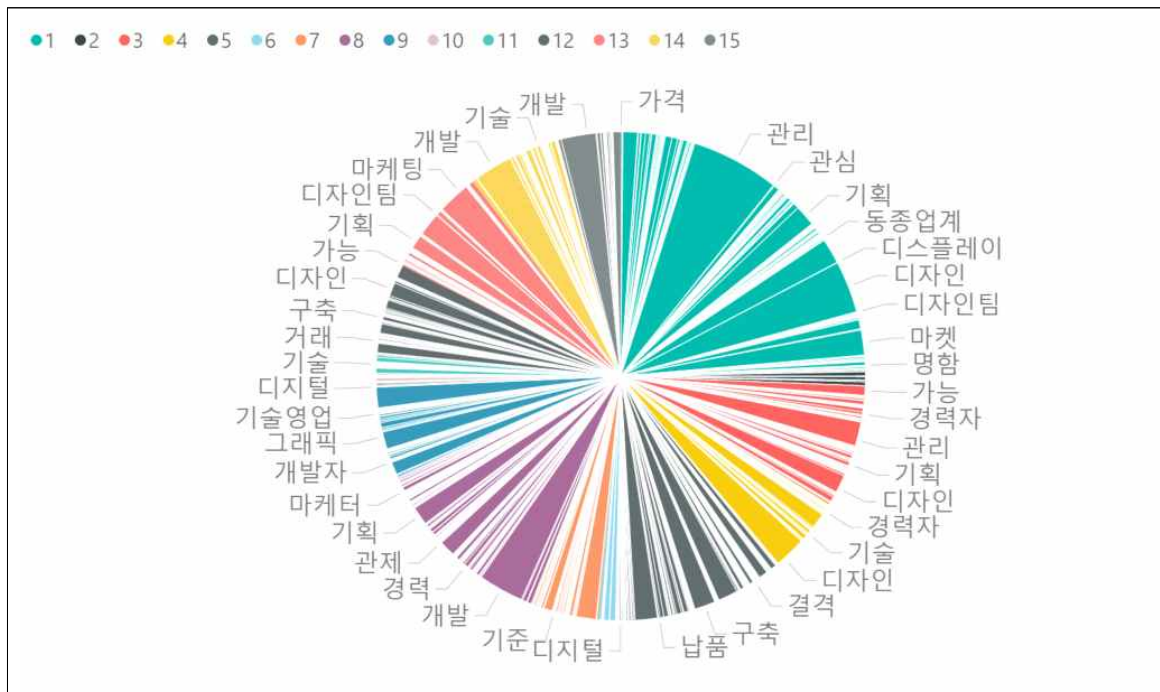
[그림 2-60] 워크넷 분류 체계 직무코드별 채용공고 수



[그림 2-61] 워크넷 분류 체계 직무코드별 핵심단어 수



[그림 2-62] 워크넷 분류 체계 직무 코드별 단어 Chord Chart



[그림 2-61]은 분석한 직무코드별 단어의 수이다. 분석에 사용한 단어는 채용공고에 등장하는 제목, 단어, 직무내용에 해당하는 단어를 추출한 것이다. 직무 코드 1번에 해당하는 단어수가 1,289개로 가장 많았다. 가장 적은 283개의 단어수를 갖는 직무코드는 2번이며, 평균적으로 직무별 653건의 단어가 추출되었다.

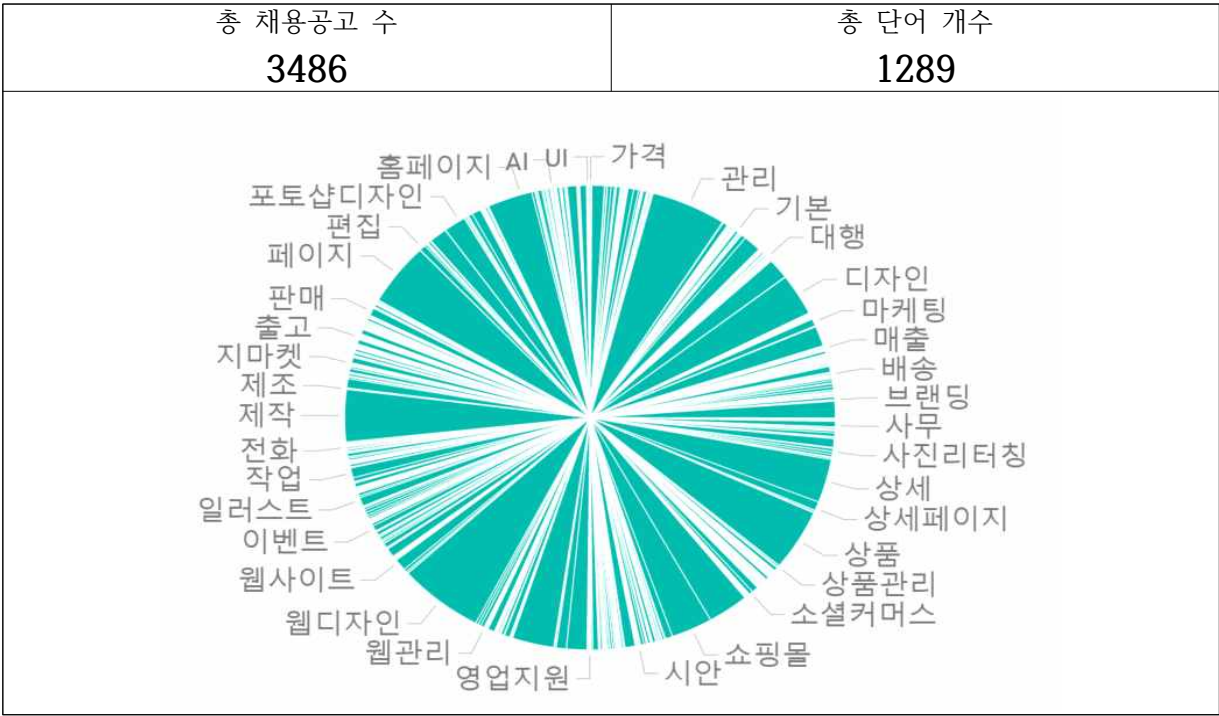
[그림 2-62]는 분류체계에 맞춰 분류된 토픽 별로 출현하는 전체 단어들을 코드 번호에 따라 색상을 달리하여 분류된 토픽별 주요 출현하는 상위 단어들을 시각화 하였다. 그 결과 가장 면적이 큰 1번 직무에는 관리, 관심, 기획, 디스플레이와 같은 단어들이 제일 많이 나타났으며 4번 직무, 13번 직무 등의 경우에도 전체 단어들 중 많은 단어들을 포함하는 군집으로 나타났다. 그에 따라 각각의 직무 군집을 바탕으로 핵심단어들을 통해 직무별 정의 및 특성을 개발한 후 직무 개발에 대한 유효성을 상호코더간신뢰성을 통해 확인하였다. 이후 직종 군집별 정의 및 특성을 도출하고 직종 개발에 대한 유효성을 상호코더간신뢰성으로 확인하였다.

2.2) LDA 결과 기반 직종 군집 별 정의 및 특성 개발

2.2.1) 웹 관리자

직무 1에 해당 하는 채용공고는 [그림 2-63]과 같이 총 3,486건으로 1,289개의 단어가 등장하는 특성을 가지고 있는 군집이다. 전체 채용공고 내의 단어 내용을 바탕으로 빈번하게 등장하는 단어는 ‘웹디자인’, ‘디자인’, ‘관리’ 이다. 또한 ‘상품’, ‘제작’, ‘홈페이지’ 와 같은 관련된 단어들이 종종 등장하고 있다. 또한 직무 1에 대한 채용공고에서의 모든 단어들을 기반으로 워드클라우드 시각화를 수행한 결과 웹 디자인, 상품, 관리, 제작, 웹UI 디자인 등은 해당 직무가 웹에서의 상품 관련 웹 페이지 또는 쇼핑몰의 관리 관련 직무를 요구하는 것으로 판단할 수 있다.

[그림 2-63] 직무 1의 특성



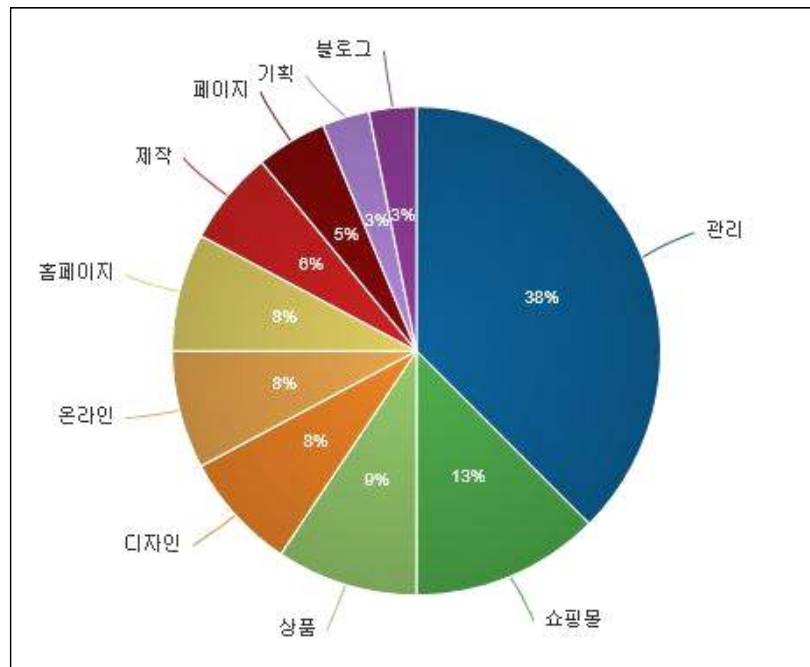
전체 채용광고에서 직무 1 군집에 대해 서로 같이 출현했던 단어들을 연결한 결과를 소셜네트워크 분석을 바탕으로 구분한 결과는 [그림 2-64]와 같다. 직무 1의 공동 출현 네트워크의 결과에서 ‘웹디자인’과 관련하여 {웹디자인-관리-페이지-상품}, {웹디자인-상품-제작} 관계가 강하게 존재하고 ‘관리’와 관련하여 {관리-홈페이지}, {관리-온라인} 관계가 존재한다. 즉, 직무 1은 쇼핑몰 홈페이지에서 상품 및 콘텐츠 디자인을 운영하고 이와 관련된 기획을 함께 하는 직무로서 확인되어 상품과 관련하여 온라인의 관리 역량을 요구하는 것으로 분석된다.

〈표 2-37〉 직무 1의 핵심단어 목록

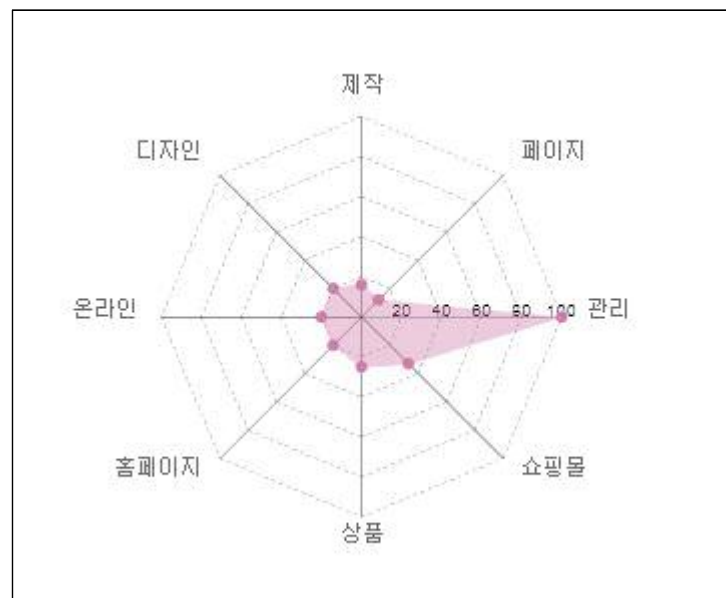
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	관리	0.2406
2	쇼핑몰	0.0832
3	상품	0.0675
4	홈페이지	0.0580
5	온라인	0.0556
6	디자인	0.0505
7	제작	0.0420
8	페이지	0.0363
9	블로그	0.0245
10	기획	0.0243

이에 따라 LDA 분석으로 도출한 직무 군집 1의 동시등장 빈도율에 대한 결과는 [표 2-37]과 같다. 이 중 동시등장 빈도율 2%이하의 값을 가지는 단어들을 제외한 상위 10개 단어는 핵심단어는 ‘관리’, ‘쇼핑몰’, ‘상품’, ‘홈페이지’, ‘온라인’, ‘디자인’, ‘제작’, ‘페이지’, ‘블로그’, ‘기획’의 단어들이 핵심단어들로 등장하고 있다. 직무 군집 1의 상위 10개 핵심단어의 비율을 계산하여 시각화한 결과는 [그림 2-65]와 같다. 결과에 따라 군집 1은 관리에 대한 강조가 많은 직무인 것을 알 수 있으며 홈페이지, 온라인, 디자인, 상품, 쇼핑몰 등의 핵심단어들이 도출됨을 알 수 있다. 이 중 홈페이지, 온라인, 디자인의 경우 비율 상 큰 차이가 없는 것으로 보인다.

[그림 2-65] 직무 1의 핵심단어 비율 그래프



[그림 2-66] 직무 1의 핵심단어 방사형 그래프



〈표 2-38〉 웹 관리자 직무 관련 국제직무분류표준

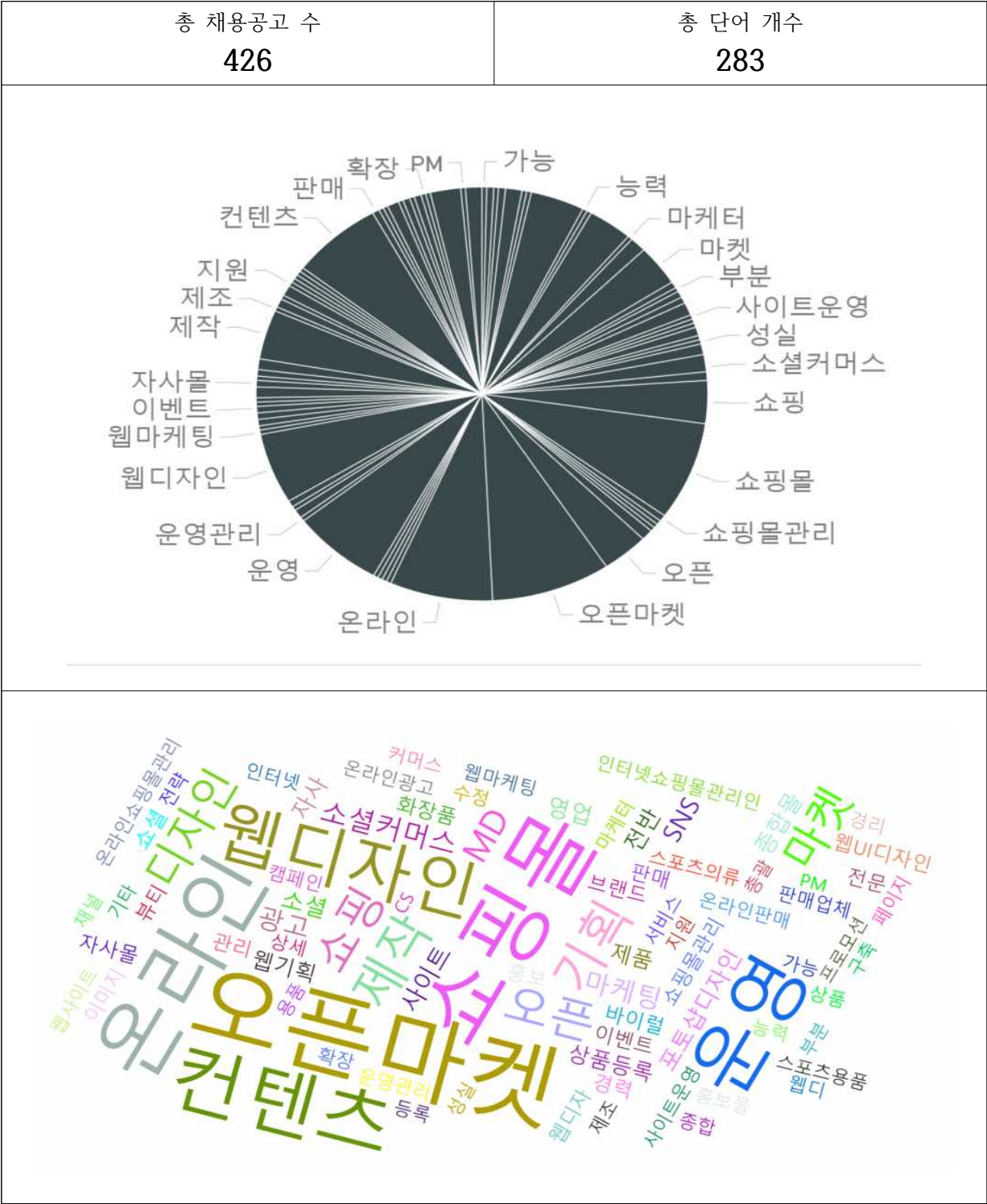
소스	분류코드	직무	정의
ISCO-08	2513	Web and multimedia developers	• 오픈 소스 콘텐츠 관리시스템 (CMS) 또는 특정 속성 SW를 사용하는 웹페이지 및 사이트를 설계, 작성, 생산, 유지보수
SOC-2016	15-1254	Web Developer	• 웹사이트를 설계 및 개발 • 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리 • 웹사이트를 위한 콘텐츠 기획
O*Net	15-1199.03	Web Administrators	• 웹 환경에 맞는 사이트 디자인, 개발 및 유지보수 관리 • 웹사이트 및 웹 콘텐츠의 품질 보증 관리 및 테스트 수행
본 보고서의 분류		웹 관리자	• 기업 웹 사이트의 게시판에 대한 관리 및 웹 사이트에서 발생하는 다양한 문제들을 해결

분석 결과에 따라 웹 관리자의 주요 직무는 다음과 같이 판단할 수 있다.

- 커뮤니티 및 쇼핑몰 사이트에서 구성원 간 참여 활동 촉진을 위한 다양한 이벤트 기획 및 고객 요구를 분석하여 상품 디스플레이와 온라인 판매촉진 또한 관장
- 웹사이트에서 고객에게 제공하기 위한 구성 내용의 기획
- 인터넷사이트 상에서의 게시판 등을 관리
- 인터넷사이트에서 제공하는 서비스의 기능을 향상시키기 위한 기획
- 인터넷사이트를 정기적으로 점검·유지
- 인터넷사이트 상에서 발생한 문제점을 확인하고 직접 해결 또는 기업 내 개발자에게 통보
- 웹사이트의 블로그 운영 및 사용자의 요구와 불만 사항을 신속히 처리한다.

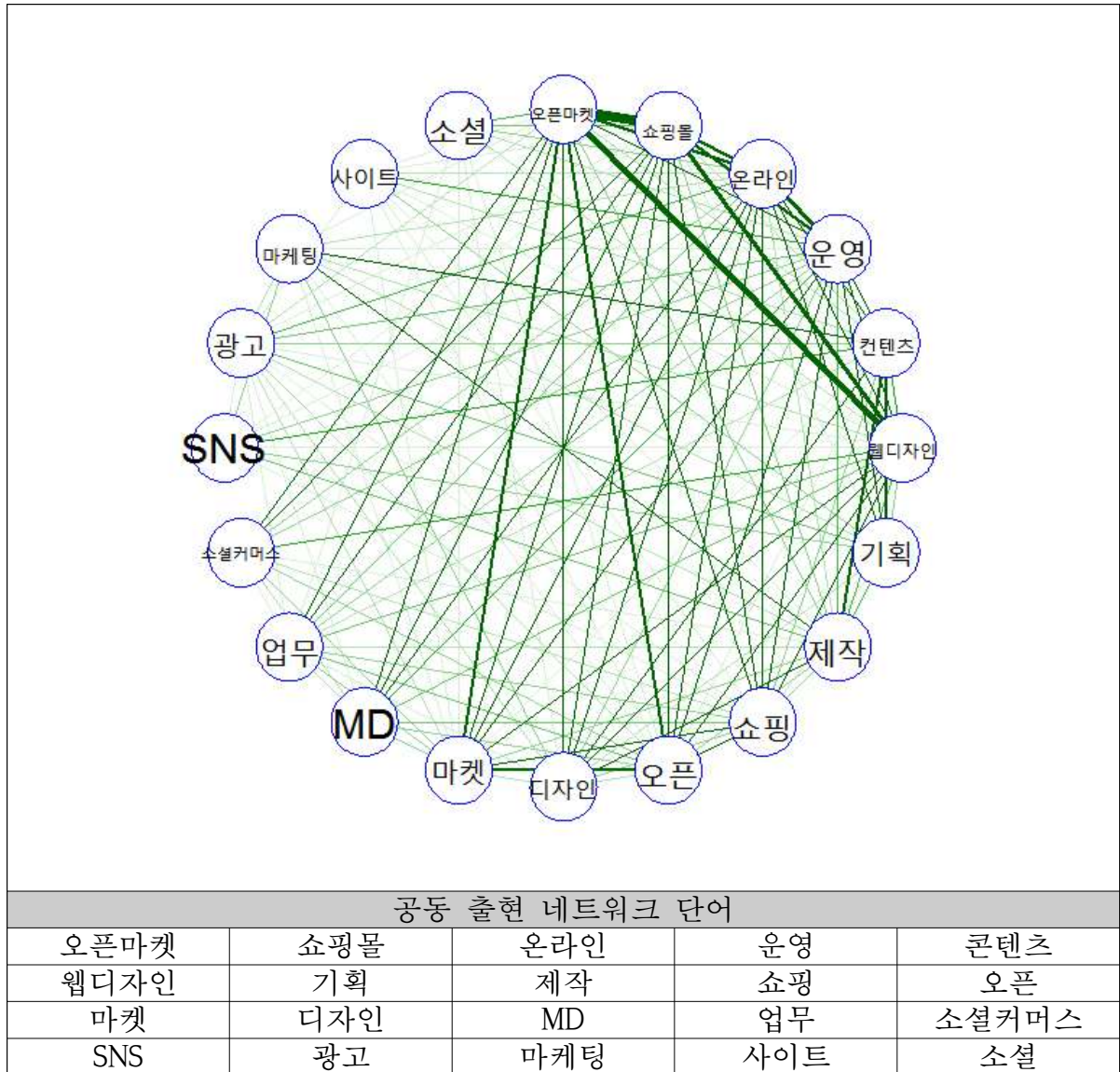
2.2.2) 웹기획자

[그림 2-68] 직무 2의 특성



직무 2에 대한 단어 데이터를 분석한 결과 채용공고는 [그림 2-68]과 같이 총 426건으로 분석으로 구분된 단어는 총 283개의 단어가 등장한다. 빈번하게 등장하는 단어는 ‘오픈마켓’, ‘쇼핑몰’, ‘운영’, ‘웹디자인’, ‘콘텐츠’, ‘온라인’이다. 또한 ‘기획’, ‘제작’, ‘쇼핑’, ‘마켓’과 같은 단어들이 등장한다. 또한 분석된 워드클라우드 결과와 같이 직무 2 군집은 오픈마켓, 쇼핑몰, 콘텐츠, 운영, 제작, 웹디자인 등의 웹 쇼핑몰 기반의 상품 운영 및 소비자에게 전달하는 직무 단어들을 확인할 수 있으며 웹기획, 마케팅, 프로모션 등 주변 단어들 또한 웹 관련 상품 및 홈페이지, 쇼핑몰 자체에 대한 기획 및 관리 관련 직무임을 알 수 있다.

[그림 2-69] 직무 2의 공동 출현 단어 네트워크

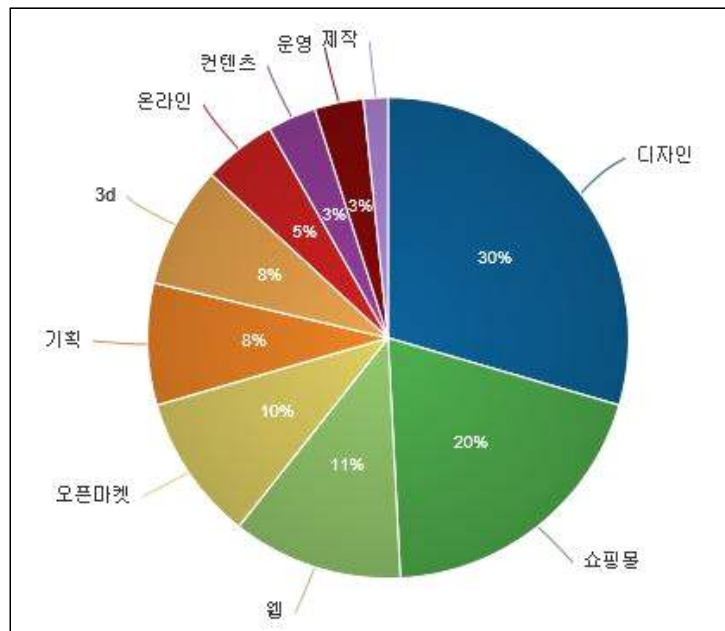


[그림 2-69]는 직무 2에 해당하는 채용공고의 단어들을 바탕으로 소셜네트워크 분석을 통한 공동 출현 네트워크이다. 분석결과 {오픈마켓-쇼핑몰-웹디자인}, {온라인-운영}, {콘텐츠-제작}, {콘텐츠-기획}의 단어 조합이 채용공고에서 주로 나타나는 것을 알 수 있다. 즉, 직무 2 군집은 쇼핑몰 및 오픈마켓에 대한 운영과 콘텐츠 제작과 관련된 관리 역량을 요구하는 것으로 분석되었다.

〈표 2-39〉 직무 2의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	디자인	0.1862
2	쇼핑몰	0.1256
3	웹	0.0701
4	오픈마켓	0.0667
5	3d	0.0543
6	기획	0.0505
7	온라인	0.0321
8	운영	0.0254
9	콘텐츠	0.0220

[그림 2-70] 직무 2의 핵심단어 비율 그래프

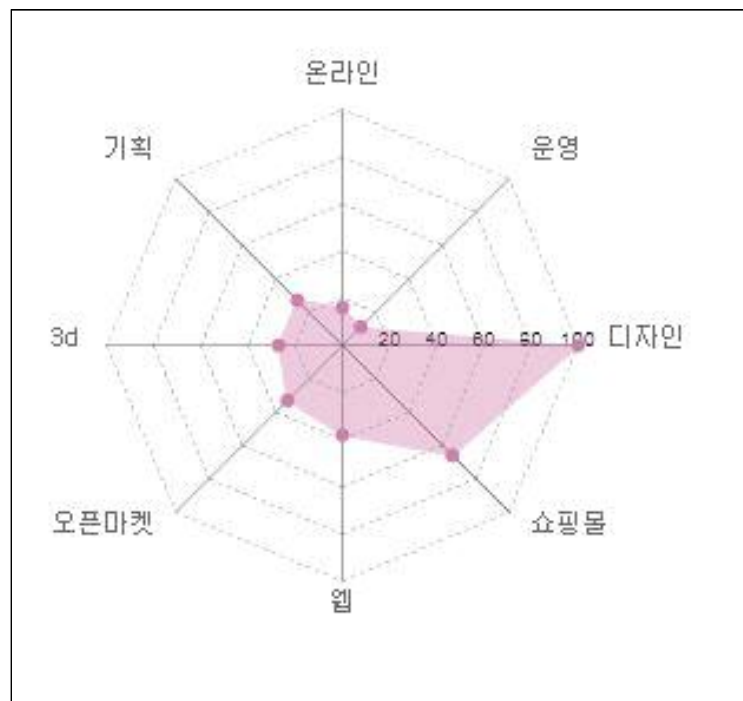


LDA 분석을 기반으로 직무 2에 대해 분류된 핵심단어들은 [표 2-39]와 같다. 분석된 핵심단어 중 동시등장 빈도율의 의한 2%이하를 제외한 단어는 9가지로 ‘디자인’, ‘쇼핑몰’, ‘웹’, ‘오픈마켓’, ‘3d’, ‘기획’, ‘온라인’, ‘운영’, ‘콘텐츠’ 순으로 핵심단어의 동시등장 빈도율이 높은 순으로 나타났다. 직무 군집 2의 경우 웹으로 제공하는 다양한 서비스 (오픈마켓, 쇼핑몰, 부동산, 웹 광고, 콘텐츠 서비스 등)의 기능과 화면을 기획하는 직무 클러스터이다. 도출된 제안

단어 중 ‘3d’의 경우 직무와 상관관계가 있는 지 확인하기 위하여 기존 수집된 채용공고 문서를 확인한 결과 오픈마켓 및 웹 광고, 콘텐츠 서비스 등에서 3d와 관련된 관리 및 운영 직무가 존재하여 핵심단어에 포함된 것으로 파악할 수 있다.

[그림 2-70]은 직무 2에 대한 핵심단어 비율 그래프이다. 시각화 자료의 결과에 따르면 디자인, 쇼핑물, 웹, 오픈마켓 등이 직무 2의 주요 단어로서 제시되고 있으며 특히 웹 사이트의 디자인, 기획 등이 주요 직무로서 제시되고 있다. 이에 따라 [그림 2-71]과 같이 직무 2에 대한 핵심단어를 100점으로 환산하여 방사형 그래프를 시각화 하였다. 그 결과 웹 쇼핑물, 오픈마켓을 대상으로 디자인, 기획 등이 주요 결과로 나타나고 있으며 위에서 설명한 3d 관련 기획 업무도 나타나고 있다.

[그림 2-71] 직무 2의 핵심단어 방사형 그래프



마찬가지로 [그림 2-72]와 같이 핵심단어들을 기반으로 워드클라우드를 시각화 하였다. 그 결과 오픈마켓, 온라인, 3D 등의 주요 요소를 바탕으로 기획 및 운영, 마케팅, 콘텐츠 등의 단어들이 제시되고 있으며 이는 웹 쇼핑물과 소셜커머스에서 콘텐츠 기획 및 운영, 퍼블리싱, 그래픽 요소의 기획, 홍보 등과 관련한 직무임을 보여주고 있다.

[그림 2-72] 직무 2의 핵심단어 워드클라우드



결과적으로 직무 2는 웹 쇼핑몰 및 오픈마켓, 소셜커머스의 운영을 기획하고, 쇼핑몰 및 오픈마켓에서 필요로 하는 디자인 및 콘텐츠를 기획하는 역할을 수행한다. 특히 <표 2-40>과 같이 국제직무분류표준에서는 주로 사용자 지원 전문가와 같은 역할을 하고 있으며 웹 사이트의 구축 관련 기획 업무 및 마케팅 등을 책임지는 전문가를 주로 제시하고 있다. 특히, 컴퓨터 사용자에게 기술적 지원을 제공할 수 있는 부분을 기획하거나 웹사이트에서 개인 고객에게 전화 또는 전자적으로 문제 해결 및 질의응답 관련 기획하는 IT 관리자 또는 컴퓨터 정보시스템 관리자로서 제시되고 있다. 추가적인 역할로는 컴퓨터 SW, 전자 우편, 운영시스템, 시스템 분석 등에 대한 지원 기획, 참여, 협력 활동에 대한 정의도 제시되어 있다.

분석 결과 직무군집 2의 경우 콘텐츠 및 웹페이지 내의 디자인에 대한 기획과 같은 홈페이지 관리를 전반적으로 담당하고 정보를 계속적으로 갱신해야 하며 고객대응 능력이 필요하다. 추가적으로 웹기획자는 웹으로 제공하는 다양한 서비스(오픈마켓, 쇼핑몰, 부동산, 웹 광고, 콘텐츠 서비스 등)의 기능과 화면을 기획하는 역할도 포함하고 있다.

〈표 2-40〉 직무군집 2의 국제직무분류표준

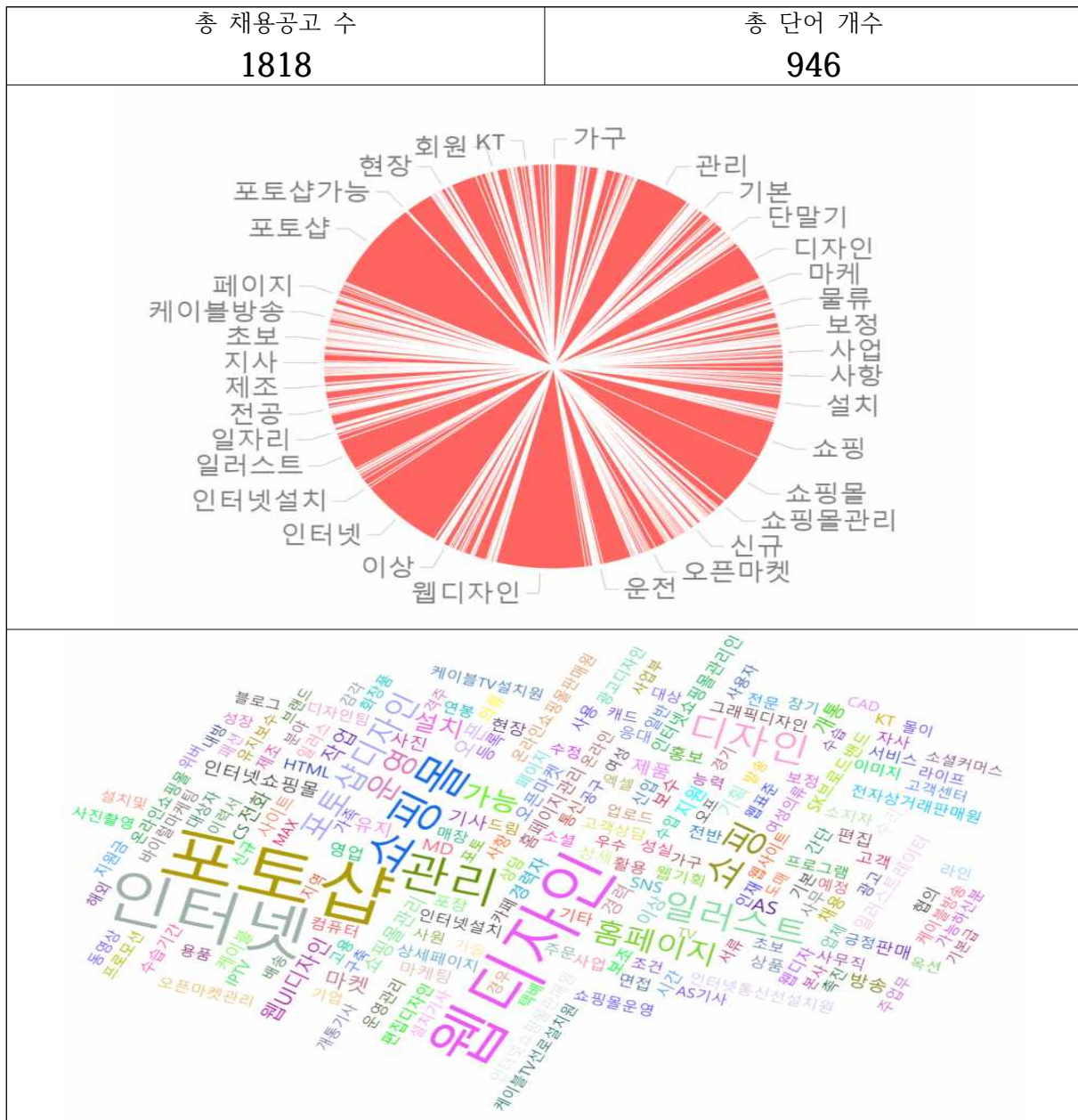
소스	분류코드	직무	정의
SOC	11-3021	IT Managers	- 컴퓨터 사용자에게 기술적 지원을 제공할 수 있는 부분 기획 - 웹사이트에서 개인 고객에게 전화 또는 전자적으로 문제 해결 및 질의응답 관련 기획
O*Net	11-3021.00	Computer and Information Systems Managers	- 컴퓨터 SW, 전자 우편, 운영시스템, 시스템 분석 등에 대한 지원 기획, 참여, 협력 활동 “Network and Computer Systems Administrators” 직무는 제외 (15-1142).
본 보고서의 분류		웹기획자	- 웹기획자는 웹 사이트의 구축에 있어서 기획과 업무의 진행 및 마케팅 등을 총괄하는 전문가를 말한다. - 콘텐츠 및 디자인 기획 등 전반적인 홈페이지 관리를 담당하고, 제공되는 정보를 계속적으로 갱신해야 하며 고객관리 능력이 필요하다.

따라서 본 보고서에서는 직무 2 군집을 ‘웹기획자’로 명칭 하였으며 그 정의는 웹 사이트의 구축에 있어서 기획과 업무의 진행 및 마케팅 등을 총괄하는 전문가로 정의하였다. 이러한 분석 결과에 따라 웹기획자의 주 업무는 다음과 같이 제시될 수 있다.

- 웹사이트와 사이트의 콘텐츠 디자인 업무를 위한 웹 관리자, 웹 프로그래머, 웹 디자이너 등과 협업
- 웹사이트의 자료를 지속적으로 갱신
- 사이트에 방문하는 사용자 성향과 접속 현황 등의 자료를 분석하여 웹사이트 콘텐츠의 수정 및 보완
- 외부업체에 웹사이트 구축을 진행할 때 외부 개발자들과 밀접한 의사소통으로 웹 사이트 구축의 목적이 명확하게 구현되도록 작업 진행 과정과 결과물을 확인

2.2.3) 웹 콘텐츠 디자이너

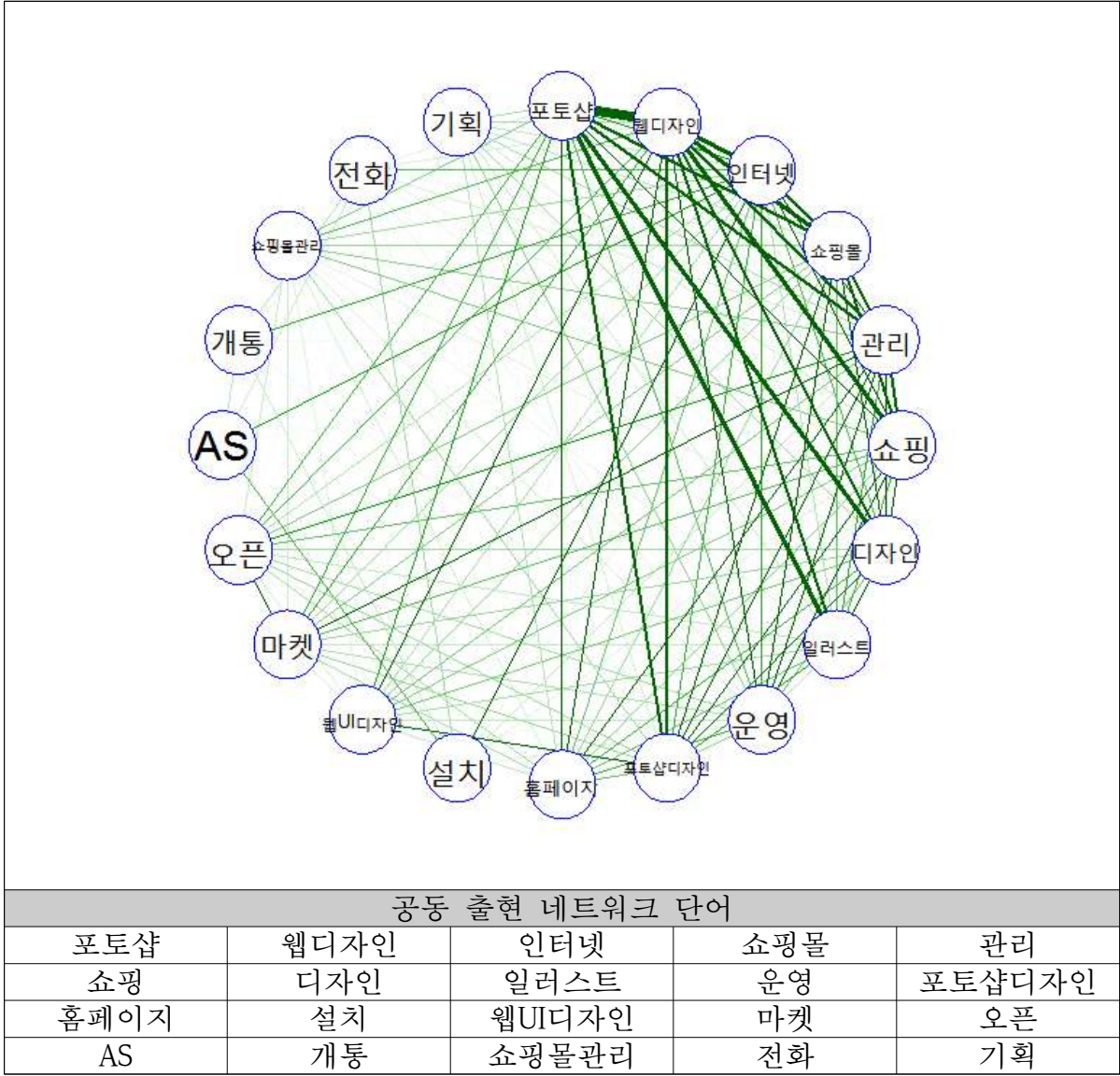
[그림 2-73] 직무 3의 특성



직무군집 3에 대한 채용공고 단어를 분석한 결과는 [그림 2-73]과 같다. 이에 따라 직무 3 군집의 채용공고 단어를 분석한 결과 총 채용공고 수는 1818개이고 제시된 단어수는 946개이다. 이 중 빈번하게 등장하는 단어는 ‘웹디자인’, ‘인터넷’, ‘포토샵’이다. 또한 ‘쇼핑몰’, ‘쇼핑’, ‘디자인’과 같은 단어들이 등장한다.

워드클라우드의 결과 총 단어들 중에서는 웹디자인, 포토샵, 인터넷, 쇼핑몰, 관리 등의 단어들이 많이 등장하는 것으로 제시되고 있다. 그러나 각 단어들 간의 연계성 및 핵심적인 단어들을 도출하기 위하여 우선적으로 [그림 2-74]와 같이 소셜네트워크 분석을 바탕으로 총 단어 중 동시에 나타나는 단어 조합을 확인하였다.

[그림 2-74] 직무 3의 단어 공동 출현 네트워크



직무군집 3의 공동 출현 네트워크를 도출한 결과 {웹디자인-포토샵}의 강한 관계가 형성되어있고 {웹디자인-인터넷-포토샵}, {웹디자인-쇼핑몰}의 관계가 다수 함께 나타나는 것을 확인할 수 있다. 또한 {포토샵-일러스트}의 관계가 존재하는 것을 확인하였다. 즉, 직무군집 3은 웹 디자인의 영역에서 포토샵, 일러스트와 관련된 도구를 사용하는 역량을 요구하는 것으로 확인할 수 있다.

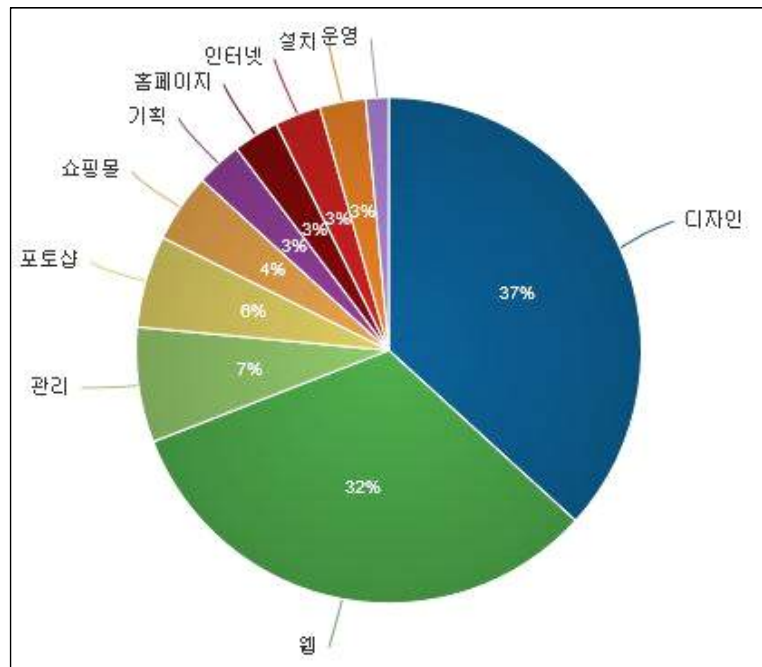
〈표 2-41〉 직무 3의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	디자인	0.2590
2	웹	0.2222
3	관리	0.0564
4	포토샵	0.0480
5	쇼핑몰	0.0398
6	설치	0.0286
7	인터넷	0.0276
8	홈페이지	0.0250
9	기획	0.0237

보다 명확한 핵심단어를 도출하기 위하여 직무 내 동시등장 빈도율이 2%이하에 해당하는 단어들을 배제한 후 직무군집 3의 동시등장 빈도율이 큰 단어부터 확인할 때 〈표 2-41〉와 같이 ‘디자인’, ‘웹’, ‘관리’, ‘포토샵’, ‘쇼핑몰’, ‘설치’, ‘인터넷’, ‘홈페이지’, ‘기획’의 순으로 나타났다.

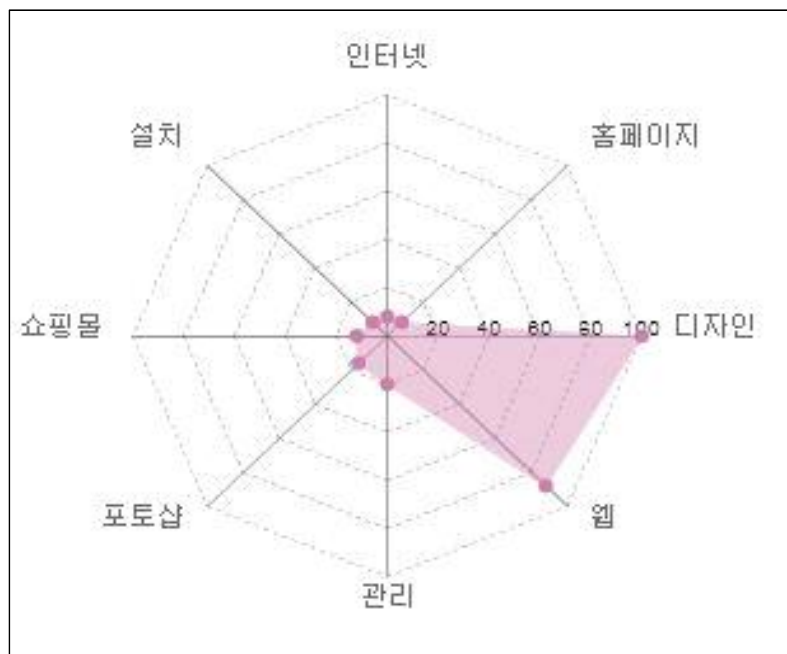
도출된 핵심단어 중 ‘설치’가 직무와 어떤 상관관계가 있는 지 확인하기 위하여 원 상태의 채용공고 문서를 확인한 결과 콘텐츠 배포와 같은 의미로 사용되며 웹 콘텐츠 디자이너 직무의 일부분으로 요구되는 경우가 있었음을 확인하였다. 따라서 직무군집 3은 주로 온라인 쇼핑몰 및 기업 홈페이지에서 필요로 하는 디자인을 포토샵 등의 그래픽 도구를 통해 제작하며, 쇼핑몰 콘텐츠 디자인을 기획하고 운영하는 역할을 수행하는 직무라고 할 수 있다.

[그림 2-75] 직무 3의 핵심단어 비율 그래프



이 결과를 시각화 한 결과는 [그림 2-75]와 같다. [그림 2-75]는 직무 3의 핵심 단어 비율 그래프이며 이 결과 웹과 디자인, 관리의 측면이 크게 나타났다.

[그림 2-76] 직무 3의 핵심단어 방사형 그래프



[그림 2-76]은 직무군집 3에 대한 핵심단어의 동시등장 빈도율을 100점으로 환산한 방사형 그래프이다. 시각화 된 결과를 보면 주로 디자인, 웹, 관리 등의 핵심단어들이 제시됨을 확인할 수 있다. 마찬가지로 [그림 2-77]과 같이 직무 3 군집에 대한 핵심단어 기반의 워드클라우드 시각화를 정리하였다. 결과 웹 쇼핑몰에 대한 디자인을 직무로 하고 있지만 웹 페이지의 제작, 운영, 사진, 블로그 등이 제시됨에 따라 일반적인 디자이너의 직무보다는 웹 표준에 맞는 콘텐츠를 운영 및 관리하는 직무라고 판단할 수 있다.

[그림 2-77] 직무 3의 핵심단어 워드클라우드



분석 결과를 바탕으로 직무 3의 정확한 직무 명칭과 직무 내용을 규정하기 위하여 <표 2-42>와 같이 국제직무분류표준의 정의를 비교하였다. 이 중 SOC와 O*Net에서 웹개발자(Web Developer)의 직무에 일부로서 콘텐츠 및 디자인 요소 등에 대한 디자인 수정 및 유지보수를 진행하는 형태로 합쳐진 직무로 설명되어 있다. 또한 ISCO는 멀티미디어 전문가로서 정보를 인쇄, 영화, 오디오 시각적 매체의 디지털 및 기타 전자적 형태로 전달하기 위한 청각 및 시각적 콘텐츠를 디자인하는 직무이며 컴퓨터 게임, 영화, 뮤직 비디오, 인쇄 매체와 광고에 사용하기 위해 그래픽 특수 효과, 애니메이션 또는 기타 시각 이미지를 디자인하는 직무를 포괄하는 것으로 제시되고 있다.

〈표 2-42〉 웹 콘텐츠 디자이너 직무 관련 국제직무분류표준

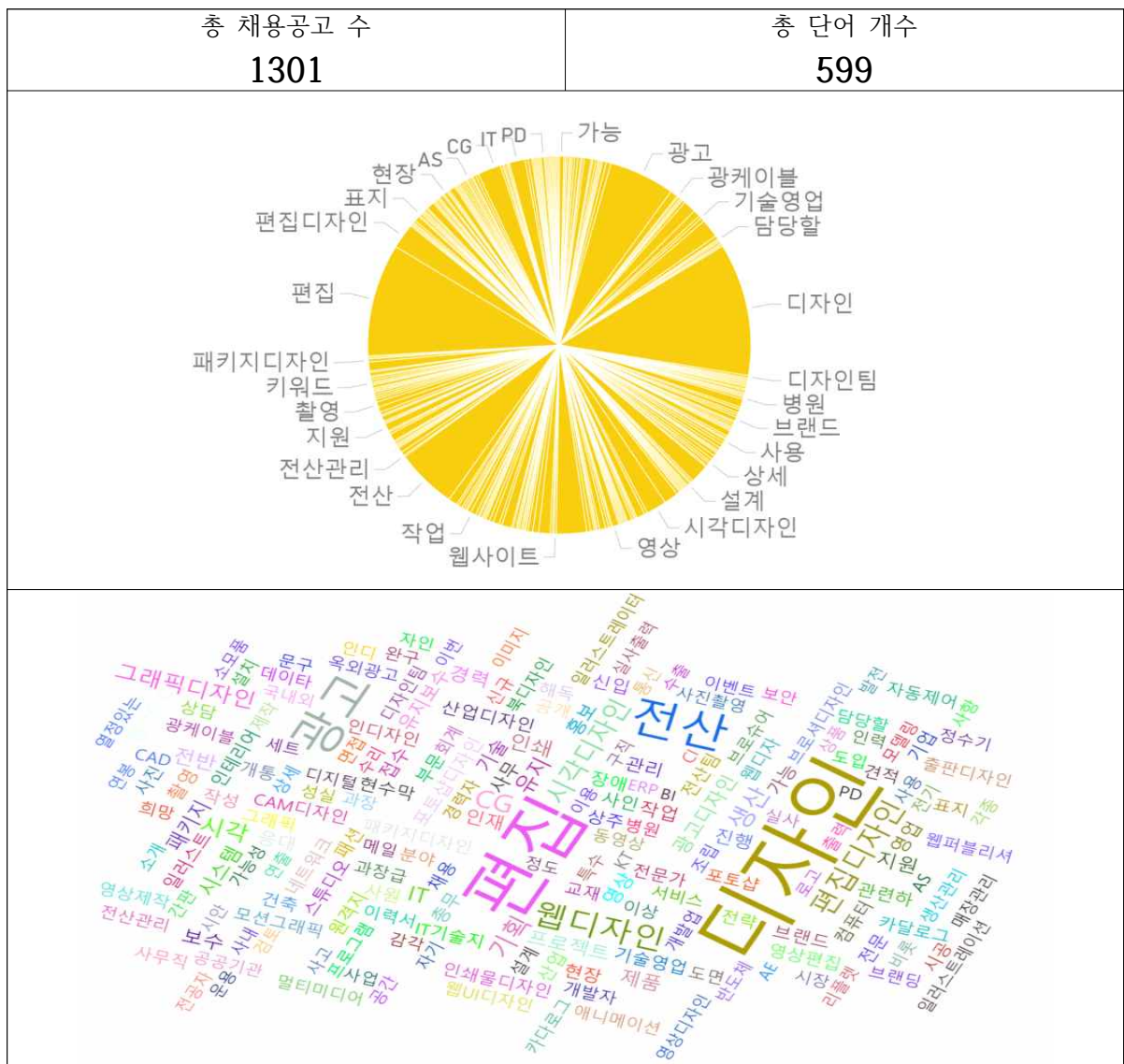
소스	분류코드	직무	정의
ISCO-08	2166	Graphic and Multimedia Designers	- 정보를 인쇄, 영화, 오디오 시각적 매체의 디지털 및 기타 전자적 형태로서 전달하기 위한 청각 및 시각적 콘텐츠를 디자인 - 컴퓨터 게임, 영화, 뮤직 비디오, 인쇄 매체와 광고에 사용하기 위해 그래픽 특수 효과, 애니메이션 또는 기타 시각 이미지를 디자인함
SOC-2016	15-1254	Web Developer	- 웹 사이트, 웹 응용 프로그램, 응용 프로그램 데이터베이스, 쌍방향 웹 인터페이스를 개발 - 브라우저와 장치에 호환 산업 표준에 충족하는 구조화된 코드를 평가 - 서버 측 코드와 프로세스, 웹사이트 성능, 확장성을 최적화 함 - 웹 사이트 인프라 구조를 개발하고 다른 컴퓨터 응용 프로그램과 웹 사이트를 통합
O*Net	15-1134.00	Web Developers	- 웹사이트를 설계, 개발, 수정 - 웹사이트 콘텐츠, 그래픽, 성능, 용량 개선을 위한 사용자 수요 분석 - 웹사이트와 다른 컴퓨터 애플리케이션 통합 - 호환되는 웹 형식으로 SW 웹과 멀티미디어 콘텐츠의 작성이 용이하도록 설계된 오디오 및 비디오 그래픽 요소들을 변환 “Multimedia Artists and Animators” (27-1014) 직무와는 별개임.
본 보고서의 분류		웹 콘텐츠 디자이너	SW를 활용하여 웹에서 활용하는 그래픽, 오디오, 비디오 관련 요소들에 대한 웹 포맷을 디자인하는 일을 담당한다.

그러나 기존 직무 내용이 작성된 시기(2010년)와 본 보고서 데이터의 SW 분야에 대한 채용공고 수집 자료의 시기(2016)가 다르며 웹 관련 영역의 활동이 활발한 상태에서 직무분야에 대한 구분이 이루어진 것으로 판단되어 이 중 직무군집 3은 세부

전문가로서 직무 4 군집의 내용과는 별도의 직무로서 구분하였다. 본 보고서에서는 비교 분석 결과에 따라 직무 3을 ‘웹 콘텐츠 디자이너’로 정의하였고 SW를 활용하여 웹에서 활용하는 그래픽, 오디오, 비디오 관련 요소들에 대한 웹 포맷을 디자인하는 일을 담당하는 직무로서 정의하였다.

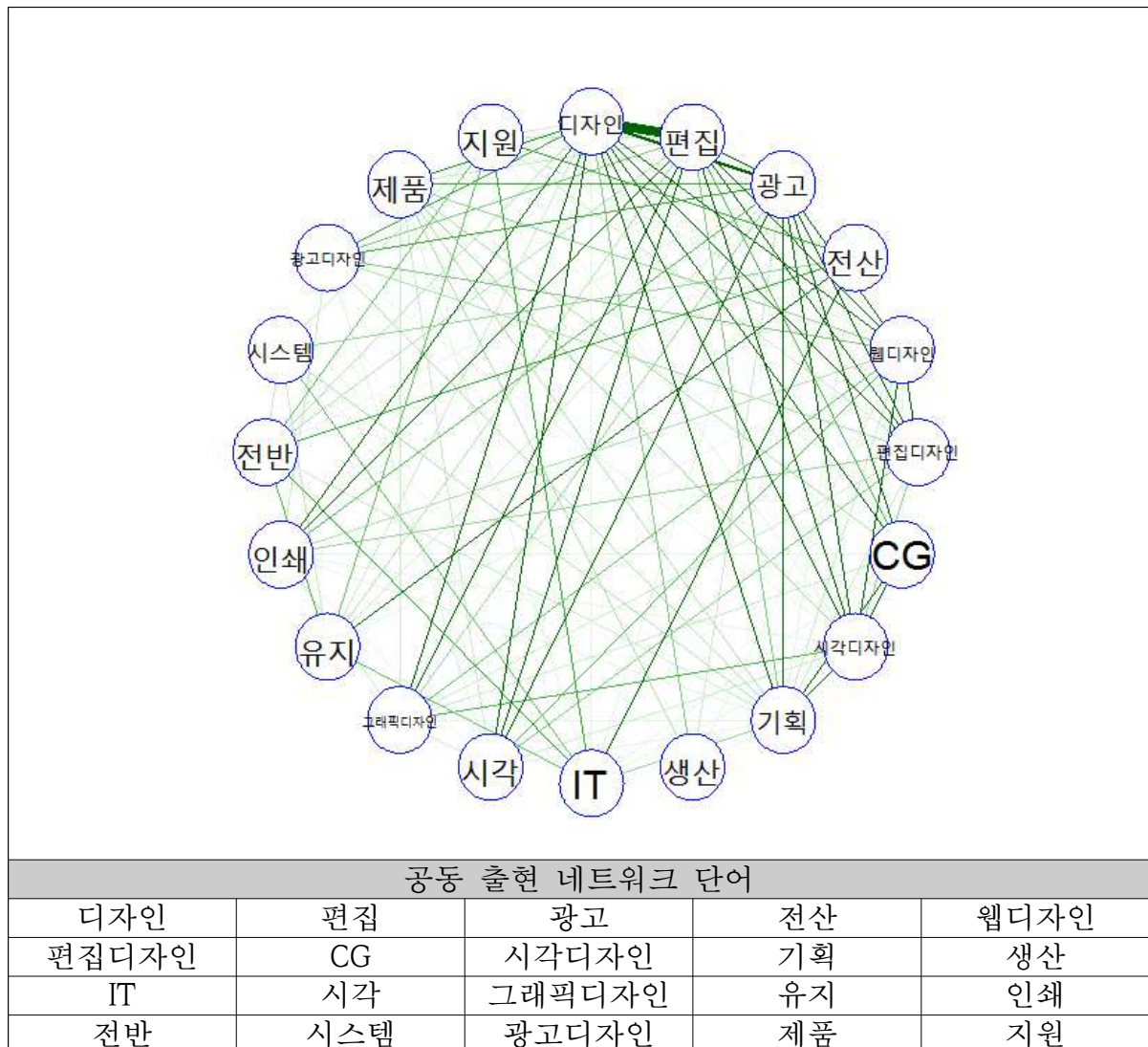
2.2.4) 그래픽 디자이너

[그림 2-78] 직무 4의 특성



직무 4에 해당하는 채용공고는 [그림 2-78]과 같이 총 1301건으로 559개의 단어가 등장한다. 전체 채용공고 내 직무 4에 대한 빈번하게 등장하는 단어는 ‘디자인’, ‘편집’, ‘광고’, ‘전산’ 이다. 특히 ‘디자인’ 과 관련하여 ‘웹디자인’, ‘그래픽디자인’, ‘편집디자인’ 등 다양한 디자인 단어들이 등장한다. 또한 직무 4에 대한 채용공고의 출현단어들을 바탕으로 워드클라우드를 확인한 결과 편집, 디자인, 광고, 웹디자인, 그래픽디자인, CAM디자인, 일러스트 등의 다양한 디자인 관련 단어들이 출현함에 따라 웹에서의 디자인 관련 직무임을 확인할 수 있다.

[그림 2-79] 직무 4의 단어 공동 출현 네트워크



그래픽 디자이너에 대한 연관 단어들을 추가로 확인하기 위하여 [그림 2-79]와 같이 소셜네트워크 분석을 통한 공동 출현 네트워크를 도출하였다. 그 결과 전체 단어 중에서는 {디자인-편집}, {디자인-광고}의 연관 관계가 강하게 존재한다. 이는 주요 디자인 관련 편집 및 광고 디자인 관련 공고가 많은 것으로 판단할 수 있다. 즉, 직무 4는 편집과 광고에 대한 역량을 기반으로 시각, 웹, 편집 등을 하는 직무를 갖는 것으로 보인다. 특히 ‘생산’의 경우 그래픽 요소 및 광고 디자인 결과물을 생산한다는 의미로 활용 되었으며 그에 따라 그래픽 디자이너는 주로 온라인 광고, 컴퓨터 그래픽, 제품 디자인을 기획하고 제작한다는 점을 확인할 수 있다.

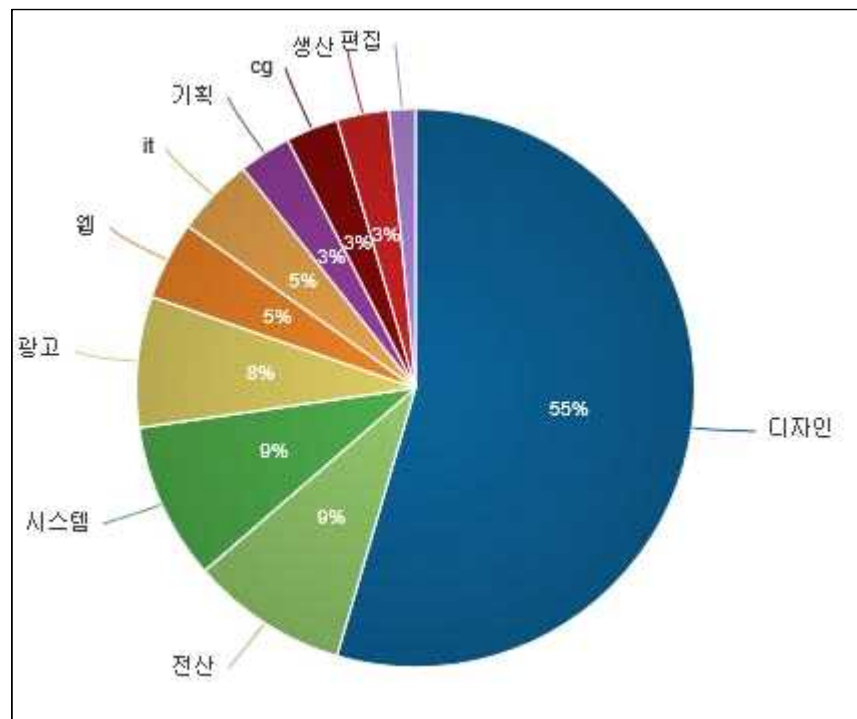
〈표 2-43〉 직무 4의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	디자인	0.2590
2	시스템	0.0222
3	전산	0.0627
4	광고	0.0577
5	IT	0.0369
6	웹	0.0344
7	생산	0.0300
8	cg	0.0294
9	기획	0.0241

보다 명확한 직무의 특성 확인을 위하여 직무군집 4에 대한 LDA 분석 결과를 확인해보면 〈표 2-43〉와 같이 동시등장 빈도율에 따른 핵심단어들을 확인할 수 있다. 이 중 직무 내 동시등장 빈도율이 2%이하인 단어들은 군집내 큰 영향력을 발휘하지 못하는 것으로 판단하여 핵심단어로 사용하지 않았다. 이 때 동시등장 빈도율이 큰 단어부터 확인할 때 ‘디자인’, ‘시스템’, ‘전산’, ‘광고’, ‘IT’, ‘웹’, ‘생산’, ‘CG’, ‘기획’이 주요 핵심단어인 것으로 판단할 수 있다. 이 때 동시등장 빈도율에 등장한 ‘생산’의 경우 직무와 어떤 상관관계가 있는 지 확인하기 위하여 원 상태의 채용공고 문서를 확인하였다. 그 결과 기계, 설비, 생활용품, 가전, 가구 등 다양한 분야의 제품을 생산하는 회사에서 제품을 그래픽 디자인하여 홍보이미지를 만들고자 하는 수요가 많다. 그 사례는 ‘자사에서 생산 중인 XX 제품 디자인, 생산하는 YY를 디자인 등’의 문건과 같아 직무와 관계있는 단어임을 판단하였다.

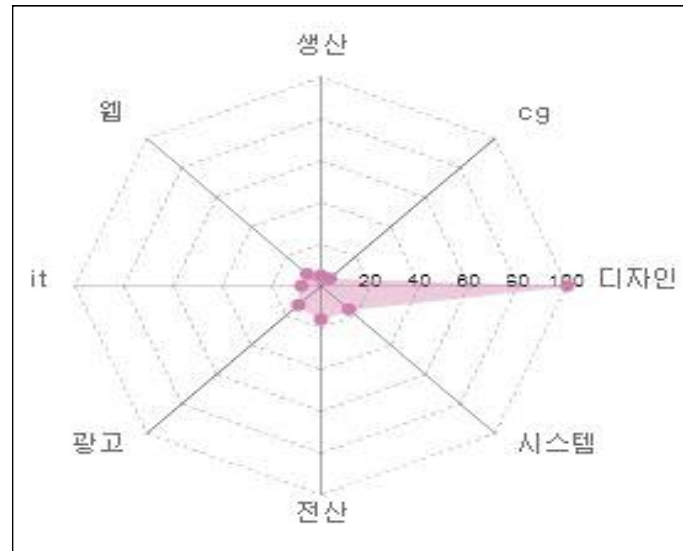
이 결과를 바탕으로 [그림 2-80]과 같이 직무군집 2-2에 대한 핵심단어 비율 그래프를 확인하였다. 결과에서 보이는 것과 같이 직무군집 4의 핵심단어 중 디자인은 55%에 달할 만큼 주요 직무 형태인 것으로나 나타나고 있으며 주로 시스템, 광고, 전산 상에서 활용할 수 있는 디자인 직무임을 알 수 있다.

[그림 2-80] 직무 4의 핵심단어 비율 그래프



마찬가지로 핵심단어를 100점으로 환산한 직무 4의 방사형그래프는 [그림 2-81]과 같이 디자인에 매우 치중된 채용공고의 핵심단어가 도출되었음을 명확히 확인할 수 있다. 추가적으로 해당 군집의 핵심단어 자료를 바탕으로 [그림 2-82]와 같이 핵심단어의 워드클라우드 결과를 시각화 하였다. 그 결과 직무군집 4는 광고, 그래픽디자인, 시스템, 전산, CG, 편집, 포토샵, 제품디자인 등 주로 그래픽을 다루는 직무임이 명확히 제시되었다.

[그림 2-81] 직무 4의 핵심단어 방사형 그래프



[그림 2-82] 직무 4의 핵심단어 워드클라우드



따라서 직무 4는 웹 페이지 및 웹 광고에서 필요로 하는 다양한 그래픽 디자인을 제작하는 역할을 수행한다고 판단할 수 있다. 그에 따라서 직무군집 4에 대한 직무명과 특징을 정의하기 위하여 국제직무분류표준 자료를 바탕으로 해당 군집에 대한 특성을 유추하였다. <표 2-44>와 같이 국제직무분류표준에서는 그래픽 및

멀티미디어 디자이너 직무로서 제시되고 있으며 정보를 인쇄, 영화, 오디오 시각적 매체의 디지털 및 기타 전자적 형태로 전달하기 위한 청각 및 시각적 콘텐츠를 디자인하는 직무를 의미하고 있다. 또한 컴퓨터 게임, 영화, 뮤직 비디오, 인쇄 매체와 광고에 사용하기 위해 그래픽 특수 효과, 애니메이션 또는 기타 시각 이미지를 디자인하는 것으로 직무의 특성이 제시된다.

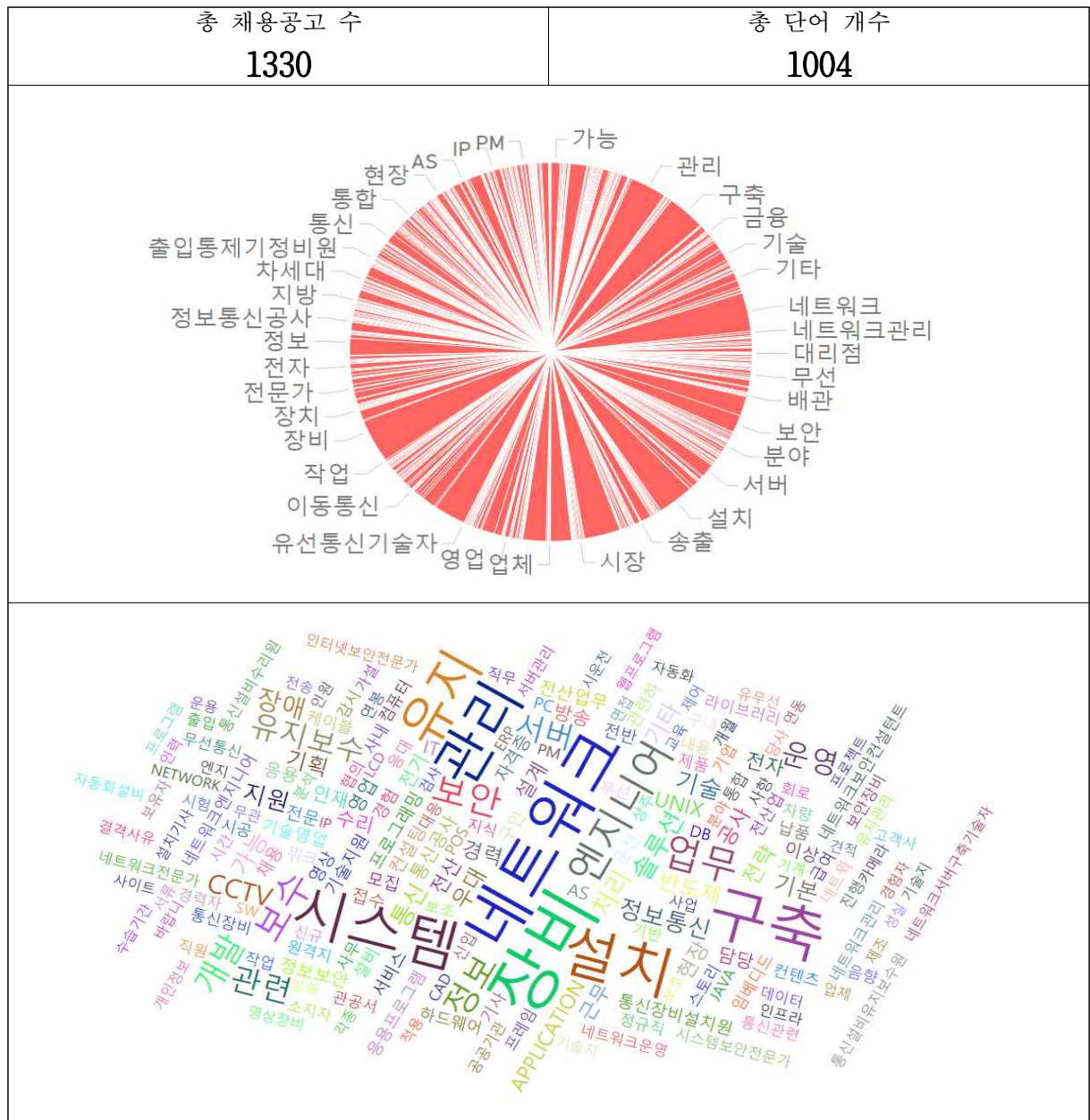
기존 분류의 형태와 본 연구의 핵심단어들의 특성이 일치됨을 바탕으로 본 보고서에서는 직무 4를 ‘그래픽 디자이너’로 명칭을 확정하였다. 그에 따라 그래픽 디자이너 직무의 정의는 특정 거래 및 촉진 활동을 위한 로고, 그래픽 등을 디자인 및 창작하는 일을 담당하는 직무로 확정하였다.

〈표 2-44〉 그래픽 디자이너 직무 관련 국제직무분류표준

소스	분류코드	직무	정의
ISCO-08	2166	Graphic and Multimedia Designers	- 정보를 인쇄, 영화, 오디오 시각적 매체의 디지털 및 기타 전자적 형태로 전달하기 위한 청각 및 시각적 콘텐츠를 디자인함 - 컴퓨터 게임, 영화, 뮤직 비디오, 인쇄 매체와 광고에 사용하기 위해 그래픽 특수 효과, 애니메이션 또는 기타 시각 이미지를 디자인함
SOC-2016	27-1024	Graphic Designer	- 패키징, 진열, 로고와 같은 특정 거래 및 촉진을 위한 그래픽을 디자인 및 개발함 - 장식 및 예술적 효과를 달성하기 위해 다양한 매체들을 활용함
O*Net	27-1024.00	Graphic Designers	
본 보고서의 분류		그래픽 디자이너	특정 거래 및 촉진 활동을 위한 로고, 그래픽 등을 디자인 및 창작하는 일을 담당한다.

2.2.5) 컴퓨터 네트워크 관리자

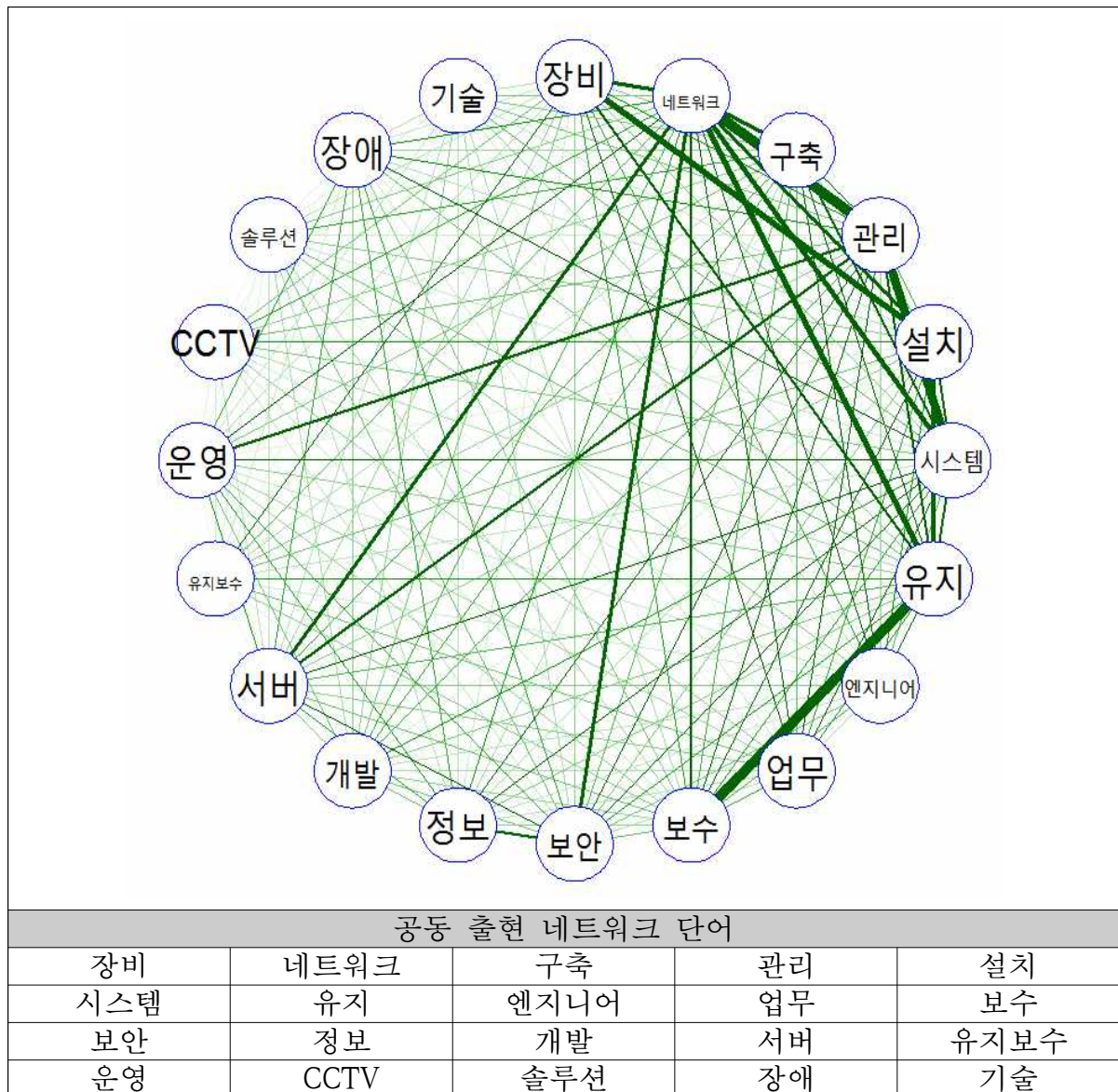
[그림 2-83] 직무 5의 특성



분석결과에서 제시된 직무 5에 해당하는 채용공고의 수는 [그림 2-83]과 같이 총 1330건으로 1004개의 단어가 등장하고 있다. 이 중에서 전체 단어들 중 등장 빈도가 가장 빈번한 순으로는 ‘네트워크’, ‘장비’, ‘시스템’, ‘설치’, ‘구축’ 등의 단어가 존재한다. 직무 5에 대한 워드클라우드 결과 장비, 네트워크, 설치, 구축,

시스템, 보안 등과 같은 웹 및 네트워크 관련 직무 단어가 출현하였음을 알 수 있다. 직무 5 군집에 대한 전체 단어의 빈도수를 조사한 결과 관리, 네트워크, 장비, 정보, 보안 등의 단어들이 높은 비율을 차지하는 것을 확인할 수 있다. 그와 함께 해당 직무군집에서 사용된 전체 단어를 대상으로 빈도에 대한 워드클라우드를 확인한 결과 네트워크, 시스템, 장비, 엔지니어, 설치, 구축, 관리, 보안 등의 단어들이 주로 발생하는 것을 판단할 수 있으며 이는 네트워크 관리와 관련한 직무임을 보여주고 있다.

[그림 2-84] 직무 5의 단어 공동 출현 네트워크

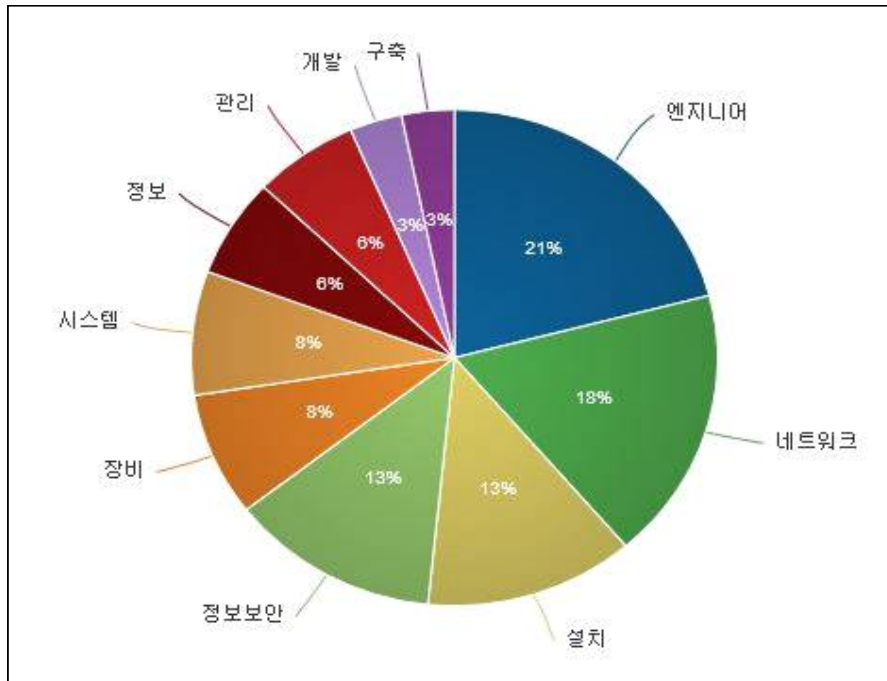


직무군집 5에 대한 출현한 전체의 단어를 기반으로 [그림 2-84]와 같이 소셜네트워크 분석을 수행한 공동 출현 네트워크 결과를 확인하였다. 분석 결과 ‘네트워크’ 단어를 중심으로 주로 {네트워크-장비}, {네트워크-구축}, {네트워크-관리-설치}, {네트워크-시스템}, {네트워크-서버}와 같은 관계가 함께 등장하는 것으로 나타났다. 또한 {유지-보수}, {네트워크-관리}의 관계도 매우 강한 관계가 형성되어 있다. 따라서 직무군집 5의 경우 네트워크와 관련한 서버와 시스템을 구축하고 관리하며 유지 보수하는 역량을 요구하는 것으로 판단할 수 있다.

〈표 2-45〉 직무 5의 핵심단어 목록

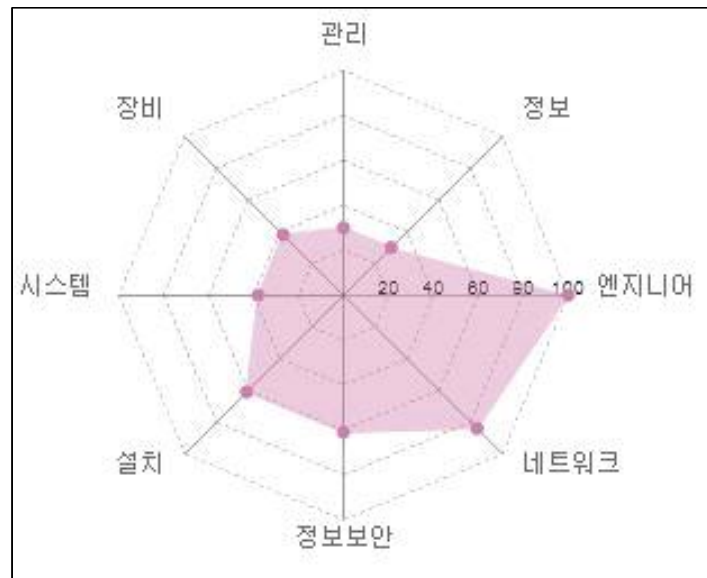
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	엔지니어	0.1336
2	네트워크	0.1117
3	정보보안	0.0899
4	설치	0.0854
5	시스템	0.0557
6	장비	0.0506
7	관리	0.0497
8	정보	0.0474
9	구축	0.0265
10	개발	0.0227

[그림 2-85] 직무 5의 핵심단어 비율 그래프



도출된 전체 단어 자료들을 바탕으로 직무 5에 대한 LDA 분석 결과를 활용하여 군집 5내에서의 동시등장 빈도율에 따른 핵심단어들을 <표 2-45>와 같이 도출하였다. 이 중 직무 내 동시등장 빈도율이 2%이하인 단어들은 역시 사용하지 않는 범위에서 동시등장 빈도율이 큰 단어부터 확인할 때 ‘엔지니어’, ‘네트워크’, ‘정보보안’, ‘설치’, ‘시스템’, ‘장비’, ‘관리’, ‘정보’, ‘구축’, ‘개발’의 순으로 주요 핵심단어가 나타났다. 핵심단어 결과와 관련하여 시각화를 진행한 결과 [그림 2-85]와 같이 제시되었다. 군집 5의 경우 엔지니어, 네트워크, 설치와 정보보안에 요구가 큰 핵심단어들이 제시되었다.

[그림 2-86] 직무 5의 핵심단어 방사형 그래프



이를 바탕으로 [그림 2-86]과 같이 핵심단어들을 100점으로 환산한 방사형 그래프를 도출하였다. 그 결과 직무 5는 엔지니어 직무이지만 그 중에서도 주로 네트워크와 정보보안, 시스템, 설치 등과 관련한 직무 내용이 큰 것으로 나타났다.

[그림 2-87] 직무 5의 핵심단어 워드클라우드



마찬가지로 [그림 2-87]과 같이 핵심단어 워드클라우드 시각화를 진행한 결과 시스템, 정보보안, 정보, 네트워크 구축 등과 같은 직무의 특성이 반영되는 결과가 나온 것을 확인하였다.

그에 따라 본 보고서는 직무 5의 명칭 및 정의를 확정하기 위하여 <표 2-46>와 같이 국제직무분류표준을 활용하였다. 네트워크 및 정보보안과 관련하여 주로 컴퓨터 네트워크 전문가 직무들이 유사 정의를 차용하고 있었다. ISCO의 경우 네트워크 하드웨어와 SW를 개발, 개선, 관리 및 유지보수 직무로 정의하고 있었으며 SOC와 O*NET의 경우 구성된 조직의 지역 네트워크(LAN), 광역 통신망(WAN), 인터넷 시스템 또는 네트워크 시스템의 세그먼트 지원 설치, 모니터 네트워크와 필요한 정비 네트워크 가용성을 지원하기 위해 규정을 수행할 수도 모든 시스템 사용자에게 네트워크 가용성과 보안사항 보장 등을 주요 업무로 지정하고 있다.

따라서 본 보고서에서도 유사 직무 개념을 확인한바 직무 5를 ‘컴퓨터 네트워크 관리자’ 직무로 명칭을 확정 하였다. 따라서 ‘컴퓨터 네트워크 관리자’는 네트워크에 대한 운영과 관리 지침을 구성하고 효율적 관리를 수행한다. 또한 전체 네트워크 시스템의 구성 현황을 파악하고 실시간으로 시스템의 운영 상태를 모니터링 하는 직무로서 네트워크 관련 HW 및 SW 자원을 효과적으로 관리하고 네트워크에서의 보안정책을 수립하고 외부로부터 해킹이 시도될 수 없도록 체계적 운영을 하는 역할을 수행하는 직무로서 정의한다.

〈표 2-46〉 컴퓨터 네트워크 관리자 관련 국제직무분류표준

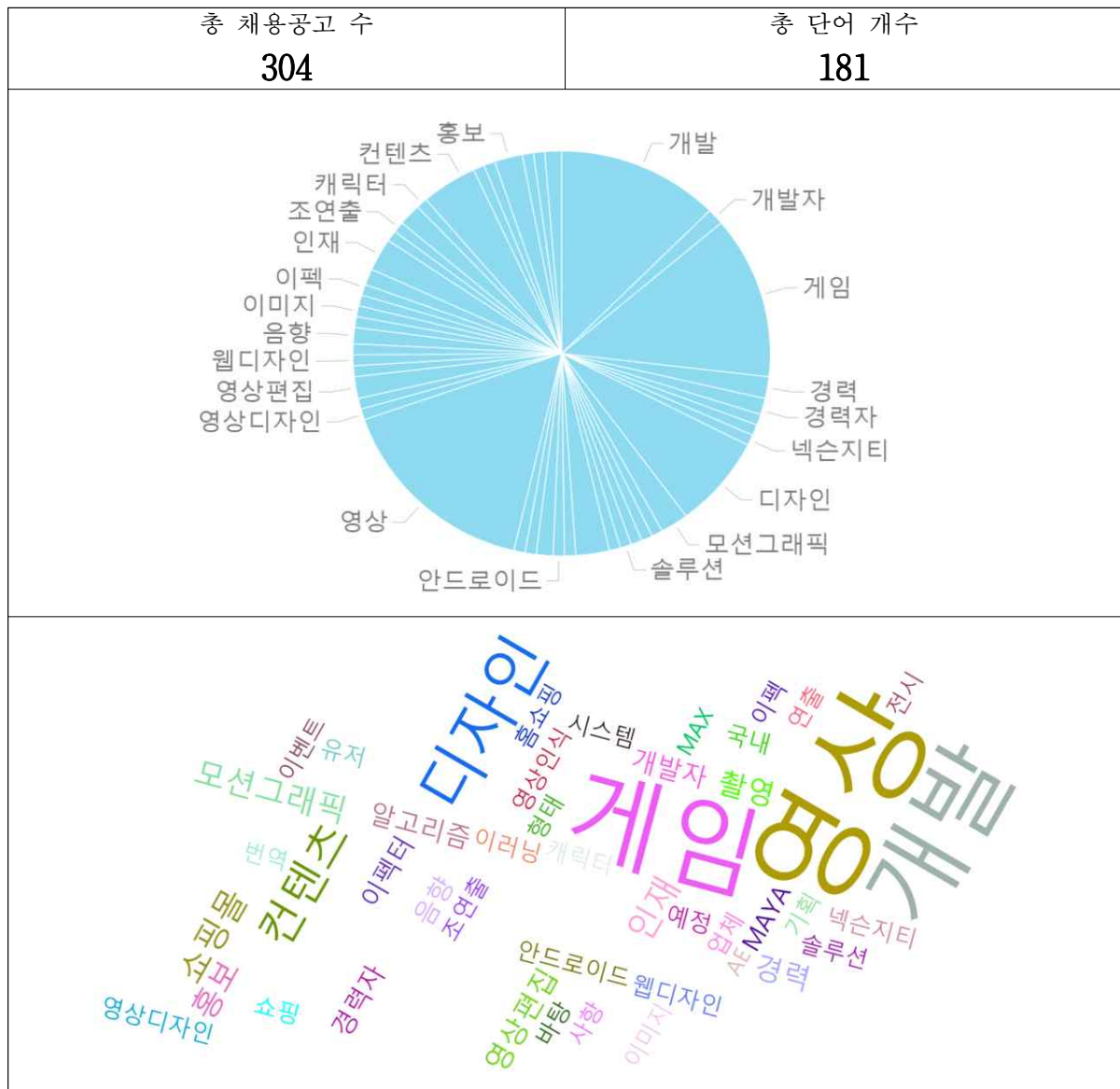
소스	분류코드	직무	정의
ISCO-08	2523	Computer network professionals	• 네트워크 설계를 위한 전략 연구, 분석 및 추천 • 네트워크 하드웨어와 SW를 개발, 개선, 관리 및 유지보수 • 네트워크에 대한 모니터링, 문제해결, 성능 최적화
SOC-2016	15-1244	Network and Computer Systems Administrators	• 구성한 조직의 지역 네트워크(LAN), 광역 통신망(WAN), 인터넷 시스템 또는 네트워크 시스템의 세그먼트 지원 설치 • 모니터 네트워크와 필요한 정비 네트워크 가용성을 지원하기 위해 규정을 수행할 수도 모든 시스템 사용자에게 네트워크 가용성을 보장 • 웹사이트의 오류 없는 정확한 운영을 보장하기 위한 웹사이트 성능 모니터링 및 테스트
O*Net	27-1024.00	Network and Computer Systems Administrators	• 네트워크 모델링, 분석, 기획, 네트워크 및 데이터 통신 하드웨어와 SW 간 협력을 지원함 • 컴퓨터 사용자 지원 전문가들과 컴퓨터 네트워크 지원 전문가들을 감독 • 관리자 네트워크의 보안 관리
본 보고서의 분류		컴퓨터 네트워크 관리자	• 네트워크 관련 HW 및 SW 자원을 효과적으로 관리하고 네트워크에서의 보안정책을 수립하고 외부로부터 해킹이 시도될 수 없도록 체계적 운영을 하는 역할을 수행함.

이와 관련하여 컴퓨터 네트워크 관리자의 직무 특성을 정리하면 다음과 같다.

- 네트워크의 사용자 컴퓨터, 서버 및 주변기기 등에 대한 네트워크 주소체계를 설정·부여하고 주소의 재설정을 조정 관리
- 실시간 네트워크 트래픽을 분석하여 업그레이드 계획 반영
- 네트워크 시스템 및 장비의 고장 및 이상 발견 시 문제점 파악·분석·복구
- 네트워크 관리 시스템의 기능과 특성을 중심으로 성능 및 시스템 구성과 오류를 관장함
- 네트워크에서 활용되는 프로토콜과 네트워크 서비스를 지속적으로 갱신하여 네트워크 성능과 기능 향상
- 네트워크 시스템의 보안 관련 정책 수립 및 해킹방지 관련 기술적 보안과 관리적 보안 이행
- 네트워크 이용자 정보의 등록, 부여, 삭제 등을 관리

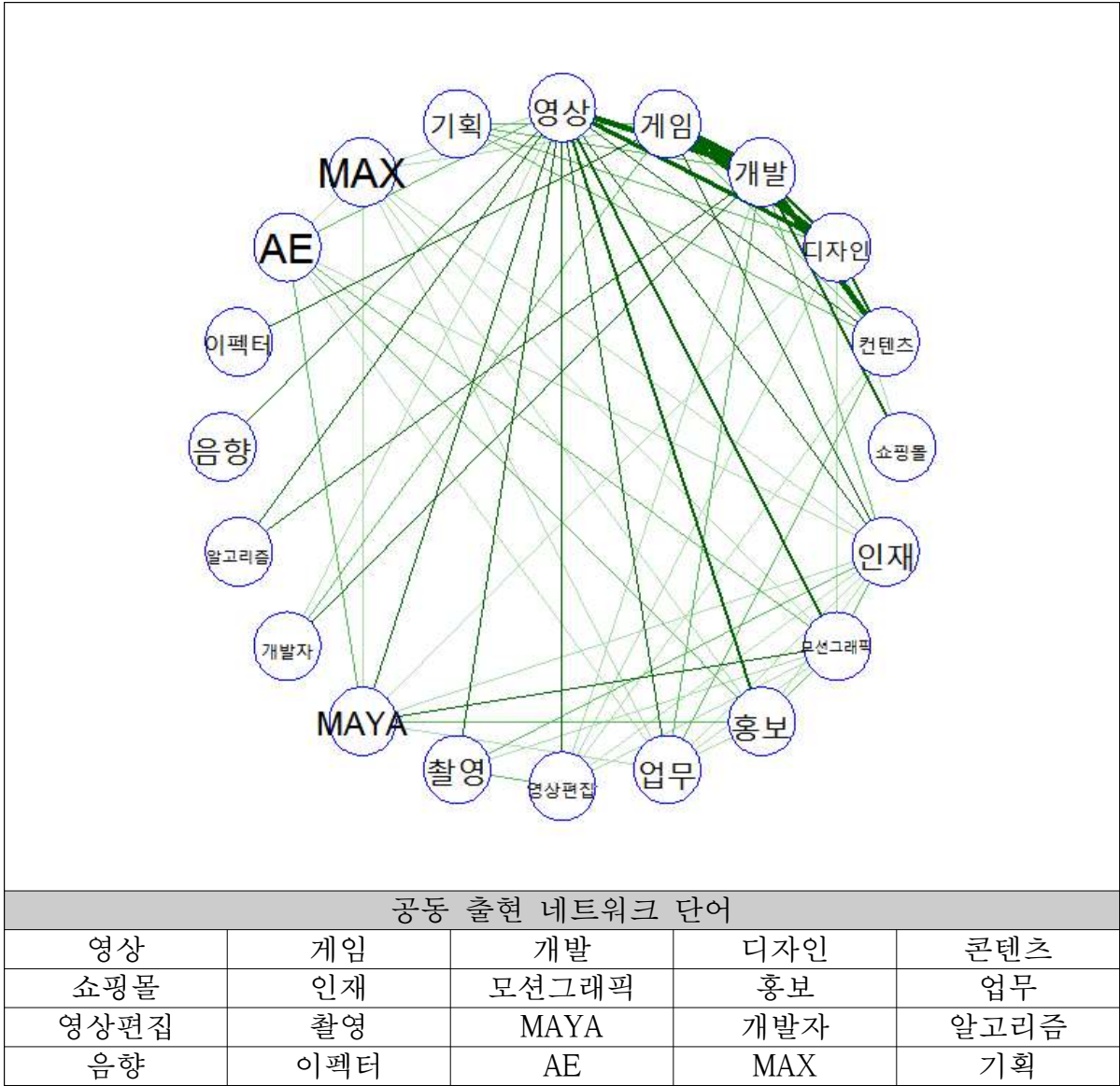
2.2.6) 소프트웨어 개발자

[그림 2-88] 직무 6의 특성



SW 개발자에 해당 하는 채용공고는 [그림 2-88]과 같이 총 304건으로 181개의 단어가 등장한다. 이 공고 중 빈번하게 등장하는 단어는 ‘게임’, ‘영상’, ‘개발’, ‘디자인’ 이다. 이와 마찬가지로 전체 출현 단어에 대한 워드클라우드 결과를 보면 게임, 영상, 개발, 디자인, 콘텐츠 등 다양하게 출현하고 있으며 주로 디자인, 게임 등에 대한 개발이 주를 이루는 것으로 보인다.

[그림 2-89] 직무 6의 단어 공동 출현 네트워크

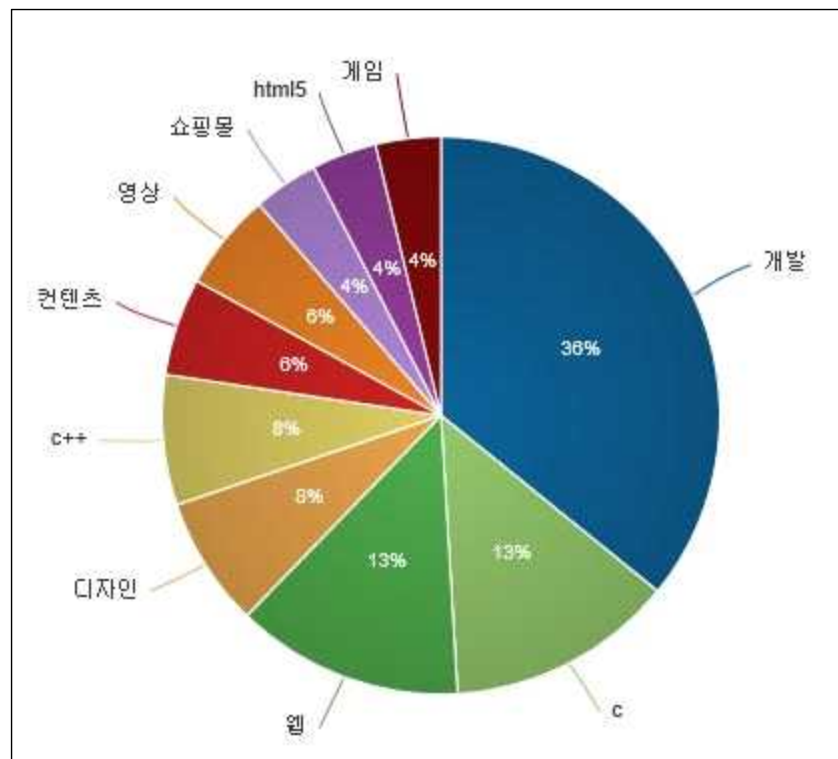


[그림 2-89]는 SW 개발자의 공동 출현 네트워크이다. {영상-개발-디자인}, {게임-개발}, {콘텐츠-개발}의 관계가 존재한다. 따라서 직무 6의 경우 영상과 관련된 게임 개발이나 디자인, 콘텐츠 개발의 역량을 요구하는 것으로 보인다.

〈표 2-47〉 직무 6의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	개발	0.0192
2	웹	0.0761
3	c	0.0703
4	c++	0.0454
5	디자인	0.0426
6	영상	0.0395
7	콘텐츠	0.0321
8	게임	0.0266
9	html5	0.0263
10	쇼핑몰	0.0257

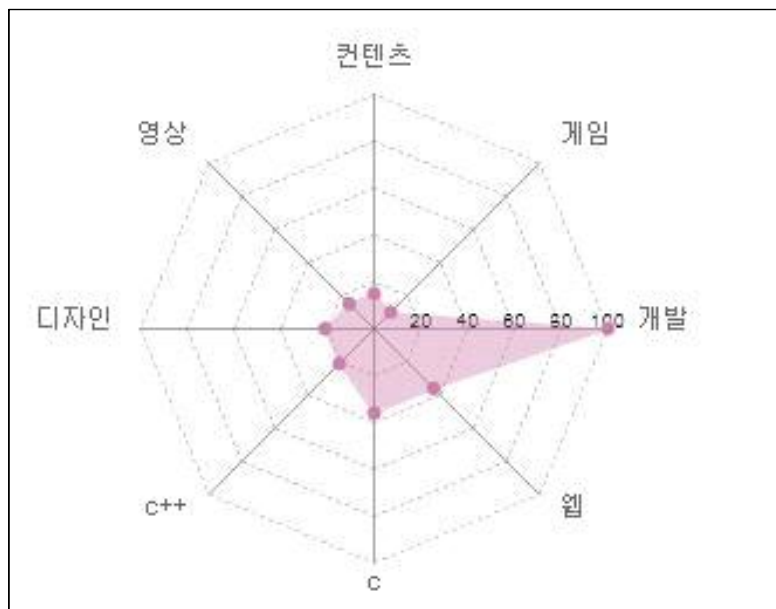
[그림 2-90] 직무 6의 핵심단어 비율 그래프



컴퓨터 네트워크 관리자에 대한 직무 내 분류를 위한 LDA 분석 결과를 확인해보면 동시등장 빈도율에 따른 핵심단어들을 〈표 2-47〉과 같이 확인할 수 있다. 이 중 직무 내 동시등장 빈도율이 2%이하인 단어들을 활용하지 않고 동시등장 빈도율이 큰 단어부터 핵심단어로 사용할 때 ‘개발’, ‘웹’, ‘c’, ‘c++’,

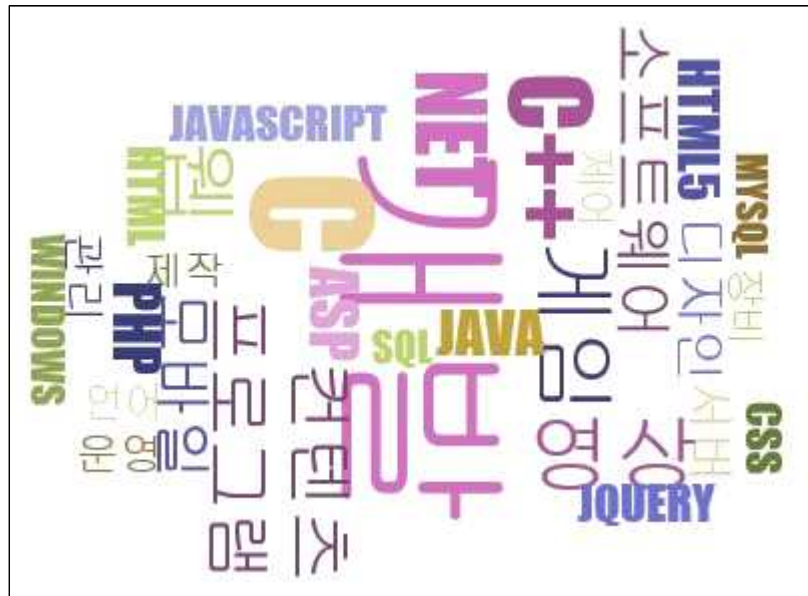
‘디자인’, ‘영상’, ‘콘텐츠’, ‘게임’, ‘html5’, ‘쇼핑몰’의 순으로 주요 요구 역량인 핵심단어가 나타났다. 이를 바탕으로 직무 6 군집에 대한 핵심단어를 바탕으로 비율 그래프를 시각화 하였다. 시각화 결과는 [그림 2-90]과 같이 나타났으며 주로 개발, c, 웹과 같은 핵심단어들이 주를 이루고 있다.

[그림 2-91] 직무 6의 핵심단어 방사형 그래프



그에 따라 직무 6 군집에 대한 핵심단어들을 100점으로 환산한 방사형 그래프를 도출하였다. 이 결과는 [그림 2-91]과 같이 나타났으며 주로 웹, C, C++과 같은 웹과 프로그래밍 언어가 주를 이루고 있음을 확인할 수 있다. 이와 마찬가지로 직무 6에 대한 핵심단어들을 워드클라우드로 시각화 한 결과는 [그림 2-92]와 같다. 워드클라우드 결과는 개발, 콘텐츠, C, ASP, JAVA, JQUERY 등 소프트웨어 개발 관련 핵심단어들이 주로 중요한 것으로 나타나고 있다. 따라서 직무 6은 온라인 쇼핑몰이나 웹 페이지 등에서 소프트웨어를 개발하고 C 또는 JAVA 등과 같은 프로그래밍을 통한 소프트웨어 개발자로서의 직무로서 판단할 수 있다.

[그림 2-92] 직무 6의 핵심단어 워드클라우드



분석 결과와 함께 본 연구과제는 국제직무분류표준을 활용하여 직무 6이 어떤 직무를 나타내는 것인지 판단하기 위하여 <표 2-48>와 같이 ISCO, SOC, O*Net의 분류를 참고하였다. ISCO의 경우 해당 유형의 직무를 소프트웨어 개발자로 정의하고 있으며 주로 SW 개발자 연구, 분석하기 위하거나 새로운 SW 응용 프로그램과 기존 운영 시스템의 요건을 평가하고 개발 요구사항을 충족하기 위한 SW 솔루션의 설계, 개발, 테스트 및 유지 보수하는 직무로 정의하고 있다. SOC와 O*Net의 경우도 소프트웨어 개발자로서 유사한 정의를 채택하고 있어 본 보고서에서는 직무 6을 소프트웨어 개발자로 정의하였으며 기업이나 개인 등이 사용할 수 있는 오피스 프로그램, 회계 프로그램, 데이터베이스 툴, 통계 프로그램 등 각종 SW를 개발하고 컴퓨터시스템의 사용 환경에 따라 SW의 환경을 변경하는 일을 담당하는 직무로서 정의하였다.

〈표 2-48〉 소프트웨어 개발자 직무 관련 국제표준

소스	분류코드	직무	정의
ISCO-08	2131, part 2512	Software developers	• SW 개발자 연구, 분석하기 위하여나 새로운 SW 응용 프로그램과 기존 운영 시스템의 요건을 평가함 • 개발 요구사항을 충족하기 위한 SW 솔루션의 설계, 개발, 테스트 및 유지보수
SOC-201 6	15-1252	Software Developers	• 의료 산업, 군사, 통신, 항공, 사업, 과학, 그리고 일반적인 컴퓨팅 활동을 위한 운영시스템 수준의 SW, 컴파일러, 네트워크 분배 SW를 연구, 설계, 개발 및 테스트함
O*Net	15-1133. 00	Software Developers, Systems Software	• 운영상의 특수성을 설정하고 SW 요구사항을 분석 및 공식화 함 • 임베디드 시스템 SW를 설계하고 컴퓨터 공학, 공학 및 수학적 분석을 위한 원칙 및 기술 적용 관리
본 보고서의 분류		SW 개발자	• 기업이나 개인 등이 사용할 수 있는 오피스 프로그램, 회계 프로그램, 데이터베이스 툴, 통계 프로그램 등 각종 SW를 개발한다. • 컴퓨터시스템의 사용 환경에 따라 SW의 환경을 변경하는 일을 담당한다.

SW 개발자는 새로운 SW나 기존의 응용프로그램을 분석하고 평가하여 사용자의 요구사항을 파악하고, 사용자의 요구사항을 충족시키기 위한 SW를 기획하거나 개발하고, 테스트 및 운영하는 역할을 수행한다. 또한 업이나 개인 등이 사용할 수 있는 오피스 프로그램, 회계 프로그램, 데이터베이스 툴, 통계 프로그램 등 다양한 SW를 개발하고 컴퓨터시스템의 사용 환경에 따라 SW의 환경을 변경하는 일을 담당한다. 따라서 기업이나 개인 등이 사용할 수 있는 워드프로세서에서부터 회계 관리, 데이터베이스, 통계처리, 문서결재 프로그램 등 각종 SW를 개발하고 시스템의 환경에 따라 SW 환경 변경의 업무를 담당하며 직무에 대한 요약은 하면 다음과 같다.

- SW 기능 · 성능 및 사용자의 요구사항을 분석하고 정의
- 요구사항을 기본으로 응용SW 설계
- 응용SW를 제작하고 테스트
- 프로세스나 기계제어, 계측, 로봇, 통신, 환경감시, 원격감지, 의료 및 관련 공학, 과학 및 산업 애플리케이션을 위한 실시간 환경에서 컴퓨터기술의 적용요건을 분석
- 프로그래머의 작업을 감독
- 오퍼레이션 명세서를 설정하고 응용SW 요건을 체계화
- 사용자의 요구사항을 조사하고 기록하는 응용SW 개발자의 일 보조
- 응용SW 설명서를 작성하는데 참여하며 컴퓨터 프로그램을 설계, 작성하고 유지.
- 응용SW 개발자가 명기한 목적 및 문제들을 분석
- 응용SW 개발자가 제시한 형식화된 솔루션을 상세한 프로그램 명세서로 전환하고, 프로그램의 타당성과 로직을 검사하고 수정
- 컴퓨터 장비 공급자, 시스템 설계자, 프로그래머 및 컴퓨터 오퍼레이터들이 사용하는 시스템 SW의 상태, 작동 및 유지관리에 관한 보고서, 매뉴얼 및 문서를 작성

2.2.7) 시스템 개발자

[그림 2-93] 직무 7의 특성

총 채용공고 수

1313

총 단어 개수

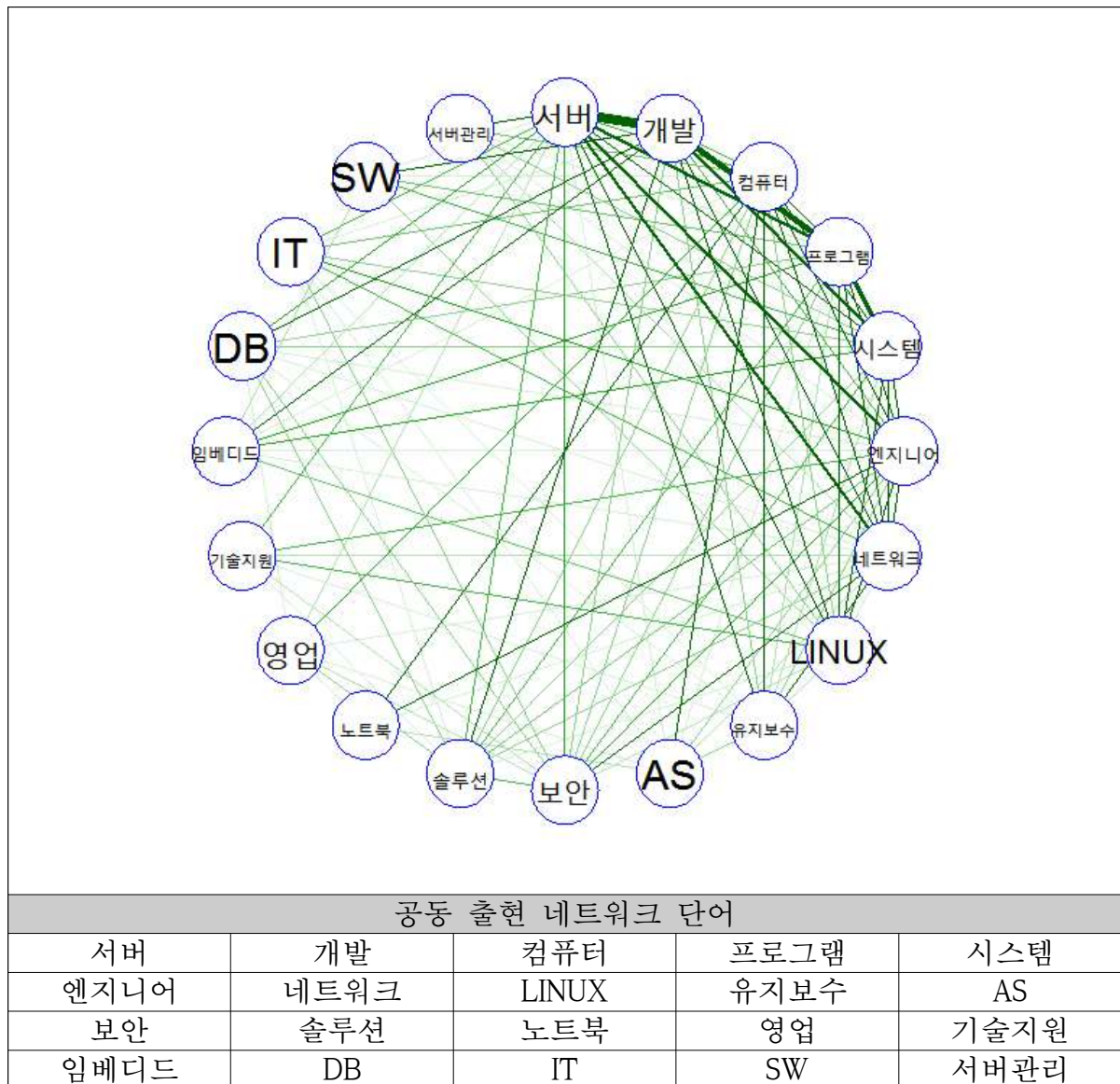
670

태그 클라우드 (태그 개수)

태그	개수
엔지니어	100
서버	80
컴퓨터	70
프로그래밍	60
네트워크	50
시스템	40
개발	30
인력	20
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10
프로젝트	10
프로그램	10
컴퓨터수리기사	10
컴퓨터	10
중급	10
제안	10
전기	10
인력	10
유지	10
엔진	10
소프트	10
스토리지	10
시스템개발	10
서버개발	10
서버	10
분석	10
보안	10
리눅스	10
노트북	10
기술영업	10
고객	10
개발	10
SI	10
MFC	10
IT	10
ENGINEER	10
AS	10
형태	10

직무 7 군집에 해당 하는 채용공고는 [그림 2-93]과 같이 총 1313건으로 670개의 단어가 등장한다. 해당 군집내에서의 전체 단어 중 빈번하게 등장하는 단어는 ‘서버’, ‘컴퓨터’, ‘프로그램’, ‘개발’, ‘시스템’, ‘엔지니어’ 이다. 또한 ‘OS’, ‘LINUX’, ‘UNIX’ 등 시스템 수준의 단어들이 등장함에 따라 해당 직무의 특성이 시스템의 개발 관련 직무임을 확인할 수 있다. 이와 함께 워드클라우드 결과에서도 서버, 프로그램, 개발, 컴퓨터, 시스템, 엔지니어 등과 같은 단어들이 출현하여 시스템 및 서버 관련 개발 직무에 가까움을 확인할 수 있다.

[그림 2-94] 직무 7의 단어 공동 출현 네트워크



[그림 2-94]는 직무 7 군집 내 전체 단어에 대한 공동 출현 네트워크이다. ‘개발’ 과 관련하여 {서버-개발}, {프로그래밍 개발}의 연관 관계가 존재하며, ‘서버’ 와 관련하여 {서버-프로그램}, {서버-엔지니어}, {서버-네트워크}의 연관 관계가 형성되어 있다. 따라서 직무 7의 경우 시스템 수준의 서버, 네트워크를 개발하는 역량과 OS관련 역량을 요구하는 것으로 보인다.

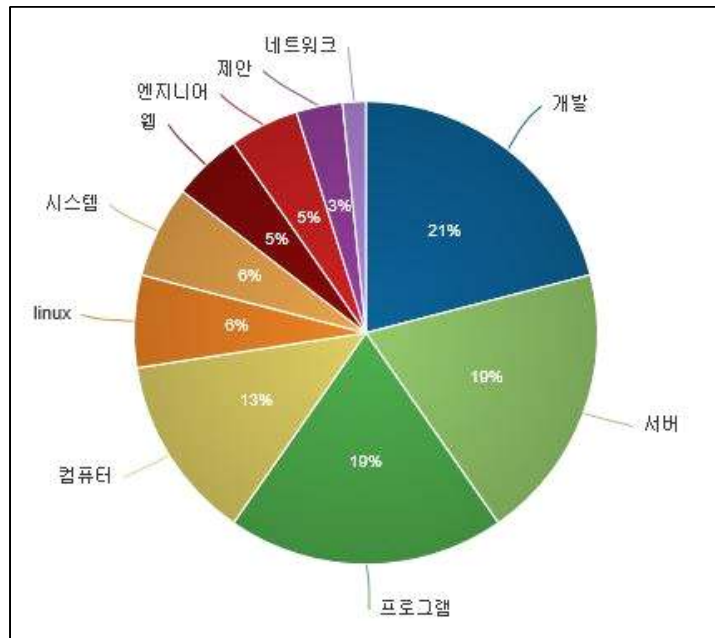
〈표 2-49〉 직무 7의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	개발	0.1327
2	프로그램	0.1293
3	서버	0.1225
4	컴퓨터	0.0837
5	시스템	0.0452
6	linux	0.0442
7	엔지니어	0.0368
8	웹	0.0333
9	제안	0.0216
10	네트워크	0.0206

직무 7 군집에 대한 LDA 분석 결과를 확인해보면 〈표 2-49〉과 같이 직무 내 동시등장 빈도율이 2%이하인 값은 역시 영향력이 크지 않아 누락한 후 동시등장 빈도율이 큰 단어부터 확인할 때 ‘개발’, ‘프로그램’, ‘서버’, ‘컴퓨터’, ‘시스템’, ‘LINUX’, ‘엔지니어’, ‘웹’, ‘제안’, ‘네트워크’의 순으로 주요 요구 역량인 핵심단어가 나타났다. 이와 관련하여 전반적인 직무 7 군집에 해당하는 직무내용은 컴퓨터 시스템의 자체기능 수행명령체계인 시스템SW를 연구 및 개발하고 설계하며, 시스템과 관련한 프로그램을 작성하는 업무를 수행하는 것으로 정리하였다. 이 결과는 [그림 2-95]의 핵심단어 비율 그래프에서도 유사하게 나타나고 있다. 핵심단어에 대한 비율 그래프 시각화에서는 주로 개발, 서버, 프로그램 등과 같은 시스템 관련 핵심 단어가 주로 나타나고 있으며 단순 소프트웨어가 아닌 시스템 운영을 위한 컴퓨터, Linux와 같은 용어도 함께 등장하고 있어 소프트웨어 개발자와는 약간의 차이를 보이고 있다.

핵심단어를 100점으로 환산하여 [그림 2-96]과 같이 핵심단어 방사형 그래프를 시각화 하였다. 그 결과 개발, 서버, 컴퓨터, 프로그램의 용어들이 주로 이루고 있으며 시스템과 Linux에 대한 요구사항이 함께 도출되고 있음을 알 수 있다.

[그림 2-95] 직무 7의 핵심단어 비율 그래프



[그림 2-96] 직무 7의 핵심단어 방사형 그래프



이와 마찬가지로 직무 7 군집에 대한 핵심단어들의 워드클라우드를 시각화 한 결과는 [그림 2-97]과 같다. 분석 결과 서버, 엔지니어, 개발, Linux, 시스템이 주요 핵심단어로서 제공되고 있으며 UNIX나 구축 등 시스템 개발 및 운영에 관련된 핵심단어들이 도출되고 있음을 알 수 있다.

[그림 2-97] 직무 7의 핵심단어 워드클라우드



이 결과에 따라 본 보고서는 직무 7은 주로 시스템 엔지니어와 소프트웨어 개발자 직무의 일부를 유사 정의로서 활용하는 범주 내에 있는 것으로 확인하였다. <표 2-50>와 같이 국제직무분류표준을 통해 직무를 비교한 결과 O*Net의 경우 시스템 엔지니어의 직무로서 정의되고 있으며 복잡한 응용문제, 시스템 관리 이슈, 네트워크 이슈 등에 대한 솔루션 설계 및 개발하고 시스템 관리 및 기능적 통합을 수행하는 직무로서 구분하고 있다. ISCO와 SOC의 경우 소프트웨어 개발자로서 직무를 통합설명하고 있으나 O*Net과 유사 정의로서 정의되고 있었다. 그에 따라 본 보고서는 직무 7을 ‘시스템 개발자’로 정의하였으며 컴퓨터 시스템의 자체 기능 수행명령체계인 시스템SW를 연구 및 개발하고 설계하며 시스템 개발과 관련된 프로그램을 작성하는 업무를 담당하는 직무로서 정리하였다.

〈표 2-50〉 시스템 개발자 직무 관련 국제직무분류표준

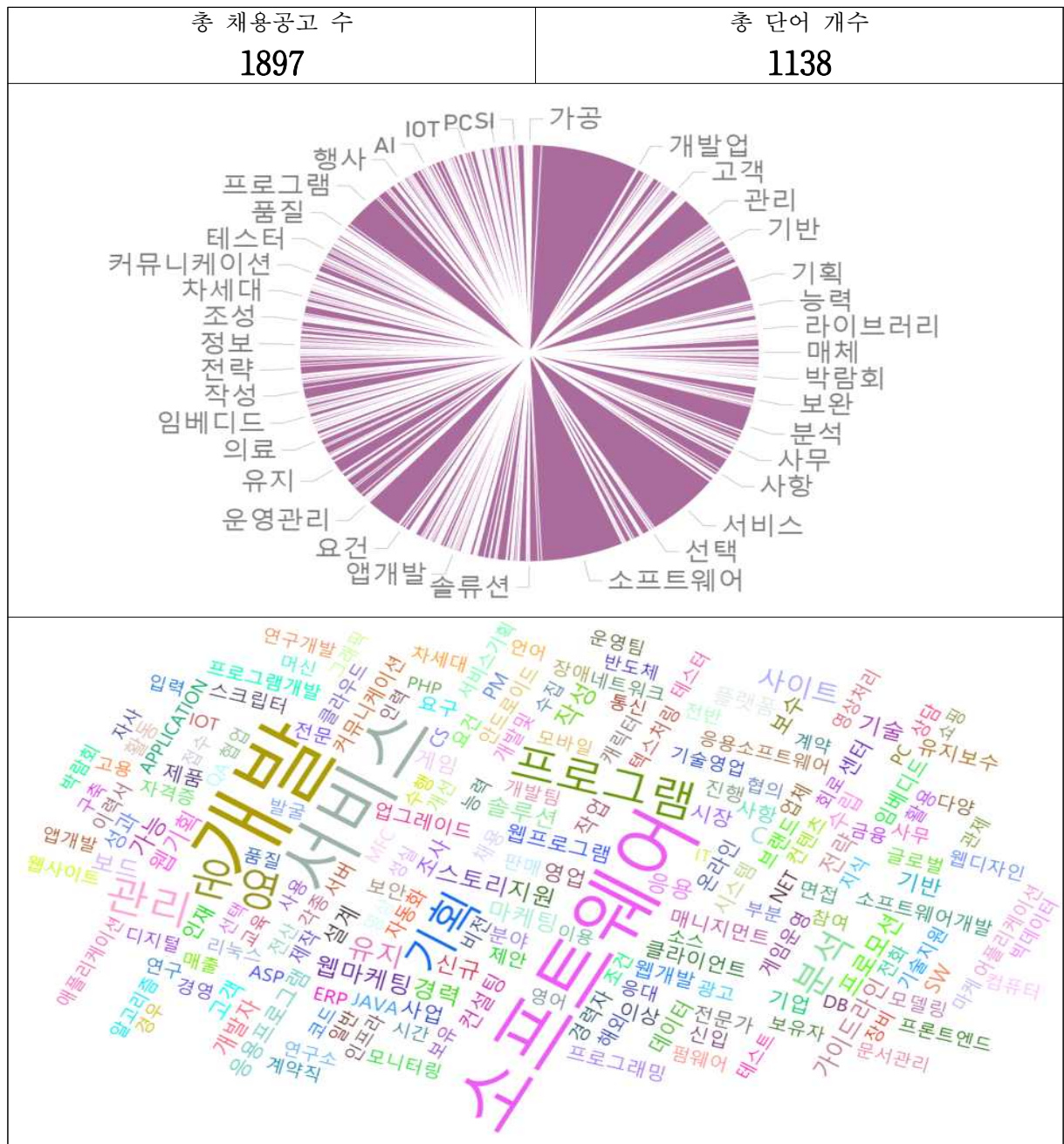
소스	분류코드	직무	정의
ISCO-08	2512	Software developers	- SW 개발자 연구, 분석하기 위하여나 새로운 SW 응용 프로그램과 기존 운영 시스템의 요건을 평가함 - 개발 요구사항을 충족하기 위한 SW 솔루션의 설계, 개발, 테스트 및 유지보수
SOC	15-1133	Software Developers, Systems Software	- 복잡한 응용문제, 시스템 관리 이슈, 네트워크 이슈 등에 대한 솔루션 설계 및 개발 - 시스템 관리 및 기능적 통합을 수행함
O*Net	15-1199.02	Computer Systems Engineers/Architects	
본 보고서의 분류		시스템 개발자	컴퓨터 시스템의 자체기능 수행명령체계인 시스템SW를 연구 및 개발하고 설계하며, 시스템과 관련한 프로그램을 작성하는 업무를 담당한다.

따라서 시스템 개발자는 새로운 서버용 시스템이나 기존의 서버 프로그램을 분석하고 평가하여 사용자의 요구사항을 파악하고, 사용자의 요구사항을 충족시키기 위한 시스템 SW를 기획하거나 개발하고, 테스트 및 운영하는 역할을 수행한다. 해당 직무에 대한 세부 사항은 다음과 같이 정리할 수 있다.

- 시스템 SW의 기능 및 사용자의 요구사항을 분석 및 시스템SW 설계 및 테스트
- 웹 서버 운영 체계 및 컴퓨터 운용 프로그램과 LINUX와 같은 시스템SW 패키지를 설계, 개발, 조정
- 컴퓨터언어 및 시스템SW 패키지를 설계, 개발, 조정
- 컴퓨터 장비 공급자, 시스템 설계자, 프로그래머 및 컴퓨터 오퍼레이터들이 사용하는 시스템 SW의 상태, 작동 및 유지관리에 관한 보고서, 매뉴얼 및 문서를 작성

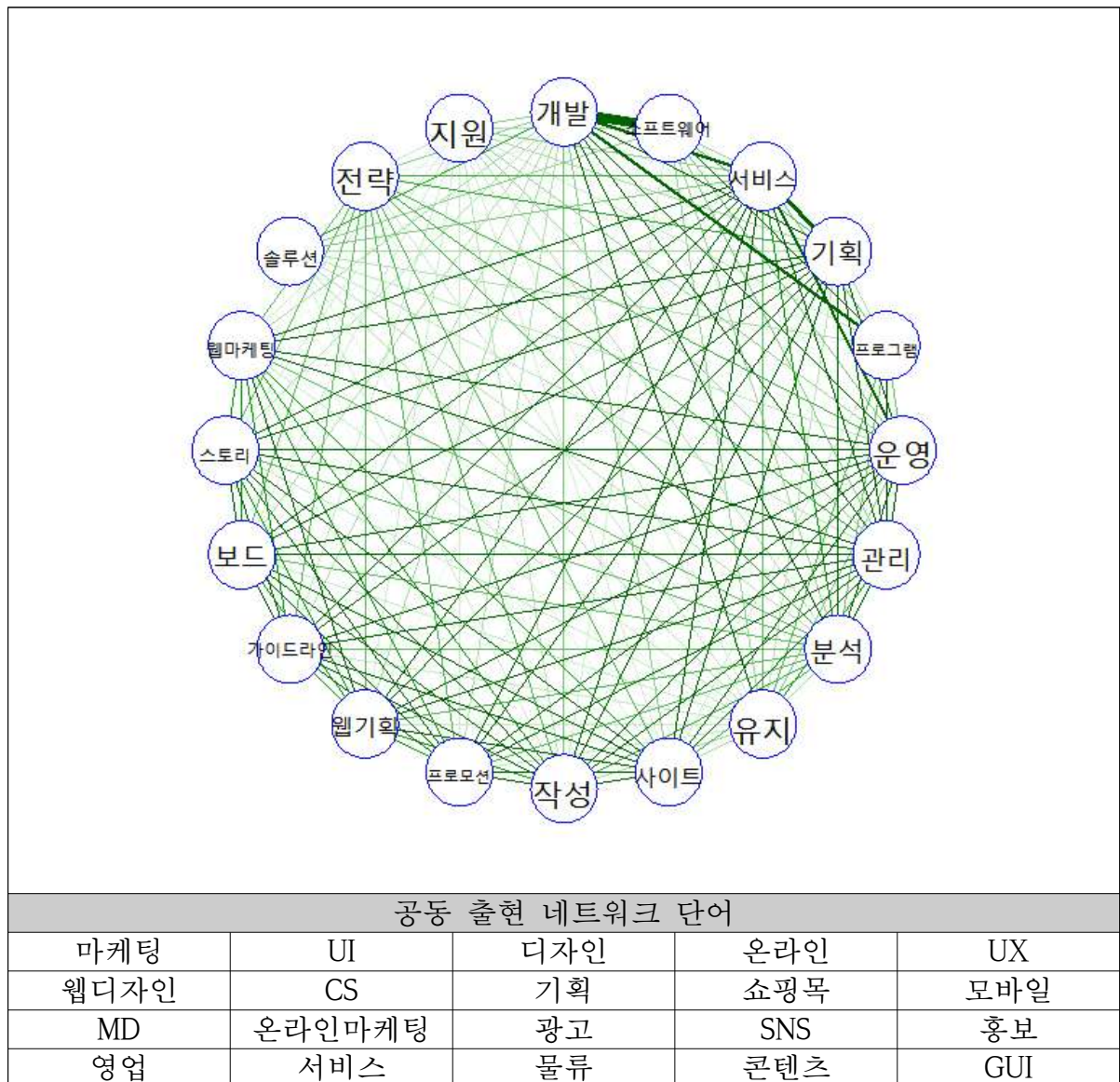
2.2.8) 소프트웨어 분석가

[그림 2-98] 직무 8의 특성



직무 8 군집에 해당 하는 채용공고는 [그림 2-98]과 같이 총 1493건으로 616개의 단어가 등장한다. 빈번하게 등장하는 단어는 ‘SW’, ‘개발’, ‘서비스’ 이다. 또한 ‘프로그램’, ‘기획’, ‘운영’, ‘분석’ 과 같은 단어들이 등장한다. 또한 워드클라우드 결과를 확인하면 소프트웨어, 개발, 기획 등을 바탕으로 소프트웨어 개발 직무와 유사하지만 연구개발, 분석, 운영, 조사 등의 특정 단어들이 출현함에 따라 일반적인 소프트웨어 개발과는 다른 직무임을 확인할 수 있다.

[그림 2-99] 직무 8의 단어 공동 출현 네트워크



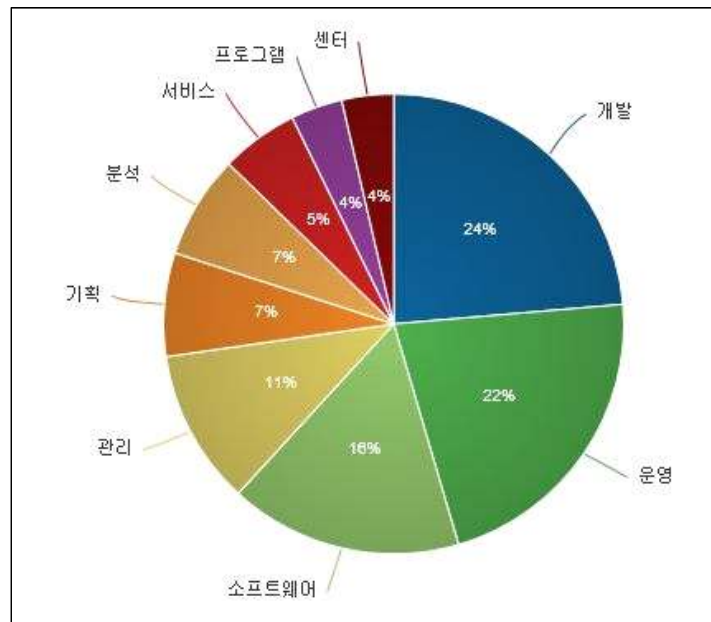
[그림 2-99]는 직무 8의 공동 출현 네트워크이다. {개발-SW}의 단어 조합이 강한 연관 관계로 채용공고에서 자주 출현하는 것으로 나타나고 있으며 {개발-서비스}, {개발-프로그램}, {서비스-기획}의 연관 출현 관계 또한 단어 조합으로서 존재한다. 즉, 직무 7은 소프트웨어 개발과 관련된 역량을 요구하지만 따른 특정 프로그래밍 언어나 기술적인 역량 보다는 분석, 관리, 유지 등의 별도 능력을 요구하고 있는 직무로 판단된다.

이와 관련하여 직무 8 군집에 대한 LDA 분석 결과를 판단하면 <표 2-51>와 같다. 직무 내 동시등장 빈도율이 2% 이상인 동시등장 빈도율이 큰 단어부터 확인할 때 ‘개발’, ‘운영’, ‘SW’, ‘관리’, ‘분석’, ‘기획’, ‘서비스’, ‘그래픽’, ‘센터’, ‘프로그램’의 순으로 주요 요구 역량인 핵심단어가 나타났다. 이 결과는 [그림 2-100]과 같이 직무 8의 핵심단어 비율그래프 시각화에서도 개발, 운영, 소프트웨어가 가장 크게 나타났으나 관리, 기획, 분석 등 순수 개발자와 다른 핵심단어가 출현함을 확인하였다.

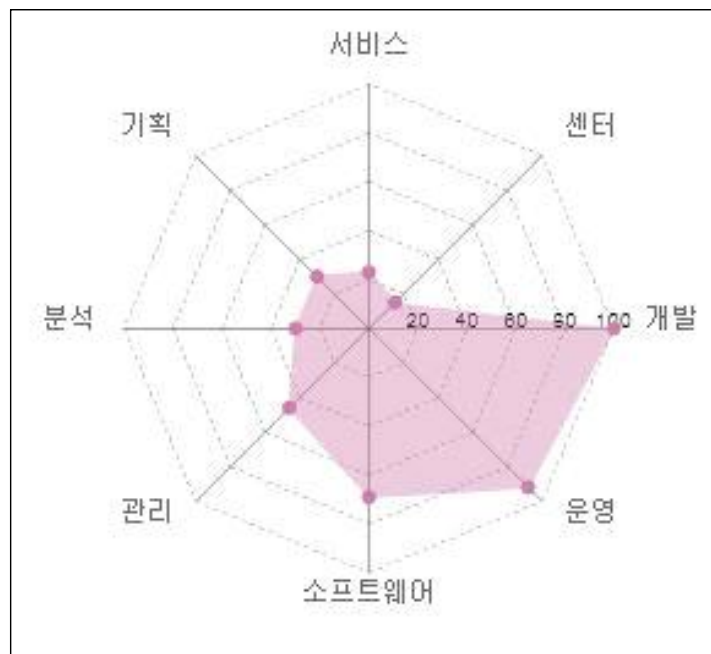
<표 2-51> 직무 8의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	개발	0.1302
2	운영	0.1286
3	소프트웨어	0.0938
4	관리	0.0609
5	분석	0.0445
6	기획	0.0421
7	서비스	0.0356
8	그래픽	0.0287
9	센터	0.0278
10	프로그램	0.0268

[그림 2-100] 직무 8의 핵심단어 비율 그래프



[그림 2-101] 직무 8의 핵심단어 방사형 그래프



그에 따라 직무 8 군집에 대한 핵심단어를 100점 기준으로 환산하여 방사형 그래프를 시각화 한 결과는 [그림 2-101]과 같다. 주요 핵심단어가 개발, 운영, 소프트웨어 이외에도 분석의 직무를 겸하고 있는 것을 확인할 수 있고 분석 및 관리에 관련된 직무 형태가 존재함을 확인할 수 있다. 그에 따라 [그림 2-102]와 같이 직무 8에 대한 핵심단어 워드클라우드를 시각화 한 결과 소프트웨어, 운영, 관리, 및 분석, 통합, 기획 등의 단어가 주요 핵심단어로서 제시됨을 알 수 있어 직무 5의 소프트웨어 개발자와는 차이가 있음을 확인할 수 있다.

[그림 2-102] 직무 8의 핵심단어 워드클라우드



이 결과를 바탕으로 직무 8 군집에 ‘그래픽’ 단어가 분류체계에 등장함에 따라 공고를 확인한 결과 채용공고 중 ‘SW 개발자를 모집합니다./프로그램 개발자를 모집합니다.’와 같이 전반적인 개발자 직무에 해당하는 직무로 나타났다. 그러나 실제 공고의 세부 직무기술 내용을 확인해보면 SW 분석가로서 분석 및 기획, 프로그래밍 관련 보조 업무를 동시에 운영하는 인력을 찾고 있는 공고로서 확인할 수 있다. 또한 그와 관련하여 ‘그래픽’ 프로그래밍과 결과물에 대한 분석 및 기획과 관련하여 SW 분석가 직무의 일부분으로서 요구되는 경우가 있음을 확인하였다.

〈표 2-52〉 SW 분석가 직무 관련 국제표준

소스	분류코드	직무	정의
ISCO-08	2519	Software and Applications Developers and Analysts Not Elsewhere Classified	- Minor Group 251: Software and Applications Developers and Analysts에 분류되지 않는 SW 및 응용 SW 개발자 및 분석가 - SW 시험, 사업적 분석을 포함한 품질 보증을 위해 직무가 특화됨
SOC-2016	15-1253	Software Quality Assurance Analysts and Testers	- SW 문제 및 관련 요인을 찾아내기 위해 SW 테스트를 수행 - SW 개선을 준비하기 위하여 시스템 수정사항을 시험함 - 문서 SW와 응용 프로그램 결함 버그 추적 시스템과 SW나 웹개발자들에게 결함 보고 - 알려진 결함에 고려하여 데이터베이스 개발 및 유지보수 수행 - 기능적 요구사항, 운영적 특성, 설계 및 일정에 제공하기 위한 SW 설계에 참여
O*Net	15-1121.02	Computer Systems Analysts	정보기술 시스템에 대한 연구, 계획, 설계, 작성, 시험 및 IT 시스템에 대한 조언 제공 및 개선
본 보고서의 분류		SW 분석가	SW의 설계, 작성, 실험, 개선된 성능을 제공하기 위한 목적을 달성하고 특성 응용을 위한 SW의 계획, 설계, 평가 및 실험 등의 업무를 담당한다.

이와 같은 결과에 따라 본 보고서는 직무 8 군집이 SW의 설계, 작성, 실험, 개선된 성능을 제공하기 위한 목적으로 SW 및 시스템을 설계, 계획, 평가하고 시스템에 적합한 SW를 계획, 설계, 평가 및 실험 등의 업무를 담당하는 것으로 확인하였다. 〈표 2-52〉와 같이 ISCO의 경우 소프트웨어 개발자 및 분석가, SOC는 소프트웨어 품질 보증 및 분석가, O*Net에서는 컴퓨터 시스템 분석가로서 분류가 되었으며 각 직무분류는 주로 SW 시험, 사업적 분석을 포함한 품질 보증, 기능적 요구사항, 운영적 특성, 설계 및 일정에 제공하기 위한 SW 설계에 참여, 정보기술 시스템에

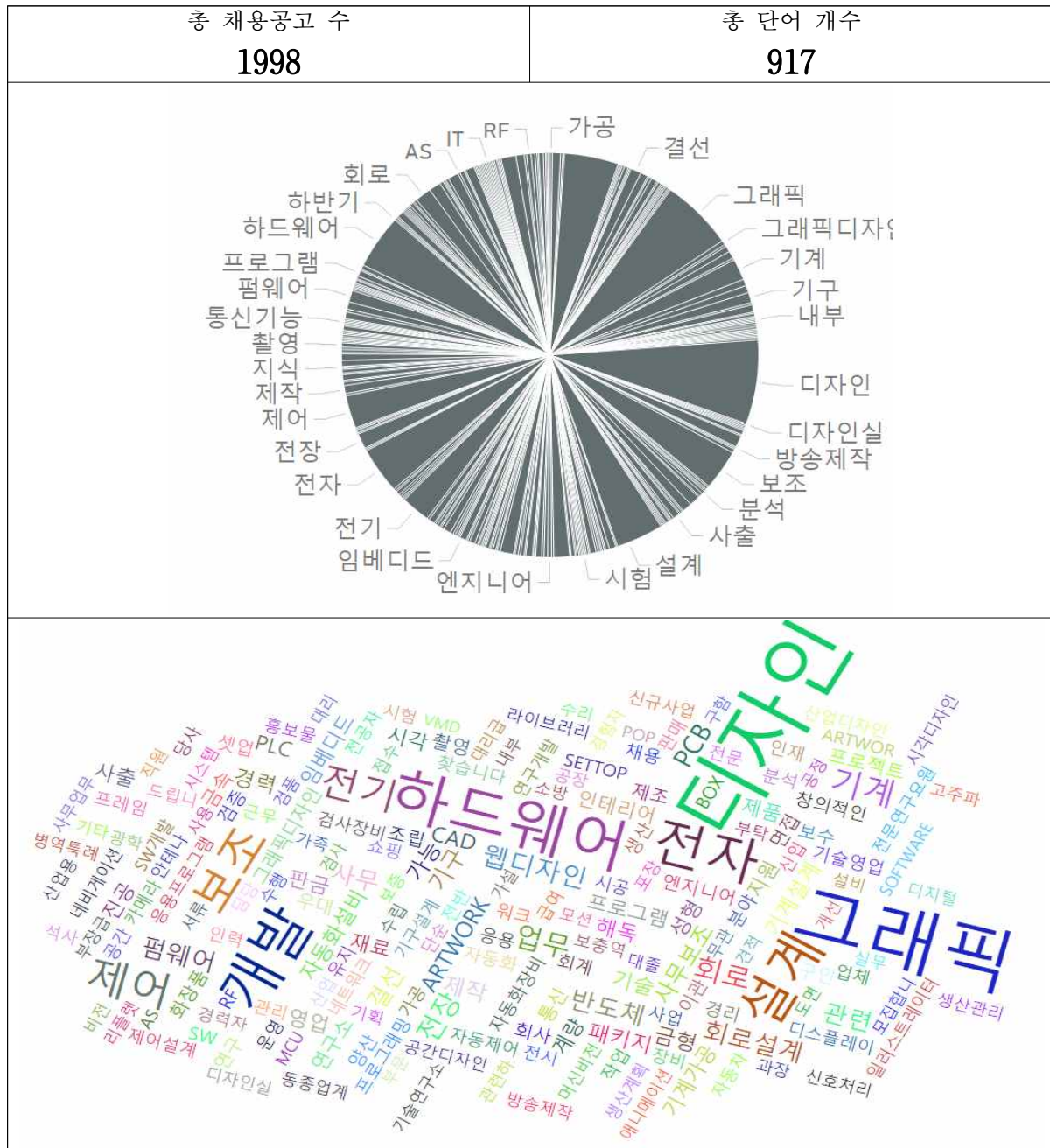
대한 연구, 계획, 설계, 작성, 시험 및 IT 시스템에 대한 조언 제공 및 개선 등의 정의를 제시하고 있어 개발자 집단에 주로 포함되어 있으며 정의 또한 광의의 개념으로 SW 개발자의 직무로서 제시되어 있다. 그러나 해당 직무의 응용 직업을 살펴보면 SW 분석가(Quality assurance analyst), SW 테스터(Software tester), 시스템 테스터(Systems tester)의 역할을 하는 것으로 제시되고 있다. 본 연구과제는 텍스트 데이터의 분석 결과에 따라 직무 8 군집을 ‘소프트웨어 분석가’로 명명하고 SW의 설계, 작성, 실험, 개성된 성능을 제공하기 위한 목적을 달성하고 특정 응용을 위한 SW의 계획, 설계, 평가 및 실험 등의 업무를 담당하는 분석가로 정의하였다.

소프트웨어 분석가의 직무 내용은 다음과 같이 정리 할 수 있다.

- 비즈니스 관점에서 SW의 사용이 최대의 효율성을 확보하기 위한 연구 수행
- 컴퓨터 사용 및 SW 운영과 관련하여 이론적 접근 및 운영 방식 등을 분석 및 연구 수행
- SW 활용 사항에 대하여 테스트 및 분석을 통해 향후 SW 운영 및 관리와 관련하여 기획
- 시스템 및 SW 성능에 대한 계획, 설계, 평가 관점에서 프로그램들을 테스트 및 유지보수 하며 비즈니스 목적에 맞는 인터넷, 인트라넷, 멀티미디어 시스템 전체를 계획, 설계, 평가함

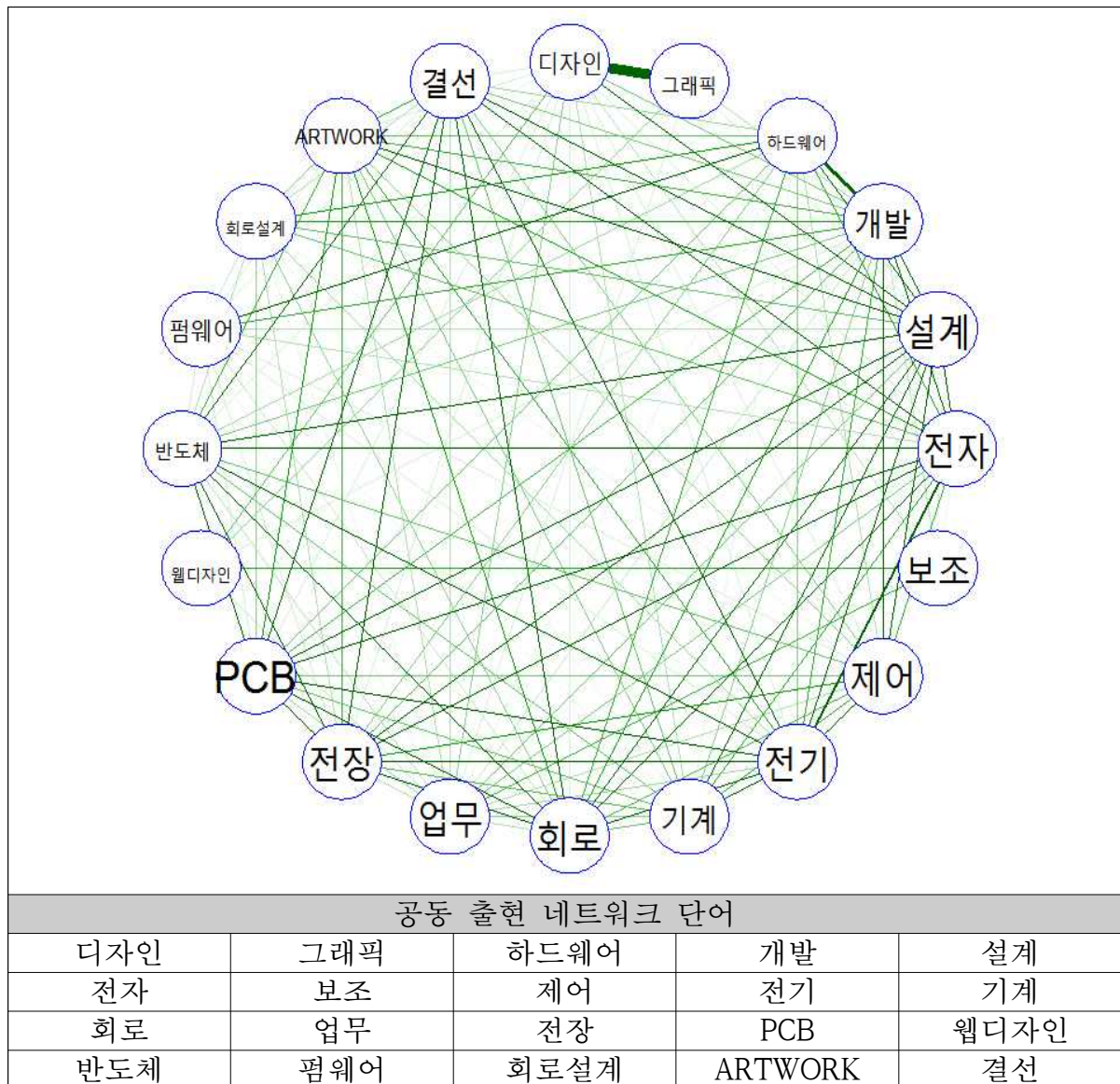
2.2.9) 시스템 설계분석가

[그림 2-103] 직무 9의 특성



직무 9 군집에 해당하는 채용공고는 [그림 2-103]과 같이 총 1998건으로 772개의 단어가 등장한다. 이 중 빈번하게 등장하는 단어는 ‘디자인’, ‘개발’, ‘그래픽’, ‘하드웨어’, ‘설계’, ‘전자’, ‘제어’이다. 또한 ‘펌웨어’, ‘회로’, ‘PCB’ 등 회로설계와 관련된 단어들이 등장함에 따라 시스템의 설계 및 제어와 관련된 직무임을 확인할 수 있다. 이에 대한 워드클라우드 결과를 확인하면 하드웨어, 그래픽, 설계, 개발, 제어 등의 단어들이 발생함에 따라 기업의 특성에 따른 시스템 개발 및 설계를 위한 직무 특성을 확인할 수 있다.

[그림 2-104] 직무 9의 단어 공동 출현 네트워크



그에 따라 직무 9에 대한 전체 공고 내의 단어들을 바탕으로 소셜네트워크 분석을 수행하였다. 분석 결과를 바탕으로 도출된 공동 출현 네트워크는 [그림 2-104]와 같다. 분석된 단어 들 중 채용공고에서 함께 다루고 있는 단어들은 주로 {디자인-그래픽}, {하드웨어-개발}, {전자-전기}의 단어 조합이 채용공고에서 관계가 높은 빈도의 연관 관계로서 확인되었다.

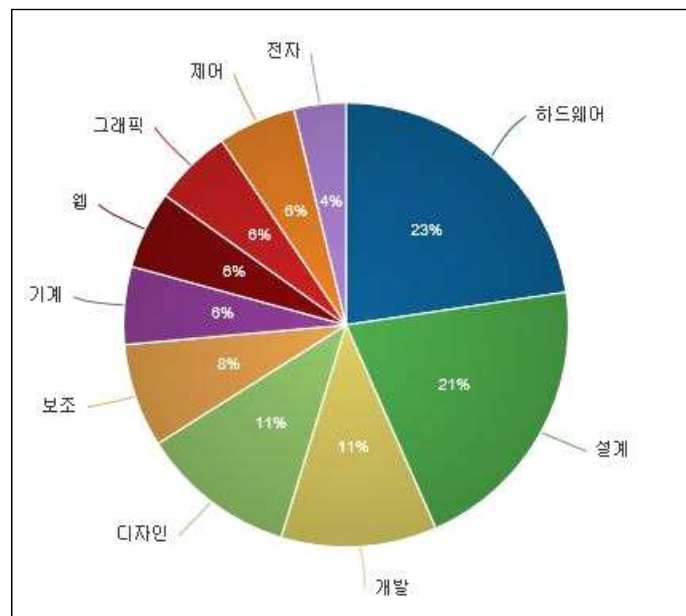
〈표 2-53〉 직무 9의 핵심단어 목록

순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	하드웨어	0.1267
2	설계	0.1154
3	디자인	0.0696
4	개발	0.0687
5	보조	0.0473
6	제어	0.0360
7	그래픽	0.0355
8	웹	0.0313
9	기계	0.0301
10	전자	0.0296

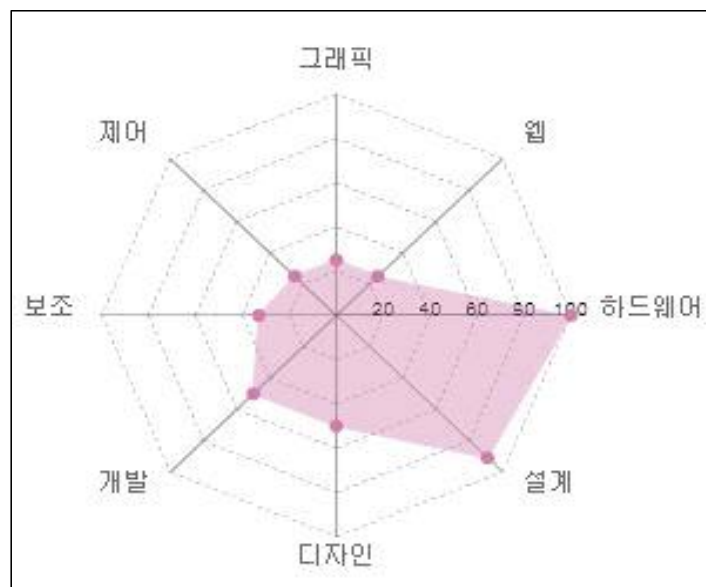
직무 9 군집에 대한 LDA 기반의 동시등장 빈도율에 따른 핵심단어들을 〈표 2-53〉과 같이 도출하였다. 이 중 직무 내 동시등장 빈도율이 2%이하인 단어들을 누락시킨 후 동시등장 빈도율이 큰 단어부터 확인할 때 ‘하드웨어’, ‘설계’, ‘디자인’, ‘개발’, ‘보조’, ‘제어’, ‘그래픽’, ‘웹’, ‘기계’, ‘전자’의 차례로 주요 요구 역량인 핵심단어가 나타났다. 이 중 ‘전자’라는 단어가 도출되어 직무와 연관 관계 확인을 위하여 초기 채용공고 원본들을 확인하였다. 그 결과 주로 ‘전자’와 관련된 업체의 기업들의 공고가 나타났다. 예를 들어 ‘00전자기기에 들어가는 펌웨어 개발 및 분석가 모집’ 등과 같은 공고가 발견되었으며 ‘전자 시스템’ 또는 ‘전자전기 전공’ 인력, ‘전자 문서’, ‘기계’에 들어가는 펌웨어 분석 등의 분야에 대한 구인하는 공고들이 나타남에 따라 해당 결과가 제시된 것을 확인할 수 있었다.

이와 관련하여 직무 9의 핵심단어 비율 그래프는 [그림 2-105]와 같다. 시각화 결과 하드웨어, 설계, 개발 등이 주로 나타나 시스템 개발자와 유사하게 보이지만 보조, 제어, 디자인 등 시스템의 구성과 운영 및 분석 설계에 대한 분석가 직무임을 확인할 수 있다.

[그림 2-105] 직무 9의 핵심단어 비율 그래프



[그림 2-106] 직무 9의 핵심단어 방사형 그래프



또한 직무 9 군집의 핵심단어의 값을 100점으로 변환하여 방사형 그래프로 시각화 표현한 결과는 [그림 2-106]과 같다. 방사형 그래프는 주로 하드웨어와 관련하여 설계 및 디자인, 개발 등 시스템의 설계와 관련된 개념들이 주를 이루고 있다.

[그림 2-107] 직무9의 핵심단어 워드클라우드



직무 9에 대한 핵심단어들을 워드클라우드 시각화한 결과는 [그림 2-107]과 같다. 결과 설계, 개발, 제어 등이 제시되고 있으며 회로, 장비 등의 단어들이 주를 이루고 있으며 시스템의 설계 과정에서 활용 가능한 핵심단어들이 제시되고 있다.

그에 따라 본 보고서는 직무 9의 경우 시스템 설계 및 분석에 해당 하는 측면에서 SW 분석가와 유사하지만 실제 개발자로서의 직무와는 다르며 일반적인 IT 프로젝트 전문가의 역할과 유사하다는 점을 확인하였다. <표 2-54>와 같이 국제직무분류표준에서도 각 분류체계별로 명칭 또한 다르게 지칭되고 있으나 연구, 분석하기 위한 및 클라이언트 정보 기술 요건 절차 및 문제들을 평가, 현존하는 컴퓨터 시스템에 대한 개선과 시스템의 기능적 통합 및 관리 수행, IT기술 관련 프로젝트의 계획, 개시, 관리 등의 관점을 정의하고 있었다.

일반적인 SW 관련 개발 보조 및 시스템, SW, 하드웨어 등 전반적인 개념 구성을 수립해야 한다는 점에서 본 보고서는 직무 9 군집을 ‘시스템 설계분석가’로 분류를 확정하였다. 결과적으로 시스템 설계분석가는 시스템을 설계 및 분석하고 시스템 용량과 개발 작업 절차 및 일정을 검토하는 전반적 관리를 책임지게 된다. 따라서 전체 시스템 설계를 시행하는 직무를 담당하고 있으며 세부적으로는 구입 할 하드웨어와 SW 및 개발 소요기간 등을 산정하고 비용 및 총괄 설계안을 통해 시스템 설계를 기획한다.

〈표 2-54〉 시스템 설계분석가 직무 관련 국제표준

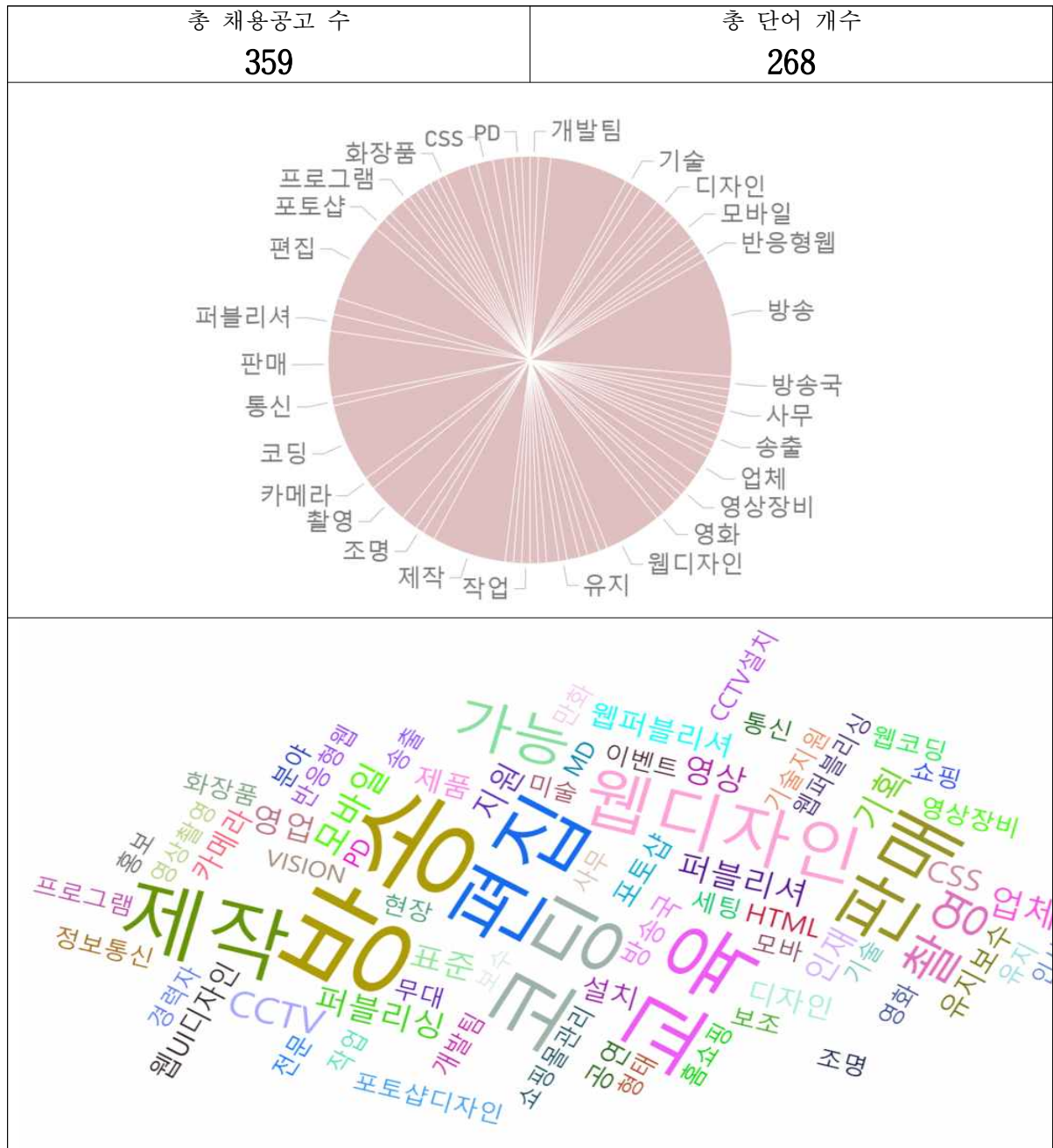
소스	분류코드	직무	정의
ISCO-08	2511	Systems analysts	- 연구, 분석하기 위한 및 클라이언트 정보 기술 요건 절차 및 문제들을 평가 - 현재 및 향후 정보 시스템에 대한 개선 계획 수립 및 추천
SOC-2016	15-1211	Computer Systems Analysts	- 복잡한 응용문제, 시스템 관리 이슈 및 네트워크 문제에 대한 솔루션 개발 및 개선을 위한 분석 - 경영상 데이터 프로세스 이슈 분석 - 현존하는 컴퓨터 시스템에 대한 개선 - 시스템의 기능적 통합 및 관리 수행 - 시스템 용량 및 제반사항 검토 - 유용한 SW 분석 및 추천
O*Net	15-1199.09	Information Technology Project Managers	- IT기술 관련 프로젝트의 계획, 개시, 관리 - 기술 인력의 업무에 대한 지침 작성 - 기업 경영 및 기술적 문제에 대한 통합 관리 제공 - 프로젝트 단계 기획 및 각 단계별 경영상 시사점 도출 - 일정관리 모니터링 및 비용 관리
본 보고서의 분류		시스템 설계분석가	- 컴퓨터시스템을 설계 및 분석하고, 컴퓨터시스템의 용량, 작업 절차 및 일정을 검토하여 전체적인 컴퓨터시스템 설계를 시행하는 업무를 담당한다. - 또한 구입해야 할 하드웨어 및 SW, 소요기간 등을 산정하고 설계안을 제출하여 설계를 기획한다.

특히 시스템 설계 분석가의 경우 개발자 역량을 요구하기 보다는 시스템 설계를 평가하고 통합적으로 시스템을 제어 및 통제, 분석하는 직무를 주로 하는 것을 확인할 수 있다. 또한 기업체의 고객관리시스템, 종합병원의 의료정보시스템, 정부 부처의 행정정보시스템, 대학교의 사이버교육시스템 등 각종 컴퓨터시스템을 설계 및 분석하고, 컴퓨터시스템의 용량, 작업 절차 및 일정을 검토하여 전체적인 컴퓨터 시스템 설계를 시행한다. 따라서 시스템 설계분석가의 경우 다음과 같은 직무 내용으로 정리할 수 있다.

- 프로그래머의 작업을 감독
- 오퍼레이션 명세서를 설정하고 SW 요건을 체계화
- 시스템 개발자가 명기한 목적 및 문제들을 분석
- 시스템 개발자가 제시한 형식화된 솔루션을 상세한 프로그램 명세서로 전환하고, 프로그램의 타당성과 로직을 검사하고 수정
- 컴퓨터 장비 공급자, 시스템 설계자, 프로그래머 및 컴퓨터 오퍼레이터들이 사용하는 시스템 SW의 상태, 작동 및 유지관리에 관한 보고서, 매뉴얼 및 문서를 작성
- 구입해야 할 하드웨어 및 SW, 소요기간 등을 산정하고 설계안을 제출하여 설계를 기획한다.
- 설계안이 확정되면 실제 시스템 구축작업에 들어간다.
- 시스템구축을 통해 높은 생산성을 올릴 수 있도록 고객의 요구사항이 설계단계에 체계적으로 반영한다.
- 데이터베이스를 분석하고 컴퓨터시스템의 환경을 최적상태로 하는 SW 튜닝을 실시한다.

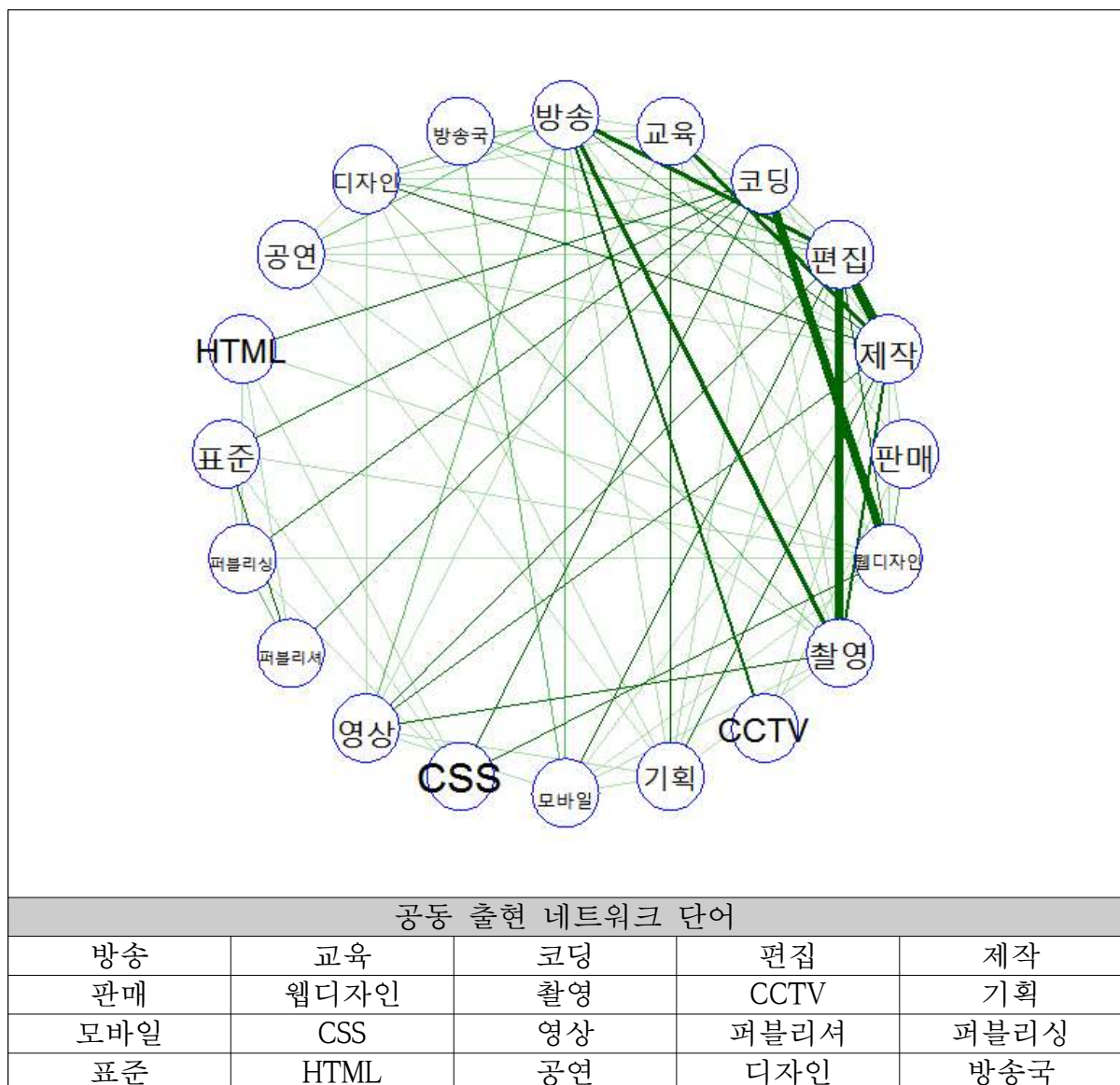
2.2.10) 미디어 콘텐츠 창작자

[그림 2-108] 직무 10의 특성



직무 10 군집에 해당 하는 채용공고는 총 359건으로 268개의 단어가 등장하였다. 이 중 해당 직무 군집내의 모든 채용공고에 대해 주요하게 등장하는 단어는 [그림 2-108]과 같이 ‘방송’, ‘편집’, ‘코딩’, ‘교육’, ‘웹디자인’, ‘제작’, ‘판매’ 이다. 이와 함께 해당 군집의 모든 채용공고에 등장하는 단어들을 바탕으로 워드클라우드를 수행한 결과 방송, 편집, 코딩, 교육, 제작, 웹디자인 등의 단어들이 주로 출현함에 따라 콘텐츠 창작 및 방송, 교육제작 등의 콘텐츠 관련 직무라는 점을 유추할 수 있었다.

[그림 2-109] 직무 10의 단어 공동 출현 네트워크



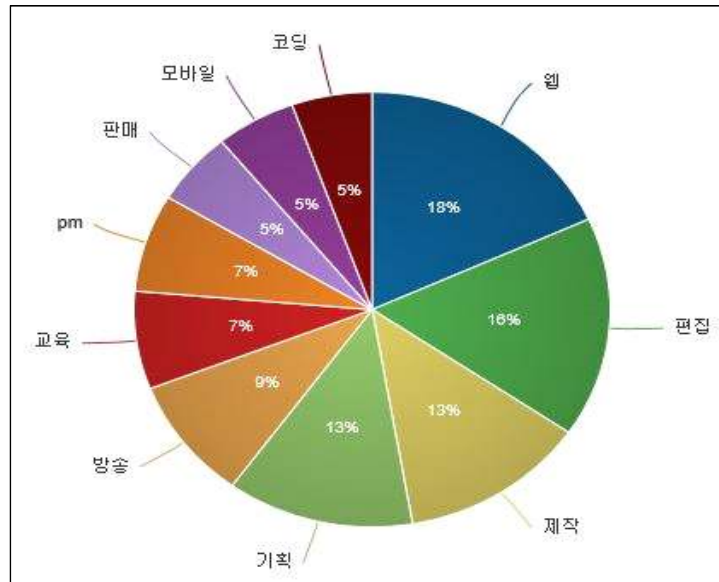
[그림 2-109]는 직무 10의 채용공고에 출현한 모든 단어들을 바탕으로 작성한 소셜네트워크 분석 기반의 공동 출현 네트워크이다. 분석 결과 {방송-편집-촬영}, {코딩-웹디자인}, {교육-제작}, {편집-제작-촬영}의 관계가 높은 연관관계를 가지고 나타난 단어 조합으로 보인다. 즉, 직무 10 군집은 방송과 관련하여 제작, 편집, 촬영에 대한 역량과 SW 개발과 관련된 역량을 요구하는 것으로 보인다.

〈표 2-55〉 직무10의 핵심단어 목록

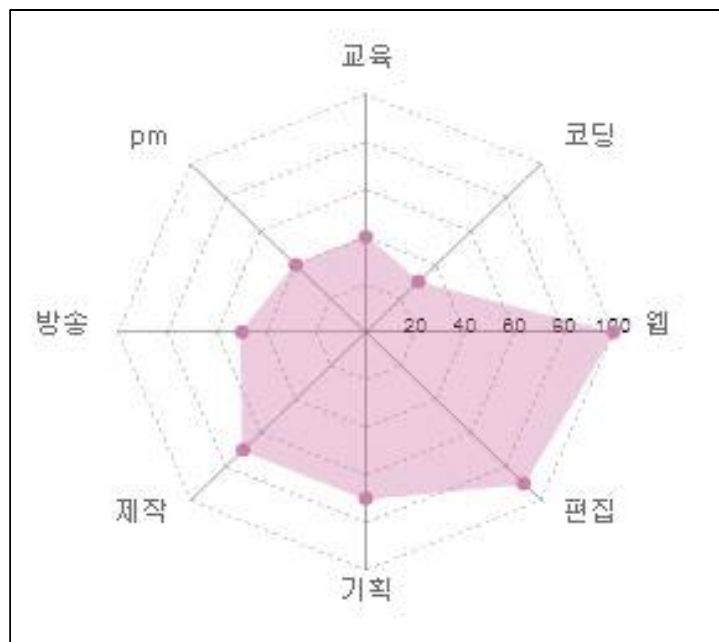
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	웹	0.1090
2	편집	0.0913
3	기획	0.0754
4	제작	0.0726
5	방송	0.0515
6	프로젝트 매니저	0.0437
7	교육	0.0404
8	코딩	0.0391
9	모바일	0.0386
10	판매	0.0328

보다 명확한 직무 내 특성을 파악하기 위하여 직무 10 군집에 대한 LDA 분석을 실시하였다. 분석 결과 해당 직무 내 동시등장 빈도율 2% 이하 단어를 탈락시킨 후 동시등장 빈도율이 큰 단어부터 확인할 때 나타난 핵심단어는 〈표 2-55〉과 같이 ‘웹’, ‘편집’, ‘기획’, ‘제작’, ‘방송’, ‘프로젝트 매니저’, ‘교육’, ‘코딩’, ‘모바일’, ‘판매’의 순서로 나타났다. 이 중 ‘교육’라는 단어가 도출되어 다른 단어들과의 연관 관계 확인을 위하여 초기 채용공고들을 모두 확인하였다. 그 결과 주로 ‘교육’은 ‘교육 영상 콘텐츠 제작 및 영상 편집자 모집’과 같은 채용공고가 많았으며 ‘교육 콘텐츠 기획’, ‘이러닝 콘텐츠 기획’ 등의 구직 사항이 제시되었다. 또한 ‘판매’ 단어가 도출되어 역시 직무와 상관관계를 확인하기 위하여 채용공고를 확인하였다. 그 결과 주로 생산/판매하는 제품에 대한 홍보 콘텐츠와 관련된 공고를 의미하였다. 결론적으로 도출된 핵심단어들은 해당 직무와 관련된 주요 단어들을 확인할 수 있었다.

[그림 2-110] 직무 10의 핵심단어 비율 그래프



[그림 2-111] 직무 10의 핵심단어 방사형 그래프



이 결과와 마찬가지로 핵심단어 비율 그래프 결과인 [그림 2-110]을 보면 웹, 편집, 제작, 기획, 방송 등 콘텐츠에 대한 제작, 편집, 기획, 방송 등의 전반적인 콘텐츠 서비스의 형태를 수행하는 직무임을 확인할 수 있다.

[그림 2-112] 직무 10의 핵심단어 워드클라우드



[그림 2-111]은 직무 10 군집에 대한 핵심단어 값을 100점을 기준으로 변환했을 때 시각화된 방사형 그래프이다. 시각화 결과는 주로 웹, 편집, 기획, 제작, 방송, pm, 교육, 코딩의 단어들이 주요 핵심단어로서 제시되고 있음을 확인할 수 있다. 그에 따라 직무 10의 핵심단어들을 대상으로 워드클라우드 시각화를 수행한 결과는 [그림 2-112]와 같다. 그 결과 영상, 편집, 기획, 제작 등의 단어들이 주요 핵심단어로 제시되고 있으며 특히 콘텐츠와 판매, 운영, 홍보, 광고 등의 단어들도 함께 제시되고 있다.

이 결과는 바탕으로 본 보고서는 기존 국제직무표준분류에서 유사 직무를 도출하고자 살펴보았으나 해당 직무에 대하여 찾아볼 수 없고 <표 2-56>와 같이 유사 직무만 제시되고 있는 실정이며 이 결과는 ISCO나 SOC에서는 작성 기간이 각각 2012, 2010으로 포함되지 못한 신규 직무이기 때문으로 판단된다. 그러나 해당 직무는 최근 미국 내에 커리어 전문 기업인 Linked-in에서는 이미 Content Creator의 채용공고가 매우 활발하게 게시되고 있으며³⁰⁾ 보편화 되고 있는 직무 중 하나인 ‘미디어콘텐츠창작자’와 매우 흡사하게 판단된다. 미디어콘텐츠창작자는 광고기반 플랫폼에 개인의 영상 콘텐츠를 제작하여 업로드 하고 이를 통해 수익을 창출하는 일을 한다. 기획안 구성부터 연출, 촬영, 편집 등을 종합적으로 소화하기 때문에 PD, 배우, 작가의 역할을 총망라하여 담당하며 개인이 활동하거나 팀으로 활동하며,

30) <http://www.contentengine.tv/what-does-it-mean-to-be-a-content-creator/>

창의적인 영상 콘텐츠를 기획, 연출, 그리고 구성 및 편집하여 완성된 영상 콘텐츠를 업로드 하는 직무의 특성을 가지고 있다. 그에 따라 본 보고서에서는 직무 10을 ‘미디어콘텐츠창작자’로 명칭을 정하고 그에 대한 주요 직무 내용은 아래와 같다.

〈표 2-56〉 미디어 콘텐츠 창작자 직무 관련 국제직무분류표준

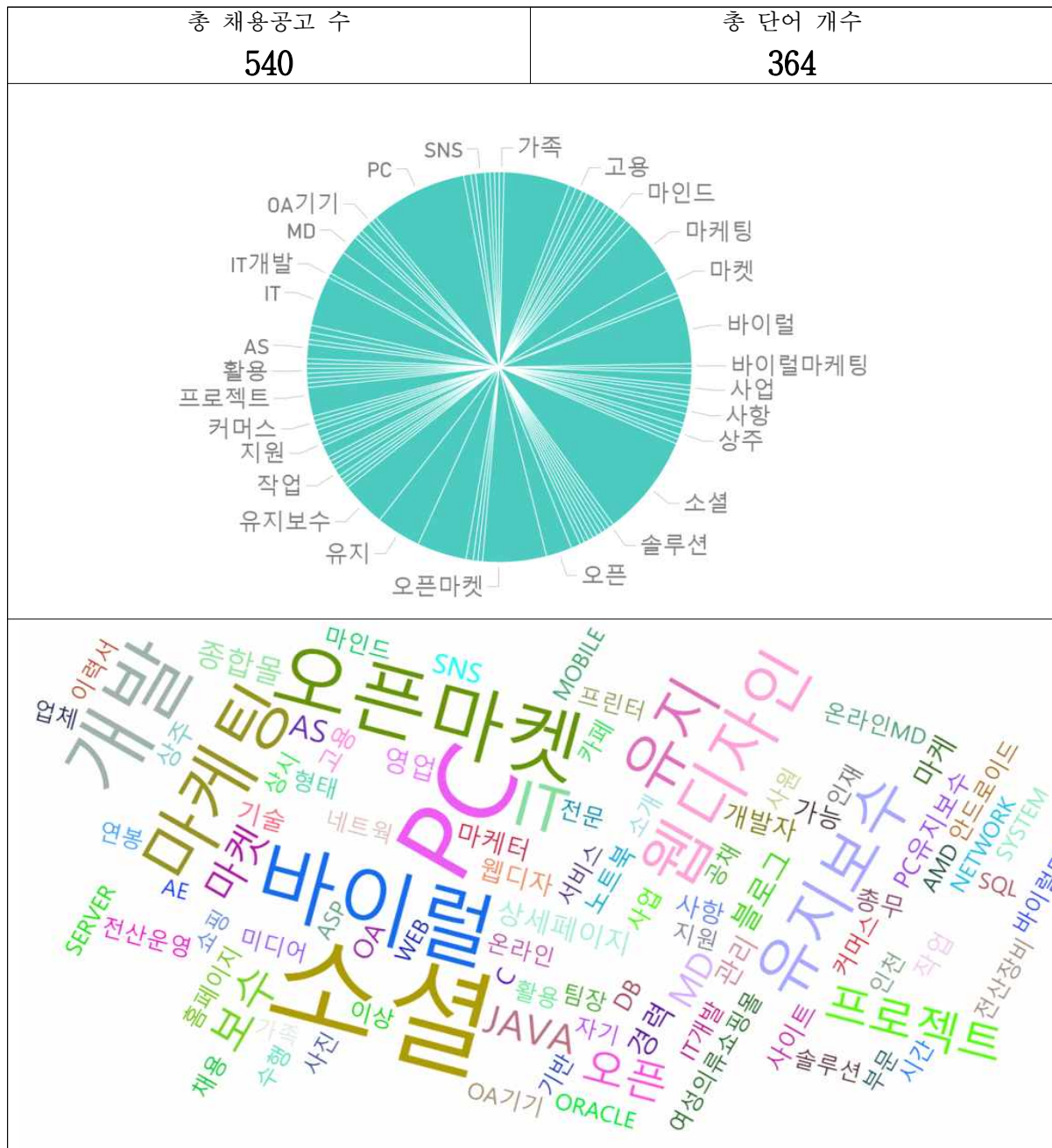
소스	분류코드	직무	정의
SOC-2016	27-1014	Special Effects Artists and Animators	컴퓨터 게임, 영화, 뮤직비디오 그리고 상거래 활동에서 활용되는 제품 또는 창조물에 대한 영상, 전자적 툴과 미디어를 사용한 특수효과, 애니메이션 및 시각적 이미지를 만들어냄
O*Net	27-101400	Multimedia Artists and Animators	
Linked-In		Content Creator	- 하나 이상의 미디어 플랫폼에 직접 작성한 콘텐츠를 창조 및 출판함 - 문자, 영상, 음향 및 애니메이션 콘텐츠를 개발함
본 보고서의 분류		미디어 콘텐츠 창작자	- 광고기반 플랫폼에 개인의 영상 콘텐츠를 제작하여 업로드 하고 이를 통해 수익을 창출하는 일을 한다. - 기획안 구성부터 연출, 촬영, 편집 등을 종합적으로 소화하기 때문에 PD, 배우, 작가의 역할을 총망라하여 담당한다. - 개인이 활동하거나 팀으로 활동하며, 창의적인 영상 콘텐츠를 기획, 연출, 그리고 구성 및 편집하여 완성된 영상 콘텐츠를 업로드 한다.

다시 말해서 미디어콘텐츠창작자는 웹 페이지나 방송, 모바일 페이지에서 필요로 하는 콘텐츠를 기획하고 제작하는 프로젝트의 프로젝트 매니저 역할을 수행하며, 콘텐츠의 기획, 제작 및 편집 등의 전반적인 과정을 관리 및 콘텐츠를 기획하는 직무 군집으로 확인되었다. 세부 직무 내용은 다음과 같다.

- 미디어 콘텐츠 창작자는 광고기반 플랫폼에 개인의 영상 콘텐츠를 제작하여 업로드 하고 이를 통해 수익을 창출하는 일을 한다. 기획안 구성부터 연출, 촬영, 편집 등을 종합적으로 소화하기 때문에 PD, 배우, 작가의 역할이 총망라되어 있다고 할 수 있다.
- 찍고 싶은 영상 관련 주제를 수집하고, 이를 기반으로 주제와 내용을 정리해 세부적인 대사, 출연진, 촬영구도 등을 직접 기획, 구성한다.
- 편집 및 영상 업로드 작업을 한다.
- 개인이 활동하거나 팀으로 활동하며, 창의적인 영상 콘텐츠를 기획, 연출, 그리고 구성 및 편집하여 완성된 영상 콘텐츠를 업로드 한다.
- 연출, 촬영, 편집 등을 혼자 처리해야 하므로 동영상 제작에 대한 기본적인 지식과 기술이 필요하다. 무엇보다 새로운 콘텐츠에 대한 호기심이 많아야 하고, 인기 콘텐츠를 면밀히 분석해서 적용하기 위한 남다른 노력이 필요하다.

2.2.11) 소셜 미디어 관리자

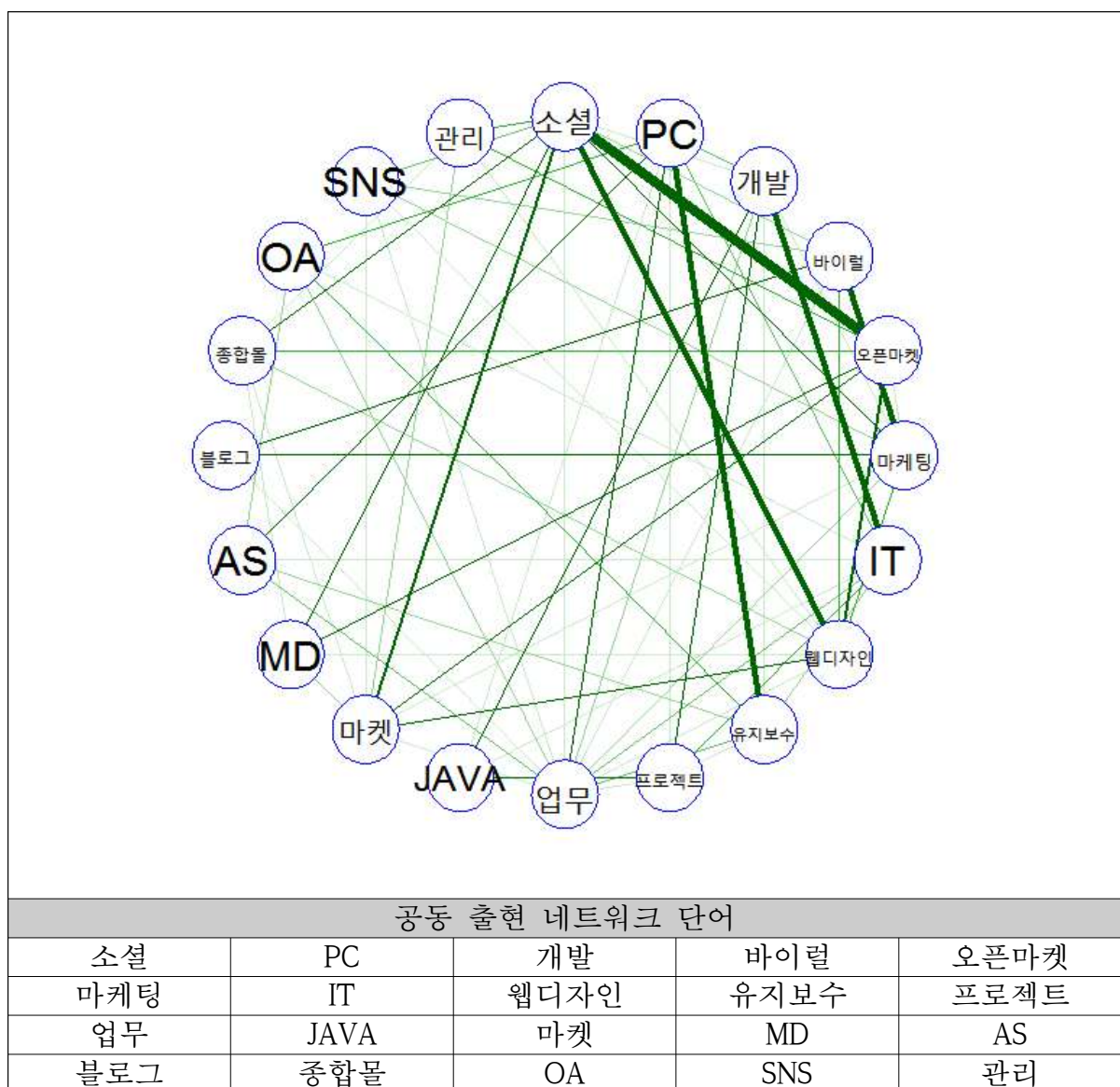
[그림 2-113] 직무 11의 특성



직무 11 군집에 해당 하는 채용공고는 [그림 2-113]과 같이 총 540건으로 364개의 단어가 등장한다. 해당 직무 군집에 해당하는 구직 공고에서 빈번하게 등장하는

단어는 ‘PC’, ‘오픈마켓’, ‘소셜’, ‘바이럴’, ‘마케팅’, ‘웹디자인’, ‘개발’로 나타나 주로 소셜 미디어 관리 또는 운영 관련 직무임을 확인할 수 있으며 이는 워드클라우드 결과에서도 유사하다. 오픈마켓, 바이럴, 소셜, 개발, 마케팅, 유지보수 등의 단어들을 주로 사용하는 채용광고들이 대다수라는 점에서 직무 11은 소셜 미디어 기반의 직무인 것으로 판단되었다.

[그림 2-114] 직무 11의 단어 공동 출현 네트워크



그에 따라 소셜네트워크 분석을 통한 직무 11 군집의 공동 출현 네트워크는 [그림 2-114]와 같다. {소셜-오픈마켓-웹디자인}, {PC-유지보수}, {개발-IT},

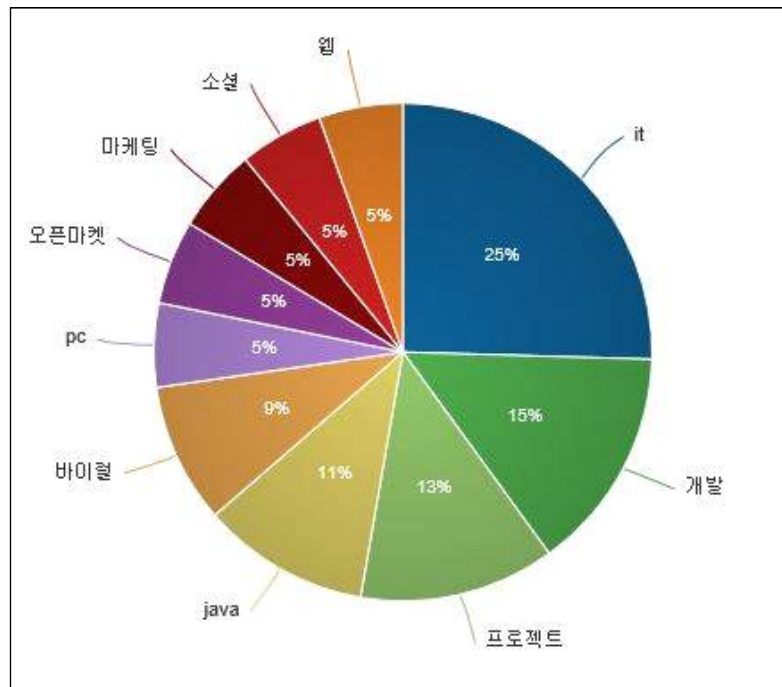
{오픈마켓-마케팅-바이럴}의 관계가 주로 함께 활용되는 단어들이므로 나타난다. 즉, 직무 11 군집은 소셜네트워크서비스를 기반으로 마케팅과 관련한 역량을 요구하는 것으로 판단된다.

〈표 2-57〉 직무 11의 핵심단어 목록

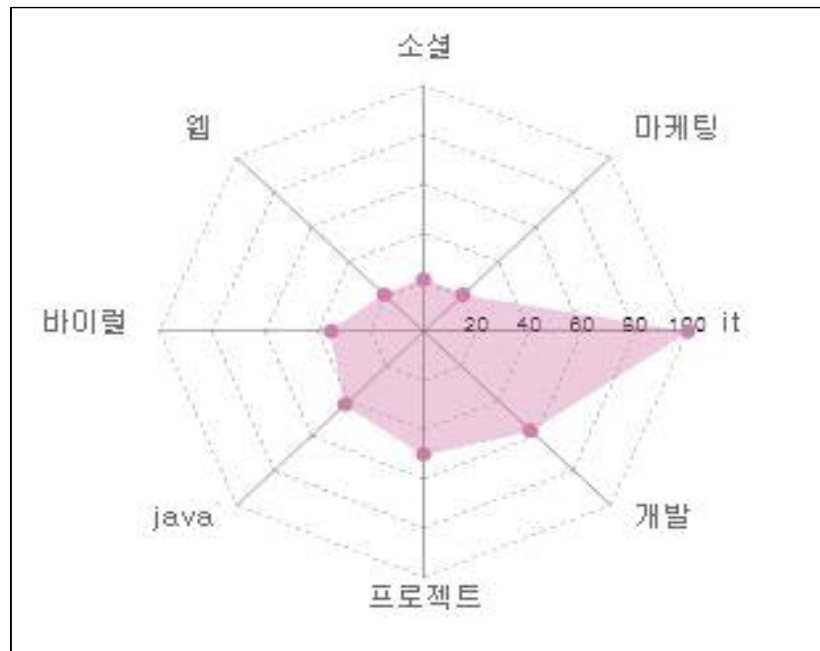
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	IT	0.1466
2	개발	0.0891
3	프로젝트	0.0759
4	java	0.0699
5	바이럴	0.0515
6	웹	0.0398
7	소셜	0.0373
8	마케팅	0.0370
9	오픈마켓	0.0346
10	PC	0.0338

본 보고서의 직무 11 군집에 대한 기초 단어분석 결과에 추가적으로 직무 11에 대한 LDA 분석을 통해 동시등장 빈도율을 바탕으로 〈표 2-57〉과 같이 동시등장 빈도율이 2%이상인 단어들만을 핵심단어로서 선정하였다. 그에 따라 해당 직무 내에서 동시등장 빈도율이 큰 단어부터 확인할 때 나타난 핵심단어는 ‘IT’, ‘개발’, ‘프로젝트’, ‘java’, ‘바이럴’, ‘웹’, ‘소셜’, ‘마케팅’, ‘오픈마켓’, ‘PC’의 순서로 나타났다. 그와 함께 직무 11에 대한 핵심단어 비율 그래프를 시각화 한 결과는 [그림 2-115]와 같다. 전반적으로 IT, 개발, 프로젝트와 함께, 바이럴, java를 다루는 요건을 보이고 있으나 미디어콘텐츠창작자 직무와는 다르게 마케팅, 소셜 등의 핵심단어들이 나타나고 있음을 확인할 수 있다. 마찬가지로 [그림 2-116]과 같은 직무 11에 대한 핵심단어 동시등장 빈도율을 100점으로 환산하여 시각화 한 방사형 그래프 결과에서도 기본적인 정보기술 및 프로젝트 직무를 포함하고 있으나 주로 바이럴, 소셜 등과 같은 소셜네트워크서비스 기반의 직무로서 채용공고 핵심단어들이 제시되고 있음을 확인할 수 있었다.

[그림 2-115] 직무 11의 핵심단어 비율 그래프



[그림 2-116] 직무 11에 대한 핵심단어 방사형 그래프



[그림 2-117] 직무 11의 핵심단어 워드클라우드



[그림 2-117]과 같이 핵심단어 워드클라우드를 분석한 시각화 결과는 IT, 개발과 같은 개발자 업무와 함께 바이럴, 마케팅, 소셜, 서비스, 쇼핑몰, 오픈마켓, 관리와 같은 소셜 미디어 기반 마케팅 활동을 주로 진행하는 직무임을 확인할 수 있다.

분석 결과와 함께 국제직무분류표준을 바탕으로 직무 11에 대한 정의를 비교 정리하기 위해 기존 체계를 확인한 결과 ISCO에서는 해당 직무는 등장하지 않으며 SOC와 O*NET에서는 Online Merchant 또는 Marketing Manager에 해당하고 있으나 소셜미디어 환경에 적합한 직무는 존재하지 않아 신규 직무로서 판단해야 할 직무 군집으로 판단되었다.

이에 따라 현재 해외에서 핵심단어들을 통해 확인한 채용광고에서는 일반적으로 많이 광고되는 직무로서 판단되었고 Linked-In에서도 소셜 미디어 관리자로서 활발한 구직광고가 제시되고 있다. Linked-In의 구직 광고들을 종합하면 소셜 마케터는 웹 페이지나 오픈 마켓 등의 마케팅 활동의 일환으로 소비자들의 구전 효과를 일으키는 소셜 마케팅을 수행하는 역할을 한다. IT 기술을 활용하여 구전효과를 일으킬 수 있는 콘텐츠를 기획, 제작하고 직접 마케팅 활동을 수행한다.

〈표 2-58〉 소셜 미디어 관리자 직무 관련 국제표준

소스	분류코드	직무	정의
SOC-2016	13-1161	Market Research Analysts and Marketing Specialists	- 지역, 광역, 국제적, 국가적 또는 온라인 시장에 대한 현황 연구 - 마케팅 및 광고 캠페인에 대한 계획 또는 제품이나 서비스에 대한 판매 가능성에 대한 정보 수집 - 마케팅 및 유통 관련 경쟁자, 가격, 판매 및 방식에 대한 정보 수집 - 검색 마케팅 전략의 적용, 웹 시장 현황 분석, 검색엔진 순위 향상 계획 수립 및 표적 시장에 대한 개발
O*Net	13-1199.06	Online Merchants	- 기업, 경쟁자 및 잠재고객과 관련된 제품 및 서비스 수요 결정과 같은 마케팅 정책에 대한 기획 - 웹에서의 기업 이익 극대화에 대한 목적을 위한 가격 전략 개발 및 고객 만족을 통한 시장 점유율 개선 방안 마련 - 신제품 및 서비스에 대한 수요 측정을 위한 트렌드 모니터링 및 제품 개발
Kruse Control Inc ³¹⁾		Social Media Manager	- 기업의 소셜 미디어 마케팅 및 광고 전략에 대한 관리 - 소셜 미디어 마케팅 기획의 수립 및 개선
본 보고서의 분류 ³²⁾		소셜 미디어 관리자	- 소셜미디어를 기반으로 하는 기획 및 운영 등 소셜미디어와 관련된 업무를 수행하는 관리자로 페이스북, 트위터, 유튜브, 블로그 등을 활용하여 마케팅 및 미디어 전략을 담당한다. - 기업의 소셜미디어 계정을 기획 운영하며 소셜콘텐츠 기획 및 관리, 소셜마케팅, 소셜데이터 분석, 소셜위기관리 등 소셜미디어 관련 업무를 수행한다.

31) <https://www.krusecontrolinc.com/social-media-manager-job-description-complete-guide/>

32) Parveen et al. (2015)의 정의 참고

또한 해당 직무는 해외 연구내용에서도 소개되고 있는데 기업의 소셜미디어 계정을 기획 운영하며 소셜콘텐츠 기획 및 관리, 소셜마케팅, 소셜데이터 분석, 소셜위기관리 등 소셜미디어 관련 업무를 수행하는 직무로서 제시되고 있다 (Parveen et al., 2015). 따라서 본 보고서에서는 직무 11에 대하여 ‘소셜 미디어 관리자’로 정의하고 아직 구분되지 않았으나 채용 과정이 활발한 신규 직무로서 규정하여 독립적 직무로서 정의하였다.

참고적으로 <표 2-58>와 같이 동일한 직무는 아니지만 소셜미디어 관리자와 관련된 직무들을 기반으로 판단할 때, 소셜미디어 관리자는 소셜미디어 기반의 기획 및 운영 등 소셜미디어와 관련된 업무를 수행하는 관리자로 페이스북, 트위터, 유튜브, 블로그 등을 활용하여 마케팅 및 미디어 전략을 담당하는 직무로서 판단할 수 있다.

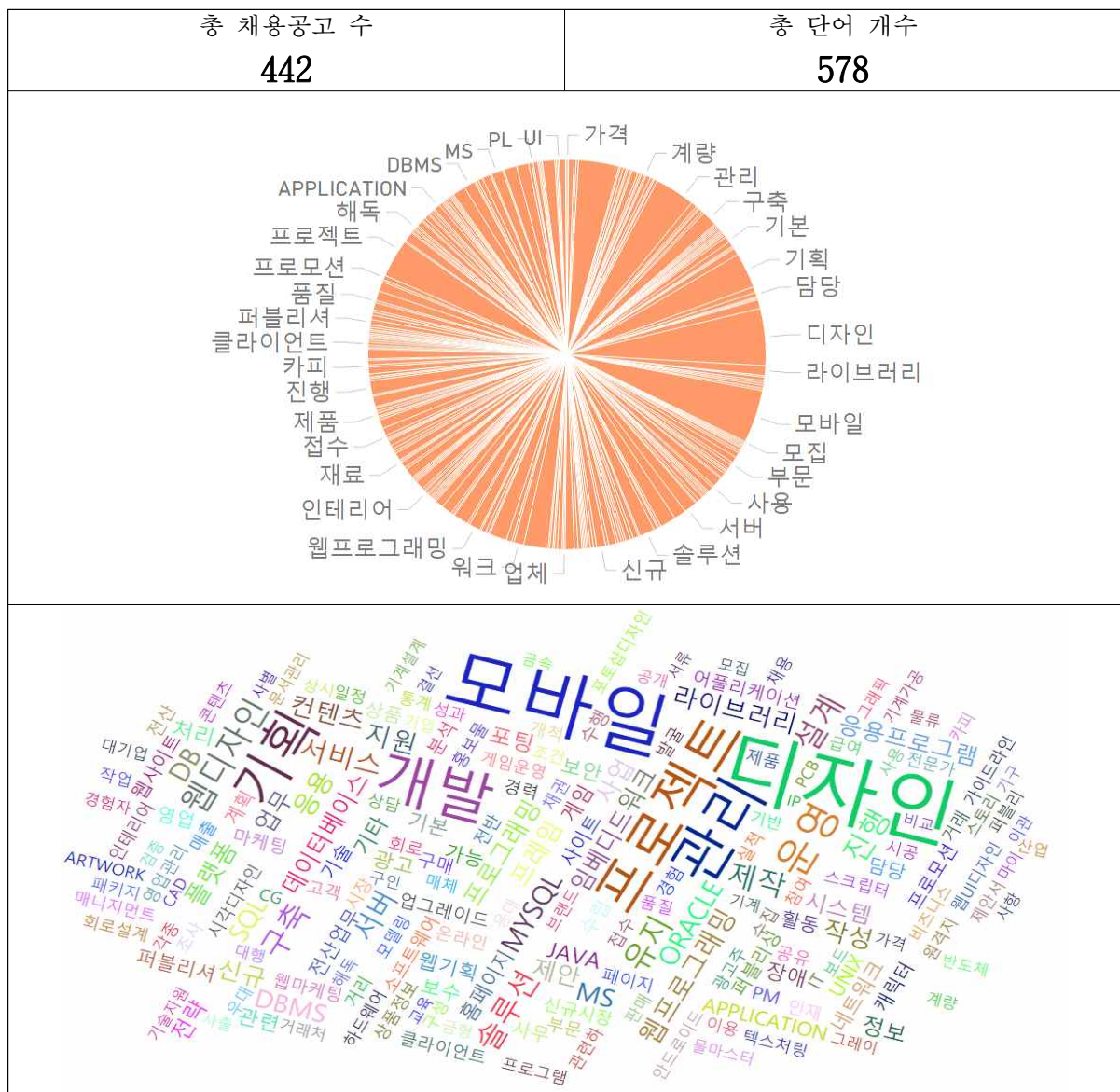
이에 따라 본 보고서에서 소셜 미디어 관리자에 대한 정의는 소셜 미디어를 기반으로 하는 기획 및 운영 등 소셜 미디어와 관련된 업무를 수행하는 관리자로 페이스북, 트위터, 유튜브, 블로그 등을 활용하여 마케팅 및 미디어 전략을 수행 (Parveen et al., 2015)하는 직무로서 판단할 수 있다. 특히 기업의 소셜 미디어 계정을 기획 운영하며 소셜 콘텐츠 기획 및 관리, 소셜 마케팅, 소셜 데이터 분석, 소셜 위기관리 등 소셜 미디어 관련 업무를 수행할 수 있다. 그에 따른 소셜미디어 관리자의 직무 내용을 정리하면 다음과 같다.

- 오픈마켓 기업 관련 소셜 미디어(Facebook, Twitter, Youtube, Blog, Linked-In 등) 계정 생성
- 자사 및 경쟁사에 대한 시장 및 수요조사 실시를 바탕으로 소셜미디어를 활용한 마케팅 전략 기획
- 바이럴 마케팅 수행, 소셜 커뮤니티 관리, 소셜 미디어 콘텐츠 (체험후기, 동영상, 게시물, 이미지, 정보, 뉴스 등)를 기획 및 제작
- 분석을 위한 기초자료 수집 및 적절한 소셜 데이터 분석 방법을 결정
- 이슈 단어 중심의 산업 분야 별 이슈 도출 및 데이터분석 툴 (R-Studio, hadoop, sass 등) 활용 및 마케팅 기획
- SNS 모니터링 수행으로 산업분야별로 핵심 단어 도출 및 소셜미디어 내 위기관리 업무 수행 (부정적 구전, 온라인 악용, 사이버폭력 등)

2.2.12) 데이터베이스 관리자

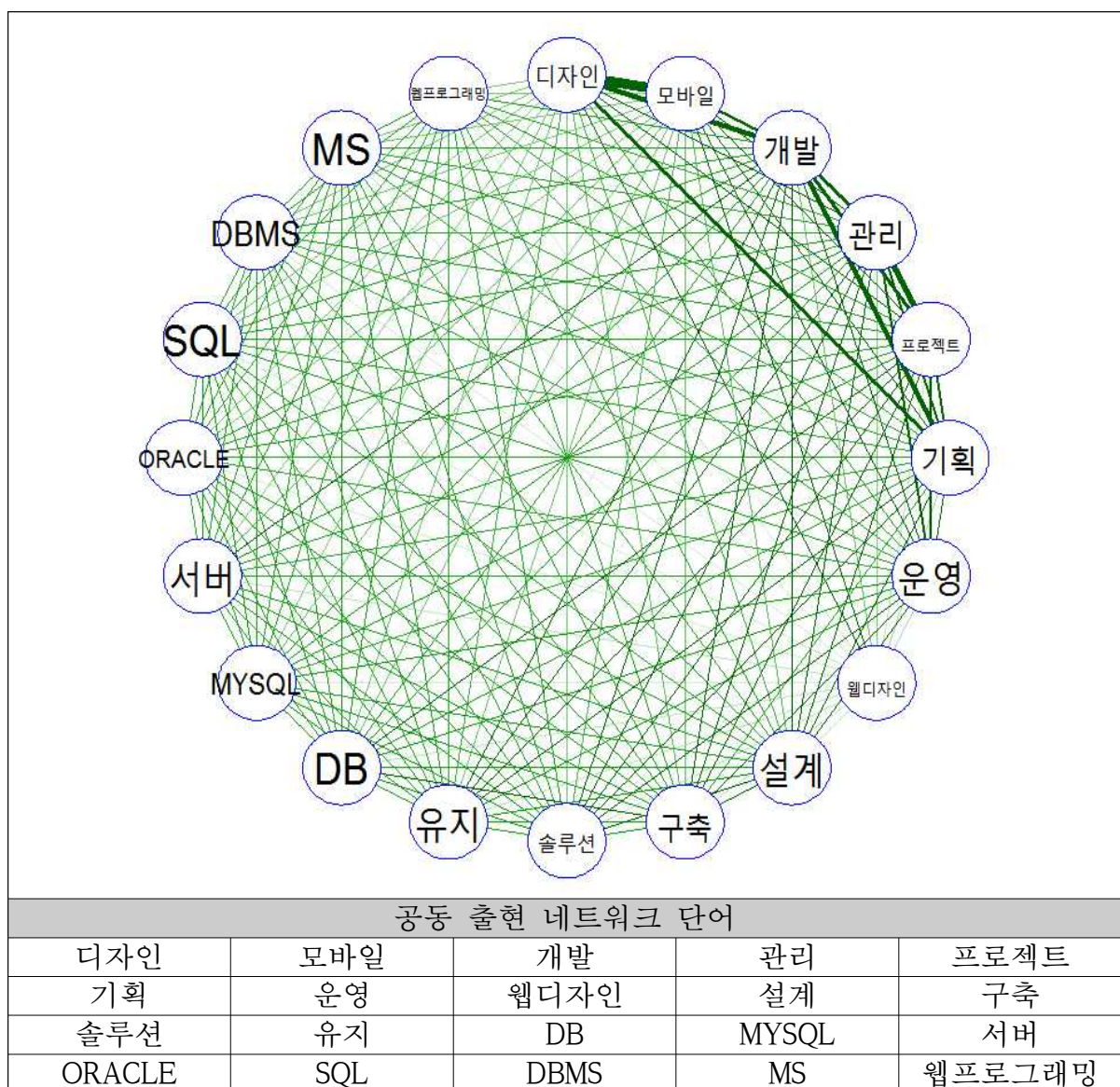
직무 12 군집에 해당 하는 채용공고는 [그림 2-115]와 같이 총 442건으로 578개의 단어가 등장한다. 해당 군집의 모든 채용공고를 대상으로 빈번하게 등장하는 단어는 ‘모바일’, ‘디자인’, ‘프로젝트’, ‘개발’, ‘기획’, ‘관리’ 이다. 또한 ‘데이터베이스’, ‘DBMS’ 등의 단어들이 등장하며, 세부적인 프로그래밍 언어로 ‘SQL’, ‘MYSQL’, ‘ORACLE’ 이 등장한다.

[그림 2-118] 직무 12의 특성



또한 직무 12의 채용공고에서 나타난 모든 단어들에 대한 워드클라우드 결과 모바일, 프로젝트, 디자인 등의 단어가 도출되고 있으나 Oracle, Java, Unix, MySQL 등 데이터베이스 관련 프로그램 단어들이 제시되고 있으며 운영, 설계, 관리 등의 직무 관련 단어들이 발생하고 있다. 따라서 데이터베이스의 관리 또는 유지를 위한 직무임을 유추해볼 수 있었다.

[그림 2-119] 직무 12의 단어 공동 출현 네트워크



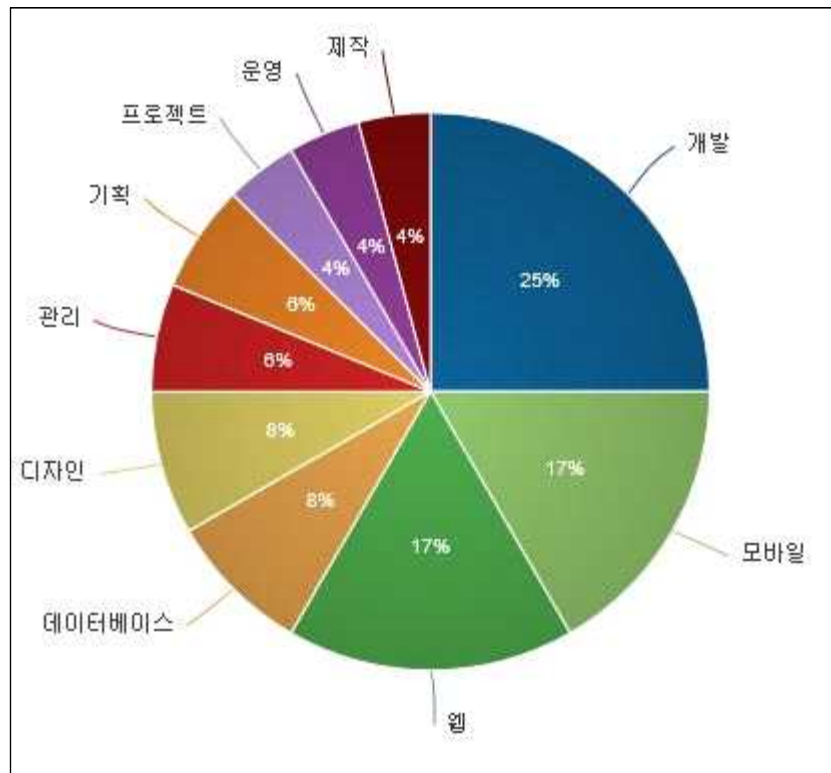
그에 따라 [그림 2-119]는 직무 12 군집의 공동 출현 네트워크를 소셜네트워크 분석을 기반으로 도출하였다. 분석 결과 {디자인-모바일-개발}, {개발-관리-프로젝트}, {디자인-개발-기획}에 대한 높은 연관 관계가 존재한다. 이는 다른 직무 군집들과 다르게 직무 12 군집의 경우 공동 출현 네트워크 관계가 몇 가지 관계를 제외하면 상호 균일하게 등장하는 특징도 확인할 수 있었다.

〈표 2-59〉 직무 12의 핵심단어 목록

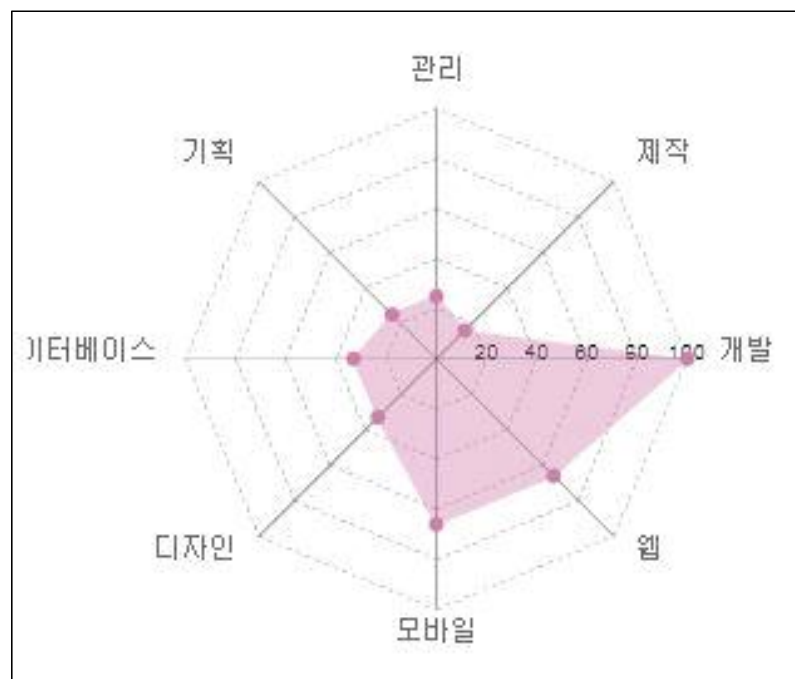
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	개발	0.1263
2	웹	0.0875
3	모바일	0.0852
4	디자인	0.0473
5	데이터베이스	0.0423
6	기획	0.0345
7	관리	0.0301
8	제작	0.0232
9	운영	0.0231
10	프로젝트	0.0219

데이터 분석 기반의 구분된 직무 체계를 바탕으로 데이터베이스 관리자에 대한 LDA 분석을 수행한 결과는 〈표 2-59〉과 같다. 도출된 핵심단어의 동시등장 빈도율에 따른 주요 핵심단어를 확인하였으며 해당 직무 내 동시등장 빈도율 2%이상에 해당하지 않는 단어들은 직무 내 핵심단어에서 탈락시켰다. 그에 따라 해당 직무 내에서 동시등장 빈도율이 큰 단어부터 확인할 때 나타난 핵심단어는 ‘개발’, ‘웹’, ‘모바일’, ‘디자인’, ‘데이터베이스’, ‘기획’, ‘관리’, ‘제작’, ‘운영’, ‘프로젝트’의 순서로 나타났다. 또한 직무 12의 핵심단어 비율 그래프는 [그림 2-120]과 같이 도출되었다. 주요 핵심단어는 개발, 모바일, 웹 등이지만 데이터베이스를 요구하는 핵심단어 배열임을 확인할 수 있었다. 즉, 직무 12의 경우 데이터베이스 관리에 대한 범용 개발 및 관리를 할 수 있는 역량을 요구한다.

[그림 2-120] 직무 12의 핵심단어 비율 그래프



[그림 2-121] 직무 12의 핵심단어 방사형 그래프



직무 12에 대한 핵심단어들의 동시등장 빈도율을 100점으로 환산하여 [그림 2-121]과 같이 방사형 그래프를 시각화 하였다. 그 결과 개발, 웹, 모바일, 디자인 등이 제시되고 있으며 데이터베이스, 기획, 관리 등의 핵심단어 영역이 발견되었다.

[그림 2-122] 직무 12의 핵심단어 워드클라우드



직무 12의 핵심단어 워드클라우드를 [그림 2-122]와 같이 시각화 한 결과 데이터베이스 프로그램인 MySQL이나 Oracle, 서버, SQL, JAVE, 라이브러리 등의 핵심단어들이 출현하여 데이터베이스의 구축, 운영, 관리, 기획, 설계 등을 행하는 직무 군집임을 판단할 수 있었다. 그에 따라 직무 12 군집은 데이터베이스 성능 최적화와 보안 유지를 위해 데이터베이스를 디자인하고 개발, 관리하는 역할을 수행한다고 볼 수 있었다.

직무 12 군집의 직무 내용을 명확히 하기 위하여 <표 2-60>과 같이 국제직무분류표준의 정의를 비교 분석하였다. 그 결과 ISCO, SOC, O*Net에서 모두 데이터베이스 관리자로서의 직무 분류 및 정의가 되어 있음을 확인하여 해당 직무와 본 보고서의 핵심단어들을 비교하였다. 그 결과 본 보고서의 분석 결과에서도 동일 직무 능력을 핵심단어로서 제시되는 것으로 판단하였다.

그에 따라 본 보고서는 직무 12 군집을 ‘데이터베이스 관리자’로 정의하였다. 데이터베이스 관리자는 데이터베이스 시스템의 분석과 운영 관리 및 튜닝의 업무를 맡는 직무로서 정의하였다. 또한 해당 직무는 데이터베이스의 보안 정책을 계획하고 이에 따른 보안 관리(자료보안, 사용자보안)를 준수하며 전체적인 데이터베이스 시스템의 구조를 분석하고 평가하는 역할도 겸하는 것으로 판단하였다.

〈표 2-60〉 데이터베이스 관리자 직무 관련 국제표준

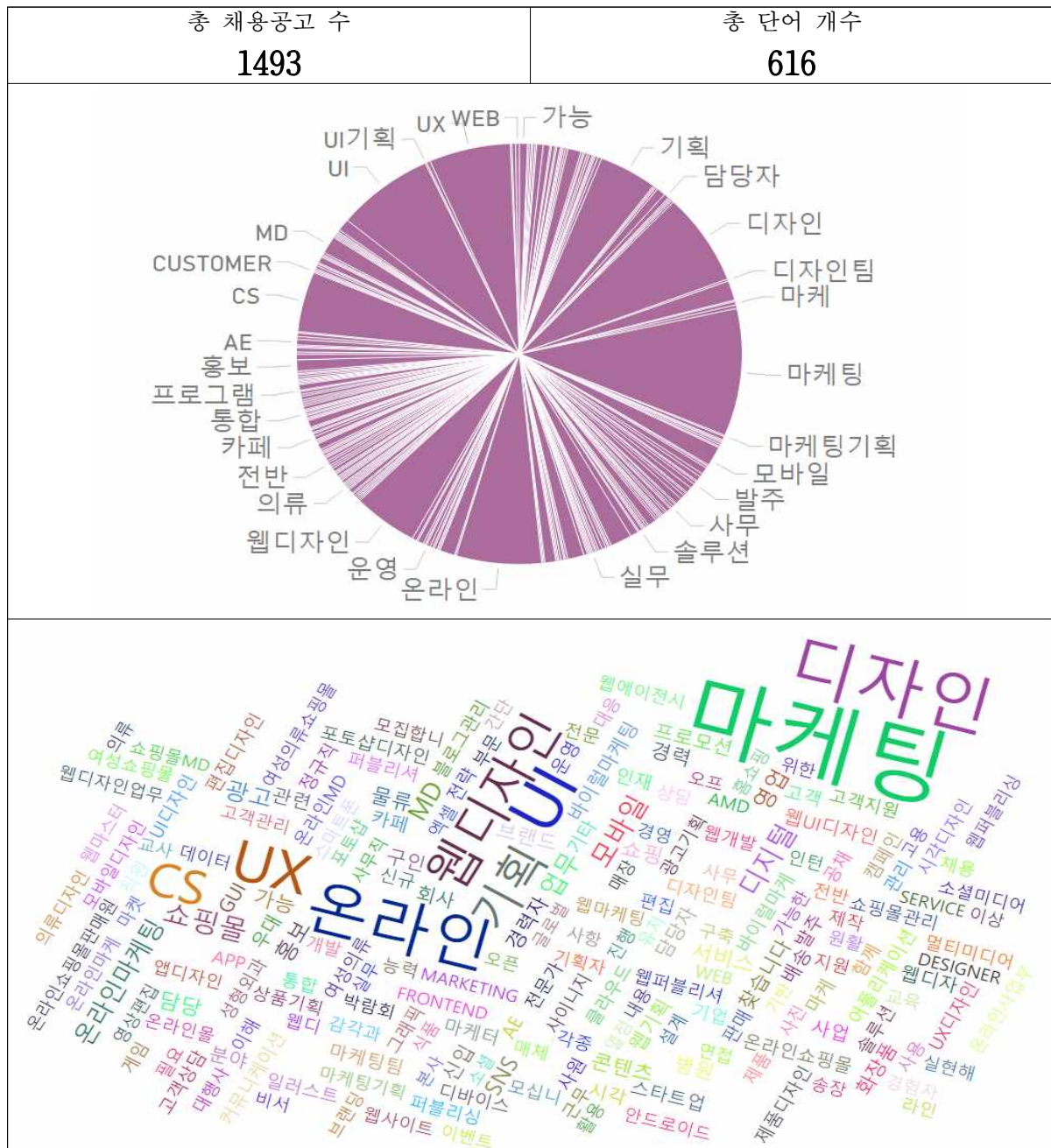
소스	분류코드	직무	정의
ISCO-08	2521	Database designers and administrators	데이터베이스의 최적화된 성능과 보안을 설계, 개발, 통제 및 유지보수 지원
SOC-2016	15-1242	Database Administrators	데이터베이스 관리 시스템의 지식을 바탕으로 데이터베이스에 대한 관리, 시험 및 개선
O*Net	15-1141.00	Database Administrators	- 컴퓨터 데이터베이스 수정에 참여 - 데이터베이스 안전관리를 위한 보안사항 기획 및 개선
본 보고서의 분류		데이터베이스 관리자	- 데이터베이스 시스템의 분석과 운영 관리 및 튜닝의 업무의 전문가를 의미한다. - 데이터베이스의 보안 정책을 계획하고, 이에 따른 보안 관리(자료보안, 사용자보안)를 준수해 나간다. 전체적인 데이터베이스 시스템의 구조를 분석하고 평가

이에 따라 데이터베이스 관리자의 직무 내용은 다음과 같이 요약할 수 있다.

- 전체 데이터베이스 시스템 구조를 분석하고 평가.
- 데이터베이스의 저장공간, 접근방법 등과 같은 논리적 및 물리적 구조 분석 및 디자인
- 데이터베이스 사용자 등록 및 모니터링 수행
- 데이터베이스 자료에 대한 백업 및 복구절차를 제작하고 백업주기에 따른 데이터의 백업 실시
- 데이터베이스의 보안 정책 계획 및 자료와 사용자 대상 보안 관리 수행
- 데이터베이스 성능을 모니터링하고 및 모바일 등 환경 변화에 따른 성능향상을 위한 시스템 튜닝 실시
- 기존 데이터베이스 업그레이드 내지는 새로운 데이터베이스 시스템의 개발 구축 시 프로젝트 개발팀에 합류되어 데이터베이스 분석, 설계, 테스트
- 데이터베이스 시스템 고장 및 사용상의 문제가 발생 시 최단 복구 및 차후 시스템 변경 및 업그레이드에 반영

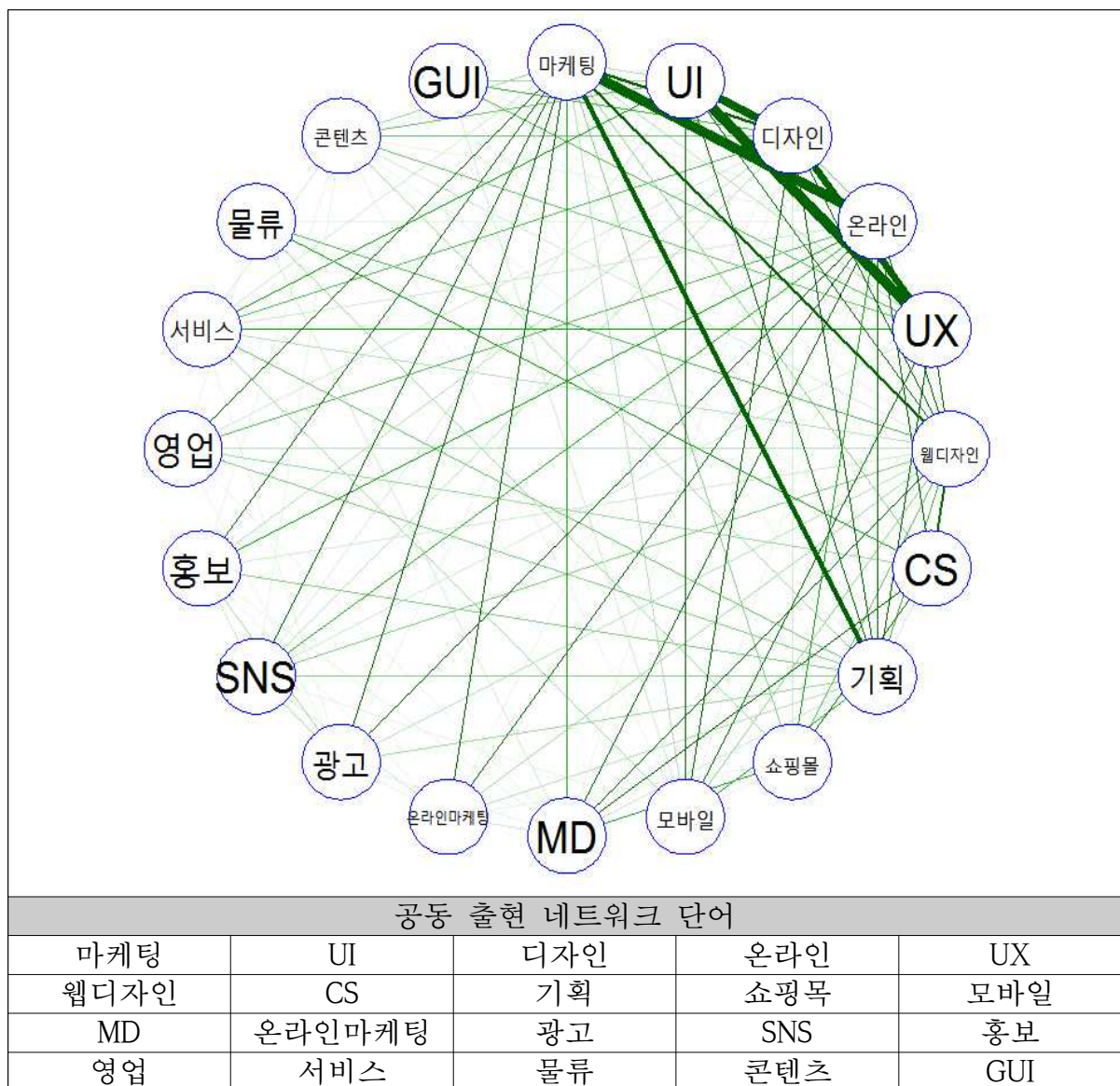
2.2.13) HCI 컨설턴트

[그림 2-123] 직무 13의 특성



직무 13 군집에 해당 하는 채용공고는 [그림 2-123]과 같이 총 1493건으로 616개의 단어가 등장한다. 해당 군집에 해당하는 채용공고에 대한 단어들 중 빈번하게 등장하는 단어는 ‘UX’, ‘UI’, ‘웹디자인’, ‘온라인’, ‘기획’, ‘디자인’, ‘마케팅’, ‘CS’ 이다. 해당 직무의 전체 단어들을 대상으로 워드클라우드 결과를 확인한 결과 디자인, 온라인, UX, UI, MD 등의 직무 관련 단어들이 출현함에 따라 그래픽 디자이너와는 다른 사용자 관련 인터페이스와 마케팅 관련 전문 직무임을 유추할 수 있었다.

[그림 2-124] 직무 13의 단어 공동 출현 네트워크



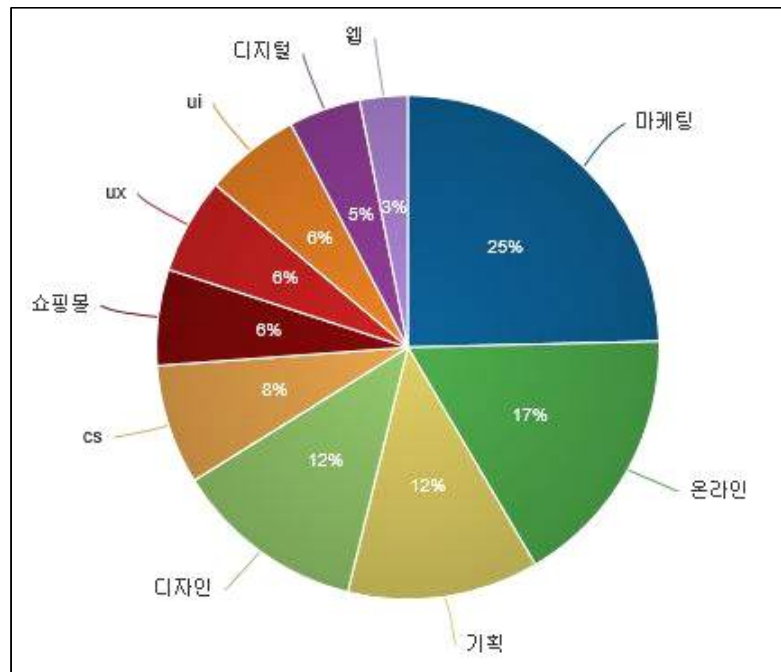
[그림 2-124]는 직무 13 군집에 대한 전체 채용공고에 출현한 단어 조합을 소셜네트워크 분석으로 도출한 공동 출현 네트워크이다. 분석 결과 주로 {UI-UX-디자인-온라인}, {마케팅-온라인}의 강한 연관관계를 가지는 단어 조합이 발견되었고 {마케팅-기획}에 대한 관계가 존재함을 확인하였다.

〈표 2-61〉 직무 13의 핵심단어 목록

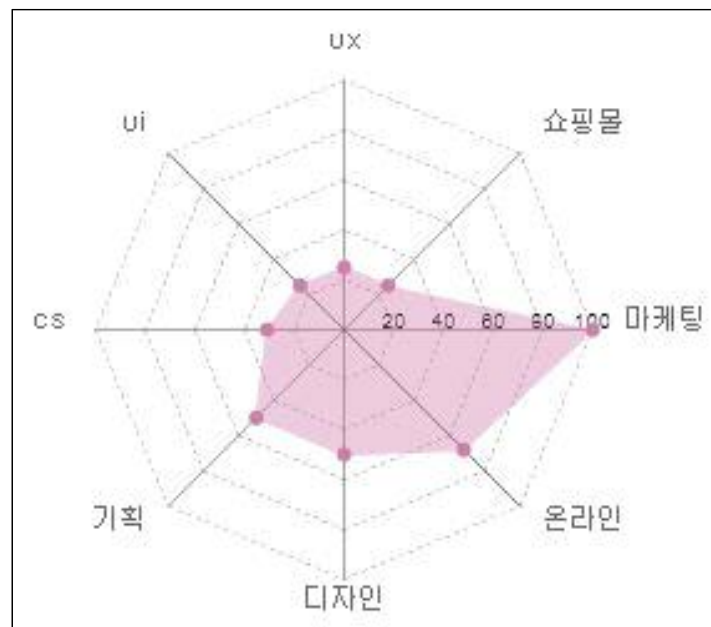
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	마케팅	0.1693
2	온라인	0.1110
3	디자인	0.0869
4	기획	0.0826
5	cs	0.0530
6	UI	0.0462
7	UX	0.0422
8	쇼핑몰	0.0411
9	디지털	0.0307
10	웹	0.0293

직무 체계 기반으로 직무 13에 대한 LDA 분석 결과를 확인한 결과는 〈표 2-61〉과 같다. 도출된 동시등장 빈도율에 따라 직무 내 동시등장 빈도율이 2% 이하인 값을 포함하지 않고 동시등장 빈도율이 큰 단어부터 확인할 때 나타난 핵심단어는 ‘마케팅’, ‘온라인’, ‘디자인’, ‘기획’, ‘CS’, ‘UI’, ‘UX’, ‘쇼핑몰’, ‘디지털’, ‘웹’의 순서로 나타났다. 따라서 본 직무는 사용자 경험을 반영할 수 있는 디자인을 역량을 요구하며, 마케팅과 기획의 관점과 관련된 역량을 요구하고 있다. [그림 2-125]는 직무 13에 대한 핵심단어 비율 그래프이다. 시각화 결과 마케팅, 온라인과 관련하여 기획 및 디자인 직무를 요구하고 있는 직무임을 확인하였고 cs, UX, UI와 같은 특성화 직무가 도출되어 UX와 UI를 기반으로 사용자 대상 마케팅을 진행하는 HCI 관련 직무들이 도출되었음을 확인하였다.

[그림 2-125] 직무 13의 핵심단어 비율 그래프



[그림 2-126] 직무 13의 핵심단어 방사형 그래프



[그림 2-126]은 직무 13 군집에 대한 핵심단어들을 바탕으로 동시등장 빈도율을 100점으로 환산하여 방사형 그래프를 시각화 하였다. 그 결과 마케팅, 온라인, 디자인, 기획 등의 업무를 수행하며 cs, UX, UI를 활용하는 직무임이 나타났다.

이와 마찬가지로 [그림 2-127]과 같이 직무 13 군집에 대한 핵심단어 워드클라우드를 시각화 하였다. 시각화 결과를 살펴보면 온라인, 콘텐츠, 쇼핑몰 등의 단어가 주요하게 나타나고 있으며 이와 함께 UX, UI, CS 등의 직무 특성이 반영된 핵심 단어의 점유율이 높게 나타나 주로 HCI 영역의 직무 단어들이 출현하고 있음을 확인하였다.

[그림 2-127] 직무 13의 핵심단어 워드클라우드



특히 최근 Linked-In에서 다수의 HCI 컨설턴트 공고를 확인할 수 있으며 매우 활발한 채용공고가 진행되고 있다. 특히 해외에서는 HCI Consultant 또는 HCI Specialist로 공고가 많이 게시되고 있으며 UX 및 UI를 책임지는 전문가로서의 역할을 하는 신규 직무이다. 그에 따라 본 보고서에서는 직무 13 군집에 대하여 ‘HCI 컨설턴트’로 명칭 하였다. HCI 컨설턴트는 livari (2006)의 연구를 참고하여 개발자들과 함께 웹 디자인 설계 및 디자인 의사결정에 참여하며 사용자에게 적합한 컨설팅을 제시하는 집단으로서 정의할 수 있다. 특히 인간과 컴퓨터의 상호작용(Human Computer Interaction, 이하 HCI) 을 토대로 사용자 친화적이고 접근성이 확보된 컴퓨터시스템 및 웹사이트 등을 개발하고 컨설팅 하는 영역을 담당하고 있는 직무이다.

〈표 2-62〉 HCI 컨설턴트 직무 관련 국제표준

소스	분류코드	직무	정의
SOC-2016	15-1255	Web or Digital Interface Designers	• 웹사이트 및 디지털 사용자 인터페이스 설계 • 웹 브라우저 및 장치에 대한 사용성 개선을 위한 레이아웃, 인터페이스, 기능성 및 내비게이션에 대한 개발 및 시험 • 마케팅 가능성 및 검색엔진 순위에 대한 웹사이트 최적화 • 웹 사용 방식 분석 및 웹 접속 표준 설계 및 평가 • HCI 응용 인터페이스 설계 및 시험 • 디지털 장치, 웹사이트, SW의 사용성 최대화 및 상호작용성 설계 및 개선 • 웹사이트 콘텐츠 및 링크 관리 및 웹사이트 내 그래픽 요소 개발
O*Net	27-1014.00	Multimedia Artists and Animators	• 컴퓨터 게임, 영화, 뮤직비디오 그리고 상거래 활동에서 활용되는 제품 또는 창조물에 대한 영상, 전자적 톨과 미디어를 사용한 특수효과, 애니메이션 및 시각적 이미지를 만들어냄
Hedberg & livari (2009)		HCI Consultant	• 사용성, 사용자 경험, 사용자 인터페이스 등을 기반으로 상호작용 설계를 위한 전문가
본 보고서의 분류 ³³⁾		HCI 컨설턴트	• 개발자들과 함께 웹 디자인 설계 및 디자인 의사결정에 참여하며 사용자에게 적합한 컨설팅을 제시하는 집단이다. • 인간과 컴퓨터의 상호작용을 토대로 사용자 친화적이고 접근성이 확보된(HIC : Human Computer Interaction) 컴퓨터시스템 및 웹사이트 등을 개발하고 컨설팅 한다. • 이미 해외에서는 HCI Consultant 또는 HCI Specialist로 공고가 많이 게시되고 있으며 UX, UI를 책임지는 전문가로서의 역할을 하는 직무이다.

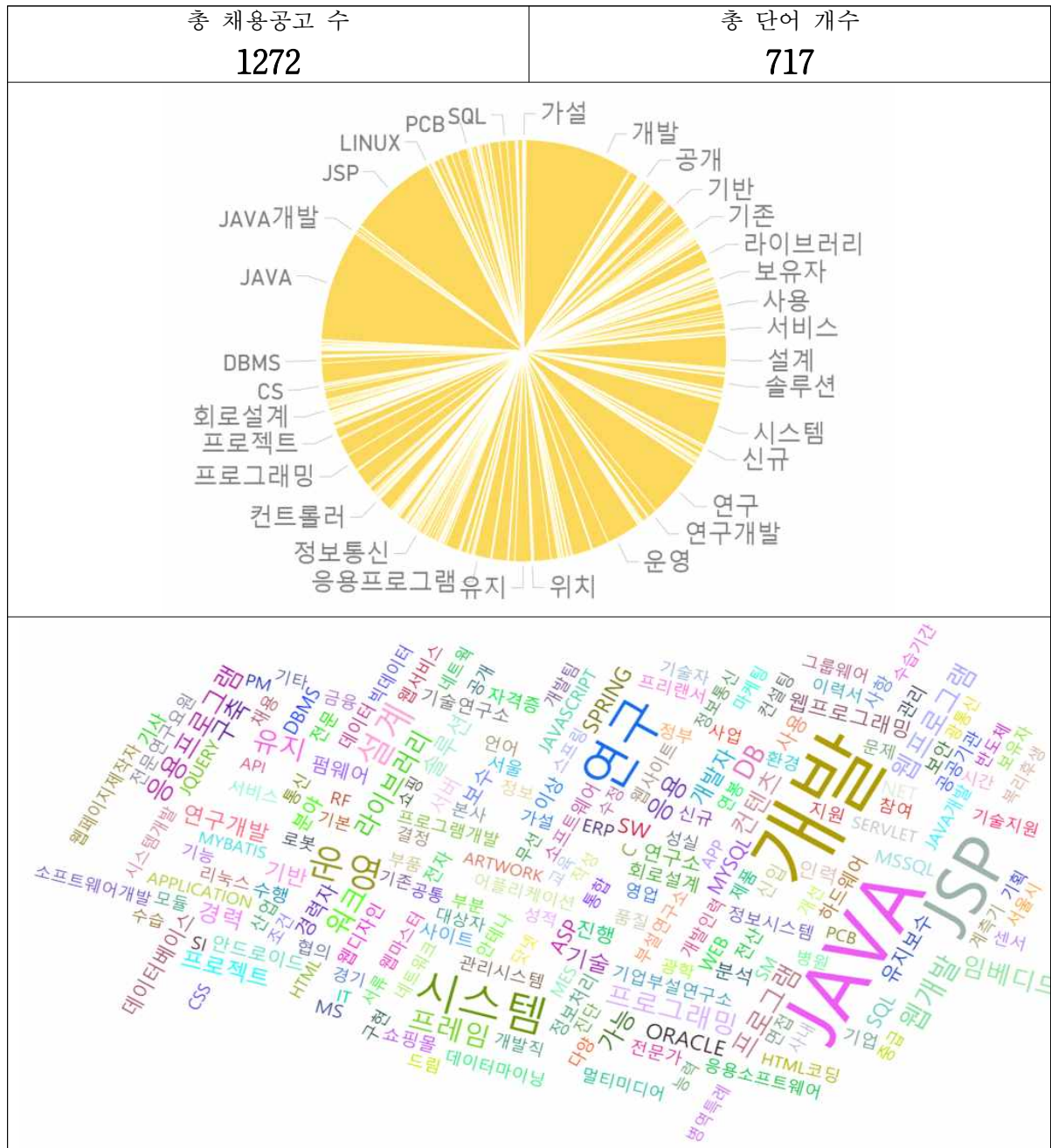
33) Hedberg & livari (2009)와 livari (2006)의 정의를 참고하였음

특히 HCI 컨설턴트는 <표 2-62>와 같이 멀티미디어 예술가 및 애니메이터 직무에 속하는 개념으로 SOC 2016버전에서 HCI를 다루는 전문가로서 웹 또는 디지털 인터페이스 디자이너로 제시되고 있다. 따라서 HCI 컨설턴트는 UX 또는 UI를 직접 다루는 직무로 볼 수 있다. 따라서 기존 국제 분류 체계에서는 공식적으로 직종이 제시되고 있지 않지만 HCI컨설턴트는 사용성, 사용자 경험, 사용자 인터페이스 등을 기반으로 상호작용 설계를 위한 전문가로서 정의할 수 있다(Hedberg & livari, 2009). 그에 따라 HCI 컨설턴트에 대한 직무에 대한 요약은 다음과 같다.

- 사용자 친화적 웹사이트와 컴퓨터부품 등을 개발하기 위해 문제점 분석
- 고객을 대상으로 사용자 친화성을 평가하기 위해 심층 인터뷰 진행 및 UX와 UI
- 고객에게 개선된 사이트를 사용하게 하는 사용성 테스트 실시
- 웹사이트 관련 고객관계 및 만족도에 대한 태도 및 상황적 인지 등을 평가
- 인간-컴퓨터의 상호작용(HCI)이 원활한 사이트가 되게 하는 방안 기획 및 제시를 바탕으로 마케팅 활동 지원 및 UI 디자인, CS활동 등 수행

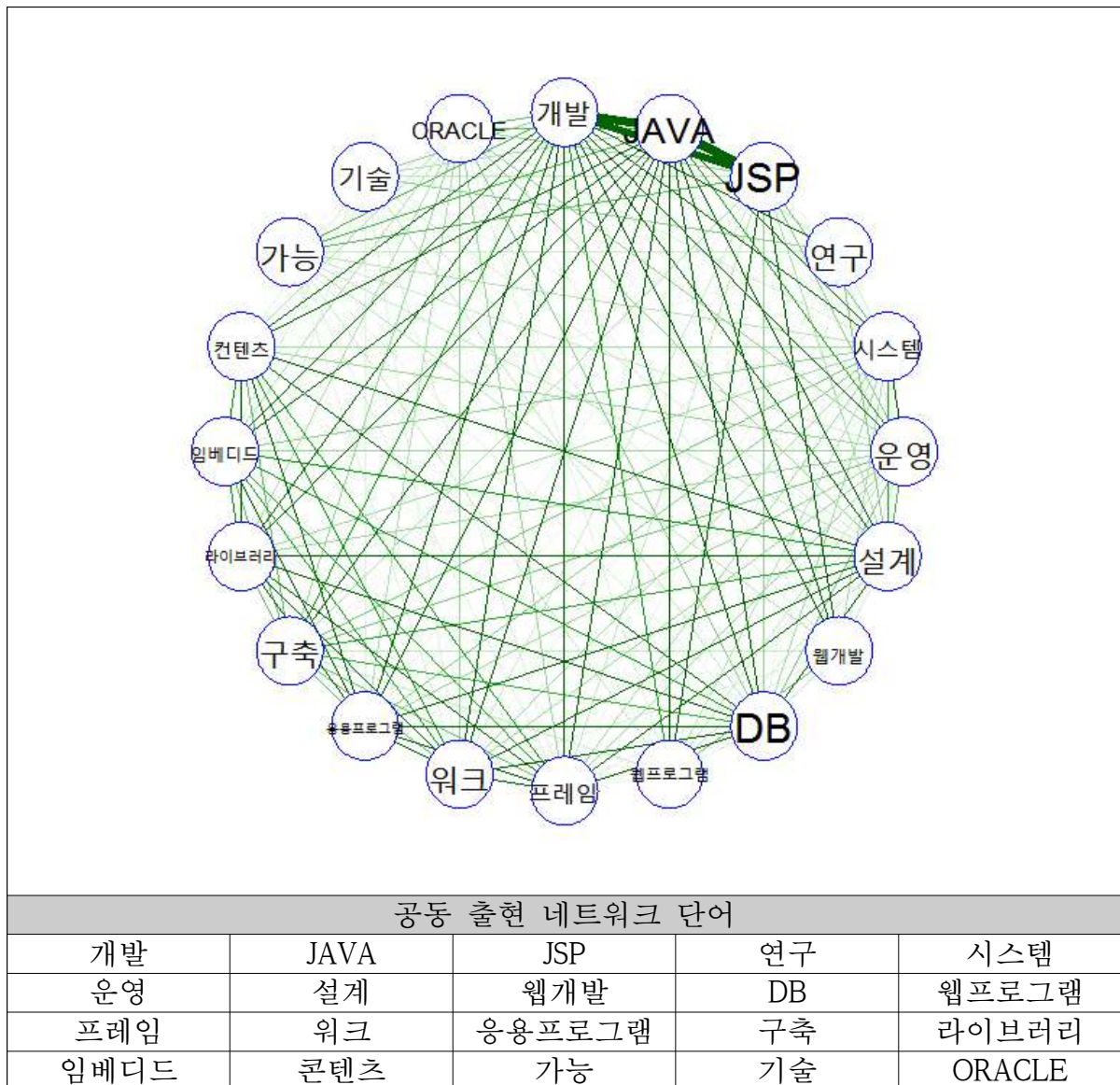
2.2.14) 애플리케이션 개발자

[그림 2-128] 직무 14의 특성



직무 14 군집에 해당 하는 채용공고는 [그림 2-128]과 같이 총 1272건으로 717개의 단어가 등장한다. 이 중 직무 14 군집과 관련된 전체 채용공고에서 가장 빈번하게 등장하는 단어는 ‘개발’, ‘JAVA’, ‘JSP’, ‘시스템’, ‘연구’ 이다. 이와 함께 해당 직무에 해당하는 전체 채용공고에 대한 출현단어를 워드클라우드를 도출한 결과 연구 개발, Java, JSP 등의 프로그램 관련 직무단어가 도출됨을 확인할 수 있었다. 또한 웹개발, 프로그래밍, JAVASCRIPT, DB와 소프트웨어, 시스템 개발 등의 단어들이 출현함에 따라 응용 소프트웨어와 관련된 개발 직무임을 확인할 수 있었다.

[그림 2-129] 직무 14의 단어 공동 출현 네트워크



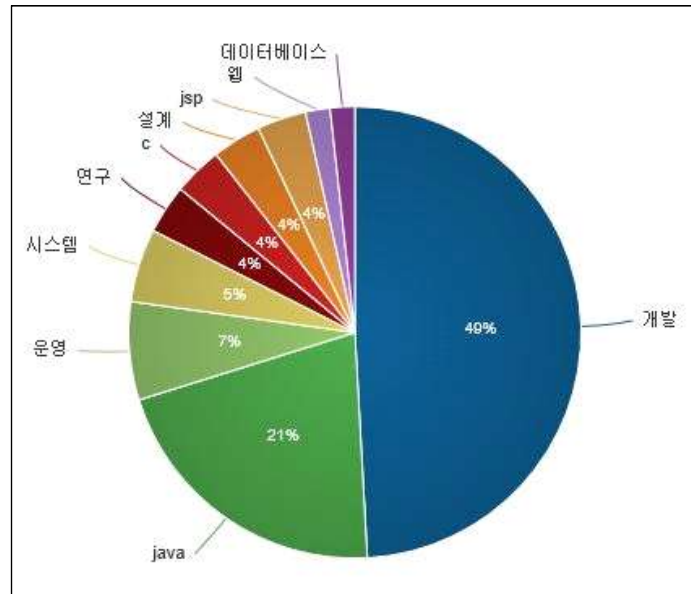
직무 14 군집에 대한 소셜네트워크 분석 기반의 공동 출현 네트워크는 [그림 2-129]와 같다. 이 결과 중에서도 {개발-JAVA-JSP}의 단어 간의 강한 연관관계가 존재한다. 그에 따라 직무 14 군집의 경우 DB, 임베디드 등 다양한 연구, 개발과 관련이 있지만 JAVA, JSP 기반의 연구 개발의 역량을 요구하는 것으로 확인되었다.

〈표 2-63〉 직무 14의 핵심단어 목록

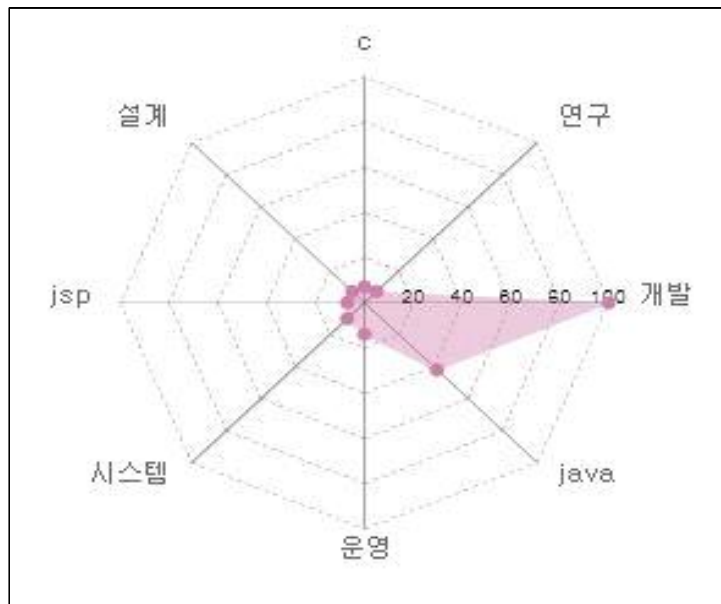
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	개발	0.2862
2	java	0.1262
3	운영	0.0483
4	시스템	0.0374
5	jsp	0.0274
6	설계	0.0250
7	c	0.0217
8	연구	0.0212
9	데이터베이스	0.0200

직무 14 군집의 LDA 분석을 수행한 결과 해당 직무 내 동시등장 빈도율 2%이상인 핵심단어들만으로 구성했을 때 도출된 핵심단어는 〈표 2-63〉과 같이 ‘개발’, ‘JAVA’, ‘운영’, ‘시스템’, ‘JSP’, ‘설계’, ‘c’, ‘연구’, ‘데이터베이스’의 순서로 도출되었다. 이 중 ‘연구’ 단어가 직무와 어떤 상관관계에 있는 지 확인할 필요가 있어 채용공고 원본을 확인하였다. 그 결과 새로운 애플리케이션 시스템 개발이나 기존 애플리케이션 시스템의 개선을 위한 ‘연구’ 공고가 주요한 것으로 나타났다. 결과적으로 도출된 핵심단어들은 응용소프트웨어 개발과 관련된 단어들로 판단하였다. 이와 관련하여 직무 14에 대한 핵심단어 기반의 비율 그래프를 시각화 한 결과는 [그림 2-130]과 같다. 도출된 시각화 그래프를 보면 절반정도의 비율로 개발 직무를 요구하는 군집이며 주로 JAVA와 운영뿐만 아니라 시스템, 설계, jsp 등을 활용하는 직무라는 것을 확인할 수 있다. 마찬가지로 직무 14에 대한 동시등장 빈도율을 100점 환산한 결과를 바탕으로 방사형 그래프를 도출한 결과는 [그림 2-131]과 같다. 시각화 결과를 살펴보면 주로 개발 및 JAVA와 관련된 응용 프로그래밍에 중점을 두고 핵심단어들이 제시되고 있음을 확인할 수 있었고 그 외 활용 기술인 c와 jsp 등을 바탕으로 웹 기반 응용소프트웨어 개발 직무에 해당하는 결과임을 유추할 수 있었다.

[그림 2-130] 직무 14의 핵심단어 비율 그래프



[그림 2-131] 직무 14의 핵심단어 방사형 그래프



[그림 2-132]는 직무 14 군집에 대한 핵심단어들을 바탕으로 워드클라우드 시각화를 도출한 결과이다. 그 결과 JAVA와 개발이 주요 핵심단어로 나타났으며 C++ 및 JSP를 바탕으로 하드웨어 및 소프트웨어를 통한 프로그래밍 및 설계와 연구 직무를 주로 요구하는 직무 군집임을 확인하였다.

[그림 2-132] 직무 14의 핵심단어 워드클라우드



그에 따라 직무 14 군집은 소프트웨어 응용프로그램에 대한 기술적인 지침과 기획을 바탕으로 Java, C++ 같은 프로그래밍 언어를 활용하여 코드를 작성하고 관리하며 애플리케이션 및 응용프로그램을 개발하고 운영하는 역할을 수행하는 것으로 판단하였다. 분석결과에 따라 직무 14 군집에 대한 구체적인 직무명과 직무 내용을 확정하기 위하여 국제직무분류표준을 비교 분석하였다. <표 2-64>에서 보이는 것과 같이 기존 국제 분류 체계에서도 응용 소프트웨어 개발자로서 정의를 내리고 있으며 직무의 주요 내용을 확인해보면 SW 애플리케이션 및 운영 시스템에 대한 기술적 지침 및 특수사항에 대한 프로그램 코드 작성 및 유지보수 및 일반적인 애플리케이션 SW 또는 특수 프로그램에 대한 개발, 작성 및 수정을 주요 직무로 판단하고 있다. 또한 SW 솔루션에 대한 개발 및 사용자 수요 분석까지도 직무의 영역으로 정의되고 있다. 그에 따라 본 보고서에서는 직무 14 군집에 대하여 ‘애플리케이션 개발자’ 직무로 정의하였으며 애플리케이션 개발자를 고객의 요구에 적합한 애플리케이션을 개발하고 유지·관리하는 직무로서 정의하였다. 특히 해당 직무는 새로운 애플리케이션 시스템 개발 혹은 기존 애플리케이션 시스템의 개선을 위한 업무분석 및 설계 업무에 참여하는 전문가로 판단할 수 있다.

〈표 2-64〉 애플리케이션 개발자 직무 관련 국제표준

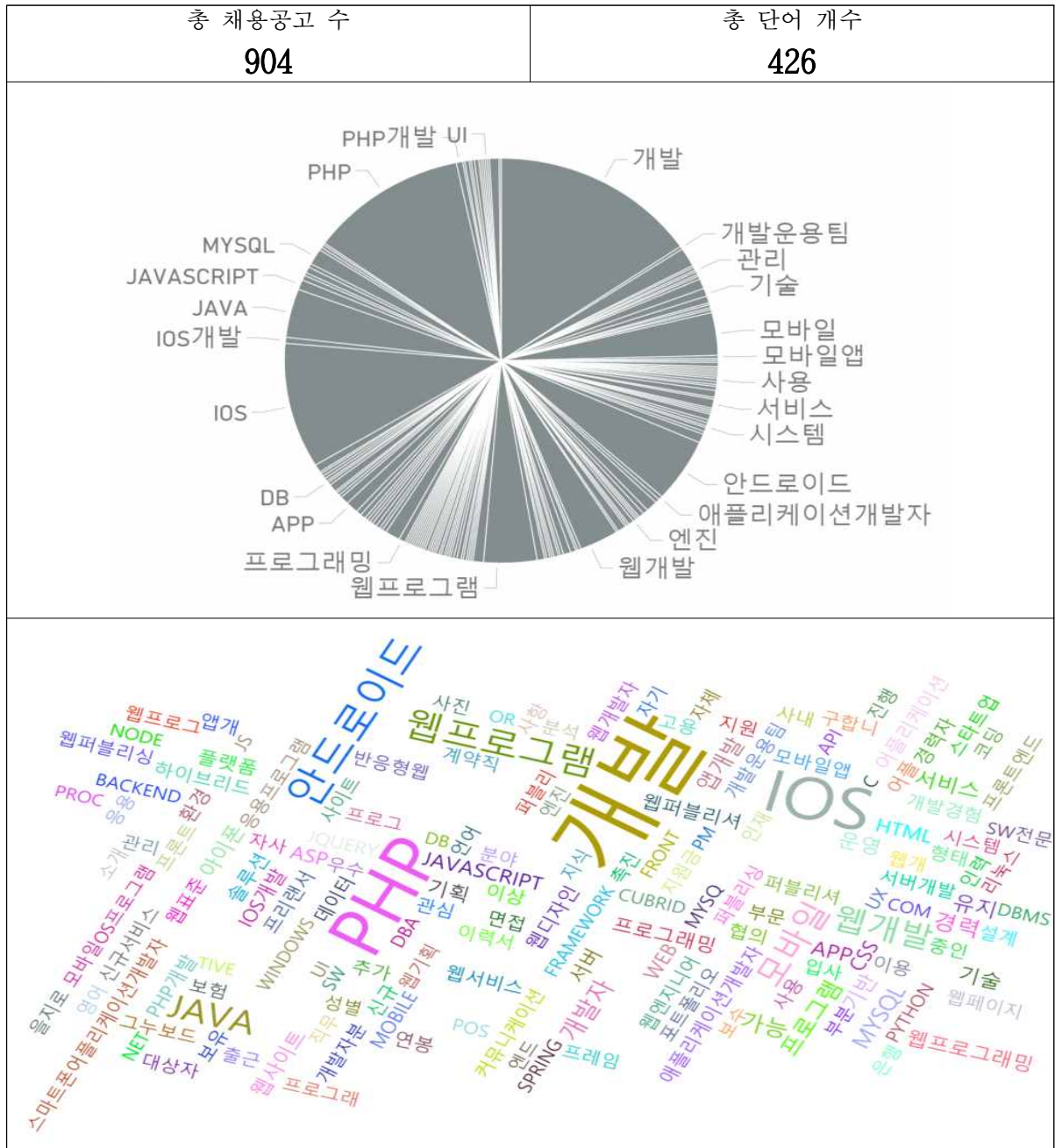
소스	분류코드	직무	정의
ISCO-08	2514	Applications programmers	SW 애플리케이션 및 운영 시스템에 대한 기술적 지침 및 특수사항에 대한 프로그램 코드 작성 및 유지보수
SOC	15-1132	Software Developers, Applications	- 일반적인 애플리케이션 SW 또는 특수 프로그램에 대한 개발, 작성 및 수정 - SW 솔루션에 대한 개발 및 사용자 수요 분석
O*Net	15-1132.00	Software Developers, Applications	- 운영상 효율성 최적화의 목적에서 사용자 중심의 맞춤형 SW 설계 - 애플리케이션 SW 영역의 데이터베이스 분석 및 설계 - 일반 프로그래머 감독 및 지원
본 보고서의 분류		애플리케이션 개발자	- 고객의 요구에 적합한 애플리케이션을 개발하고 유지·관리한다. - 새로운 애플리케이션 시스템 개발 혹은 기존 애플리케이션 시스템의 개선을 위한 업무분석 및 설계 업무에 참여한다.

이와 관련하여 애플리케이션 개발자에 대한 역량을 요약하면 다음과 같다.

- 애플리케이션의 성능 검사 및 실행시간 개선 및 개발과 관련하여 실제 사용자와의 창구 역할 수행
- 애플리케이션 관련 고객 요구사항 수집·분석 / 문제점 및 대책 확인·지원
- 신규 애플리케이션 시스템 개발이나 기존 애플리케이션 시스템의 개선을 위해 문제점을 찾고 개선
- 애플리케이션 시스템에 필요한 사용자 교육, 장비, 하드웨어, SW 설치 주도

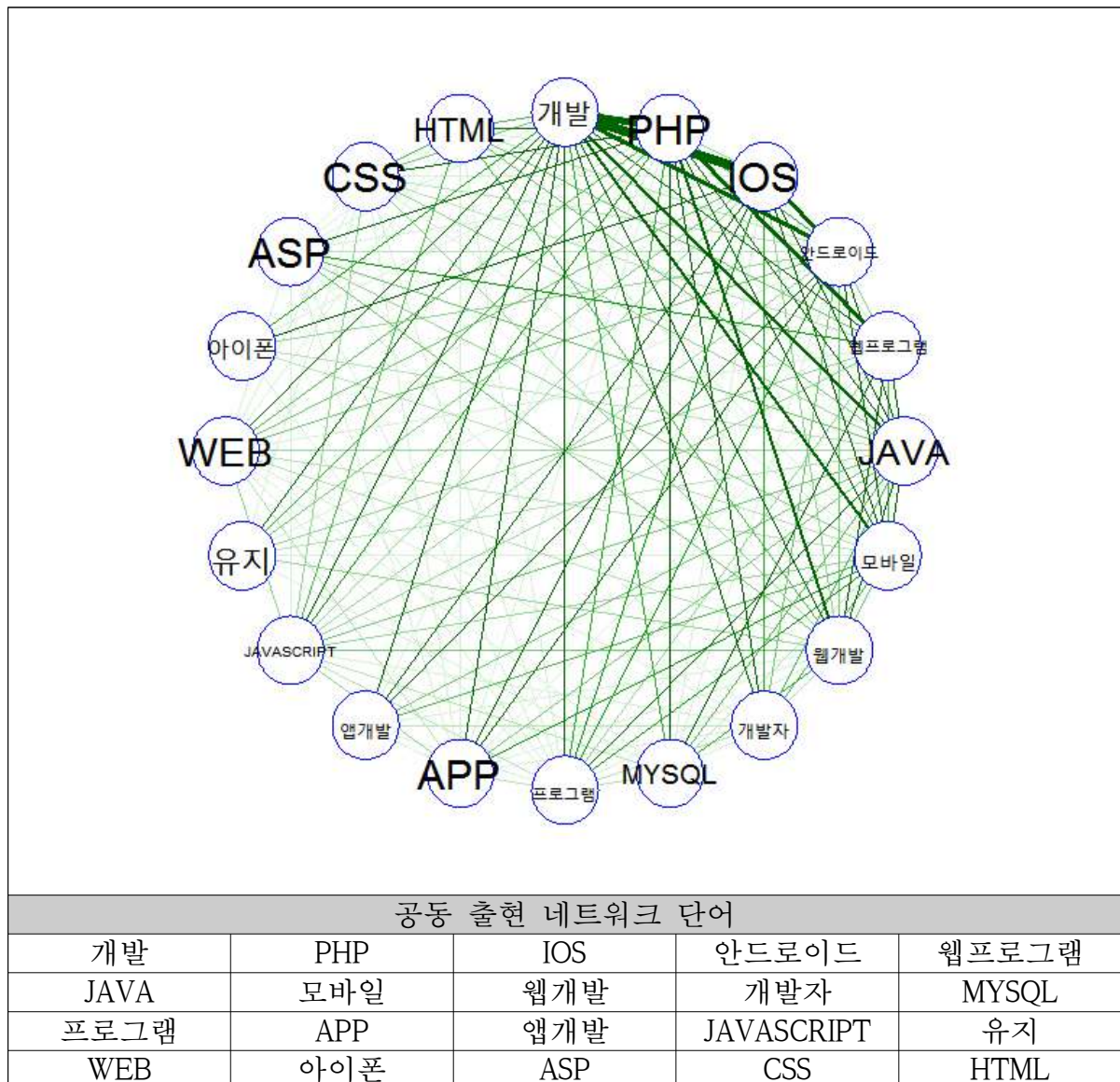
2.2.15) 모바일 개발자

[그림 2-133] 직무 15의 특징



직무 15 군집에 해당 하는 채용공고는 [그림 2-133]과 같이 총 904건으로 426개의 단어가 등장한다. 이 중 직무 15에 해당하는 채용공고에서 가장 빈번하게 등장하는 단어는 ‘PHP’, ‘ios’, ‘안드로이드’, ‘개발’ 이다. 마찬가지로 직무 15와 관련된 채용공고의 전체 단어를 바탕으로 워드클라우드 시각화를 한 결과 개발, Android, ios, PHP, 모바일, 스마트폰애플리케이션개발자 등과 같은 모바일 기반의 직무 내용이 도출됨에 따라 개발자 직무 중에서도 모바일에 특화된 직무 군집임을 알 수 있다.

[그림 2-134] 직무 15의 단어 공동 출현 네트워크



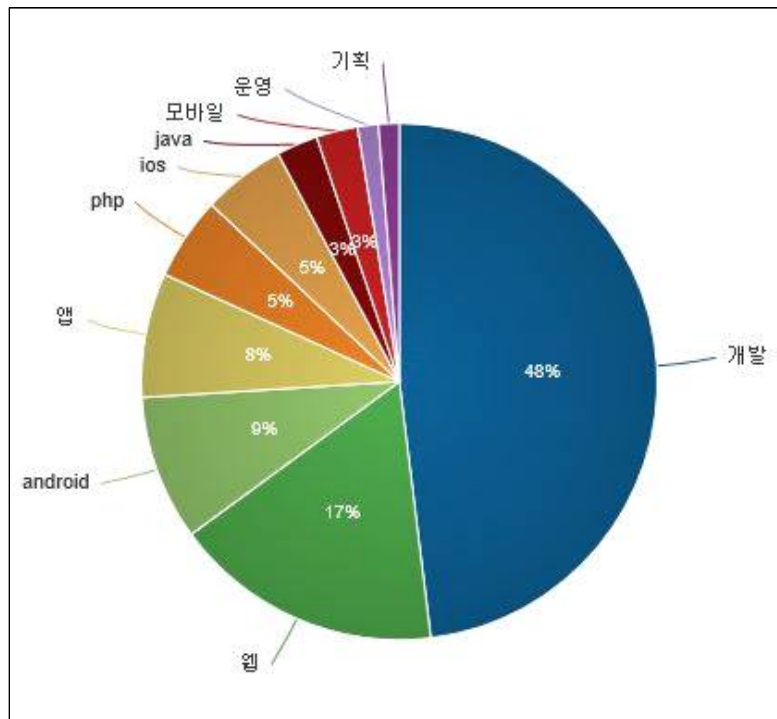
직무 15 군집에 대한 소셜 네트워크 분석을 통해 도출된 공동 출현 네트워크는 [그림 2-134]와 같다. 이 결과에서 주로 {개발-PHP}, {개발-IOS}의 단어 조합이 가장 채용공고에 함께 출현하는 관계로서 나타났으며 {개발-IOS-안드로이드}, {개발-자바}, {PHP-웹개발}의 관계도 주요 출현하는 단어 조합으로 제시되었다. 즉, 직무 15 군집의 경우 ios, 안드로이드와 같은 모바일 운영체제에서의 개발 역량과 PHP-웹개발 역량을 요구하는 것을 확인할 수 있다.

〈표 2-65〉 직무 15의 핵심단어 목록

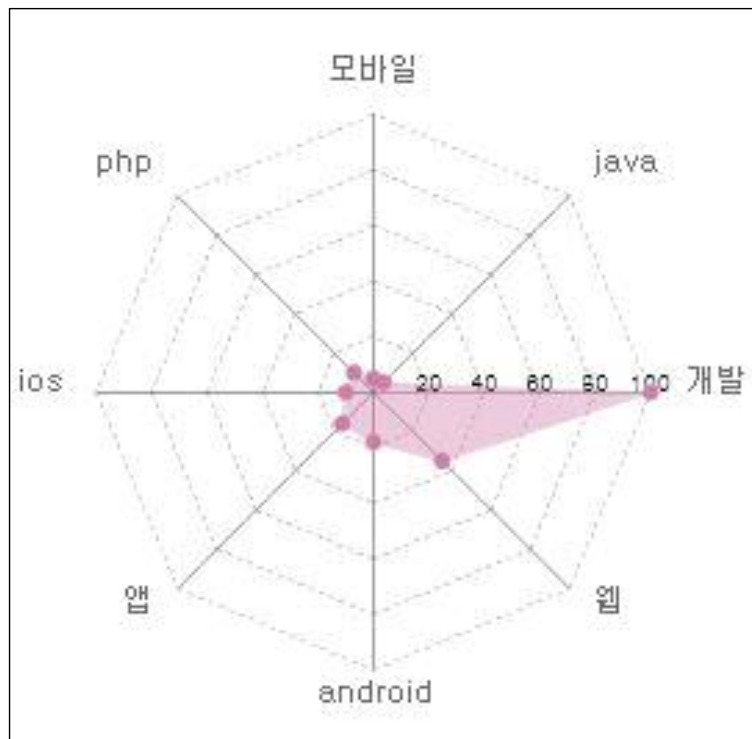
순위	핵심단어	동시등장 빈도율 (0.02 이상)
1	개발	0.3717
2	웹	0.1351
3	android	0.0709
4	앱	0.0687
5	ios	0.0499
6	php	0.0410
7	모바일	0.0271
8	java	0.0243

직무 15 군집에 대해 LDA 분석을 기반으로 도출된 직무 체계 기반 단어들의 비교를 수행하였다. 분석 결과 〈표 2-65〉와 같이 해당 직무 내 동시등장 빈도율 2%이상인 단어들만을 바탕으로 해당 직무 내에서 동시등장 빈도율이 큰 단어부터 도출된 핵심단어는 ‘개발’, ‘웹’, ‘android’, ‘앱’, ‘ios’, ‘php’, ‘모바일’, ‘java’의 순서로 도출되었다. 해당 핵심단어에 포함되지 않았으나 ‘기획’ 및 ‘운영’ 단어가 도출되어 이에 대한 채용공고 원본을 확인하였다. 그 결과 모바일 개발 기획 및 모바일 프로그램의 운영과 관련하여 해당 단어들이 공고에 제시된 것을 확인할 수 있었으며 결과적으로 도출된 핵심단어들은 해당 직무 군집과 연관된 핵심단어임을 확인하였다. 이를 바탕으로 [그림 2-135]와 같이 직무 15 군집에 대한 핵심단어 비율 그래프를 도출하였다. 그 결과 48%가 개발로 나타났으며 웹, Android, 앱, php와 같은 모바일 웹과 연관된 개발 직무 내용이 도출됨을 확인할 수 있다.

[그림 2-135] 직무 15의 핵심단어 비율 그래프



[그림 2-136] 직무 15의 핵심단어 방사형 그래프



직무 15 군집의 핵심단어들의 동시등장 빈도율을 100점으로 환산하여 방사형 그래프를 도출한 결과는 [그림 2-136]에서 확인할 수 있다. 시각화 자료에서 보이는 것과 같이 주로 개발, 웹, Android와 앱, ios 관련 핵심단어들이 전체적으로 제시되고 있음을 파악할 수 있다. 마찬가지로 직무 15 군집의 핵심단어들을 바탕으로 워드클라우드를 시각화 한 결과 [그림 2-137]과 같은 결과를 획득하였다. 워드클라우드 결과 개발, Android, ios, 모바일, php, Java, 앱 등의 주요 핵심단어들이 직무의 핵심단어를 주도함을 확인함에 따라 직무 15는 모바일에 특화된 개발자 직무임을 다시 한 번 확인할 수 있다.

[그림 2-137] 직무 15의 핵심단어 워드클라우드



그에 따라 직무 15 군집은 모바일 앱이나 웹에 관한 기술적인 지침과 기획서를 바탕으로 Java, PHP 같은 프로그래밍 언어를 활용하여 코드를 작성하고 관리하며, Android와 ios를 위한 앱이나 모바일 웹페이지를 개발하고 운영하는 역할을 수행하는 직무로서 정의할 수 있다.

직무 15 군집에 대한 직무명과 직무 내용을 명시하기 위하여 본 보고서는 국제직무분류표준을 활용하였다. 그러나 <표 2-66>과 같이 기존 국제 분류 체계에서는 ISCO, SOC, O*Net 모두 응용 소프트웨어 프로그래머로만 구분되어 있으며 모바일 환경을 포함하는 분류는 나타나지 않는다. 주요 정의 내용은 일반적인

애플리케이션 SW 또는 특수 프로그램에 대한 개발, 작성 및 수정 직무로 정의되고 있으며 SW 애플리케이션 및 운영 시스템에 대한 기술적 지침 및 특수사항에 대한 프로그램 코드 작성 및 유지보수 직무로서 분류되어 있다.

모바일이 별도의 분류로 구분되지 못한 이유는 모바일 웹의 활성화 시기가 장기간에 걸쳐 발현된 것이 아닌 단기간에 걸친 확장으로 발생되어 왔기 때문에 주요 부흥 시기는 2014년 이후부터 활성화 되고 있는 분야라는 점이다. ISCO, SOC 등의 국제직무분류체계에서는 기술 시점 또한 2010년 및 2016년으로 분류되고 있으나 대부분의 경우 단순 웹 프로그래머의 직무로서 소개 되고 있다. 그러나 현재 웹과 모바일 웹 환경에서는 직무상 요구하는 능력이 극명히 구분되기 때문에 본 보고서에서는 기존 애플리케이션 개발자 직무와는 구분되어야 하는 신규 직무로서 제시되었다.

〈표 2-66〉 모바일 개발자 직무 관련 국제표준

소스	분류코드	직무	정의
ISCO-08	2514	Applications programmers	SW 애플리케이션 및 운영 시스템에 대한 기술적 지침 및 특수사항에 대한 프로그램 코드 작성 및 유지보수
SOC	15-1132	Software Developers, Applications	- 일반적인 애플리케이션 SW 또는 특수 프로그램에 대한 개발, 작성 및 수정 - SW 솔루션에 대한 개발 및 사용자 수요 분석
O*Net	15-1132.00	Software Developers, Applications	- 운영상 효율성 최적화의 목적에서 사용자 중심의 맞춤화 SW 설계 - 애플리케이션 SW 영역의 데이터베이스 분석 및 설계 - 일반 프로그래머 감독 및 지원
본 보고서의 분류		모바일 개발자	- 현재로서는 웹과 모바일 웹의 구분으로 직무가 구분되어야 하는 신규 직무 - 특히 모바일 특성상 휴대폰에 모바일 게임이나 음악서비스, 소셜결제 시스템 등의 적용이 가능하도록 프로그램을 작성하거나 애플리케이션을 개발

이에 따른 모바일 개발자에 대한 주요 직무 사항을 요약하면 다음과 같다.

- 모바일 앱 기반의 게임, 음악, 소액결제 서비스 등 다방면의 분야에 적용 가능한 프로그램 또는 앱을 개발
- 웹프로그램(ASP, php 등)을 근간으로 모바일 기반의 프로그램을 개발
- 모바일 웹기획을 통해 모바일 웹 DB 설계
- 응용프로그램(JAVA, VC) 및 모바일 프로그램을 이용한 개발
- 모바일 기반의 서버 프로그램에 대한 유지·보수 관리 수행
- 휴대폰의 기능 검토 및 활용 가능한 콘텐츠 개발 방안 연구

2.3) 직종 군집 별 정의 및 특성 개발

2.3.1) 웹 기획 관리자 (Web Administrators and Planner)

직종 1 군집의 경우 직무 1과 직무 2의 하위 군집을 포함하는 직종 군집으로 분류된다. 직무 1과 2의 분류 결과에 따라 도출된 핵심단어는 주로 웹사이트의 게시판 또는 사이트의 운영 속도 등과 같은 전반적인 기술적 성능을 관리, 기획, 유지 보수하는 단어들이 출현한다.

〈표 2-67〉 웹기획 관리자 직종 관련 국제표준

소스	분류코드	직종	정의
ISCO	251	Software and Applications Developers and Analysts	• SW 개발자 및 분석가 직종으로 정보시스템에 대한 조사, 기획, 설계, 작성의 업무를 가지고 있으며 하드웨어 및 SW뿐만 아니라 기타 영역의 관련 직무를 모두 포함함 • 업무 영역에서 활용 가능한 정보기술 조사 및 효율성 측정 • 컴퓨터 사용에 대한 이론적, 실무적 방법에 대한 조사 및 적용 직종 • 특정 환경적 요구 사항에 대한 컴퓨터 프로그래밍의 설계, 기획 • 인터넷, 인트라넷, 멀티미디어 시스템에 대한 평가, 기획, 설계 업무에 대한 직종으로 시스템 분석, 시스템 개발, 웹과 멀티미디어 개발자, 응용 프로그래머, SW 및 응용 SW 개발자, 분석가 등 관련 직무를 모두 포함함
SOC-201 634)	15-1250	Software and Web Developers, Programmers, and Testers	• 웹사이트를 설계 및 개발 • 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리 • 웹사이트를 위한 콘텐츠 기획
O*Net	15	Information Support and Services	• 웹 환경에 맞는 사이트 디자인, 개발 및 유지보수 관리 • 웹사이트 및 웹 콘텐츠의 품질 보증 관리 및 테스트 수행

그에 따라 국제직무분류표준인 ISCO, SOC로부터 관련 업무와 관련된 직종의 형태를 비교한 결과 <표 2-67>과 같이 소프트웨어 및 웹개발자 직종 영역에 대해 제시되어 있다. 기존 직종의 분류에서는 주로 소프트웨어와 웹개발자, 프로그래머, 테스터를 동일 직종으로 묶어 설명하고 있는 직종으로서 정보시스템에 대한 조사, 기획, 설계, 작성의 업무를 가지고 있으며 HW 및 SW 뿐만 아니라 기타 영역의 관련 직무를 모두 포함하여 웹사이트를 설계 및 개발하는 역할을 한다. 또한 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리하고 웹사이트를 위한 콘텐츠 기획하는 직종을 총괄하는 분류이다. 본 연구과제는 데이터 분석 결과에서 제시된 2개 직무 군집의 핵심단어들의 조합을 바탕으로 ‘웹기획 관리자’로 직종을 구분하였다.

특히 웹기획 관리자 직종은 온라인 쇼핑몰, 오픈 마켓, 블로그, 웹 페이지 등의 온라인 채널을 운영하기 위해서 필요한 콘텐츠, 디자인, 서비스를 기획하고 운영을 관리하는 역할을 담당한다. 웹마스터, IT 관리자, 웹사이트 관리자 등의 업무 특성이 해당 직업에 속하며 세부 특성에 따라 웹 직무 1과 직무 2의 두 가지 하위 직무 군집으로 구분되었으며 각 직무 군집의 정의 및 특성은 아래에 각 직무군집별로 자세히 설명하였다.

34) SOC-2016에서는 Web administrator의 경우 Web developer와 Digital Interface Designer가 구분되어 있기 때문에 별도의 직종으로 정하지 않고 Software and Web Developers, Programmers, and Testers로 단일하게 합친 후 세무 직무만 구분할 계획이다.

2.3.2) 디지털 콘텐츠 디자이너 (Digital Content Designers)

직종 2는 LDA 분석의 직무군집 분류 결과에 따라 제시된 핵심단어에 따라 직무 3과 직무 4를 포괄하는 상위 직종 군집으로 분류된다. 직무 군집 3과 4의 핵심단어 구성을 살펴보면 주로 기업의 웹 쇼핑몰 및 홈페이지 등의 온라인 채널에서 필요로 하는 그래픽과 콘텐츠 디자인을 기획하고 제작하는 역할을 수행하는 용어들로 정리되어 있다.

그에 따라 본 연구는 <표 2-68>과 같이 국제직무분류표준을 참고하여 유사 정의에 해당하는 직종 분류를 적용하여 직종 2의 명칭과 특성을 정리하였다. 직종군집 2와 관련된 기존 분류의 경우 멀티미디어 전문가로서의 정의를 제시하고 있으며 역시 특수효과, 온라인 콘텐츠의 형태 및 내용을 개발한다. 그에 따라 본 보고서는 직종군집 2를 ‘디지털 콘텐츠 디자이너’ 직종으로 정의하고 다양한 웹에서의 온라인 광고, 컴퓨터 그래픽, 제품 디자인 등과 같은 요소들을 디자인하는 역할을 맡는 직종으로 구분하였다.

<표 2-68> 디지털 콘텐츠 디자이너 직종 관련 국제표준

소스	분류코드	직종	정의
ISCO-08	216	Architects, Planners, Surveyors and Designers	- 정보 의사소통을 위한 시각 및 청각적 콘텐츠 및 인테리어, 익스테리어 등을 계획 및 디자인함 - 스케치, 다이어그램, 삽화, 애니메이션, 차트 등을 디자인하는 분야의 영역에서 직무를 수행함
SOC-2016	27-0000	Arts, Design, Entertainment, Sports, and Media	컴퓨터 게임, 영화, 뮤직비디오, 전자상거래 환경에서 제품과 같은 제품 또는 창조물에 사용하기 위하여 동영상 및 전자적 툴을 사용한 특수효과, 애니메이션 또는 기타 시각적 이미지 디자인 및 개발
O*Net	27	Audio and Video Technology and Film	시각 이미지 디자인 및 디자인, 저널리즘, 엔터테인먼트와 관계된 모든 영역에서 멀티미디어 콘텐츠를 디자인, 생산, 전시, 운영, 작성, 출판하는 영역의 직종을 의미함

2.3.3) 컴퓨터 네트워크 관리자 (Computer Network Administrators)

직종군집 3은 단일한 하위군집인 직무 5에 대한 직종으로 분류된다. LDA 분석에 의해 출현한 직무 5의 핵심단어 상에서는 주로 컴퓨터 네트워크 구조와 구현을 위한 연구와 분석을 수행하고 컴퓨터 네트워크 구조 디자인을 운영하며 구현된 네트워크를 테스트하는 역할과 관련된 핵심단어들이 나타나고 있음이 위의 직무 5의 설명에서 기술된 바 있다.

그에 따라 직종 3에 대한 직종명과 특성을 구성하기 위하여 기존 국제직무분류표준 자료들의 정의와 직무 5에 대한 핵심단어들을 조합할 때 <표 2-69>와 같이 직종 3의 경우 네트워크 하드웨어를 구축하고 관리하며, 정보보안의 역할을 수행하는 역할을 하는 것으로 판단되었으며 본 보고서에서 직종 3은 컴퓨터 네트워크 관리자로서 명칭을 정하였다.

<표 2-69> 컴퓨터 네트워크 관리자 직종 관련 국제표준

소스	분류코드	직종	정의
ISCO-08	252	Database and Network Professionals	- 정보시스템에 대한 최적의 성능 및 보안을 위한 전문적 설계, 개발, 운영, 유지, 지원을 위한 직종 - 데이터베이스 설계자 및 관리자, 시스템 관리자, 컴퓨터 네트워크 전문가, 기타 데이터베이스 및 네트워크 전문가 직무를 포함함
SOC-2016	15-1240	Database and Network Administrators and Architects	- 구성한 조직의 지역 네트워크(LAN), 광역 통신망(WAN), 인터넷 시스템 또는 네트워크 시스템의 세그먼트 지원 설치 - 모니터 네트워크와 필요한 정비 네트워크 가용성을 지원하기 위해 규정을 수행할 수도 모든 시스템 사용자에게 네트워크 가용성을 보장
O*Net	27	Information Support and Services	- 웹사이트의 오류 없는 정확한 운영을 보장하기 위한 웹사이트 성능 모니터링 및 테스트 - 컴퓨터 네트워크 아키텍처, 데이터베이스 관리자, 데이터베이스 통합 아키텍처, 네트워크와 컴퓨터 시스템 관리자 등이 하위 직무로 구분됨

그에 따른 컴퓨터 네트워크 관리자의 정의는 네트워크 관련 HW 및 SW 자원을 효과적으로 관리하고 네트워크에서의 보안정책을 수립하고 외부로부터 해킹이 시도될 수 없도록 체계적 운영을 하는 직종으로 구분하였다.

2.3.4) 시스템 소프트웨어 개발자 (Software and Applications Developers)

직종군집 4는 직무 6, 7 그리고 8의 군집을 포함하는 상위 직종 군집으로 분류되었다. LDA 분석에 의한 직무 6, 7, 8의 핵심단어들의 도출 형태를 보면 개발, 웹 서버, 프로그램, 분석, 관리 등의 단어 들이 3가지 하위 직무군집으로 구분되고 있다. 도출되는 핵심단어들을 바탕으로 <표 2-70>과 같이 국제직무분류표준의 직종 정보를 비교하였다. 그 결과 관련 직종의 경우 주로 소프트웨어 개발자로 명시하고 있으나 주로 시스템 및 소프트웨어 개발자와 시스템 개발을 위한 분석 직무를 포함하고 있다.

그에 따라 본 보고서는 국제직무분류표준의 정의를 기반으로 직종군집 4를 ‘시스템 소프트웨어 개발자’로 정하고 새로운 서버용 시스템, 웹 페이지, SW 나 기존의 프로그램을 분석하고 평가하여 사용자의 요구사항을 파악하고, 사용자의 요구사항을 충족시키기 위한 시스템 SW, 웹 페이지, SW 및 응용프로그램을 기획하거나 개발하고, 테스트 및 운영하는 역할을 수행하는 직종으로 구분하였다.

그에 따라 본 직종의 경우 기존 분류 체계에서는 매우 광범위한 직무 내용을 포함하고 있기 때문에 위에서 설명한 소프트웨어 개발자(직무 6), 시스템 개발자(직무 7), 소프트웨어 분석가(직무 8)의 3가지 군집을 포함하는 직종으로 구분되었다. 특히 본 직종의 경우 컴퓨터 하드웨어 및 시스템의 동작, 제어 및 관리에 연관성이 있는 SW들을 개발함에 따라 각종 컴퓨터시스템의 OS 체제인 UNIX, LINUX, WINDOW NT등과 같은 운영체제시스템을 설계하고 개발한다. 또한 주로 활용하는 컴퓨터 프로그램 언어인 FORTRAN, C/C++, JAVA와 같은 프로그래밍 언어 및 컴파일러를 설계하며 응용SW를 개발하기 위한 개발용 툴을 설계하고 개발하는 직종으로 분류된다.

그 외에도 네트워크에서 운영되는 프로토콜 처리에 관련된 SW (TCP/IP, NETWARE 등) 설계 및 보안에 활용되는 보안 관련 SW (방화벽, 인증 프로그램 등)를 설계 및 개발한다. 또한 시스템 영역에서는 컴퓨터시스템의 각종 주변 장치의 제어 및 인터페이스에 관련된 SW와 컴퓨터 시스템에서 운용되는 각종 유틸리티 SW를 설계하고 개발하며 이와 관련된 시스템은 주로 파일 관리, 데이터베이스 시스템, 네트워크 시스템 등이 있다.

〈표 2-70〉 시스템 SW 개발자 직종 관련 국제직무분류표준

소스	분류코드	직종	정의
ISCO	251	Software and Applications Developers and Analysts	• SW 개발자 및 분석가 직종으로 정보시스템에 대한 조사, 기획, 설계, 작성의 업무를 가지고 있으며 하드웨어 및 SW뿐만 아니라 기타 영역의 관련 직무를 모두 포함함 • 업무 영역에서 활용 가능한 정보기술 조사 및 효율성 측정 • 컴퓨터 사용에 대한 이론적, 실무적 방법에 대한 조사 및 적용 직종 • 특정 환경적 요구 사항에 대한 컴퓨터 프로그래밍의 설계, 기획 • 인터넷, 인트라넷, 멀티미디어 시스템에 대한 평가, 기획, 설계 업무에 대한 직종으로 시스템 분석, 시스템 개발, 웹과 멀티미디어 개발자, 응용 프로그래머, SW 및 응용 SW 개발자, 분석가 등 관련 직무를 모두 포함함
SOC-2 016	15-1250	Software and Web Developers, Programmers, and Testers	• 웹사이트를 설계 및 개발 • 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리 • 웹사이트를 위한 콘텐츠 기획
O*Net	15	Software Developers, Systems Software	• 의료 산업, 군사, 통신, 항공, 사업, 과학, 그리고 일반적인 컴퓨팅 활동을 위한 운영시스템 수준의 SW, 컴파일러, 네트워크 분배 SW를 연구, 설계, 개발 및 테스트함

2.3.5) 시스템 설계분석가 (System Architects / Analysts)

직종군집 5는 직무 9에 대한 상위 군집으로 분류된 직종군집이다. LDA 분석결과에 의해 도출된 직무 9의 핵심단어들을 확인해보면 컴퓨터 시스템의 구성요소 중 하나인 하드웨어를 기계적 혹은 전자적으로 분석하여 적합한 하드웨어 디자인을 설계하고 개발하는 역할을 수행하는 컴퓨터 시스템 설계자와 그 외의 각종 시스템 설계 및 분석을 담당하는 인력이 포함되는 단어들이 제시되고 있다.

특히 하드웨어나 설계 등의 용어가 제시됨에 따라 본 연구는 국제직무분류표준의 정의를 바탕으로 유사 직종들을 탐색 및 비교하였으며 그 결과는 <표 2-71>과 같다. 비교 결과에서 보이는 것과 같이 해당 직종은 소프트웨어 개발자, 분석가, 또는 IT 프로젝트 관리자 직종에 해당하는 직종에 속하는 것을 알 수 있다. 유사한 직종으로는 SW 개발자 및 분석가 직종으로 볼 수 있으며 정보시스템에 대한 조사, 기획, 설계, 작성의 업무를 가지고 있다.

기존 정의에 따르면 주로 HW 및 SW 뿐만 아니라 기타 영역의 관련 직무로 정의되고 있으나 특히 개발자 직무 능력에 집중되는 것과 다르게 IT기술 관련 프로젝트의 계획, 개시, 관리와 기업 경영 및 기술적 문제에 대한 통합 시스템 관리를 제공하는 직무로서 유사함을 확인하였다. 따라서 본 보고서는 직종 5 군집을 ‘시스템 설계분석가’ 직종으로 분류하였다.

〈표 2-71〉 시스템 설계분석가 직종 관련 국제직무분류표준

소스	분류코드	직종	정의
ISCO-08	251	Software and Applications Developers and Analysts	• SW 개발자 및 분석가 직종으로 정보시스템에 대한 조사, 기획, 설계, 작성의 업무를 가지고 있으며 하드웨어 및 SW뿐만 아니라 기타 영역의 관련 직무를 모두 포함함 • 업무 영역에서 활용 가능한 정보기술 조사 및 효율성 측정 • 컴퓨터 사용에 대한 이론적, 실무적 방법에 대한 조사 및 적용 직종 • 특정 환경적 요구 사항에 대한 컴퓨터 프로그래밍의 설계, 기획 • 시스템 분석, 시스템 개발, 웹과 멀티미디어 개발자, 응용 프로그래머, SW 및 응용 SW 개발자, 분석가 등 관련 직무를 모두 포함함
SOC-2016	15-1210	Computer and Information Analysts	• 복잡한 응용문제, 시스템 관리 이슈 및 네트워크 문제에 대한 솔루션 개발 및 개선을 위한 분석 • 경영상 데이터 프로세스 이슈 분석 • 현존하는 컴퓨터 시스템에 대한 개선 • 시스템의 기능적 통합 및 관리 수행 • 시스템 용량 및 제반사항 검토 • 유용한 SW 분석 및 추천 • 하위 직무로는 컴퓨터 시스템 분석가, 정보 보안 분석가가 있음
O*Net	15-1199.09	Information Technology Project Managers	• IT기술 관련 프로젝트의 계획, 개시, 관리 • 기술 인력의 업무에 대한 지침 작성 • 기업 경영 및 기술적 문제에 대한 통합 관리 제공 • 프로젝트 단계 기획 및 각 단계별 경영상 시사점 도출 • 일정관리 모니터링 및 비용 관리

2.3.6) 콘텐츠 크리에이터 (Content Creators)

직종 6 군집은 직무 10과 11 군집을 포함하는 상위 군집으로 분류되고 있다. 직무 10과 11 군집은 본 보고서의 분류 상 미디어 콘텐츠 창작자와 소셜 미디어 관리자에 해당하는 직무이다. LDA 분석 결과 도출된 직무 10과 11의 주요 구성 핵심단어들이 웹, 편집, 기획, 방송 또는 IT, 바이럴, 소셜, 마케팅 등의 단어들이 제시되고 있어 직접적인 분류를 확인하고자 하였다.

이를 위하여 기존 국제직무분류표준에서 유사 직종을 구분하고자 하였으나 유사한 분류들을 직종 단위에서 확인할 수 없다는 결론에 따라 해외 관련 단어들 기반의 직무 공고 및 학술연구들을 확인하였다. 탐색 결과 콘텐츠 크리에이터 분야가 확인되었고 해당 직종은 최근 1~2년 사이에 그 수요가 폭발적으로 증가하고 있는 직종 중 하나이다. 주요 업무로는 웹 페이지나 방송, 모바일 페이지에서 필요로 하는 콘텐츠를 기획하고 제작하는 프로젝트의 프로젝트 매니저 역할을 수행하는 직종으로 구분되고 있다.

추가적으로 학술적 연구를 탐색한 결과 Albuquerque et al. (2012)에 따르면 <표 2-72>와 같이 콘텐츠 크리에이터 (Content Creators)는 구전 마케팅 및 마케팅 캠페인의 형태로서 플랫폼 콘텐츠에 대한 마케팅 콘텐츠 들을 개발하는 전문가를 의미한다. 특히 콘텐츠 크리에이터는 사용자가 직접 생성하는 콘텐츠 웹사이트의 개발 및 마케팅 자원 할당 계획 수립 시 관리자로서의 역할을 겸하는 것으로 제시되었다. 또는 온라인 채널이나 콘텐츠에 대해 구전효과를 일으키는 프로젝트를 수행하는 프로젝트 매니저의 역할을 한다. 그에 따라 필요한 콘텐츠를 직접 기획하고 제작하는 과정을 관리하며 소셜 마케팅을 위한 콘텐츠를 기획하고 제작하는 과정을 관리하는 역할을 한다. 그에 따라 본 보고서에서는 직무 6을 ‘콘텐츠 크리에이터’ 직종으로 구분하였다.

<표 2-72> 콘텐츠 크리에이터의 역할

출처	정의와 역할
Albuquerque et al. (2012)	- 마케팅 캠페인과 조회 수 향상의 목적을 위하여 플랫폼 콘텐츠에 대한 자유로운 마케팅 요소 개발 - 대다수 사용자가 만들어낸 콘텐츠에 대한 개발에 영향력을 미치며 마케팅 자원으로 활용하기 위한 콘텐츠 할당 및 관리

2.3.7) 데이터베이스 관리자 (Database Administrators)

직종 7 군집의 경우 LDA 분석에 의한 분류결과 직무 12를 포함하는 단일 직종 군집으로 형성되었다. 따라서 LDA 분석에 의해 도출된 직무 12의 핵심단어 구성을 조합을 확인하여 유사 직무분류를 확인하고자 하였다. 특히 직무 7은 개발, 웹, 데이터베이스, 디자인 등의 단어들이 제시됨에 따라 데이터베이스 관련 분야임을 확인하고 관련 국제직무분류표준의 데이터베이스 직종에 대한 정의를 참고하고자 하였다. 그에 따라 본 보고서는 <표 2-73>과 같이 직종 7에 대한 유사 직종의 정의들을 확인하고 그 정의의 큰 차이가 존재하지 않아 직종 7을 데이터베이스 관리자 직종으로 구분하였다.

<표 2-73> 데이터베이스 관리자 직종 관련 국제직무분류표준

소스	분류코드	직종	정의
ISCO-08	252	Database and Network Professionals	- 정보시스템에 대한 최적의 성능 및 보안을 위한 전문적 설계, 개발, 운영, 유지, 지원을 위한 직종 - 데이터베이스 설계자 및 관리자, 시스템 관리자, 컴퓨터 네트워크 전문가, 기타 데이터베이스 및 네트워크 전문가 직무를 포함함
SOC-2016	15-1240	Database and Network Administrators and Architects	- 구성한 조직의 지역 네트워크(LAN), 광역 통신망(WAN), 인터넷 시스템 또는 네트워크 시스템의 세그먼트 지원 설치 - 모니터 네트워크와 필요한 정비 네트워크 가용성을 지원하기 위해 규정을 수행할 수도 모든 시스템 사용자에게 네트워크 가용성을 보장 - 웹사이트의 오류 없는 정확한 운영을 보장하기 위한 웹사이트 성능 모니터링 및 테스트 - 컴퓨터 네트워크 아키텍처, 데이터베이스 관리자, 데이터베이스 통합 아키텍처, 네트워크와 컴퓨터 시스템 관리자 등이 하위 직무로 구분됨
O*Net	15	Information Support and Services	

2.3.8) HCI 컨설턴트

직종군집 8은 LDA 분류결과 직무 13을 포함하는 단일한 군집으로 도출되었다. 직무 13에 출현하는 핵심단어들을 확인한 결과 주로 마케팅, 디자인 등과 함께 UX, UI, CS 등의 단어들이 도출됨에 따라 본 보고서는 직종 8의 경우 UX(User Experience, 사용자경험) 또는 UI(User Interface, 사용자인터페이스)와 관련된 직종임을 확인하였다.

이와 관련하여 유사 직종에 대한 조사를 한 결과 <표 2-61>과 같이 기존 국제 직무분류표준에서의 직종을 확인할 수 있었다. ISCO의 경우 분류자체가 밀접한 직종이 없으며 SOC의 경우 소프트웨어 및 웹개발자, 프로그래머 및 시험자 직종에 일부로서 포함된 점을 확인하였다. 또한 O*Net의 경우 오디오 및 비디오 기술과 영상 직종에 해당 내용들이 포함된 것을 발견할 수 있었다. 그러나 해당 직종들은 분류를 통해 나타난 핵심단어들보다는 광의의 개념으로 판단됨에 따라 관련 개념들을 연구 내용에서 보완하고자 하였다.

그 결과 <표 2-74>과 같이 기존의 애니메이션 및 그래픽 디자이너와는 다른 직종으로 UX 및 UI와 같은 인간-컴퓨터 상호작용(Human-Computer Interaction, 이하 HCI)로 구분되어 기존 분류를 참고하여 Hedberg & livari (2009)의 정의에 따라 본 보고서에서는 인간과 컴퓨터의 상호작용을 토대로 사용자 친화적이고 접근성이 확보된(HIC : Human Computer Interaction) 컴퓨터시스템 및 웹사이트 등을 개발하고 컨설팅 하는 직종으로 구분하여 ‘HCI 컨설턴트’로서 직종 명을 구분하였다. 특히 HCI컨설턴트는 현재 구인 사이트인 Linked-In에서도 매우 쉽게 찾아볼 수 있는 직종명이며 UX 및 UI분야의 직무를 책임지는 전문 직종으로서 구분되어 있음을 확인하였다.

〈표 2-74〉 HCI 컨설턴트 직종 관련 국제표준

소스	분류코드	직종	정의
SOC-201 6	15-1250	Software and Web Developers, Programmers, and Testers	• 웹사이트를 설계 및 개발 • 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리 • 웹사이트를 위한 콘텐츠 기획
O*Net	15	Audio and Video Technology and Film	• 시각 이미지 디자인 및 디자인, 저널리즘, 엔터테인먼트와 관계된 모든 영역에서 멀티미디어 콘텐츠를 디자인, 생산, 전시, 운영, 작성, 출판하는 영역의 직종을 의미함
Hedberg & livari (2009)		HCI Consultant	• 사용성, 사용자 경험, 사용자 인터페이스 등을 기반으로 상호작용 설계를 위한 전문가
본 보고서의 분류 ³⁵⁾		HCI 컨설턴트	• 개발자들과 함께 웹 디자인 설계 및 디자인 의사결정에 참여하며 사용자에게 적합한 컨설팅을 제시하는 집단 • 인간과 컴퓨터의 상호작용을 토대로 사용자 친화적이고 접근성이 확보된 (HIC : Human Computer Interaction) 컴퓨터시스템 및 웹사이트 등을 개발하고 컨설팅 • 이미 해외에서는 HCI Consultant 또는 HCI Specialist로 공고가 많이 게시되고 있으며 UX, UI를 책임지는 전문가로서의 역할을 하는 직무

35) Hedberg & livari (2009)와 livari (2006)의 정의를 참고하였음

2.3.9) 응용소프트웨어 개발자 (Applications Programmers)

직종 군집9의 경우 직무 14와 15를 하위 군집으로 두고 있는 직종 군집이다. LDA 모델링을 통해 직무 14와 직무 15의 핵심단어들을 확인한 결과 주로 개발, jsp, ios, php 등의 핵심단어들이 제시되는 직종으로 응용소프트웨어와 관련된 직종으로 분류할 수 있다.

그에 따라 직종 9에 대하여 유사 의미를 포함하고 있는 직종 정의와 비교하기 위하여 국제직무분류표준을 비교한 결과 <표 2-75>과 같이 기능, 성능요건 및 운영 환경(OS, 네트워크, 사용자 인터페이스 등)에 대한 종합적인 검토 분석 하에 응용SW의 개발 범위와 목표를 설정하는 직무로서 볼 수 있다.

ISCO의 경우 소프트웨어 및 애플리케이션 개발자 및 분석가 직종으로 구분되며 SOC와 O*Net은 각각 소프트웨어 및 웹개발자 또는 인터랙티브 미디어 직종으로 구분되고 있다. 주요 직무는 ISCO의 경우 인터넷, 인트라넷, 멀티미디어 시스템에 대한 평가, 기획, 설계 업무에 대한 직종으로 시스템 분석, 시스템 개발, 웹과 멀티미디어 개발자, 응용 프로그래머, SW 및 응용 SW 개발자, 분석가 등 관련 직무를 모두 포함하는 광의의 개념을 제시하고 있다. 또한 SOC와 O*Net의 경우 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리하고 웹사이트를 위한 응용 소프트웨어 및 콘텐츠 기획하는 직종으로 제시되었다.

기존 직종에 대한 정의와 본 보고서의 분석 결과에 따라 군집 13은 웹 또는 모바일 웹에서의 응용소프트웨어 분석을 수행하는 개발자 직종으로서 구분하여 직무 13을 ‘응용소프트웨어 개발자’ 직종으로 명명하고 각종 응용분야의 컴퓨터 SW를 전문적으로 설계하고 개발하는 일을 담당하는 직종으로서 정의하였다. 따라서 응용소프트웨어 개발자 직종은 각종 응용분야의 컴퓨터 SW를 전문적으로 설계하고 개발하는데 종사하는 직종으로 애플리케이션 개발자(직무14)와 모바일 개발자(직무15)의 상위 직종으로서 정리되었다.

〈표 2-75〉 응용소프트웨어 개발자 직종 관련 국제직무분류표준

소스	분류코드	직종	정의
ISCO	251	Software and Applications Developers and Analysts	• SW 개발자 및 분석가 직종으로 정보시스템에 대한 조사, 기획, 설계, 작성의 업무를 가지고 있으며 하드웨어 및 SW뿐만 아니라 기타 영역의 관련 직무를 모두 포함함 • 업무 영역에서 활용 가능한 정보기술 조사 및 효율성 측정 • 컴퓨터 사용에 대한 이론적, 실무적 방법에 대한 조사 및 적용 직종 • 특정 환경적 요구 사항에 대한 컴퓨터 프로그래밍의 설계, 기획 • 인터넷, 인트라넷, 멀티미디어 시스템에 대한 평가, 기획, 설계 업무에 대한 직종으로 시스템 분석, 시스템 개발, 웹과 멀티미디어 개발자, 응용 프로그래머, SW 및 응용 SW 개발자, 분석가 등 관련 직무를 모두 포함함
SOC-2016	15-1250	Software and Web Developers, Programmers, and Testers	• 웹사이트를 설계 및 개발 • 웹사이트 구성과 사이트의 기능적 구성에 책임이 있으며 사이트의 성능 및 속도 측정 등을 관리
O*Net	15	Interactive Media	• 웹사이트를 위한 응용 소프트웨어 및 콘텐츠 기획

특히 응용소프트웨어 개발자는 요구사항에 대한 조사 및 분석을 바탕으로 응용 SW 개발을 수행한다. 그에 따라 기능, 성능요건 및 운영 환경(OS, 네트워크, 사용자 인터페이스 등)에 대한 종합적인 검토 분석 하에 응용SW의 개발 범위와 목표를 설정한다. 예를 들어 모바일 또는 특정 OS환경에 따른 애플리케이션 및 모바일 웹 환경에 대한 충분한 이해를 바탕으로 SW를 개발하게 된다.

- SW를 개발·완성을 위한 전체적인 개발 계획과 자원 조달 편성
- 개발에 사용할 웹 및 SW 언어 (C, PHP 등) 및 환경 파악 (ios, Android 등), 프로그램 언어 (JAVA, JSP 등) 설정
- 응용시스템에 대한 정보보호의 방법과 계획을 설정한다.
- 상세설계에 입각하여 단위 프로그램 개발 및 테스트를 통한 버그 수정
- 개발된 여러 프로그램들을 모아 응용시스템으로 결합 및 다각적인 시험 과정을 통해서 기능 및 성능을 평가·분석
- 패키지 형태의 개발 SW에 대해 버전관리 유지

이 결과를 바탕으로 본 직종 군집별 분류가 정확한 의미를 통해 구분되었는지 확인하기 위하여 직종 분류에 대한 신뢰성 조사를 수행하였다. 신뢰성 조사는 SW분야 전문가 5인(교수 3인, 산업계 전문가 2인)으로 구성된 대상을 바탕으로 상호코더신뢰성 분석을 수행하였다. 상호코더신뢰성 분석을 바탕으로 각 직종 군집 명에 대해 각 코더들이 이름을 구성하도록 한 결과를 바탕으로 본 보고서에서 개발한 분류체계와 비교한 결과 상호코더간신뢰성 결과는 88.3%의 일치도를 보이는 것을 확인하였다. 따라서 본 보고서에서 제시하는 직종명과 각 직종별 특성은 유효성이 높다는 점을 확인하였다.

최종적으로 지금까지의 각 직무 및 직종 별 분류를 바탕으로 <표 2-76>과 같이 총 9개 직종 15개 직무를 구분하여 SW 산업 분야의 수요를 반영한 데이터 분석 기반의 직무분류체계를 개발하였다. 각 직무는 국제직무분류표준 등의 정의를 참고하여 정리되었으며 신규 개발된 직종인 시스템 설계분석가, 콘텐츠 크리에이터, HCI 컨설턴트를 개발하였다. 또한 미디어 콘텐츠 창작자, 소셜 미디어 관리자, 모바일 개발자, HCI 컨설턴트, 시스템 설계분석가와 같은 세부 직종들도 신규 도출함에 따라 SW산업 수요에 의거한 데이터 기반의 분류체계를 개발하였다.

〈표 2-76〉 데이터 분석 기반 SW 산업 직무·직종 분류체계

명칭			정의	관련 국제직업분류코드
직종 1	웹 기획 관리자 (Web Administrators and Planners)		웹 기획 관리자는 웹사이트의 게시판 또는 사이트의 운영 속도 등과 같은 전반적인 기술적 성능을 관리, 기획, 유지 보수하는 일을 주 업무로 삼고 있다.	ISCO 251, SOC 2016 15-1250, O*Net 15
	직무1	웹 관리자 (Web Administrator)	웹 관리자는 기업 웹 사이트의 게시판에 대한 관리 및 웹 사이트에서 발생하는 다양한 문제들을 해결하는 직무를 가진다.	ISCO-08 2513, SOC-2016 15-1254, O*Net 15-1199.03
	직무2	웹기획자 (Web Planner)	웹 기획자는 웹 쇼핑몰 및 오픈마켓 운영을 기획하고, 쇼핑몰 및 오픈마켓에서 필요로 하는 디자인 및 콘텐츠를 기획하는 역할을 수행한다.	SOC 11-3021, O*Net 11-3021.00
직종 2	디지털 콘텐츠 디자이너 (Digital Content Designers)		디지털 콘텐츠 디자이너 직종은 기업의 웹 쇼핑몰 및 홈페이지 등의 온라인 채널에서 필요로 하는 그래픽과 콘텐츠 디자인을 기획하고 제작하는 역할을 수행한다.	ISCO-08 216, SOC-2016 27-0000, O*Net 27
	직무3	웹 콘텐츠 디자이너 (Web Content Designer)	웹 콘텐츠 디자이너는 소프트웨어를 활용하여 웹에서 활용하는 그래픽, 오디오, 비디오 관련 요소들에 대한 웹 포맷을 디자인하는 일을 담당하는 직무로서 정의할 수 있다.	ISCO-08 2166, SOC-2016 15-1254, O*Net 15-1134.00
	직무4	그래픽 디자이너 (Graphic Designer)	그래픽 디자이너는 웹 페이지에서 필요로 하는 다양한 그래픽 디자인을 제작하는 역할을 수행한다.	ISCO-08 2166, SOC-2016 27-1024, O*Net 27-1024-00
직종 3	컴퓨터 네트워크 관리자 (Computer Network Administrator)		컴퓨터 네트워크 관리자는 네트워크 관련 HW 및 SW 자원을 효과적으로 관리하고 네트워크에서의 보안정책을 수립하고 외부로부터 해킹이 시도될 수 없도록 체계적 운영을 하는 역할을 수행한다.	ISCO-08 252, SOC-201 15-1240, O*Net 27
	직무5	컴퓨터 네트워크 관리자 (Computer Network Administrator)	컴퓨터 네트워크 관리자는 네트워크 관련 HW 및 SW 자원을 효과적으로 관리하고 네트워크에서의 보안정책을 수립하고 외부로부터 해킹이 시도될 수 없도록 체계적 운영을 하는 역할을 수행하는 인력을 의미한다.	ISCO-08 2523, SOC-2016 15-1244, O*Net 27-1024-00
직종 4	시스템 소프트웨어 개발자 (System and Software Developers)		시스템 소프트웨어 개발자는 새로운 서버용 시스템, 웹 페이지, 소프트웨어나 기존의 프로그램을 분석하고 평가하여 사용자의 요구사항을 파악하고, 사용자의 요구사항을 충족시키기 위한 시스템 소프트웨어, 웹 페이지, 소프트웨어 및 응용프로그램을 기획하거나 개발하고, 테스트 및 운영하는 역할을 수행한다.	ISCO 251, SOC-201 15-1250, O*Net 15
	직무6	소프트웨어 개발자 (Software Developer)	소프트웨어 개발자는 새로운 소프트웨어나 기존의 응용프로그램을 분석하고 평가하여 사용자의 요구사항을 파악하고, 사용자의 요구사항을 충족시키기 위한 소프트웨어를 기획하거나 개발하고, 테스트 및 운영하는 역할을 수행한다.	ISCO-08 2131 part 2512, SOC-2016 15-1252, O*Net 15-1133.00
	직무7	시스템 개발자 (System Developer)	시스템 개발자는 새로운 서버용 시스템이나 기존의 서버 프로그램을 분석하고 평가하여 사용자의 요구사항을 파악하고, 사용자의 요구사항을 충족시키기 위한 시스템 소프트웨어를 기획하거나 개발하고, 테스트 및 운영하는	ISCO-08 2512, SOC-2016 15-1133, O*Net 15-1199.02

			역할을 수행한다.	
	직무8	소프트웨어 분석가 (Software Analyst)	소프트웨어 분석가는 소프트웨어의 설계, 작성, 실험, 개선된 성능을 제공하기 위한 목적으로 소프트웨어 및 시스템을 설계, 계획, 평가하고 시스템에 적합한 소프트웨어를 계획, 설계, 평가 및 실험 등의 업무를 담당한다.	ISCO-08 2519, SOC-2016 15-1253, O*Net 15-1121.00
직종 5		시스템 설계분석가 (System Analyst)*³⁶⁾	시스템 설계분석가는 컴퓨터시스템을 설계 및 분석하고, 컴퓨터시스템의 용량, 작업 절차 및 일정을 검토하여 전체적인 컴퓨터시스템 설계를 시행하는 업무를 담당한다. 또한 구입해야 할 하드웨어 및 소프트웨어, 소요기간 등을 산정하고 설계안을 제출하여 설계를 기획한다.	ISCO-08-251, SOC-2016 15-1210, O*Net 15-1199.09
	직무9	시스템 설계분석가 (System Analyst)	시스템 설계분석가는 컴퓨터시스템을 설계 및 분석하고, 컴퓨터시스템의 용량, 작업 절차 및 일정을 검토하여 전체적인 컴퓨터시스템 설계를 시행하는 업무를 담당한다. 또한 구입해야 할 하드웨어 및 소프트웨어, 소요기간 등을 산정하고 설계안을 제출하여 설계를 기획한다.	ISCO-08 2511, SOC-2016 15-1211, O*Net 15-1199.09
직종 6		콘텐츠 크리에이터 (Content Creators)*	콘텐츠 크리에이터는 최근 1-2년 사이 수요가 증가하고 있는 직종 중 하나로서 주로 웹 페이지나 방송, 모바일 페이지에서 필요로 하는 콘텐츠를 기획하고 제작하는 프로젝트의 프로젝트 매니저 역할을 수행한다.	Albuquerque et al. (2012)
	직무10	미디어 콘텐츠 창작자 (Media Content Creator)*	미디어콘텐츠창작자는 광고기반 플랫폼에 개인의 영상 콘텐츠를 제작하여 업로드 하고 이를 통해 수익을 창출하는 일을 한다. 기획안 구성부터 연출, 촬영, 편집 등을 종합적으로 소화하기 때문에 PD, 배우, 작가의 역할을 총망라하여 담당한다.	SOC-2016 27-1014, O*net 27-1014.00
	직무11	소셜 미디어 관리자 (Social Media Manager)*	소셜 미디어를 기반으로 하는 기획 및 운영 등 소셜 미디어와 관련된 업무를 수행하는 관리자로 페이스북, 트위터, 유튜브, 블로그 등을 활용하여 마케팅 및 미디어 전략을 수행 한다.	SOC-2016 13-1161, O*Net 13-1199.06
직종 7		데이터베이스 관리자 (Database Administrator)	데이터베이스 관리자는 데이터베이스 성능 최적화와 보안유지를 위해 데이터베이스를 디자인하고 개발, 관리하는 역할을 수행한다.	ISCO-08 252, SOC-2016 15-1240, O*net 15
	직무12	데이터베이스 관리자 (Database Administrator)	데이터베이스 관리자는 데이터베이스 성능 최적화와 보안유지를 위해 데이터베이스를 디자인하고 개발, 관리하는 역할을 수행한다.	ISCO-08 2521, SOC-2016 15-1242, O*Net 15-1141.00
직종 8		HCI 컨설턴트 (HCI Consultant)*	HCI컨설턴트는 사용성, 사용자 경험, 사용자 인터페이스 등과 관련하여 사용자 친화적이고 접근성이 확보된 컴퓨터시스템 및 웹사이트 등을 개발하고 컨설팅 하는 역할을 수행한다.	SOC-2016 15-1250, O*Net 15, Hedberg & livari (2009)
	직무13	HCI 컨설턴트 (HCI Consultant)	HCI컨설턴트는 사용성, 사용자 경험, 사용자 인터페이스 등과 관련하여 사용자 친화적이고 접근성이 확보된 컴퓨터시스템 및 웹사이트 등을 개발하고 컨설팅 하는 역할을 수행한다.	SOC-2016 15-1255, O*Net 27-1014.00, Hedberg & livari (2009)
직종 9		응용 SW 개발자 (Applications Developers)	응용소프트웨어 개발자는 각종 응용분야의 컴퓨터 소프트웨어를 전문적으로 설계하고 개발하는 일을 담당한다.	ISCO 251, SOC-2016 15-1250, O*Net 15
	직무14	애플리케이션 개발자 (Software)	애플리케이션 개발자는 소프트웨어 응용프로그램에 대한 기술적인 지침과 기획을 바탕으로 Java, C++ 같은	ISCO-08 2514, SOC 15-1132,

		Applications Developer)	프로그래밍 언어를 활용하여 코드를 작성하고 관리하며 애플리케이션 및 응용프로그램을 개발하고 운영하는 역할을 수행한다.	O*Net 15-1132.00
	직무15	모바일 개발자 (Mobile Applications Developer)*	모바일 개발자는 모바일 앱이나 웹에 관한 기술적인 지침과 기획서를 바탕으로 ASP, Java, PHP 같은 프로그래밍 언어를 활용하여 코드를 작성하고 관리하며, Android와 iOS를 위한 앱이나 모바일 기반 프로그램을 개발하고 운영하는 역할을 수행한다.	ISCO-08 2514, SOC 15-1132, O*Net 15-1132.00

2.4) 개발된 직무 직종 분류체계의 특성

분석된 분류체계를 기반으로 전체 직종별 / 직무별 특성을 확인하기 위하여 추가적인 시각화 그래프 도출을 바탕으로 직종 및 직무 간의 비교 분류를 수행하였으며 그 결과는 다음과 같다. <표 2-77>과 같이 시각화 도출을 위한 데이터의 활용 기간은 2016년 9월 1일부터 2016년 12월 25일까지의 채용공고 등록건수를 기준으로 개발되어 있으며 총 41,432건의 채용공고 정보가 구성되어 있다.

<표 2-77> 주별 채용공고 등록 건수

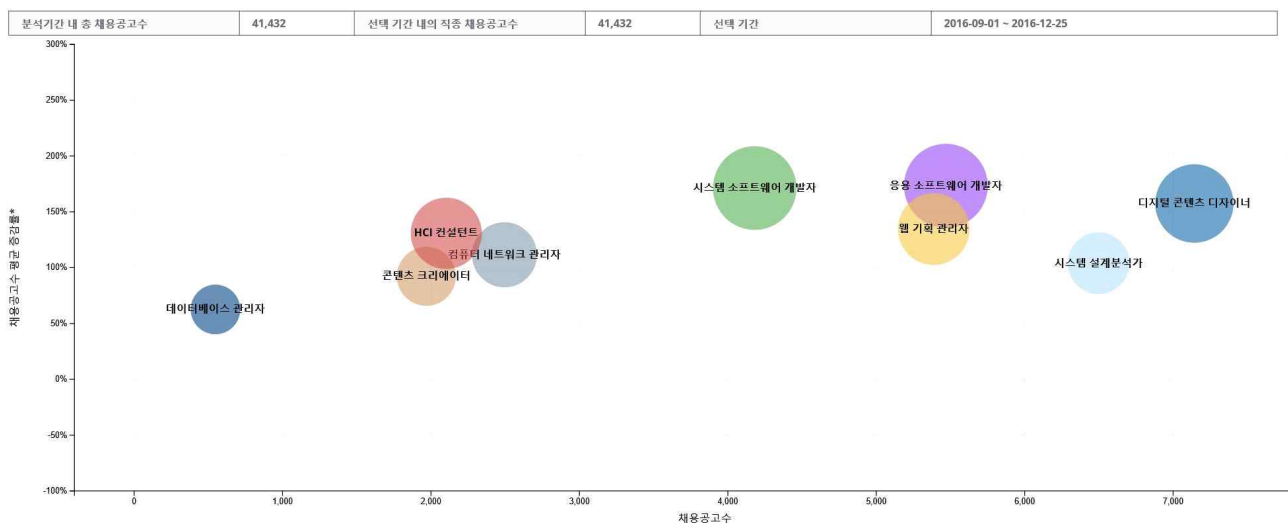
	첫째 주	둘째 주	셋째 주	넷째 주
9월	2	7	149	136
10월	176	279	681	665
11월	864	857	1,950	2,568
12월	3,853	6,666	11,394	11,185

[그림 2-138]과 같이 직종 군집들의 채용공고 개수(X축)와 채용공고 개수 증감률(Y축)을 2차원 그래프로 시각화 그래프를 도출하였다. 분석 데이터의 기간은 2016년 9월 1일부터 2016년 12월 25일까지이다. 채용공고 개수(X축)는 선택된 기간 동안의 각 군집의 채용공고 개수를 총합한 값이다. 채용공고수 평균 누적증감률(Y축)은 설정된 분석기간 내에 등록된 채용공고 개수의 증감률의 평균값을 의미하여 이는

36) 신규 개발된 직종 및 직무의 경우 *로 표기함

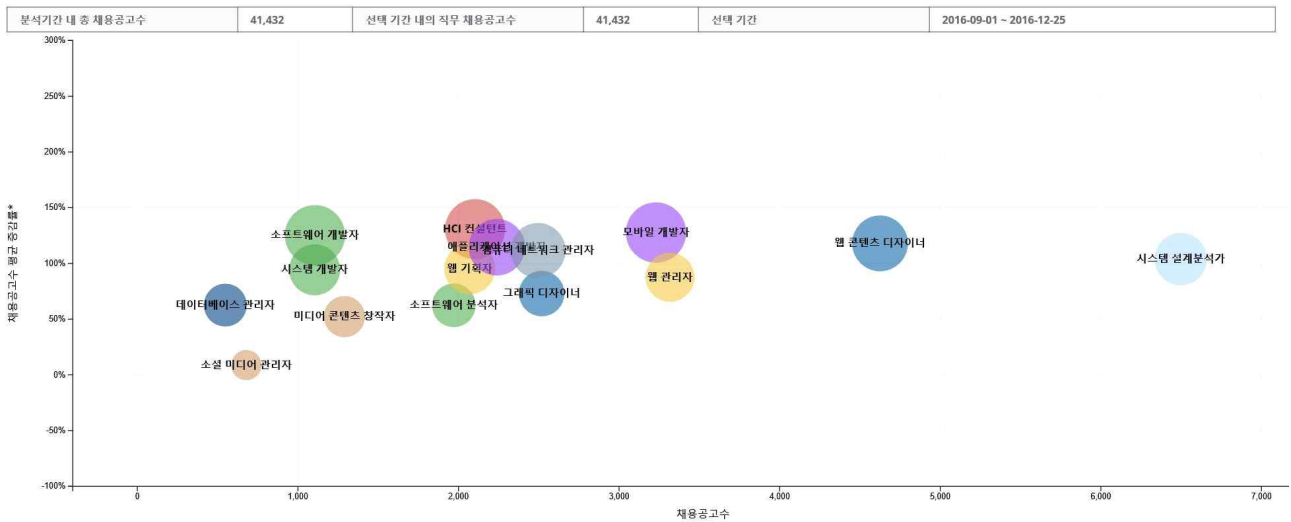
전체 설정기간의 전주 기간 대비 주별 채용공고 개수 차이의 총합을 주 평균 채용공고 개수로 나눈 것이다. 또한 각 직종 군집들은 원의 크기가 미세하게 다른 것을 확인할 수 있는데 이는 군집의 반지름을 통해 채용공고 개수의 증감 변동수치를 적용하였기 때문이다. 채용공고수 평균 누적변동률은 설정된 분석기간 내에 등록된 채용공고 개수의 주단위 증감률의 평균값 대비 증감 변동(Fluctuation)에 대한 크기 값으로 이 전주 대비 현재주 채용공고 개수의 증감총합(절대값)을 주 평균 채용 공고 개수로 나눈 값이다.

[그림 2-138] 직종 채용공고수 증감 그래프



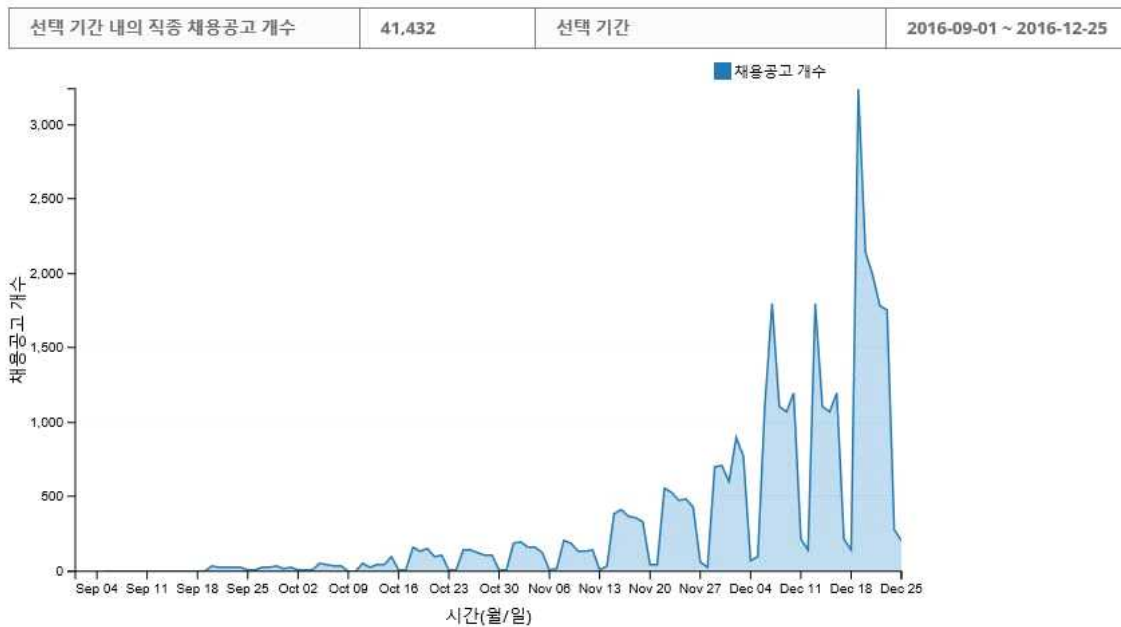
시각화 결과 총 41,432개 공고 개수 중 채용공고 증가율은 응용소프트웨어 개발자 직종과 디지털 콘텐츠 디자이너, 시스템 소프트웨어 개발자 직종이 높은 증감률을 보이고 있다. 그러나 데이터베이스 관리자와 콘텐츠 크리에이터 직종의 경우 큰 증감률을 보이지 않는 것으로 나타나 채용공고 기간 동안 직종별로 산업계 수요가 매우 다른 양상을 보이고 있다.

[그림 2-139] 직무 채용공고수 증감 그래프



마찬가지로 각 직무 군집별로 채용공고 개수(X축)와 채용공고수 평균 증감률(Y축)을 2차원 그래프로 시각화 그래프를 도출한 결과는 [그림 2-139]와 같다. 분석 데이터의 기간은 [그림 2-138]과 같이 2016년 9월 1일부터 2016년 12월 25일까지이며 각 축의 정의도 동일하다. 시각화 결과를 보면 직무별로 SW업계의 수요가 분석 데이터 기간 동안 차이가 나타나고 있다. 특히 소프트웨어 개발자 및 시스템 개발자, HCI 컨설턴트, 소프트웨어 개발자와 모바일 개발자 등의 직무들에 대한 채용공고의 증감률이 높게 나타나고 있으며 소셜미디어 관리자, 미디어 콘텐츠 창작자 등의 전문 SW 역량이 필요 없는 직무의 경우 증감률이 크지 않았다. 이는 SW 전문 역량을 보유한 인력에 대한 산업계 수요가 높음을 알 수 있다. 그러나 해당 기간 동안 모바일 개발자 등의 공고 증감률은 매우 크게 나타나 스마트폰의 보급에 따른 SW 수요가 전통적인 PC에서 스마트폰으로 변화하고 있는 것을 확인할 수 있다.

[그림 2-140] 직무·직종 공고 기반 채용공고 변화에 대한 시계열 그래프



[그림 2-140]은 시간에 따른 채용공고 개수의 변화를 보여준다. 시각화 결과는 2016년 9월부터 12월 25일의 기간 중 채용공고의 변화를 의미하며 시간이 지남에 따라 점차적으로 그래프가 점차 증가하는 모양을 확인할 수 있다. 각 채용공고의 일자별 분석을 통하여 특정 기간 내에 전체 채용공고의 개수 변화 추세를 확인할 수 있으며 시간이 지날수록 점차적으로 채용공고의 수가 늘어나고 있음을 알 수 있다.

2.5) SW직무직종 분류체계의 특성 및 향후 보완사항

빅데이터 기반의 직무분류체계 개발을 정리한 결과 <표 2-76>과 같이 총 9개 직종 15개 직무를 구분하였으며 분류를 기반으로 각 직종 및 직무의 정의를 새로 진행하였다. 특히 기존 국제 분류체계에서 제공하고 있는 분류 중에서 2가지 직종(콘텐츠 크리에이터, HCI 컨설턴트)과 4가지 직무(미디어 콘텐츠 창작자, 소셜미디어 관리자, HCI 컨설턴트, 모바일 개발자)의 경우 신규 개발하여 제시하였다.

분류된 직무분류체계는 SW관련 분야의 산업계 전문가와 SW관련 전공 교수진의 참여를 바탕으로 직무 분류에 대한 신뢰성을 판단하여 코더 간 상호 신뢰성 분석 결과 수용가능한 수준의 분류임을 확인하였다.

특히 본 분류체계의 경우 SW산업계의 채용공고 자료에 의거하여 분석된 분류체계라는 점에서 기존 워크넷과 NCS의 분류체계 간 미일치성으로 인한 불안정성을 보완하고 산업계 특성을 적극 반영한 직무·직종 분류가 가능하였다. 추가적으로 각 분류 내 핵심단어(동시등장 빈도율)를 확인함으로써 현업에서 요구되거나 중요도가 높은 역량 및 분야에 대한 정보를 도출할 수 있다는 점에서 그 시사점이 매우 크다고 볼 수 있다.

그럼에도 불구하고 본 분류체계는 채용공고 자료가 추가되고 시간이 흐를수록 직무·직종 분류가 변화할 수 있으며 그에 따라 현재 분류된 직무·직종 분류 체계는 6개월에서 1년 내에 재분류 과정을 거쳐 신규 분류체계를 수립하고 그에 따른 각 직무별 핵심단어와 핵심단어의 동시등장 빈도율을 다시 구해야 할 필요성을 고려하여 보완할 필요가 있다. 또한 시각화에 따른 분석 결과 또한 좀 더 장기간에 걸친 데이터 수집을 통해 명확한 변화 추이를 파악할 필요가 있다. 마지막으로 본 연구는 워크넷 데이터만을 가지고 분석이 이루어졌기 때문에 추후 대기업의 공채 정보 등을 포함시켜 전체 산업계의 수요 변화를 파악한다면 좀 더 확실한 결과 도출이 가능할 것이다.

3. 빅데이터 분석 기반의 직무·직종 분류 체계의 시사점

본 절은 빅데이터 분석 기반의 직무·직종 분류체계를 개발하였으며 기존 워크넷 분류체계와 NCS 분류체계의 불명확한 분류기준을 보다 현실화하기 위한 체계를 개발하였다. 그에 따라 본 보고서의 장점은 다음과 같다. 첫째, SW분야 기업들의 실제 채용공고를 전수 활용하여 산업계의 인력 수요에 대하여 확인할 수 있는 체계를 개발했다는 점은 본 보고서만의 독창적인 장점으로 활용할 수 있다. 이는 기존 분류체계인 NCS와 워크넷이 보다 더 광의의 개념 및 모든 직업 분류를 구분하는 목적을 가지고 있었다는 점에서 SW분야에 특화된 체계를 제시할 수 없었던 반면 본 보고서의 결과는 SW산업 수요에 특화된 데이터를 활용했다는 장점을 갖는다.

둘째, 본 보고서에서는 기존 직무·직종 분류체계들이 보유하고 있는 논리적 직무분류 방식에 추가적으로 빅데이터 분석을 활용하여 최신 기술을 적용한 데이터 기반의 SW 직무·직종 분류체계를 개발했다는 점이 큰 장점이다. 특히 비정형 데이터인 채용광고 자료들을 분석하기 위하여 전체 데이터들을 단어로 구분하여 일괄 자료화 하였다. 이렇게 텍스트마이닝 기법을 적용하여 형태소 분석을 통해 정제된 단어 자료들을 바탕으로 본 연구는 직무·직종 체계에 대한 토픽 모델링을 분석하기 위하여 LDA 알고리즘을 사용한 분류 모델링을 활용하였다.

셋째, 본 보고서에서 제시된 직무·직종 분류체계는 데이터를 통해 분류된 직무 및 직종에 대하여 SW 분야의 학계와 산업계 전문가의 상호 일관된 직무 및 직종에 대한 평가들을 바탕으로 직무 및 직종에 대한 구분 및 정의를 수행하였다. 또한 마찬가지로 국제직무분류체계인 ISCO, SOC-2016, O*Net의 기존 활용되는 정의를 바탕으로 직무 및 직종에 대해 정의하고 해당 직무의 특성을 분류 하였다. 또한 기존에 분류되지 못했던 직무 군집인 HCI 컨설턴트 및 미디어 콘텐츠 창작자, 소셜 미디어 관리자와 모바일 개발자 등 신규 직종 및 직무에 대하여 국외 학술 연구논문 및 해외 구직 사이트에서 사용하고 있는 용어들을 바탕으로 해당 직무 및 직종을 정의하였다.

이에 대한 본 과제의 한계점 및 향후 연구 방향은 다음과 같다.

첫째, 다중직무 채용공고에 대한 세분화 데이터 수집 및 분석을 수행할 필요가 있다. 본 연구의 분석 대상 데이터(20,408건)에서 다중직무 채용공고를 확인한 결과 약 156건의 공고가 존재했다. 본 과제에서 156건의 다중 공고 개수는 전체 채용공고의 0.7%인 매우 극소량의 공고 개수이기 때문에, 데이터 분석 기반 직무·직종 분류체계에 크게 영향을 미치지 못하는 것으로 판단하여 분석에 포함하여 진행하였다. 특히 워크넷 채용공고 데이터에는 한 건의 채용공고에서 여러 개의 다중직무 인력을 요구하는 경우가 있다. 대표적으로 기업에서 SW 직군을 부문별로 공개 채용하는 경우와 웹개발 및 웹디자이너를 동시에 채용하는 경우의 사례를 생각해볼 수 있다. 따라서 다중직무 채용공고는 그 수가 전체 공고에서 절대적으로 50% 이상의 많은 부분을 차지하게 된다면 상이한 직무에 대한 핵심단어가 혼재되는 결과를 만들 수 있다. 따라서 향후 다중직무 채용공고에 대하여 별도의 자연어 처리 및 데이터 정제 기법을 통해 분류체계에 반영하는 방법을 연구하여 개선할 필요가 있다.

둘째, 본 연구과제에서는 채용공고로부터 ‘웹’, ‘디자인’, ‘개발’ 등과 같은 한 가지 단어 단위의 데이터를 추출하여 등장 확률과 빈도를 분석하였다. 추가적인 향후 연구 사항으로 ‘웹 디자인’, ‘시스템 디자인’, ‘웹개발’, ‘앱개발’ 과 같은 두 가지 단어를 쌍으로 연결하는 방식의 정제방법을 활용함으로써 더 고도화 된 직무 데이터를 바탕으로 분석을 수행할 필요가 있다.

추가적으로 본 연구의 데이터 분석 기반의 SW 직무·직종 분류체계 개발을 바탕으로 현재 대학에서 SW 관련 교육 체계를 고려한다면 고질적인 SW인력의 대학간 간극의 문제를 해결할 수 있을 것이다.

제3장 시사점 및 제언

제1절 시사점 및 향후 연구 방향

본 연구과제는 SW분야의 실무적 관점에서의 인력수요를 확인하고 직무별 특성을 확인하기 위하여 채용공고 자료를 전량 수집하여 텍스트마이닝과 LDA 알고리즘을 통해 형태소가 구분된 단어 형태의 자료를 획득하였다. 이를 바탕으로 기존 워크넷의 분류체계 및 NCS 직업 분류 형태를 비교하고 빅데이터 분석 기반의 새로운 직무·직종 분류체계를 개발함으로써 산업현장의 SW 인력의 실태를 파악하고 나아가 현실적인 SW 인력의 가이드라인을 제공하기 위하여 작성되었다.

또한 새로 개발된 직무·직종 분류체계의 채용공고 데이터에 따른 시간적 변화를 볼 수 있도록 시스템을 개발하고 정보 제공 서비스를 제공하였다. 그에 따라 각 직무·직종별로 기간별 수요의 특성 및 정의, 중요 산업계 수요 요소 등을 확인할 수 있다.

본 연구과제는 다음의 3가지 시사점을 갖는다.

1. 빅데이터 분석을 통한 신규 SW 직무·직종 분류체계 확립

본 연구과제는 워크넷 채용공고 데이터를 기반으로 현장에서 요구하고 있는 전문 인력에 대한 수요를 확인하기 위하여 채용공고의 텍스트 자료를 분석함으로써 실수요자의 특성에 맞춘 직무·직종 분류체계를 개발하였다. 특히 기존광의의 의미로 활용되고 있는 분류체계를 산업 수요 관점으로 재정립함으로써 이론과 실무 관점의 직무 전문성에 대한 이질감을 축소하였으며 나아가 SW 인력 양성의 영역의 개선 발판을 마련하였다. 이는 워크넷 채용공고 데이터에 대한 텍스트마이닝 기법으로 추출 및 분석을 통하여 기존 워크넷 및 NCS에서 분류되었던 직무 및 직종의 모호함과 실제 현장과 매칭이 잘 되지 않는 점을 개선할 수 있었다. 특히 현실적인 핵심단어 기반의 SW 신(新) 분류체계를 제공함으로써 즉각적인 현장 투입 및 SW산업분야의 개선이 가능하게 되었다.

2. 빅데이터 분석을 통한 직무·직종 분류체계 개발 및 분석 방법 개발

본 연구과제는 기존 유사연구와 같이 기술적인 접근을 통하여 분석한 것이 아닌 채용공고 정보를 처리한 텍스트마이닝을 통한 단어 구분 방식을 적용하였다. 또한 구분된 단어 형태를 바탕으로 LDA 알고리즘을 통하여 기존의 직무·직종분류체계와 다르게 실무 수요 중심의 직무·직종 분류체계 개발 방법을 도출하였다. 특히 LDA 토픽모델링을 바탕으로 도출된 직무·직종체계를 바탕으로 NCS와 워크넷의 분류를 분석함으로써 기존 NCS와 워크넷 직무 기술에 관한 적합도가 비교적 떨어진다는 사실을 도출 할 수 있었고 이를 바탕으로 데이터 분석 기반의 직무·직종 체계 개발방식을 개발하여 향후 산업 수요 기반의 SW분야 개선방향의 방법을 확인하였다.

특히 LDA분석을 바탕으로 구분된 15개 직무 및 직종의 군집화 결과를 응용하여 채용공고를 토픽에 할당하면 주기적으로 각 직무 및 직종의 변화 추이의 관찰이 시간 흐름에 따라 어떻게 변화되는 지 파악이 가능하다는 점 또한 본 연구과제의 특징점이다. 하지만 LDA 기반의 군집화 결과는 채용공고의 개수가 지속적으로 축적 및 수요가 변화함에 따라 분류체계의 형태가 변화될 수 있다는 점에서 약 6개월~1년을 주기로 직무 및 직종 분류체계를 재구성하여 직무별 핵심 요구 사항을 확인할 필요가 있다. 분석에 사용되는 자료의 시간적 변화에 따라 분류체계의 재구성 결과가 분류된 핵심단어 및 직무의 핵심단어 등에 대한 변화 가능성이 있음은 향후 유의할 부분이다.

3. 산학 연계의 개선방안 제시

현재 NCS의 비교적 실제와의 이질적 부분과 나아가 대학교육에서 현장과 다른 교육 실태를 진단함에 있어 추후 SW 교육 설계에 개선점에 대한 참고자료로서 직무직종 분류체계의 이용이 가능하다. 개발된 분류체계에 따라 향후 SW산업계 수요와 직무·직종 분류체계를 고려하여 전문 인력 양성을 위한 교육 과정의 개선방향을 참고함으로써 SW 분야의 산학 연계의 교육과정 개선 정책의 활용이 가능할 것이다.

제2절 제언

1. NCS의 체계적 개선을 통하여 산학의 징검다리 역할 재확립 필요

현재까지 NCS는 다수의 산업분야를 총망라할 수 있는 직무 특성 파악에 있어 미흡한 부분이 존재한다. 이는 NCS가 제작된 지 오래되지 않았기 때문에 아직까지 실무적 부분을 체계적으로 적용하는 측면에서는 상대적으로 미흡한 부분이 없지 않다. 결과적으로 산학을 이어주는 징검다리의 역할을 해야 하는 관점에서 SW 분야에 대해 본 연구과제가 개발한 빅데이터 분석 기반의 직무·직종 분류체계와 같이 산업 수요에 의거한 특성을 보강하여 SW 인력 육성에 이바지 할 수 있는 체계의 재확립이 필요하다.

2. 실무 능력을 보유한 SW 인재양성을 위한 산업 수요에 따른 대학 교육과정 개선 필요

본 연구의 SW산업계 수요에 대한 분석 결과를 통하여 알 수 있듯이 실무 밀착형 전문가 양성을 위한 교육 체계 개선에 대한 많은 관심이 필요하다. SW 인력에 대한 전문성과 수요는 지속적으로 급증하고 있으며 많은 대학에서도 그 대응을 위한 노력이 진행되고 있다. 마찬가지로 미래창조과학부 및 교육부 등 정부 부처에서도 지속적으로 SW인력 양성을 위한 노력을 하고 있다.

따라서 SW산업계에서 요구하는 실무 기반의 요구 역량을 분야별로 분석하고 지속적 관찰을 함으로서, SW 전문 인력의 현재를 파악하고 기존 교육체계를 개선하여 실무밀착형 SW 인력을 양성해 나갈 수 있는 대학을 육성하는 방안을 마련할 수 있으며 그에 따른 정부 및 대학, 기업 간 균형에 맞출 수 있는 정책적 개선방안을 운영하는 것이 필요하다.

참 고 문 헌

국내문헌

- 강성경 (2016), ‘텍스트마이닝과 소셜 네트워크 분석을 이용한 재난대응 용어분석,’ *한국경영정보학회*, 제 18권, 1호, pp. 141-155.
- 김갑수 (2016), ‘2015 개정 교육과정의 SW 교육에 대한 초등 교사들의 인식 분석,’ *한국정보교육학회*, 제 20권, 1호, pp. 47-56.
- 김성근 (2016), ‘학술연구에서의 텍스트마이닝 활용 현황 및 주요분석기법,’ *한국EA학회*, 제 13권, 제 2호, pp. 317-329.
- 노형남 (2014), ‘빅 데이터 텍스트마이닝 - 정치 연설을 중심으로-,’ *한국화법학회*, 제 26권, pp. 289-325.
- 류지성 (2011), ‘IT 인재 양성을 위한 한국 대학교육의 과제,’ *삼성경제연구소*.
- 미래창조과학부 (2015), ‘정보통신산업의 진흥에 관한 2015 연차보고서,’ *정보통신산업연구진흥원*.
- 박만희 (2016), ‘동시단어분석을 이용한 품질경영분야 지식구조 분석,’ *한국품질경영학회지*, 제 44권, 제 2호, pp. 389-408.
- 박자현, 송민 (2013), ‘토픽모델링을 활용한 국내 문헌정보학 연구동향 분석,’ *정보관리학회*, 제 30권, 제 1호, pp. 7-32.
- 안형준 (2013), ‘데이터 마이닝 및 통계 기법을 활용한 색상과 단어 간의 관계 분석,’ *한국정보기술학회논문지*, 제 11권, 제 12호 pp. 185-192.
- 이수연 (2016), ‘텍스트마이닝 기반 토픽 분석을 통한 온라인 성차별의 이해,’ *사이버커뮤니케이션 학보*, 제 33권 제3호, pp. 159-199.
- 이순근 (2015), ‘특허 문서로부터 단어 추출을 위한 텍스트마이닝 기반 그래프 모델,’ *대한안전경영학회*, 제 17권, 제 4호, pp. 335-342.
- 오세훈 (2005), ‘우리나라 문헌정보학 학술지 논문 및 인용문헌 분석을 통한 연구동향 연구,’ *정보관리학회지*, 제 22권, 제 3호, pp. 379-408.
- 장재영 (2013), ‘텍스트마이닝을 위한 그래프 기반 텍스트 표현 모델의 연구 동향,’ *한국인터넷방송통신학회*, 제 13권, 제 5호, pp. 37-47.
- 장필성 (2016), ‘2016 다보스 포럼: 다가오는 4차 산업혁명에 대한 우리의 전략은?,’ *과학기술정책*, 제 211호, pp. 12-15.
- 정보통신정책연구원 (2014), ‘2013 경제발전경험모듈화사업: ICT 인력양성,’ *정보통신정책연구원*.
- 정보통신산업연구진흥원 (2015), ‘2015 SW산업 연간보고서,’ *정보통신산업연구진흥원*.
- 최해욱 (2015), ‘소셜미디어의 텍스트마이닝 분석을 통한 지적측량 이슈 네트워크 분석,’ *도시행정학보*, 제 28권, 제 3호, pp. 147-161.
- 한국SW산업협회 (2016), ‘SW직종별 인력수급실태조사 결과보고서.’
- 한국대학교육협의회 (2013), ‘2013년 산업계 관점 대학평가 컴퓨터(SW)분야 평가결과 종합보고서.’
- 홍장의 (2012), ‘SW 개발 역량 강화를 위한 Skill Set 기반의 직무 관리,’ *중소기업정보기술융합학회*

문지, 제 2권, 제 1호, pp. 59-67.

황성수 (2016), ‘2015 ICT·SW 전문인력 실태분석 및 전망,’ 한국직업능력개발원.

국외문헌

- Albuquerque, P., Pavlidis, P., Chatow, U., Chen, K.Y., and Jamal, Z. (2012), ‘Evaluating Promotional Activities in An Online Two-Sided Market of User-Generated Content,’ *Marketing Science*, Vol. 31, No. 3, pp. 406-432.
- Blei, D., Ng, A. and Jordan, M. (2003), ‘Latent Dirichlet Allocation,’ *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022.
- Christopher D. M. and Hinrich S. (1999), *Foundations of Statistical Natural Language Processing*, MIT Press (1999), ISBN 978-0-262-13360-9.
- Duan, Y. and Cruz, C. (2011), ‘Formalizing Semantic of Natural Language through Conceptualization from Existence,’ *International Journal of Innovation, Management and Technology*, Vol. 2, No. 1, pp. 37-42.
- Griffiths, D.M.B.T.L. and Tenenbaum, M.I.J.J.B. (2004), ‘Hierarchical Topic Models and the Nested Chinese Restaurant Process,’ *Advances in Neural Information Processing Systems 16: Proceedings of the 2003 Conference*, MIT Press.
- Harman, D. (1992), ‘Overview of The First Text Retrieval Conference(TREC-1),’ in *Proceedings of the First Text Retrieval Conference (TREC-1)*, pp. 1-20.
- Harvard University (2014), *The Unofficial Guide to Computer Science at HARVARD ver. 14*
- Hedberg, H., and Iivari, N. (2009), ‘Integrating HCI Specialists into Open Source Software Development Projects,’ in *IFIP International Conference on Open Source Systems*, pp. 251-263.
- Hutchins, J. (2005). *The history of machine translation in a nutshell*, <http://ourworld.compuserve.com/homepages/WJHutchins>.
- Iivari, J., Iivari, N. (2006), ‘Varieties of User-Centeredness,’ in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society Press, LA.
- Kishorjit, N., Vidya, R.R.K., Nirmal, Y., and Sivaji B. (2012), ‘Manipuri Morpheme Identification’, in *Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing (SANLP)*, pp. 95-108.
- Parveen, F., Jaafar, N. I., and Ainin, S. (2015), ‘Social media usage and organizational performance: Reflections of Malaysian Social Media Managers,’ *Telematics and Informatics*, Vol. 32, No. 1, pp. 67-78.
- Perreault, W. D. and Laurence, E. L.(1989), ‘Reliability of nominal data based on qualitative judgments,’ *Journal of Marketing Research*, Vol. 26, pp. 135-148.
- Ronen, F. and Sanger, J. (2007), *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, Cambridge University Press.
- Sidorov, G., Gelbukh, A., Gómez-Adorno, H. and Pinto, D. (2014), ‘Soft Similarity and Soft Cosine Measure: Similarity of Features in Vector Space Model’ . *Computación y Sistemas*, Vol. 18, No. 3, pp. 491-504.
- Sidorov, G., Velasquez, F., Stamatatos, E., Gelbukh, A. and Chanona-Hernández, L. (2014), ‘Syntactic Dependency-based N-grams as Classification Features,’ *LNAI*, Vol. 7630, pp. 1-11.

- Singhal, A. (2001), 'Modern Information Retrieval: A Brief Overview,' *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, Vol. 24, No. 4, pp. 35-43.
- Spruill, M.C. (2007), 'Asymptotic Distribution of Coordinates on High Dimensional Spheres,' *Electronic Communications in Probability*, Vol. 12, pp. 234-247.
- Sullivan, D. (2001), *Document Warehousing and Text Mining: Techniques for Improving Business Operations, Marketing, and Sales*, John Wiley & Sons, Inc.
- Tan, A.H. (1999), 'Text mining: The state of the art and the challenges,' in *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*.
- Tseng, Y. H., Lin, C. J. and Lin, Y. I. (2007), 'Text Mining Techniques for Patent Analysis,' *Information Processing & Management*, Vol. 43, No. 5, pp. 1216-1247.
- Weiss, S. M., Apte, C., Damerau, F. J., Johnson, D. E., Oles, F. J., Goetz, T., and Hampp, T. (1999). 'Maximizing Text-mining Performance,' *IEEE Intelligent systems*, Vol. 14, No. 4, pp. 63-69.

웹 문헌

- 가톨릭대학교 정시정보: <http://ipsi.catholic.ac.kr/pages/?p=16&mj=02>, p. 3.
- 가톨릭대학교 수시정보: <http://ipsi.catholic.ac.kr/pages/?p=11&mj=01>, p. 5.
- 가톨릭대학교 컴퓨터정보공학부 홈페이지, '컴퓨터정보공학부 교육목적 & 교육목표,' <http://csie.catholic.ac.kr/>
- 경북대학교 IT대학 컴퓨터학부 홈페이지, '경북대학교 교육 목표', http://computer.knu.ac.kr/01_sub/02_sub.html
- 경북대학교 IT대학 컴퓨터학부 홈페이지, '경북대학교 컴퓨터 학부 정원,' http://computer.knu.ac.kr/01_sub/04_sub.html
- 고려대학교 입학처 홈페이지, '컴퓨터학과 정원,' http://oku.korea.ac.kr/oku/admissions/early_std.jsp
- 고용노동부, '능력중심사회를 위한 직업능력개발 혁신 3개년 실천계획', 관계부처합동, 보도자료 2014.09.25.
- 고용노동부, '능력중심을 위한 직업능력개발 혁신 3개년 실천계획,' 2014, http://www.kosta.or.kr/mail/2014/download/05-SW_NCS_conference_141217_KimYoungJu.pdf
- 한국산업인력공단, '국가직무 능력표준 활용범위,' http://www.ncs.go.kr/ncs/page.do?sk=P1A1_PG01_001
- 한국산업인력공단, '국가직무능력표준 활용범위,' 국가직무능력표준 웹페이지, http://www.ncs.go.kr/ncs/page.do?sk=P1A1_PG01_001
- 한국산업인력공단, 국가직무능력표준 홈페이지, 'NCS 학습모듈의 개념', https://www.ncs.go.kr/ncs/page.do?sk=P1A1_PG01_006
- 한국산업인력공단, 국가직무능력표준 홈페이지, 'NCS 분야별 검색', https://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_001#none
- 한국산업인력공단, 국가직무능력표준 홈페이지, '특수부가통신서비스 NCS 직무정의,'

http://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_002&mk=MU000000099&uk=MU000000067
 한국산업인력공단, 국가직무능력표준 홈페이지, ‘이동통신서비스 NCS직무정의,’
http://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_002&mk=MU000000099&uk=MU000000067
 한국산업인력공단, 국가직무능력표준 홈페이지, ‘부가통신응용중계서비스 NCS 직무정의,’
http://www.ncs.go.kr/ncs/page.do?sk=P1A2_PG01_002&mk=MU000000099&uk=MU000000067
 교육과학기술부, 2010년 산업계 관점 대학평가 평가결과 보도자료(20110520, 조간)
 성균관대학교 정시정보, <http://ibook.skku.edu/Viewer/2B4Z3D3FPQ01>, p. 18
 성균관대학교 수시정보, http://ibook.skku.edu/Viewer/2017_summer p. 10,
 성균관대학교 소프트웨어대학 홈페이지, ‘소프트웨어학과 비전,’ http://cs.skku.edu/college/dsw_intro
 순천향대학교 정시 정보: <http://ipsi.sch.ac.kr/jungsi/application/handbook.html>, p. 8.
 순천향대 소프트웨어공학과 홈페이지, ‘컴퓨터소프트웨어학과 비전,’
<https://homepage.sch.ac.kr/computer/01/06.jsp>
 순천향대학교 수시 정보, http://ipsi.sch.ac.kr/susi/application/handbook_01.html, p. 8.
 순천향대학교 컴퓨터공학과 홈페이지, ‘순천향대학교 컴퓨터공학과 학과발전계획,’
<https://homepage.sch.ac.kr/xe/01/06.jsp>
 연세대학교 수시정보, <http://admission.yonsei.ac.kr/seoul/admission/html/rolling/guide.asp>
 연세대학교 정시정보, <http://admission.yonsei.ac.kr/seoul/admission/html/regular/guide.asp>
 인하대학교 입학처 홈페이지, ‘컴퓨터공학부 정원,’
http://admission.inha.ac.kr/UI/college/college_leaflet.aspx?DeptCode=20
 카이스트 신문 홈페이지, ‘입학 시 희망전공과 실제 신청전공 비교,’
<http://times.kaist.ac.kr/news/articleView.html?idxno=2050>
 통계청, 고용직업분류, KECO, https://kssc.kostat.go.kr:8443/ksscNew_web/index.jsp#
 한국과학기술원 입학처, ‘2017학년도 학사과정 모집요강,’ 2016학년도 대비 주요 변경사항
 Employment and Training Administration of the United States Department of Labor, ‘Occupational Information Net.’ <https://www.onetonline.org/>
<http://techterms.com/category/software>
<http://www.copyright.or.kr/information-materials/dictionary>
 International Labour Organization, ‘International Standard Classification of Occupation,’
<http://www.ilo.org/public/english/bureau/stat/isco/>
 Jaccard index, https://en.wikipedia.org/wiki/Jaccard_index
 U. S. Bureau of Labor Statistics, ‘Standard Occupational Classification.’ <https://www.bls.gov/soc/>
 Wikipedia, ‘잠재 디리클레 할당,’
https://ko.wikipedia.org/wiki/%EC%9E%A0%EC%9E%AC_%EB%94%94%EB%A6%AC%ED%81%B4%EB%A0%88_%ED%95%A0%EB%8B%B9

연구보고서 2016-016

**빅데이터 분석 기반 SW산업 직종·직무 분류체계
개발 및 동향 연구**

2017년 05월 인쇄

2017년 04월 발행

발행처 정보통신산업진흥원 부설 소프트웨어정책연구소
경기도 성남시 분당구 대왕판교로712번길22 A동 4층
Homepage: www.spri.kr

ISBN : 978-89-6108-377-5

주 의

1. 이 보고서는 소프트웨어정책연구소에서 수행한 연구보고서입니다.
2. 이 보고서의 내용을 발표할 때에는 반드시 소프트웨어정책연구소에서 수행한 연구결과임을 밝혀야 합니다.

ISBN : 978-89-6108-377-5



[소프트웨어정책연구소]에 의해 작성된 [SPRI 보고서]는 공공저작물 자유이용허락 표시기준 제 4유형(출처표시-상업적이용금지-변경금지)에 따라 이용할 수 있습니다.
(출처를 밝히면 자유로운 이용이 가능하지만, 영리목적으로 이용할 수 없고, 변경 없이 그대로 이용해야 합니다.)