Research Report

연구보고서 2017-012

제4차 산업혁명 시대에 대비한 SW인력 중장기 수요 전망

이호/김정민

2018.04.

SPRi 소프트웨어정책연구소

이 보고서는 2017년도 과학기술정보통신부 정보통신진흥기금을 지원받아 수행한 연구결과로 보고서 내용은 연구자의 견해이며, 과학기술정보통신부의 공식입장과 다를 수 있습니다.

목 차

제1장 서 론 ····· 1

제1절 연구 배경 ····· 1

제2절 연구 목적 ····· 6

제2장 SW 산업의 환경 분석 및 신·구 직종 및 직업 분류 ····· 7

제1절 데이터 분석 개요 ····· 7

1. SW인력 수요 빅데이터 분석 프로세스 ····· 7

제2절 거시환경 분석(PEST) ····· 9

1. 정치적(Political) ····· 9

2. 경제적(Economical) ····· 12

3. 사회문화적(Socio-cultural) ····· 14

4. 기술적(Technological) ····· 16

제3절 텍스트마이닝 분석 모델 ····· 20

1. 잠재 디리클레 할당(LDA) ····· 20

2. 다이내믹 토픽 모델링(DTM) ····· 38

제4절 SW산업 직종 및 직업 분류체계 ····· 50

1. 한국표준직업분류 (KCSO-07) ····· 50

2. ICSO-88 (국제표준직업분류) ····· 55

3. 한국고용직업분류 (KECO-09) ····· 60

4. 국가직무능력표준 (NCS) ····· 66

5. IT 분야 역량인정체계 (ITSQF) ····72
6. O*net (미국직업정보체계) ····79
7. SFIA (The Skills Framework for the Information Age) ····84
제3장 분석 결과 ····90
제1절 텍스트마이닝 분석 ····90
1. SW인력 수요 정보 데이터 ····90
2. 빅데이터 기반 분석 프로세스 ····94
3. 분석 결과 ····100
4. 분석 후 직종/직업 재배치 ····130
5. 일자리 변화상(직종/직업) ····138
제2절 거시 환경 분석과 일치성 검토 ····142
1. 거시 환경 분석(PEST)과 SW직종 재배치의 일치성 검토 ····142
2. 결과 비교 ····148
3. 거시 환경 분석과 재배치 검토의 시사점 ····159
제3절 SW직종 수요 예측을 위한 연평균복합성장률(CAGR) 분석 ····162
1. 연평균종합성장률 (CAGR) 분석 프로세스 ····162
2. 재배치된 SW 직종별 CAGR 결과 ····164
3. CAGR의 결과에 따른 SW인력 수요 전망의 시사점 ····182
제4장 시사점 및 제언 ····188
제1절 시사점 및 향후 연구 방향 ····188
제2절 제언 ····190

표 목 차

<표 1-1> 기존 산업혁명과 4차 산업혁명의 차이 ····· 1

<표 2-1> 산업별 지능화 혁신 목표 ····· 11

<표 2-2> 분할발주 대상 시범사업 개요 ····· 12

<표 2-3> 제조업의 혁신 단계 비교 ····· 14

<표 2-4> 산업별 AR과 VR 기술의 적용 ····· 16

<표 2-5> 토픽 모델링 알고리즘 방법론 비교 ····· 21

<표 2-1> LDA 알고리즘 생성과정 ····· 25

<표 2-7> 핀테크 분야 LDA 분석 결과 ····· 27

<표 2-8> 정보시스템 토픽(2010~2015) 샘플링 ····· 28

<표 2-9> 산업공학 분야 연구 동향 주요 키워드 빈도 ····· 30

<표 2-10> TF-IDF의 기본 수식 ····· 35

<표 2-11> 제7차 개정 후 분류단계별 항목 수 ····· 53

<표 2-12> 한국표준직업분류에서 SW 관련 직종 분포 ····· 53

<표 2-13> 각 단계별 그룹의 수 ISCO-88 ····· 55

<표 2-14> IT 관련 직업 분류 체계 ····· 56

<표 2-15> KECO의 구조 ····· 61

<표 2-16> KECO SW 인력관련 분류체계표 ····· 65

<표 2-17> NCS의 활용 영역 ····· 67

<표 2-18> NCS 분류체계 구조 ····· 70

<표 2-19> 20. 정보통신-01. 정보기술 SW인력 관련 직무 ····· 71

<표 2-20> ITSQF 수준표 ····· 74

<표 2-21> ITSQF 직무역량체계 ····76
<표 2-22> O*net Information Technology 직종 분류 ····83
<표 2-23> SFIA 역량 명 리스트 ····87
<표 3-1> SW인력 수요 데이터 수집 방법 ····98
<표 3-2> SW직종 토픽모델링분석을 사용하기 위한 SW 어휘사전 구성 ····99
<표 3-3> DTM 토픽모델링 분석결과 ····100
<표 3-4> 24개 직업분류를 통한 직업 이름 정의 ····128
<표 3-5> 기존 분류체계 코드 매칭 ····131
<표 3-6> 신규 일자리 정의 (12개 직종, 24개 직업) ····134
<표 3-7> 소셜 미디어 마케터 CAGR 분석결과 ····164
<표 3-8> 모바일 어플리케이션 개발자 CAGR 분석결과 ····165
<표 3-9> 멀티미디어 디자이너 CAGR 분석결과 ····166
<표 3-10> 정보 보안 전문가 CAGR 분석결과 ····166
<표 3-11> 멀티미디어 콘텐츠 관리자 CAGR 분석결과 ····167
<표 3-12> 정보 시스템 운영자 CAGR 분석결과 ····167
<표 3-13> 웹 운영자 CAGR 분석결과 ····168
<표 3-14> 사물인터넷 소프트웨어 개발자 CAGR 분석결과 ····169
<표 3-15> E-커머스 관리자 CAGR 분석결과 ····169
<표 3-16> 서버 개발자 CAGR 분석결과 ····170
<표 3-17> AR/VR 콘텐츠 개발자 CAGR 분석결과 (▲) ····171
<표 3-18> 웹 방송엔지니어 CAGR 분석결과 ····171
<표 3-19> 드론 소프트웨어 개발자 CAGR 분석결과 ····172

<표 3-20> 자동화 설계자 CAGR 분석결과 ·····173
<표 3-21> 디지털 마케터 CAGR 분석결과 ·····173
<표 3-22> 로보틱스 시스템 개발자 CAGR 분석결과 ·····174
<표 3-23> 온라인 머천다이저 CAGR 분석결과 ·····175
<표 3-24> 시스템 소프트웨어 개발자 CAGR 분석결과 ·····175
<표 3-25> 임베디드 소프트웨어 개발자 CAGR 분석결과 ·····176
<표 3-26> 모바일 게임 개발자 CAGR 분석결과 ·····177
<표 3-27> 웹 호스팅 관리자 CAGR 분석결과 ·····177
<표 3-28> 웹 표준 전문가 CAGR 분석결과 ·····178
<표 3-29> 웹 프로그래머 CAGR 분석결과 ·····179
<표 3-30> UX/UI 디자이너 CAGR 분석결과 ·····180
<표 3-31> 24개 직업별 CAGR 분석결과 간 상대적 수요 증감 비교 ·····181

그 림 목 차

[그림 1-1] 세계 IT시장의 비중 변화 및 개발원가 중 SW비용 비중 ·····3

[그림 2-1] 세계 SW산업 시장 규모 및 트렌드 ·····7

[그림 2-2] SW인력 시장에 맞는 인력 수요 예측과 새로운 직종 및 직업 분류체계 제시 9

[그림 2-3] 디지털 콘텐츠 펀드 정부출자 및 민간자금 규모 ·····10

[그림 2-4] 공유경제 경제규모 및 서비스별 성장 전망 ·····13

[그림 2-5] 모바일 온리 집단 현황 ·····15

[그림 2-6] VR/AR 시장 규모 전망 ·····17

[그림 2-7] 세계 드론 시장규모 ·····18

[그림 2-8] 세계 로봇 구매 시장 ·····18

[그림 2-9] 사물인터넷 시장 규모 ·····19

[그림 2-10] LDA의 원리 ·····22

[그림 2-11] 주제별 상위 15개의 키워드 추출 결과 ·····23

[그림 2-12] LDA 그래픽 모델 ·····24

[그림 2-13] Bi-gram과 Tri-gram의 차이 ·····26

[그림 2-14] 순수 정보시스템과 융합 정보시스템 토픽 모델링 ·····29

[그림 2-15] LDA 데이터 전처리 과정 ·····30

[그림 2-16] 문헌정보학 관련 4개 학술지에서 도출된 토픽모델링 결과 비교 · 분석 ···32

[그림 2-17] 블로그 LDA 토픽 모델링 분석 결과 각 키워드별 확률 비중치 ·····34

[그림 2-18] 특허 데이터 기반 LDA 분석 프로세스 ·····35

[그림 2-19] DTM 그래픽 모델 ·····40

[그림 2-20] 기존 LDA와 TOT, TLT의 그래프 모델 비교 ·····42

[그림 2-21] 카쉐어링 DTM 알고리즘 분석 및 오피니언 마이닝 분석 절차 ·····45
[그림 2-22] LDA, MDTM, Online DTM의 그래픽 모델 비교 ·····48
[그림 2-23] 한국표준직업분류 연혁 ·····50
[그림 2-24] 제 7차 한국표준직업분류 개정 특징 ·····52
[그림 2-25] 국가직무능력표준 개념도 ·····67
[그림 2-26] NCS분류체계 구성도(예시) ·····69
[그림 2-27] IT분야 역량체계의 개념 ·····73
[그림 2-28] ITSQF 역량의 기본적 구조 ·····75
[그림 2-29] ITSQF 직무군별 소프트웨어기술자 신고현황 ·····76
[그림 2-30] O*net Online에서 제공되고 있는 O*net 콘텐츠 구성도 ·····80
[그림 2-31] O*net Online에서 제공되고 있는 직업 정보 예시(Computer Programmers) 82
[그림 2-32] SFIA 역량 관리 프레임워크 ·····85
[그림 2-33] 역랑 난이도별 전문 역할 및 조직 역할 ·····86
[그림 3-1] 임금충족률 상위 5개 직종 ·····90
[그림 3-2] 워크넷 채용정보 직종별 'IT · 정보통신 · 웹' 검색 결과 ·····91
[그림 3-3] 워크넷 채용정보 세부항목 수집 예시 ·····92
[그림 3-4] SW 인력 수요 데이터베이스 ·····93
[그림 3-5] SW인력 수요데이터 예시 ·····93
[그림 3-6] SW인력 수요 빅데이터 분석 프로세스 ·····94
[그림 3-7] 빅데이터 기반 세부 프로세스 ·····96
[그림 3-8] SW인력 수요 데이터 수집 방법 (Python 2.7 Scrapy) ·····97
[그림 3-9] SW인력 텍스트 마이닝을 위한 사전 구성 및 추출 단어 파싱 ·····98

[그림 3-10] SW직종 분석용 데이터 생성 과정 ·····99
[그림 3-11] 소셜미디어 마케터 주요 키워드 분포 ·····104
[그림 3-12] 모바일 어플리케이션 프로그래머 주요 키워드 분포 ·····105
[그림 3-13] 멀티미디어 디자이너 주요 키워드 분포 ·····106
[그림 3-14] 정보보안 전문가 주요 키워드 분포 ·····107
[그림 3-15] 멀티미디어 콘텐츠 관리자 주요 키워드 분포 ·····108
[그림 3-16] 정보시스템 운영자 주요 키워드 분포 ·····109
[그림 3-17] 웹 운영자 주요 키워드 분포 ·····110
[그림 3-18] 사물인터넷 소프트웨어 개발자 주요 키워드 분포 ·····111
[그림 3-19] E-커머스 관리자 주요 키워드 분포 ·····112
[그림 3-20] 서버 개발자 주요 키워드 분포 ·····113
[그림 3-21] AR/VR 콘텐츠 관리자 주요 키워드 분포 ·····114
[그림 3-22] 웹 방송 엔지니어 주요 키워드 분포 ·····115
[그림 3-23] 드론 애플리케이션 개발자 주요 키워드 분포 ·····116
[그림 3-24] 자동화 설계자 주요 키워드 분류 ·····117
[그림 3-25] 디지털 마케터 ·····118
[그림 3-26] 로보틱스 시스템 개발자 주요 키워드 분포 ·····119
[그림 3-27] 온라인 머천다이저 주요 키워드 분포 ·····120
[그림 3-28] 시스템 소프트웨어 개발자 주요 키워드 분포 ·····121
[그림 3-29] 임베디드 소프트웨어 개발자 ·····122
[그림 3-30] 모바일 게임 개발자 주요 키워드 분포 ·····123
[그림 3-31] 웹 호스팅 관리자 ·····124

[그림 3-32] 웹 표준 전문가 주요 키워드 분포 ·····125
[그림 3-33] 웹 프로그래머 주요 키워드 분포 ·····126
[그림 3-34] UX/UI 디자이너 주요 키워드 분포 ·····127
[그림 3-35] SW산업 직종/직업 재배치 ·····133
[그림 3-36] 일자리 직종 변화 ·····138
[그림 3-37] 일자리 직업 변화 ·····139
[그림 3-38] 일자리 직업 변화 ·····141
[그림 3-39] SW인력 거시 환경분석 ·····142
[그림 3-40] 정치적 차원의 직업 일치성 ·····148
[그림 3-41] 경제적 차원의 직업 일치성 ·····151
[그림 3-42] 사회문화적 차원의 직업 일치성 ·····153
[그림 3-43] 기술적 차원의 직업 일치성 ·····156
[그림 3-44] 연평균 종합성장률 분석 프로세스 ·····162
[그림 3-45] 직업별 평균증감률 5개년(2017~2020, 2025) 예측 비교 ·····183
[그림 3-46] 직업별 평균변동률 5개년(2017~2020, 2025) 예측 비교 ·····184
[그림 3-47] 직업별 평균 월별 공고 수 5개년(2017~2020, 2025) 예측 비교 ·····185
[그림 3-48] 직업별 누적 월별 평균 공고 수 5개년(2017~2020, 2025) 예측 비교 ·····186

요 약 문

1. 연구 목적 및 필요성

기존 산업의 근간이 되어온 제조 분야에 대한 정보통신 기술의 결합은 사물과 사물의 연결에 의한 산업구조의 변화를 가져오고 있다. 특히 정보통신기술에 기반한 4차 산업혁명은 기존 산업 구조로 하여금 파괴적 양상으로 나타나고 있다. 이 중 가장 큰 변화는 일자리 수의 감소 또는 신규 일자리의 출현이다.

급격한 산업 환경 및 경제 환경의 변화로 야기되는 신사업 영역에서 새로운 직종의 발생 및 기존 직업에 대한 수요 전망을 확인하는 것은 매우 중요하다. 따라서 각 산업 분야별 인력 수요의 예측 및 신규 직종 및 직업을 구분해내는 것이 매우 시급한 시기이며 국가적 핵심 직업군의 육성 및 대응 전략 수립에 활용할 일자리 전략 수립을 위해 명확한 예측 정보가 필요하다.

그럼에도 불구하고 기존 직종 분류에 대한 연구들은 주로 산업 전문가 및 실무 경험자들의 정성적 평가에 의존한 경우가 많았다. 특히 데이터 수집, 가공, 분석에 소모되는 시간을 고려할 때, 기존의 분석 방식에 의한 결과 자료는 현재의 상황에 대응하기에는 약 1~2년의 시차가 발생한다. 따라서 전통적 방식의 예측방법은 현재를 확인하고 대응할 수 있는 미래 신규 일자리 수요를 정확히 예측하기 어렵다. 이 문제를 해결하기 위해서는 최신의 구인 정보를 전수 분석함으로서 2017년 현재의 시장 환경에 반영하는 것이 필요하다. 따라서 최근 활발히 도입되고 있는 빅데이터 기반의 분류 기법을 이용한 분석을 도입하는 것은 본 문제의 해결에 매우 시의적절하다.

특히 4차 산업혁명의 핵심 동인인 소프트웨어 분야의 특성을 파악하고 향후 출현할 직업 및 규모가 줄어들거나 타 직업과 융합될 직업에 대한 확인이 산업 정책의 도출을 위해 선행적으로 수행되어야 한다. 따라서 본 연구는 텍스트마이닝을 통하여 직종 및 직업 분류를 도출하고, 각각의 직종별 성장 가능성에 대한 예측 분석을 수행함으로써 다가올 4차 산업혁명 시대의 대응 전략 및 핵심 산업 육성을 위한 토대를 마련하고자 하였다.

그러므로 본 연구과제의 목적은 SW산업 관련 직종에 대한 구인정보를 바탕으로, 직종 및 직업의 재구성 및 신규 출현, 분화, 융합되는 직업을 분석함으로서 각 직종별 향후 방향성을 도출함에 있다. 연구 목적에 따른 본 연구과제의 세부 연구 사항은 다음과 같다. 첫째, 빅데이터 기반 워크넷의 구인정보 분석을 통한 현 시점의 SW분야 인력수요를 파악하여 기존·신규 직종 및 직업에 대한 데이터 기반 분류를 도출한다. 둘째, 데이터 기반 분류 결과에 기반한 정성적 검정과 직종·직업 수요의 예측결과 기반 산업 생태계 분석을 수행한다.

2. 연구 범위 및 방법

본 연구 보고서의 구성은 다음과 같다. 먼저, 1장에서는 4차 산업사회의 특징과 일자리 변화를 중심으로 연구의 배경 및 목적을 기술하였다. 2장에서는 SW산업의 환경 분석을 바탕으로 정치, 경제, 사회 및 기술적 분석을 수행하였으며 본 연구에서 활용할 빅데이터 분석 방법인 다이나믹토픽모델링(Dynamic Topic Modeling, 이하, DTM)을 제시하였다. 빅데이터 분석 방법론의 적정성 평가를 위해 DTM의 원천 알고리즘인 잠재디리클레할당(Latent Dirichlet allocation, 이하, LDA)을 간단히 소개하고 이와 DTM사이의 장단점을 비교분석하였다. 이와 더불어 기존의 SW관련 직종·직업의 분류 방법에 대한 개선 사항을 도출하였다. 3장에서는 워크넷의 SW인력 채용공고 수집 프로세스 및 실제 DTM 분석과정과 직종·직업 도출 과정을 기술하였다. 분석을 위하여 먼저 텍스트마이닝을 통해 채용공고 데이터에 대한 수집 및 단어별 전처리를 끝낸 후, DTM 기반의 분석을 수행하여 직종·직업의 분류체계를 구분하였다. 이후 도출된 신규 및 기존 직종/직업의 도출 결과를 바탕으로 각 직종의 연평균복합성장률(CAGR)을 분석하여 시사점을 도출하였다. 연구 결과에 대해서는 DTM을 통해 도출 된 핵심단어 기반 직종·직업을 기존에 존재하는 분류체계와 매핑하여 범용성을 확보하고, 거시 환경 분석과의 결과 일치성을 검토하여 신뢰성을 검증하였다. 마지막으로 4장에서는 전체 연구에서 확인된 결과의 해석 및 시사점을 도출하고 최종 정책적 제언을 제시하였다.

3. 연구 내용 및 결과

본 연구를 수행하기 위하여 SW인력 관련 고용 시장 수요 데이터를 확보하고 데이터 분석을 통하여, 텍스트마이닝을 실시하였다. 이를 통해 도출한 결과는 다음과 같다. 첫째, 수집된 내용을 토대로 빅데이터 기반 분석 프로세스를 적용하여 다이내믹 토픽 모델링 분석을 실시하고, 24개의 토픽으로 분류하여 현재 상황을 반영한 데이터 기반 채용공고 유형화를 수행하였다. 둘째, 워크넷 및 잡(Job) 채용 사이트에서 획득된 인력공급 요구 데이터와 상기 언급한 한국직업분류체계(KSCO), 미국직업분류체계(O*net), 국제직무역량표준(SFIA) 등 7개의 기존 직종 및 직업 분류체계를 기반으로 DTM을 활용한 12개의 직종, 24개의 직업을 구성하였다. 이를 통해 변화된 일자리 4가지 직종과 9개의 직업을 확인하였다. 셋째, 분석 결과를 거시 환경 분석인 PEST 모델에 맞춰 현재 시장에 알맞은 분석이 되었는지 일치성을 검토하였다. 넷째, 연평균종합성장률(CAGR) 분석을 통하여 직업별 공고 수 및 평균 증감률, 평균 변동률 이에 따른 누적 공고 수를 확인하였다.

4. 정책적 활용 내용

본 연구과제는 워크넷 데이터를 기반으로 SW산업 현장의 일자리 변화 형태를 4차 산업혁명 시대에 대비한 직종 분류를 바탕으로 재정립함으로써 이론과 실무 관점의 직종 및 직업 전문성에 대한 이질감을 축소한 점에 의의가 있으며 다음과 같은 장점이 있다. 첫째, 총 24종의 분류된 직업 중 변화된 일자리로 4개 직종 및 9개 직업을 구분함에 따라, 분화, 변화, 발견 직종을 구분하여 향후 국가 경쟁력 재고를 위한 일자리 전략 도출에 활용할 수 있다. 둘째, 24개 직종 및 직업의 군집화를 바탕으로 실시간 직종 및 직업의 변화 추이를 관찰할 수 있다. 특히 시간 흐름에 따라 직종 및 직업이 어떻게 분화, 융합, 변화, 발견 될 수 있는지 확인함에 따라 향후 직업의 변화 가능성에 시의적절한 직업 육성 전략을 수립할 수 있다.

5. 기대효과

일자리 도출에 빅데이터 기반 분석방법론을 활용하는 접근은 향후 급변하는 일자리 지형변화를 측정하기 위한 실시간 분석 도구로 활용가능하다. 특히 본 연구과제에서 다룬 분석 방식과 데이터 처리는 직종·직업 별 역량의 변화를 객관적으로 확인할 수 있다. 또한 다양한 선행연구의 단점을 보완함으로서 실시간 일자리 변화 대응에 능동적 대처가 가능할 것으로 판단된다. 특히 향후 직종·직업에 대한 능동적 대응 정책을 도출하는 것은 향후 국가 경쟁력을 강화시킬 수 있다. 또한 SW분야의 직업이 분화, 융합, 변화, 발견 등 상이한 양상으로 진화하고 있음을 확인함으로써, 4차 산업혁명을 대비하기 위한 SW인력 양성 정책의 효과적인 전략 수립에 도움을 줄 수 있을 것으로 기대한다.

SUMMARY

1. Research Purpose and Necessity

The combination of ICT technology in the manufacturing sector, which has been the basis of the existing industry, is leading to changes in the connected industrial structure. Especially, the 4th industrial revolution through ICT has a destructive effect on the existing industrial structure. The biggest change is the decrease in the number of jobs or the emergence of new jobs.

New business is being activated in a rapid change of industrial environment and economic environment. Therefore, it is very important to identify the prospects for new occupations from various job characteristics. Therefore, it is necessary to predict the demand for labor in each industry sector and to classify new occupations and jobs. This is useful for the government to develop a core workforce and to develop a response strategy.

Nevertheless, previous studies on the classification of occupations often rely on the qualitative evaluation of industry experts and practitioners. In particular, considering the time of data collection, processing, and analysis, the analysis results of previous studies show a time lag of about one or two years from the recent situation. Therefore, it is difficult to identify future demand for new jobs that can respond appropriately to the current situation. To solve this problem, it is necessary to analyze and reflect the current data of 2017 by analyzing the latest job information. Therefore, it is very timely to approach the analysis using the big data-based classification technique which is actively used recently.

In particular, this research should identify the characteristics of the software industry, which is the most important area in the fourth industrial revolution. To do this, this research need to analyze the emerging jobs and the jobs that will disappear. Therefore, this study predicted the possibility of future job growth by performing job

classification through text mining. These insights will provide the foundation for proper response strategies and upbringing of key industries in the forthcoming industrial revolution era.

The purpose of this study is to define occupational reorganization and newly emerging occupations based on the job information of SW industry related occupations and to analyze future demand direction for each job. The details of this research project are as follows. First, this study categorizes and derives existing and new occupations for SW industry at this time through analyzing Worknet's job information with big data analysis. Second, we conduct qualitative tests through expert opinions on new and existing occupations in the SW field, and analysis for forecasting the demand of each occupation.

2. Research Scope and Methodology

This research was composed as follows. In Chapter 1, we describe the background and purpose of the research based on the characteristics of the 4th industrial society and the importance of SW industry to job change. In Chapter 2, political, economic, social and technological analyzes were conducted based on the environmental analysis of the SW industry. The validity of the analytical method was evaluated by distinguishing the difference between the DTM (Dynamic Topic Modeling) presented in the big data analysis method and the Latent Dirichlet allocation (LDA) used in previous studies. In addition, we have improved the existing classification system of SW related occupations. In Chapter 3, after collecting the workforce recruitment announcements of Worknet, we rearranged the jobs and occupations based on the keywords of the announcement contents and examined the macroeconomic analysis. Moreover, this study collects and annotates the employment announcement data through text mining. And then we analyze the DTM topic modeling based on classification of job and occupation classification system. Based on deriving the new and existing jobs and occupations, the CAGR of each occupation was analyzed and implications were derived. In Chapter 4, the implications of the results and the political suggestions are presented.

3. Research Contents and Results

This study collected demand data of SW-related industry and conducted text mining. The results of this study are as follows. First, based on the collected data, DTM analysis was performed and identified that occupations were re-classified into 24 topics. Second, the data on the labor supply requirements obtained from the worknet and other recruitment sites were compared with seven job categories and job classification systems such as KSCO, O*net, and SFIA. As a result, 12 jobs and 24 professions were formed. This identified 4 jobs and 9 professions that will change. Third, according to the analysis of the PEST, which is the macroeconomic analysis, we analyzed whether the analysis was appropriate for the environment of current SW industry. Fourth, the number of announcements, average increase rate, average change rate, and cumulative announcement number were confirmed by CAGR analysis.

4. Political Implications

This study analyzed the changing patterns of the workplace in the SW industry based on the Worknet. The results are meaningful to redefine the classification of occupations in preparation for the fourth industrial revolution era. That result reduces the disparity in the professionalism of theory and practice. First, among the 24 types of classified jobs, 4 jobs and 9 professions were classified. In particular, it is possible to apply the system for deriving job strategy for future national competitiveness by classifying differentiated, changed, and appearing specific occupations. Second, we can observe periodic changes in the clustering of 24 jobs and profession. Confirming the differentiating, converging, changing, and appearing jobs and professions by time, it is possible to establish a timely career development strategy for future job change.

5. Expected Results

In the fourth industrial society, the SW field can be regarded as a core industry. Therefore, the analysis of the keyword of the Worknet presented is optimized to utilize in the rapidly changing SW industry in the process of job creation and prediction of job change. In particular, this research aims at objectively confirming the change of competency by job type and profession type. It is also possible to actively cope with changes in real-time job-changing situations by supplementing the insignificant part of the analysis methods of previous studies. As a result, using the big data analysis, it is possible to enhance the national competitiveness in the rapidly changing 4th industrial society by deducing active policies for future occupations and occupations. In addition, differentiating, converging, changing and appearing ways of jobs in SW field can be analyzed in order to prepare jobs required in whole industries in South Korea to prepare for industrial society in the fourth. Therefore, it is expected that this study provides a way to promote the change of the whole industry by deducing new jobs and observing changes in existing jobs.

CONTENTS

Chapter 1. Introduction

Section 1. Research Background

Section 2. Research Purpose

Chapter 2. The Environmental analysis of SW industry and classification of new and old occupation and job classification

Section 1. Introduction of Data Analysis

Section 2. The PEST Analysis of SW Industry

Section 3. The Methodology of Text Mining

Section 4. The Review of Previous Occupation Classifications

Chapter 3. Analysis and Improvement of SW Education in University

Section 1. The Result of Text Mining for Worknet Data

Section 2. The Validity for Mining Result with PEST analysis in SW Industry

Section 3. The CAGR results for Jobs and Professions in SW Industry

Chapter 4. Implication and Critical Review

Section 1. Implication and Future Study

Section 2. Conclusions

제1장 서 론

제1절 연구 배경

정보통신 기술의 활용처가 다양화 됨에 따라 기존 산업의 근간이 되어온 제조 분야와 정보통신 기술의 결합이 가속화 되고 있다. 특히 4차 산업혁명으로 부각된 해외 선진기업의 SW융합 선례들은 국내 사무 자동화 기반의 지식정보 혁명 시대를 넘어 사물과 사물간 연결에 의한 산업구조의 변화를 이끌어 내고 있다. 정보통신기술 기반의 4차 산업 혁명은 <표 1-1>과 같이 융합과 연결을 매개로 기술간 소통을 중심으로 발전하고 있다. 특히 기존에 없던 신규 가치가 창출되는 것이 특징이며 초연결성, 초지능성, 예측 가능성이 주요 동인으로서 제시되고 있다. 4차 산업혁명이 가져온 변화는 기존의 산업 구조에 파괴적 양상으로 나타나고 있으며 국가간, 기업간, 사회 시스템간의 변화를 야기하고 있다.

<표 1-1> 기존 산업혁명과 4차 산업혁명의 차이

구분	1차 산업혁명	2차 산업혁명	3차 산업혁명	4차 산업혁명
시기	18세기 후반	20세기 초반	1970년대 이후	2020년 이후
혁신 부문	증기 동력화	전력, 노동 분업	전자기기, ICT 혁명	ICT와 제조업 융합
커뮤니케이션 방식	책, 신문 등	전화기, TV 등	인터넷, SNS	사물인터넷, 사물서비스
생산 방식	생산 기계화	대량생산	부분 자동화	시뮬레이션을 통한 자동생산
생산통제	사람	사람	사람	기계

출처: World Economic Forum (2017)

4차 산업혁명의 영향력으로 인한 가장 큰 변화는 기존 일자리의 감소 또는 신규 발생하는 일자리의 증가이다. 최근 해외 일자리 관련 연구는 세계 고용의 65%를 이끌고 있는 15개국에서 향후 5년간 710만개의 일자리가 소멸되고 200만개의 신규 일자리가 발생할 것으로 예측하고 있다(World Economic Forum,

2017; 선대인, 2017). 한편 국내 근로자들을 대상으로 한 조사에서는 4차 산업혁명이 가져올 일자리 문제와 부의 분배에 대해 부정적으로 평가하고 있으며 10명 중 8명 이상이 일자리가 감소할 것으로 응답했다(한국언론진흥재단, 2017). 그러나 일자리 소멸 대한 불안감과는 반대로 4차 산업혁명에 의한 기대감도 동시에 증가하고 있다. 한국언론진흥재단의 분석결과는 '4차 산업혁명은 인류에게 혜택을 줄 것인가'에 대해 82.6%가 긍정적 응답을 하였으며 향후 경제의 신성장 동력이 될 것으로 제시되었다(한국언론진흥재단, 2017).

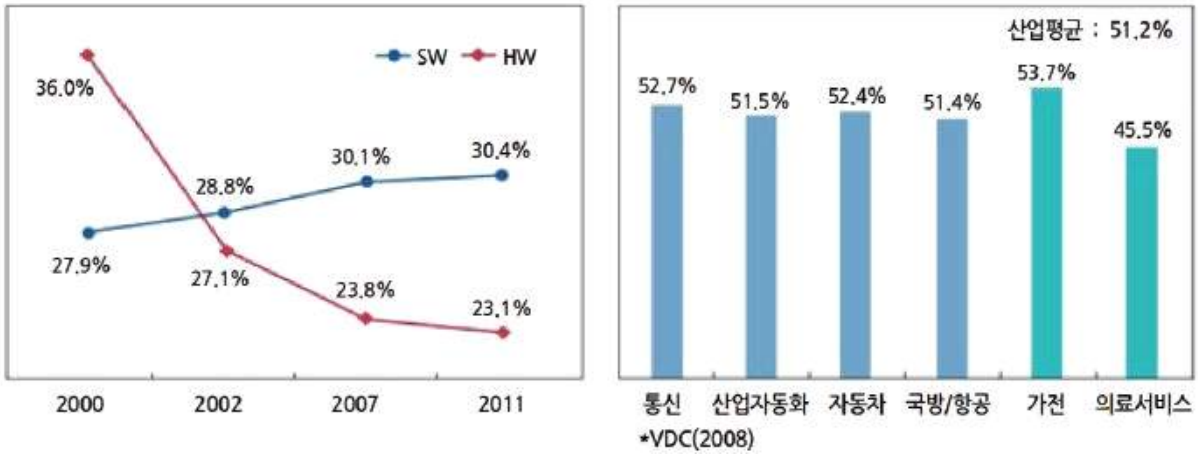
이러한 급격한 산업 환경 및 경제환경의 변화는 현재까지 등장하지 않았던 다양한 신사업의 창출로 이어지고 있다. 이에 따라 융복합적인 직무역량을 요구하는 신규 일자리의 수요가 증가하고 있는 상황이며, 기존 일자리의 변화 또한 필연적으로 다가오고 있다. 이는 4차 산업혁명으로 인해 일자리 80%가 소멸될 것이라는 부정적 전망과 사라지는 일자리 1개 당 온라인이나 미래 산업 분야에서 2.6개의 일자리가 생겨날 것이라는 긍정적 전망이 공존하는 이유를 뒷받침하고 있다(박영숙 외, 2013). 정부도 이러한 점에 착안하여 각 산업별 인력 및 직종 수요를 파악하기 위한 노력을 진행 중에 있다.

우리나라는 4차 산업혁명이 정보통신기술과 제조 환경 개선을 바탕으로 발전될 것이라 가정하고, 가장 밀접한 관련성이 있는 소프트웨어 분야의 육성에 집중해왔다. 우리나라는 하드웨어와 제조업 중심의 발전을 해왔으며, 이러한 편중된 발전을 극복하고자 SW산업을 국가 경쟁력 강화를 위해 가장 집중할 산업분야로 바라보고 있다. SW분야는 [그림 1-1]과 같이 기존 개발자 직업 이외에도 신규 출현하는 다양한 직업이 집중되는 분야이다. 이처럼 모든 산업분야에 소프트웨어가 중심이 된 IT기술이 폭넓게 접목되고 있다. 2002년을 기점으로는 소프트웨어 비중이 하드웨어 비중을 앞서기 시작했으며 2011년에는 소프트웨어비중(30%)이 하드웨어 비중(23%)을 크게 앞지르고 있다는 점이 이를 증명하고 있다.

그럼에도 불구하고 국내 소프트웨어 시장은 2013년에 약 278억 달러로 세계시장의 2.3%에 불과하며 부문별로는 임베디드 부문이 세계시장의 11.4%, IT서비스 부문은 1.1%, 패키지 부문은 0.8%를 차지하는 수준에 그치고 있다(정보통신산업진흥원, 2014). 특히 국내의 SW기술과 미국의 기술격차를 보면, 기초적인 SW기술인 운영체제는 3.27년, 빅데이터 분야는 2.63년으로

기초기술과 최신기술 전반적인 면에서 2년 이상의 격차가 나타났다 (한국산업기술평가관리원, 2013).

[그림 1-1] 세계 IT시장의 비중 변화 및 개발원가 중 SW비용 비중



출처: 미래창조과학부 보도자료(2013.10.9), 미디어잇(2014.4.22.) 재인용

앞서 언급한 격차를 극복하기 위한 하나의 방편으로 정부는 SW산업 직업별 수요를 명확히 파악하도록 직종 기준 및 역량을 분류 및 정의하고 이를 바탕으로 분야별 현황 및 수요 예측을 위한 체계를 마련하고 있다. 가령 이러한 문제의식은 한국표준직업분류(Korean Standard Classification of Occupations)와 국가직무능력표준(NCS, National Competency Standards) 등의 분류를 바탕으로 SW분야의 직업 분류 및 정의를 개선하고자 하는 유관기관의 지속적인 노력으로 설명될 수 있다. 그러나 기존 직업 분류체계는 산업현장에서 직업별 수행에 관련된 요구되는 지식, 기술, 태도 등의 내용을 소수의 전문가 의견수렴을 통해 체계화한 수준에 그치고 있다. 그러다보니 본 연구에서 수행하고자 하는 실시간 직업별 수요 파악 및 정량 데이터 분석을 위해 기존의 직업 분류 자료를 활용하기에는 현실과의 갭이 존재하는 문제점이 있다.

그간 제공되던 직업 분류 및 예측과 관련된 연구들은 주로 설문 조사를 기반으로 하는 실태조사, 전문가 협의에 의한 정성적 분류가 대부분이다. 설문조사를 통한 실태 연구는 특정 시기의 응답을 이용하여 분석이 수행되고 전문가 및 비 전문가 모두의 의견이 반영하지만 해당 조사 시기 이후의 변화상을

진단하기에는 제한점이 많다(홍성민 외, 2016). 또한 전문가 중심의 정성적 조사는 분석의 질에 따라 비교적 정확한 시사점을 제공할 수 있으나, 산업 분야에 종사하는 실무적 범주를 오인하는 경우 및 급변하는 산업의 모습을 시의적절하게 제시하지 못함에 따른 시간적 오차가 발생할 수 있다(홍성민 외, 2016).

결과적으로 기존 직종 분류 및 분석 형태의 한계는 다음과 같다.

첫째, 정량적 설문 데이터는 응답자의 특정 시기 응답을 기반으로 하기 때문에 현실에 직접 반영하기가 쉽지 않다. 특히 데이터 수집, 가공, 분석의 시간을 바탕으로 결과를 도출할 때에는 1-2년의 시차가 발생하게 된다. 결국 기존 분석 데이터는 현재 상황과는 일정 기간의 갭이 있어 신규 일자리 수요를 반영하기 어렵다. 추가적으로 직종 분류는 학문적 분류의 특성 및 국제표준의 기본 틀을 반영해야 한다는 점으로 인해 국내 특정분야에 한정하여 직접 응용하기에 어려운 측면이 있다. 둘째, 산업, 학계, 연구기관으로 구성된 전문가 협의체는 신규 발생한 산업에 대해 장기적 관점의 전략 구성이 어렵다. 셋째, 기존 분석 데이터는 현존하는 직업을 바탕으로 분류가 수행된다는 점에서 다가올 신규 산업 변화를 유추하기가 쉽지 않다. 결과적으로 신규 발생하는 직업에 대한 특성 분류가 모호한 상황에서는 향후 수요 예측 및 특성 정의, 육성이 필요한 산업에 대한 진단을 수행하기 어렵다.

2020년은 국내의 노동생산인구가 처음으로 감소하는 시기이다. 국내 실업률 상승 및 이동성 강화 등으로 실질 인구 감소가 시작되는 해다. 인구 감소가 이미시작된 선진국은 예외 없이 국력의 감소가 나타났다(박영숙 외, 2013). 결과적으로 직업별 인력 감소 및 신규 발생하는 일자리 현황을 확인하고 향후 변화 가능성을 예측하는 것은 2017년 현재 국가 경쟁력의 지속성 및 향상성 도모에 매우 중요한 사안이다. 특히 4차 산업혁명의 주요 분야인 SW분야 직업 분류 및 신규 일자리 분야의 증감을 예측하는 것이 가장 시급한 부분이다.

따라서 본 연구에서는 SW산업의 직업에 대한 현황을 파악하고 구인 정보 기반SW분야 일자리의 실제 수요를 측정하기 위한 빅데이터 기반 분석을 수행한다.

첫 번째로 SW산업의 현황과 직업 분류 환경의 특성을 진단하고 환경 변화요인을 거시환경분석 방법론을 통해 진단하였다. 두 번째로는 빅데이터

분석을 통해서 SW분야 구인공고의 특성 도출을 통해 직업을 구분하고 각 직업별 실시간 수요 변화에 따른 일자리 지형을 시각화 하였다. 결과적으로 본 연구는 SW분야 채용공고에서 획득한 빅데이터를 기반으로 정성·정량 분석을 병행함으로써, SW분야 일자리의 실수요를 반영한 신규 직업 및 기존 직업의 향후 방향성을 예측함에 그 목적을 두었다. 2절에서는 본 보고서의 연구 목적을 체계적으로 정리하였다.

제2절 연구 목적

1절에서 제시한 SW직업 및 인력 수요 현황의 특성에 따라서, 본 보고서의 목적은 다음과 같이 설정되었다. 빅데이터 분석을 기반으로 현 시점의 SW직업을 신규 분류 및 정의하고 각 직종 별 향후 수요 방향성을 분석하고자 하였다.

1. **워크넷의 구인 정보 빅데이터를 토픽모델링 방법론을 통해 유형화하고 현시점의 SW직업 명칭 및 정의**

2. **빅데이터 기반 SW분야 신규 및 기존 직업에 대한 전문가 타당성 검증과 도출된 직업별 인력 수요에 대한 중장기 예측을 위한 분석 방안 제시**

본 연구과제의 목적에 따라 2장에서는 SW인력 수요 예측과 새로운 직종 및 직업 분류 개발을 위한, SW인력 수요 빅데이터 분석 프로세스 및 일자리 환경 분석 결과를 서술하였다. 이후, 3장에서는 본 연구에서 사용된 빅데이터 분석 방법론을 설명하고 활용 방법론의 적합성을 확인한 후, 4장에서는 도출된 결과를 바탕으로 SW직종·직업 분류 및 각 직업별 정의 및 핵심단어 도출, 향후 직업별 성장 가능성에 대한 예측을 분석 및 논의하였다.

제2장 SW 산업의 환경 분석 및 신·구 직종 및 직업 분류

제1절 데이터 분석 개요

1. SW인력 수요 빅데이터 분석 프로세스

본 연구는 SW인력의 수요와 공급의 불균형과 같은 사회적 변화에 따른 SW직종·직업 변화상 도출의 필요성과 실무 현장에서 요구하는 수요를 파악하기 위하여, SW분야의 채용공고 및 관련 데이터를 수집하여 분석하고자 하였다. 2016 소프트웨어 산업 연간보고서에 따르면 [그림 2-1] 과 같이 세계적으로 SW산업 시장이 2016년 기준 약 1조 966억 달러의 시장 규모를 형성하고 있으며, 전년대비 4.5%이상의 성장률을 보이고 있다. 세계 SW시장 추이를 보면 2020년까지 1조 3,114억 이상의 시장규모를 보유할 것으로 전망하고 있다(소프트웨어정책연구소, 2016). 한국의 SW시장의 규모는 16위 수준으로 나타난다.

[그림 2-1] 세계 SW산업 시장 규모 및 트렌드



출처: 소프트웨어정책연구소 (2016)

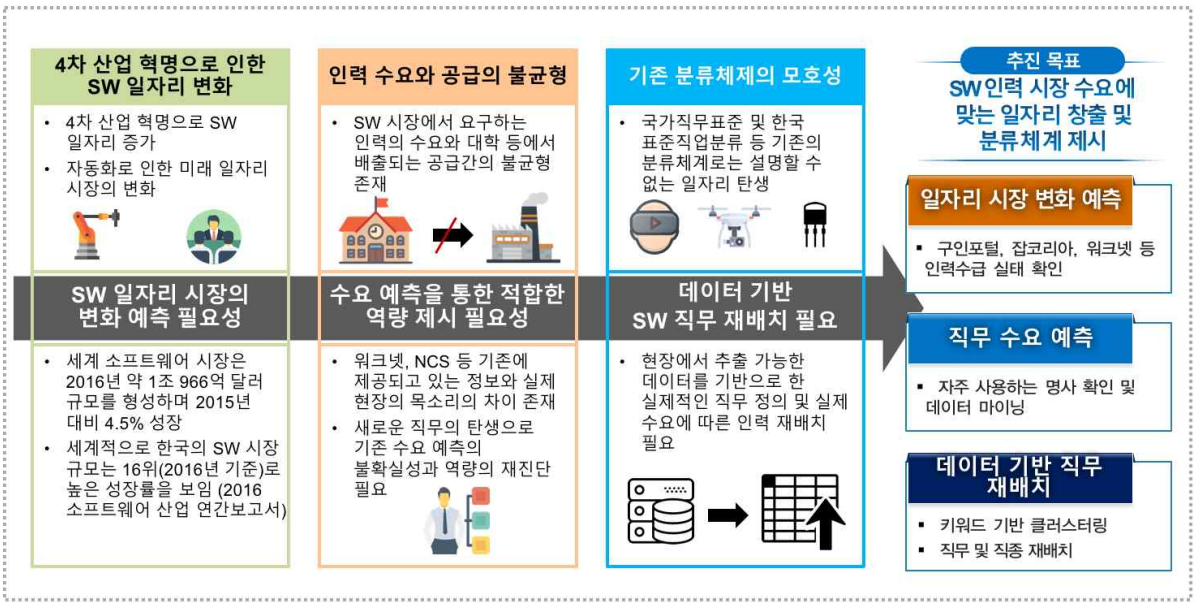
SW인력 시장의 변화를 준 요인 중 하나인 4차 산업 혁명은 공정의 자동화, 기계 학습(Machine Learning) 등의 기술 발달로 기존의 일자리가 없어지거나 새로운 일자리를 창출하여 직종의 변화를 가져왔다. 즉, 자동화가 쉽고 정형화(routinize)하기 쉬운 직종일수록 인력 수요가 줄어들고, 전반적으로 IT기술을 접목하기 어려운 직종에 대하여(예: 전문직, 기술직 등) 수요가 증가하였다(한국노동연구원, 2017). 따라서 현재 SW관련 산업은 큰 변화가 존재하며, 기존 직종 / 직업 표준에 부합하지 않는 직업들이 출현하고 있다.

SW인력 시장 수요에 부합하는 직종 및 직업에 대한 정의가 반영되어야 함에도 불구하고 현재의 워크넷, NCS 등 기존에 제공되고 있는 정보와 실제 현장의 직종별 요구사항은 다른 양상을 보인다. 시장 변화를 따라가지 못하는 기존 지표를 활용한 수요 예측의 불확실성이 증가하여 인력 수요와 공급의 불균형 또한 커지고 있다(최재원 외, 2017).

본 연구는 [그림 2-2]와 같이 기존 분류체제의 모호성을 극복하고, 빅데이터 기반으로 분석 모델링을 구축하였다. 따라서 인력 수요와 공급의 불균형을 타파하고 SW인력 시장의 실제적인 수요에 맞는 일자리를 파악 및 예측하는 것을 목표로 하였다. 이를 위해 일자리의 변화 예측에 필요한 기존의 일자리 수요 데이터를 수집하고, 자주 사용하는 명사 데이터를 파싱하여 데이터 마이닝을 실시하였다. 이후에는 실제적 데이터를 기반으로 SW분야 키워드를 추출하여 키워드 기반 클러스터링을 실시하였다. 클러스터링을 기반으로 새롭게 구성된 그룹의 직업을 재배치를 함으로써 인력 수요와 공급의 불균형을 해소하는 작업을 진행하여 새로운 SW분야 일자리 분류체계를 구성하고자 하였다.

2절에서는 기술적인 분석에 앞서, SW분야의 거시 환경의 특성을 확인하기 위해 PEST분석을 통한 경제적, 사회적, 기술적, 환경적 측면을 분석하였다. 특히 각 특성에 대한 전반적 특성에 대한 SW분야별 특징들을 배치하였다.

[그림 2-2] SW인력 시장에 맞는 인력 수요 예측과 새로운 직종 및 직업 분류체계 제시



제2절 거시환경 분석(PEST)

1. 정치적(Political)

2014년 정부는 디지털 콘텐츠 관련 업체의 창업과 성장 지원을 위해 총 4,000억 원을 투자하여 디지털 콘텐츠 운영 지원을 시작하였다. 현재 정부는 가상훈련 시스템, 콘텐츠 제작지원, 가상현실 기술의 R&D 세액공제, 가상현실 전문 펀드 조성, 서울 VR 페스티벌과 평창 동계올림픽을 통한 글로벌 홍보 등의 다양한 계획을 추진하고 있으며(과학기술정보통신부, 2016), 이를 통해 디지털 콘텐츠 관련된 분야의 제도적 지원을 확인할 수 있었다.

[그림 2-3]과 같이 정부는 2016년에서 2017년까지 가상현실 전문 펀드를 조성하기 위해 총 400억 원의 투자를 결정하였다. 그에 따라 정부는 가상현실(VR)과 증강현실(AR) 분야의 게임을 시작으로 테마파크, 교육 등 다양한 분야의 원천기술과 비즈니스 모델(BM) 등을 확보하고 있는 기업에게 제품개발 및 마케팅 등 전반적인 자금 투자 계획을 추진하고 있다. 또한 2017년 가상 증강현실 사업과 디지털 콘텐츠 기업 창업에 집중 투자를 위해 360억 원 규모의 가상현실 및 증강현실 펀드 및 창업초기 펀드의 결성과 운용을

결정하였다. 가상현실과 증강현실 관련 중소기업과 프로젝트에 총액의 60% 이상 투자할 예정이며, 창업초기 기업 및 프로젝트의 초기에는 약정 총액의 70% 이상의 예산을 투자할 예정에 있다(미래창조과학부, 2016).

[그림 2-3] 디지털 콘텐츠 펀드 정부출자 및 민간자금 규모



출처: 미래창조과학부 (2016)

4차 산업 혁명위원회의 논의를 통해 <표 2-1>과 같이 지능화를 기반으로 글로벌 경쟁력과 산업의 생산성을 확보하기 위한 '혁신 성장을 위한 사람 중심 4차 산업혁명 대응 계획'을 발표했다. 연구자를 중심으로 한 R&D 체계 혁신을 위해 총 22조원을 지능화 기술 R&D에 투자하고, 2019년 3월 5G 조기 상용화와 주요 산업별 빅데이터 전문센터 육성 등 각 분야별 규제제도 전면 재설계를 위해 혁신 모험펀드 10조원 조성을 확정하였다. 이를 통해 정부는 초연결 지능형 네트워크를 구축과 데이터 생산 공유 기반을 강화할 예정이며, 미래사회 변화에 대응하기 위한 핵심인재 14.6만 명 양성할 계획이다. 정부는 의료에서 제조업 등 다양한 분야의 산업을 2022년까지 지원할 예정이다(산업혁명위원회, 2017). 상기 사업을 추진함에 따라 의료 및 제조, 교통 등 다양한 분야에서 4차 산업에 맞춘 고도화된 기술력을 확보할 수 있는 환경을 조성하고 있다.

<표 2-1> 산업별 지능화 혁신 목표

(의료) AI 기반 신약개발 혁신, 진료정보의 전자교류 확대 및 맞춤형 정밀 진단과 치료 확산 → 건강 수명 연장 및 보건산업 수출액 30% 증가	(시티) 자율제어 기반 지능형 스마트 홈 확산과 지속가능한 스마트시티 모델 구현, → 가정 내 생활혁명 실현 및 도시문제 해결
제조) 스마트공장 확산 및 지능형 협동로봇 개발, 제조 서비스화 → 제조 생산 제고	(교통) 교통사고 위험 예측, 지능형 신호등 확산, 예보 서비스 고도화 → 도심 교통 혼잡 10% 감소 및 교통사고 5% 감소
(이동체) 산업용 드론 육성 및 자율 운항 선박 도입 →드론 시장 20배 증가	(복지) 노인 치매 생활보조 혁신으로 간병, 간호 지원 로봇 도입, → 치매 예측도 18% 증가
(에너지) 전력효율화 스마트 그리드 전국 확산 및 온실가스 저감 고효율화 기술 개발 →일반주택 지능형 전력계량기 100% 보급	(환경) 미세먼지 정밀대응, IoT활용 환경감시, 수질 최적관리 스마트 상하수도 확산 → 세계 최고의 미세먼지 예보 체계 확보
(금융,물류) 핀테크 활성화와 화물처리 자동화 및 스카트 물류센터 확산, 스마트 항만 구축 →핀테크 시장 2배 확대 및 화물 처리속도 33% 증가	(안전) 노후 시설물 관리 스마트화, AI 기반 범죄분석 →범죄 검거율 90% 증가, 해양사고 30% 감소
(농수산업) 정밀재배 스마트팜과 양식장 확산, 파종 및 수확 로봇 개발 →양식생산량 25% 증가	국방) 지능형 국방 경계감시 적용 및 인공지능 기반 지능형 지휘체계 도입 → 경계 무인화율 25% 증가 및 병력지원 감소

출처: 산업혁명위원회, 2017

또한 정부는 <표2-2>와 같이 SW산업의 전문 인력 부족 문제 등을 해결하기 위해 SW사업 분할발주 제도와 SW품질성능 평가시험(BenchMark Test, BMT) 의무화 제도를 시행하였다. 국내 SW산업 생태계를 정상화하기 위한 분할발주는 3개의 공공정보화 사업에 대해 적용되며, ‘소프트웨어 용역 계약 특수조건’, ‘SW사업 개발 공정별 표준 산출물’, ‘계약금액 조정 가이드’ 등의 발주 절차와 규정을 마련하였다. 이를 바탕으로 우수한 품질과 기술을 갖춘 국내 SW관련 제품의 공공시장 진출을 위해 도입한 ‘SW BMT 의무화’는

2016년 1월부터 시행되었으며, 공공 정보시스템을 구축하기 위해 필요한 객관적인 SW 품질정보를 제공한다. SW 관련된 기술에 대한 인증을 통하여 기존보다 공인된 기술을 활용 가능하다.

〈표 2-2〉 분할발주 대상 시범사업 개요

사업명	수요기관	진행사항	예산액	사업내용
E-발주지원 통합 관리시스템 (2차)	조달청	사전교육 공개 중	33.9억 원	제안 요청서 및 제안서 작성, 종합정보 시스템의 고도화
보험 고객 정보 통합구축	우정사업 본부	조달 요청 준비 중	142억 원	우체국 보험고객의 정보 통합 및 관리시스템 구축
한국채택국제 회계기분 통합회계 시스템	대구도시 철도공사	조달 요청 준비 중	11.8억 원	일반 기업회계에서 한국채택과 국제 회계기준으로 전환을 위한 시스템 구축

출처: 미래창조과학부, 2016

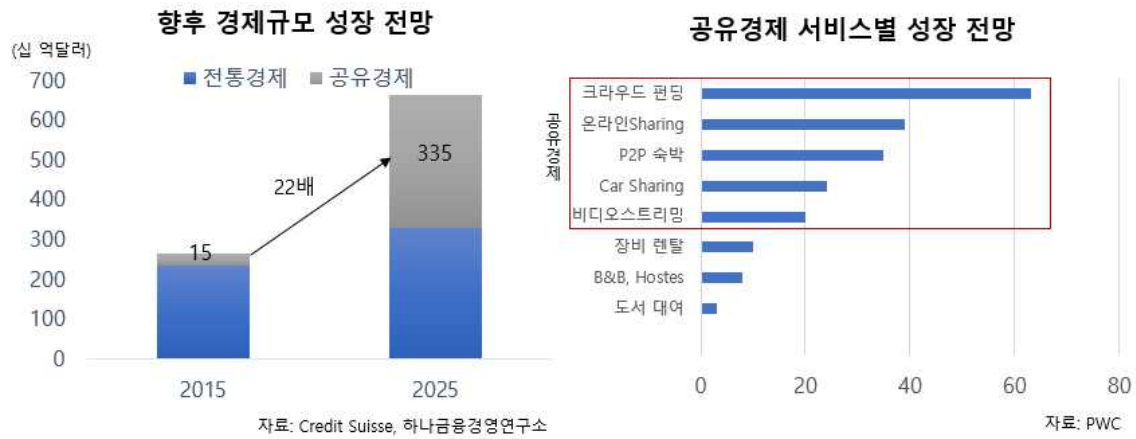
2. 경제적(Economical)

4차 산업혁명 사회로의 진입과 SNS 의 발전으로 온라인과 오프라인의 경계가 모호해지는 현상은 다양한 소비패턴의 등장을 촉진시켰다. 온라인으로 상품을 구매하기 전 오프라인 상점에서 상품을 먼저 체험한 후 구매를 결정하는 쇼루밍(showrooming), 온라인에서 상품을 파악하여 오프라인을 통해 구매하는 리버스 쇼루밍 (reverse-showrooming) 등 온라인과 오프라인을 연결하는 O2O(online to offline) 비즈니스 모델이 등장하였다(김대건, 2014). O2O 비즈니스 모델은 온라인과 오프라인의 연결을 통해 새로운 가치를 창출하고 ICT기술과 융합하여 오프라인의 고객을 온라인으로 유도한다. O2O서비스는 온라인과 오프라인 기업들의 커머스 채널 확장 유형과 플랫폼을 기반으로 한 플랫폼 사업자 비즈니스의 고도화로 구분할 수 있으며(황지현, 2015), 미래 결제 플랫폼을 통해 소비문화의 트랜드를 제시한다는 점에서 경제적 관점의 변화가 두드러지고 있다.

특히 소비자 기반 경제 환경의 변화는 소유에 대한 사회구성원의 개념이 변화하면서 소유하지 않고 공유를 추구하는 셰어리즘(sharism), 타인에 대한 기여를 고려하는 도네이즘(donaism), 타인에게 빌려서 문제를 해결하는 렌탈리즘(rentalism) 등의 공유경제가 자리 잡게 하는 원동력이 되었다(김난도 외, 2013).

공유경제는 재화, 시간, 서비스 등의 공유 방식에 따라 구분할 수 있으며, 협력적 소비의 의미와 융합되며 다양한 형태의 공유를 각각의 고유의 의미로 확장되고 있다. 특히 SNS를 기반으로 다양한 형태로 확장되고 있는 상황에서 개인적 공유로부터 사회 전반의 모든 영역으로 확대되어 가고 있다. 다시 말해서 공유경제는 지식과 경험의 공유를 통해 금전적 보상이 아닌 비금전적 교환이 이루어지는 경제적 특성을 가지고 있다. 기존에 구매를 미루던 소비자의 적극적 거래 참여로 신규시장이 창출되고, 산업화 경제시스템의 비효율을 개선하거나 유효자원을 통해 거래수익을 창출하는 등 지속으로 성장하고 있다. 특히 [그림 2-4]와 같이 공유경제로인해 온라인 직원채용, 크라우드펀딩, 카셰어링 서비스, P2P 숙박 등은 2015년 150억 달러에서 2025년 3,350억 달러로 약 20배 이상 확대되고, 일자리 및 산업 행태의 큰 변화로서 영향력을 발휘할 것으로 판단된다(Credit Suisse, 2014).

[그림 2-4] 공유경제 경제규모 및 서비스별 성장 전망



출처: 하나금융경영연구소 (2015) 및 PWC (2015)

기존의 제조업 기술과 3D프린터, IoT(Internet of Things)로봇과 빅데이터, AI(Artificial Intelligent) 등은 ICT 기술과 융합한 4차 산업혁명의 기술혁신이다. <표 2-3>과 같이 제조업은 제조경쟁력의 핵심이 되는 부품, 특허 등의 독점이 어려워지고 있다. 특히 낮은 인건비로 가격경쟁력을 갖춘 동남아시아와 중국 기업의 압박으로 선진국 간의 제조기업간 경쟁이 심화되고 있다. 독일은 4차 산업혁명을 위해 Industry 4.0을 도입하고 스마트 팩토리 구축에 집중하고 있다. 스마트 팩토리는 사물인터넷이 산업형장에서 진행된 지능화된 공장을 의미한다. 제조업은 스마트화, 개성화, 맞춤화에 대한 연구 개발을 시작으로 기존 B2C에서 C2B 비즈니스 모델로 전환하여 수요에 따른 주문제작의 산업형태로 변화할 것이다.

<표 2-3> 제조업의 혁신 단계 비교

구분	1차 산업혁명	2차 산업혁명	3차 산업혁명	4차 산업혁명
시기	18세기 후반	20세기 초반	1970년대 이후	2020년 이후
혁신	증기의 동력화	전력 노동 분업	ICT 혁명, 전자기기	제조업과 ICT의 융합
커뮤니케이션	신문, 책 등	TV, 전화기 등	인터넷, SNS	IoT, IoS
생산방식	생산 기계화	대량 생산	부분 자동화	시뮬레이션 방식의 자동 생산
생산통제	사람	사람	사람	기계

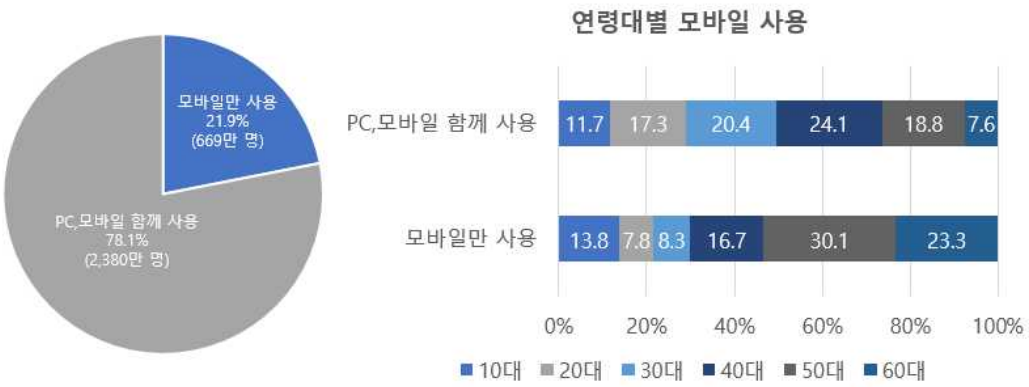
출처 : 현대경제연구원, 2013

3. 사회문화적(Socio-cultural)

국내 스마트폰 보급률이 90%를 넘어서면서 모바일 사용량이 증가하고 있다. 모바일의 빠른 확산으로 전자상거래, 영화 소비 등의 일상생활부터 비즈니스까지 모바일로 해결하는 모바일 온리(Mobile only) 집단이 등장하였다. [그림 2-5]와 같이 모바일 온리 집단의 절반 이상은 50-60대의 중장년층으로,

고연령층을 중심으로 화면이 큰 태블릿의 증가와 기술의 중간단계를 건너뛰고 적응하는 ‘립프로깅(Leapfrogging)’ 현상이 확산되고 있다. 모바일 온리 집단의 등장은 자신의 개성을 강하게 표출하는 개인이 소비의 주체로 대변되며 ‘1인 미디어’와 ‘브랜드의 몰락’ 등의 키워드가 주목받고 있다.

[그림 2-5] 모바일 온리 집단 현황



출처 : 닐슨코리아, 2017

스마트폰과 인터넷이 확산되면서 디지털 미디어 환경에서 방송 제작부터 촬영과 유통까지 스스로 진행하는 1인 미디어(Personal Media)가 도래하고 있다. 국내 1인 미디어 시장은 해외 국가들보다 확산 시기가 상대적으로 늦었다. 그러나 1인 가구의 증가와 모바일 영상 콘텐츠 시청자의 증가로 미디어 산업의 확산을 가속화하고 있으며, 스마트 플랫폼의 진화로 누구나 영상콘텐츠의 생산과 가공이 수월해지면서 최근의 1인 방송형태로 진화하고 있다. 친밀한 쌍방향 통신을 바탕으로 참여와 소통에 대한 욕구를 충족시키는 1인 미디어는 수평적인 연결 구조를 통해 사용자와의 관계를 강화하고, 접근성을 높여 더욱 개방적, 참여적 미디어 환경을 구현하고 있다.

스마트 기기의 보급과 자신의 개성을 중시하는 사회적 분위기가 자리 잡으면서 현재는 개인의 욕구를 가장 잘 보여줄 수 있는 모바일 시대가 도래하였다. 모바일 시대의 소비자는 개인화된 소비 패턴을 보이며, 소비자가 자신의 보유 데이터를 직접 비교 분석하는 등 능동적인 모바일 시장의 변화가 진행되고 있다. 최근에는 이러한 상호작용으로 브랜드에 의존하는 시대를 넘어

제품과 서비스를 소비자의 취향과 요구에 따라 제공하는 커스터마이징 서비스가 활발하게 증가하고 있다. 커스터마이징 서비스는 빅데이터를 기반으로 한 마케팅을 통해 소바자의 라이프스타일에 밀접하게 적용되고 있으며, 의학, 수공예, IT솔루션, 게임 등 다양한 분야에 이용되고 있다. 따라서 개인의 생활 및 삶의 방식을 바탕으로 다양한 직업에 대한 수요가 예측 불가할 정도로 변화되고 있다.

4. 기술적(Technological)

가상현실(VR)과 증강현실(AR) 시장은 빠르게 성장하는 시장으로 엔터테인먼트, 의료, 항공, 제조, 교육 등 다양한 산업에서 활용되고 있다. <표 2-4>는 산업별 AR과 VR 기술이 어떻게 적용되고 있는지에 대한 예시이다.

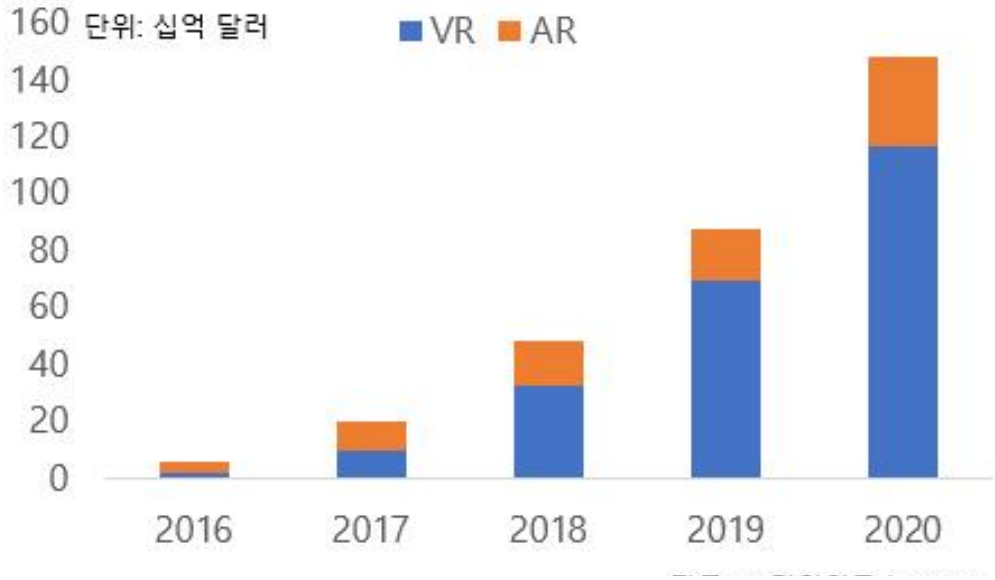
<표 2-4> 산업별 AR과 VR 기술의 적용

분야	내용
엔터테인먼트	현장감이 극대화 된 시공간적 한계를 뛰어넘는 게임이나 공연 등과 같은 체험형 콘텐츠 개발
의료, 항공, 군사	극한 상황을 대비하기 위한 반복적인 훈련 지원 및 업무 기획 수립
산업 (설계,제조)	즉각적인 설계와 도면 수정, 가상 시현 등 작업 방식의 효율성 증가
교육, 훈련	화학분자 설계, 우주탐험 등 위험비용이 많이 드는 교육이나 훈련 대체

출처 : KT경영연구소, 2016

(VR/AR) [그림 2-6]과 같이 VR과 AR은 일상생활과 다양한 산업에서 활용되면서 국내외 경제 사회에 대한 변화와 새로운 부가가치를 창출할 것으로 기대된다. 특히 페이스북, 삼성, 구글 등 ICT기업들은 VR과 AR 산업 육성 및 시장 진출을 추진하고 있으며, VR과 AR 시장 규모는 2020년까지 확대될 것으로 보인다.

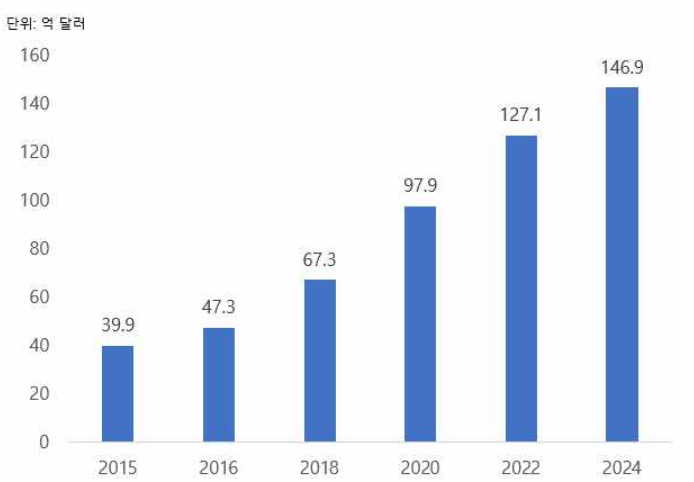
[그림 2-6] VR/AR 시장 규모 전망



출처 : KT경영연구소, 2016

(드론) 최근 드론은 군사용을 벗어나 다양한 분야에서 활용가능성이 높아지면서, 일반 사용자 시장으로 확산되고 있다. [그림 2-7]과 같이 드론은 제조와 서비스의 융합모델로 주목받아왔다. 특히 드론은 정찰과 감시 목적의 군사적 용도를 넘어, 물류운송과 농업, 정보통신 등의 다양한 산업 분야에서 활용되고 있다. 미국 방산업 컨설팅 업체인 틸그룹은 세계 드론 시장이 2024 약 147억 달러 규모의 산업 규모로 급증할 것으로 예상하고 있다.

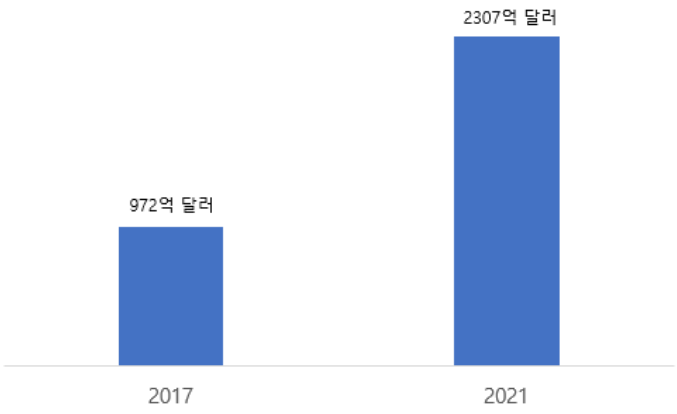
[그림 2-7] 세계 드론 시장규모



출처: 틸그룹, 2016

(로봇) 로봇은 제조업의 중심인 자동차와 정보통신(IT), 조선 산업 등의 기계화 자동화를 통해 산업성장과 함께 발전하고 있다. [그림 2-8]과 같이 국내 로봇시장은 2007년 이후 연평균 11%의 성장세를 보이며 생산성 향상을 이끌고 있다(산업통상자원부, 2017). 올해 세계 로봇시장은 2016년 보다 17.9% 증가하였으며, 연평균 22.8% 성장률을 보일 것으로 기대하고 있다.

[그림 2-8] 세계 로봇 구매 시장

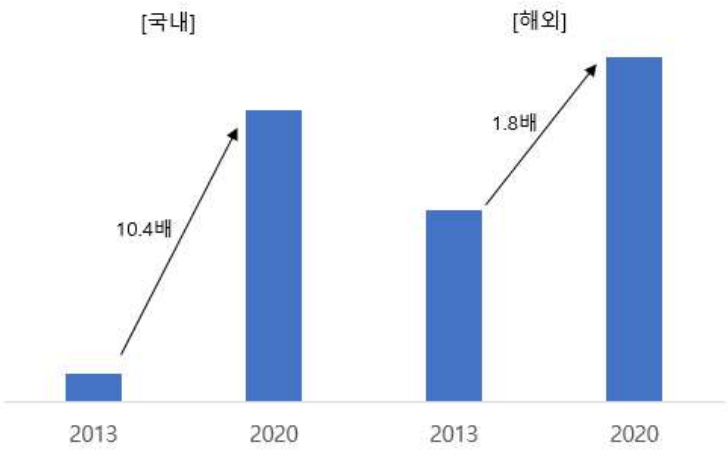


출처 : IDC 2017

로봇 시장은 ICT의 확대로 스마트공장을 시작으로 국방, 항공 등 다양한 분야에서 네트워크 기반의 로봇 서비스를 필요로 하고 있다. 특히, IoT기술 개발이 빨라지면서 전문, 상업용 서비스 분야를 시작으로 물류, 의료, 접객 등으로 다양한 직업이 요구 되고 있다.

(사물인터넷) [그림 2-9]와 같이 전자기기들의 세계적 증가와 온라인 연결의 증가는 인터넷을 기반으로 사물과 사물을 연결해주는 사물인터넷 시장의 성장이 지속될 것을 증명해준다(미래창조과학부, 2013). 커넥티드 홈, 스마트 미터, 커넥티드 자동차, 스마트 그리드 등 사물인터넷 시장이 성장하면서 관련 기업들은 사물인터넷을 이용한 플랫폼과 서비스 개발을 진행하고 있다. 최근 사물인터넷은 생활가전을 시작으로 제조산업, 유틸리티, 자동차, 보안 등 다양한 분야에서 개발되고 있으며, '소프트웨어 플랫폼'이나 '보안' 및 '데이터 분석' 같은 소프트웨어적인 요소가 점차 중요해지고 있다. 따라서 신규 확산되고 있는 사물인터넷 사업 분야도 다양한 신규 일자리 및 기술적 응용에 의한 직종 환경의 변화에 중심이 되고 있다.

[그림 2-9] 사물인터넷 시장 규모



출처 : 미래창조과학부, 2013

3절에서는 거시환경분석을 바탕으로 SW산업 분야의 데이터 분석을 위해 사용한 분석기법을 상세히 설명하였다. 특히 본 연구의 핵심인 토픽모델링 방법론을 비교 분석하여 본 연구의 분석 기법에 대한 타당성을 제시하였다.

제3절 텍스트마이닝 분석 모델

본 연구는 SW인력 수요 파악을 위하여 채용 정보 및 관련 데이터를 수집한 후 텍스트를 기반으로 하는 다이나믹토픽모델링(Dynamic Topic Modeling: 이하 DTM) 분석 방법을 활용하였다. 토픽모델링은 각각의 문서나 데이터에 있는 주제(Topic)를 확률적 혼합체로 표시하고, 문자 정보에 내포하고 있는 주제들을 찾기 위하여 개발된 통계적 분석 알고리즘이다(안정국 외, 2016).

텍스트마이닝을 위한 접근 방법으로는 두 가지 방법이 존재한다. 첫째, 규칙 기반 접근법으로 언제든 규칙을 적용할 수 있다는 장점이 존재하며, 정확도가 높지만, 데이터가 바뀔 때 마다 새롭게 규칙을 다시 지정해 주어야 때문에 지속적으로 변화하는 인력 수요에 맞게 분석하기에 어려움이 있다. 둘째, 통계 기반 접근법으로 초기의 학습해야할 데이터가 적으면 데이터 자체의 신뢰도가 떨어지는 단점이 존재하지만, 데이터의 양이 많아질수록 데이터의 예측률이 증가한다. 최근 빅데이터를 쉽게 확보할 수 있게 되면서 예측률을 올린 통계 기반의 텍스트마이닝이 널리 사용되고 있으며, 특히 기계학습(Machine learning)을 활용한 데이터 마이닝이 다양한 분야에서 활용되고 있다(김윤성 외, 2017). 본 연구는 잠재 디리클레 할당(Latent Dirichlet Allocation: 이하 LDA)에 대한 설명을 바탕으로 LDA를 기반으로 발전된 DTM을 활용한 분석 기법에 대한 타당성을 확보하기 위하여 관련 선행연구들을 확인하였다.

1. 잠재 디리클레 할당(LDA)

현재 텍스트 마이닝에서 많이 활용되고 있는 방법은 바로 잠재 디리클레 할당(Latent Dirichlet Allocation)이다. LDA와 유사한 선행 분석기법을 포함하여 주로 사용되어 온 알고리즘을 비교하면 <표 2-5>와 같다. 기존에 활용되어 온 알고리즘은 LSA(Latent Semantic Analysis: 이하 LSA)로 단어의 형태가 아닌 의미를 이용하여 동시 발생(Co-occurrence)의 내용을 유추하고 의미를 도출한다. 그러나 LSA의 경우 단어의 의미가 하나가 아닌 여러 의미로 해석되는 경우에 활용하는 의미 기반의 분석이기 때문에 본 연구의 분석에 활용이 불가능하다. 또한 다의어를 보완하고자 보완된 알고리즘인 PLSA(Probabilistic Latent Semantic Analysis: 이하 PLSA)는 기존의 LSA가 가지고 있던 다의어 문제를

해결할 수 있다. 그러나 토픽 모델링 과정에서 과적합성과 문헌에 대한 학습이 선행되어야 분석이 진행 가능하다는 점에서 한계가 존재한다. 따라서 본 연구의 데이터 속성과 가장 부합한 분석 방법으로는 LDA가 가장 적합하다.

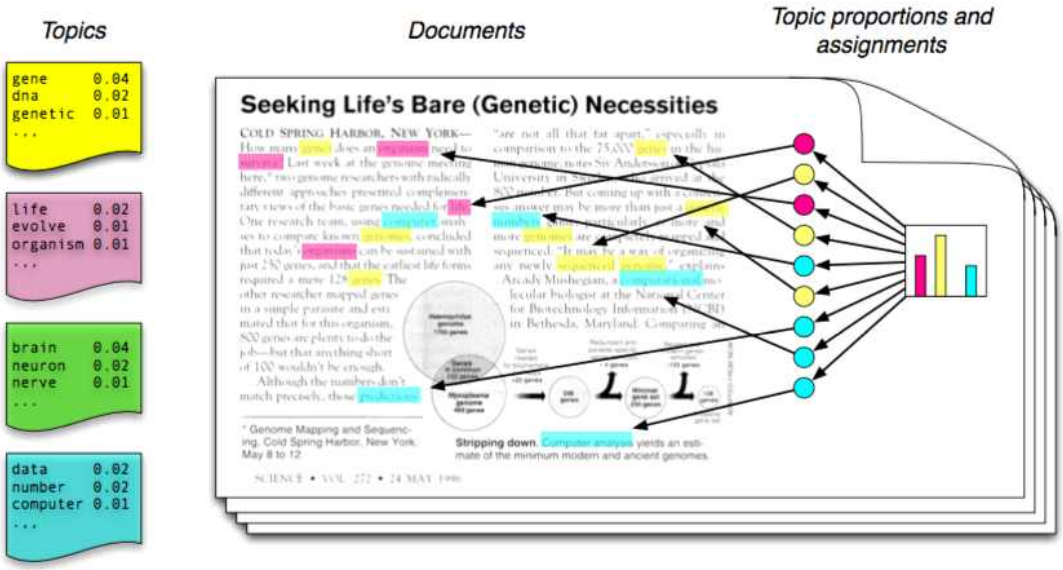
<표 2-5> 토픽 모델링 알고리즘 방법론 비교

	장점	단점
Latent Semantic Analysis(LSA)	• 차원 축소로 노이즈 제거 가능 • 토픽의 동의어 문제 등의 해결 가능	• 토픽에 다의어 존재시 문제 • 문맥이 전혀 고려되지 않음 (문서 Vector 표시) • 결과 해석이 어려운 편임
Probabilistic Latent Semantic Analysis(PLSA)	• 토픽의 다의어 문제 해결 가능	• 학습 전인 문서에 대한 일반화 불가능 • 과적합성 문제
Latent Dirichlet Allocation(LDA)	• 학습 전인 문서에 대한 일반화 • 과적합성 문제 해결 가능	• 파라미터 값의 사전 설정 필요(어휘 사전)

LDA의 분석 구조는 [그림 2-10]과 같다. 각 문헌들의 주제는 소수의 단어들의 혼합체로 구성되고, 각 단어들의 출현은 문헌들의 주제에 의하여 달라진다는 전제하에 분석이 수행된다(Blei et al., 2010). 다시 말해서 LDA는 한 문헌에 들어있는 동일한 주제에 대한 구성 단어들을 확률적으로 도출시켜 집합을 구성하게 하는 분석기법이다. 추가적으로 단어들이 모두 독립적으로 구성되어 있지 않다는 전제를 함께 응용함으로써 단어 생성의 조건에 따라 확률 추론이 가능하다는 점이 주요 특징이다(Blei et al., 2010)

LDA 알고리즘의 원리는 다음과 같다. [그림 2-10]을 보면 이 기사에서 “Seeking Life’s BARE(Genetic Necessities)”의 기사를 대상으로 LDA 원리를 설명하고 있다(Blei et al., 2010). 노란색 문서에 들어있는 주제는 유전에 대한 키워드를, 분홍색은 생명과 유기체, 파란색은 컴퓨터와 관련된 주제의 키워드, 초록색은 뇌와 관련된 주제의 단어들이 정리되어 있다. 정리하면 LDA는 기존의 주제들의 모음이 있다고 가정한 후에 주제와 관련된 명사들을 추려 내어 각자의 해당하는 주제 문서로 추출하는 분석 기법이다.

[그림 2-10] LDA의 원리



출처 : Blei et al., 2010

우선 분석을 위해 선행적으로 고정된 어휘를 통하여 주제를 공식적으로 정의해야한다. 즉, 주제에 대한 사전을 구성하여 주제별 사전에 유사성이 있는 명사를 추출 할 수 있도록 전처리 작업이 선행되어야 한다. 통계적 추론을 통하여 [그림 2-10]과 같이 자주 등장하는 주제별 15개의 단어를 추출하여 표현 가능하다(Blei et al., 2010). 상기 원리에 대한 확률적 추론 계산의 원리는 다음과 같다.

LDA 추론 계산식:

$$p(\beta 1:K,\theta 1:D,z1:D,w1:D)$$
$$=\prod_{i=1}^{K}p(\beta i)\prod_{d=1}^{D}p(\theta d)(\prod_{n=1}^{N}p(zd,n|\theta d)p(wd,n|\beta 1:K,zd,n).$$

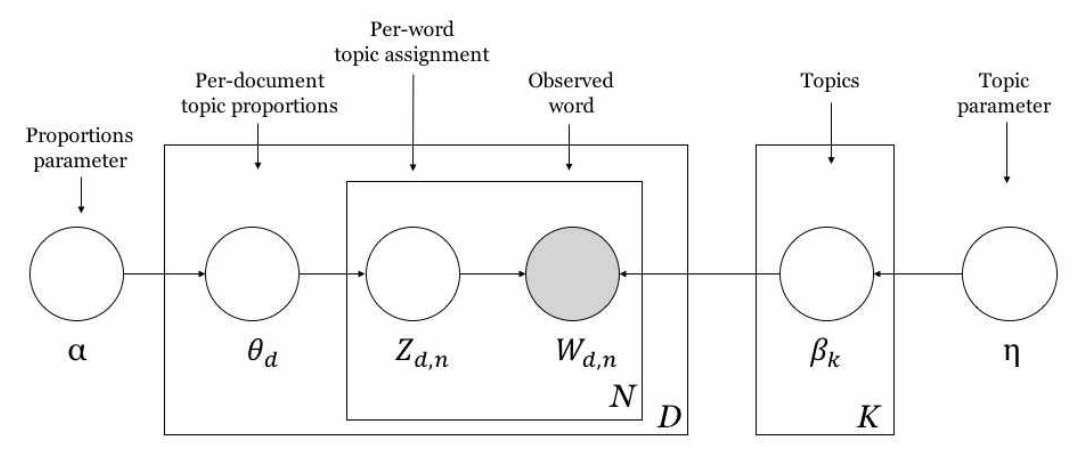
[그림 2-11] 주제별 상위 15개의 키워드 추출 결과

"Genetics"	"Evolution"	"Disease"	"Computers"
human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations

여기서, $\beta 1$은 주제이며, 각 βk는 어휘 즉, 단어에 대한 분포이다. d번째 문헌의 토픽 비율은 θd이며, 여기서 k는 토픽 k에 대한 비율이다. d번째 문서에 대한 주제 할당은 zd이며, zd의 n은 d의 n번째 단어에 대한 주제에 할당 값이다. 따라서 문서 d에 대해 관찰된 단어는 wd이며, 여기서 wd의 n은 어휘 사전의 문서 d의 n번째 단어이다.

LDA의 그래픽 모델은 [그림 2-12]와 같다. 각 노드는 랜덤 변수이며, 생성된 프로세스에 따라 레이블이 각각 지정된다. 즉, 관찰된 노드-문서의 단어로 처리되며, 직사각형은 '플레이트'로 복제된 단어를 나타낸다. N 플레이트는 문헌 내에서의 수집된 단어를 나타낸다. D 플레이트는 수집된 단어 내의 문서 모음들을 나타낸다.

[그림 2-12] LDA 그래픽 모델



출처 : 안정국 외, 2016

K : 토픽(주제)의 개수

a : 문헌별 주제 k의 디리클레 선행 가중치, θ 값을 결정하는 parameter

η : 주제별 어휘 w의 디리클레 선행 가중치, β 값을 결정하는 parameter

θ_d : 문헌별 주제의 비율

β_k : 주제별 어휘 w의 생성확률

$Z_{d,n}$: 문헌 d의 n번째 어휘의 주제(index)

$W_{d,n}$: 문헌 d의 n번째 어휘(문헌에서 관측되는 변수, index)

LDA의 알고리즘은 문헌상에서 관측되는 변수들을 통하여 보이지 않는 요인을 추론하고, 숨겨져 있는 토픽들을 찾아내는 측면에서 그 목적이 존재한다. 따라서 신문기사, 도서, 논문 등 방대한 양의 텍스트 문서들을 기반으로 의미 있는 정보를 발견하는데 탁월한 분석 기법이다(홍성연, 최재원 2017). LDA 알고리즘은 학습되지 않은 문헌에서 일반화가 가능하며 과적합의 문제를 해결할 수 있어 기존의 텍스트마이닝 기법인 LSA의 단점을 보완한다. LDA 알고리즘의 생성과정은 <표 2-1>과 같다(박대한, 2017).

<표 2-1> LDA 알고리즘 생성과정

1. 각 주제별 :
   1) 단어 위에 분포 그리기: $\phi k \sim Dir(\beta),\ k \in 1, \ldots, K$
      $Dir(\bullet)$: uniform Dirichlet distribution
      β : parameter

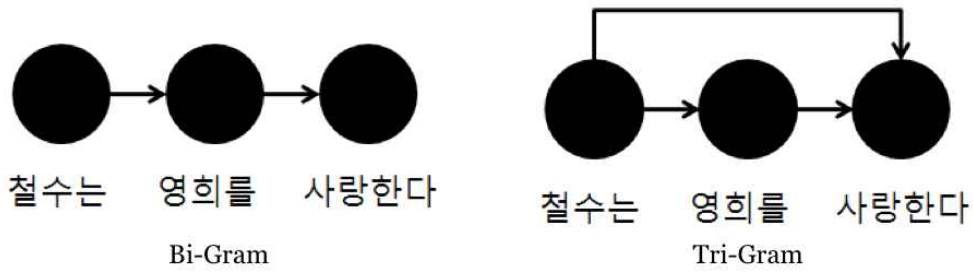
2. 각 문서 :
   1) 주제 비율의 벡터 그리기: $\theta d \sim Dir(\alpha),\ d \in 1, \ldots, D$
      α : parameter
   2) 각 단어에 따라 :
      (1) 토픽 할당 : $zd,v \sim Mult(\theta d),\ d \in 1,\ldots,D,\ v \in 1,\ldots,V$
          $Mult(\bullet)$: Multinomial distribution(ii)
      (2) 단어 그리기 : $wd,v \sim Mult\ \phi zd,n,\ d \in 1, \ldots, D,\ v \in 1, \ldots, V$

출처 : 박대한, 2017

한국어를 기반으로 한 토픽 모델링 알고리즘을 적용하기 위해서는 형태소 분석이 필요하다. 형태소 분석이란 문서의 대상의 어절의 가능한 모든 분석 결과를 도출하는 분석을 의미한다. 한글은 하나 이상의 형태소로 구성되어 있어 같은 형태로 구성된 형태소라 할지라도 결합의 방식 등의 차이에 따라서 의미가 달라진다. 기존에는 상기 문제를 해결하기 위하여 '선택적 Bigram' 모델을 활용하여 분석을 수행해왔다(허성완, 2015).

Bigram의 개념은 n-gram에서 둘이라는 의미의 'bi-'가 붙어 한 단어가 나타날 확률에 있어 선행적으로 나타난 단어에 영향을 받는 것을 가정한다. 예를 들어 [그림 2-13]과 같이 '철수는 영희를 사랑한다.'라는 문장이 있을 경우 '철수는' 어절은 '영희를' 에 영향을 주는 선행 단어이며, '영희를' 은 '사랑한다.' 에 영향을 주는 선행 단어이다. 이러한 방식은 1차 마코프 연쇄(First-order Markov chain)이라 지칭한다. 최근에는 Bi-gram 보다 좀 더 확장된 개념의 'Tri-gram' 을 활용하여 LDA 분석이 수행된다. Tri-gram은 바로 선행 단어뿐만 아니라 선행 단어의 선행된 단어까지 영향을 준다. 3개 수준의 단어에 대한 관계적 확률 값을 측정하는 것으로 이를 2차 마코프 연쇄(Second-order Markov Chain)라고 지칭한다(허성완, 2015).

[그림 2-13] Bi-gram과 Tri-gram의 차이



출처 : 허성완 2015

LDA 분석은 알고리즘 자체의 단순성과 데이터 차원 축소의 유용성을 갖게 되며 의미상으로 일관적인 주제 생산이 가능하다는 점에 분석의 의의가 있다. 특히 소셜 미디어, 저널, 신문 기사 등의 다양한 데이터 자원들을 n-gram에 맞춰 알고리즘을 적용시켜 주제를 도출함으로써 데이터 안에 있는 숨겨진 의미를 도출해 내고 다양한 의미를 이끌어 낼 수 있다. LDA 알고리즘의 이론적 내용을 기반으로 분야별 활용 사례를 소개하면 다음과 같다.

(핀테크) 최근 비트코인 등 암호화 화폐 기술의 발전으로 금융권의 판국은 직종의 변화만큼 많은 변화를 겪고 있다. 핀테크는 기술과 금융의 융합을 상징하며, 정보 기술을 활용하여 다양한 금융 상품의 관리, 판매 등을 더욱더 편리하게 이용할 수 있도록 하는 기술이다. 핀테크와 관련된 선행연구에서는 미국, 한국, 중국의 특허 데이터를 수집하여 데이터 전처리를 실시하고 데이터 토픽모델링 분석을 통해 핀테크 관련 내용을 추출하여 동향을 확인하는 분석을 실시하였다. 본 연구는 분석을 시행하기 위하여 전처리 과정에서 대문자를 소문자로 바꾸고 특수문자, 조사 등의 용어들을 제거하는 작업을 실시하였으며, 상기 과정을 통하여 만들어진 키워드들을 바탕으로 토픽모델링을 수행한 결과 샘플링을 5,000번 반복하여 핀테크에서 확인 할 수 있는 총 20개의 주제를 도출하였다(김태경, 2016).

분석 결과 <표 2-7>과 같이 핀테크는 각 주제별 키워드를 통해 20개의 그룹화를 통하여 기술적 동향 파악이 가능하다. 한국의 인기 있는 토픽으로는 [주제 9] 인증/보안, [주제 13] 모바일 결제 기술이 추출되었고, 상대적으로 인기 없는 토픽으로는 [주제 20] 온라인 상거래, [주제 1] ATM 기술이 추출되었다. 따라서 최근 핀테크는 모바일 결제와 NFC, 인증/보안, 금융 데이터분석 등이 주목받고 있다.

<표 2-7> 핀테크 분야 LDA 분석 결과

	주제별 키워드
주제 1	transfer, bank, send, machine, automat, device, money, account, ATM, machine
주제 2	trade, price, exchange, product, rate, valu, order, market, trader, match
주제 3	modul, mobil, card, nfc, function. device, equip, phone, payment, chip
주제 4	cpo, identifi, , provid, control, financi, plan, manag, offer, accept, custom
주제 5	network, support, access, embodi, software, program, secur, modul, comput, onlin
주제 6	loan, amount, mortgage, lend, , home, rate, guarante, borrow, incom, benefit
주제 7	display, interfac, imag, plural, graphic, configur, view, screen, type, machin
주제 8	custom, provid, investment, analyze, bill, databas, servic, option, allocation, monetary
주제 9	authent,secur, password, biometr, verifi, devic, code, identif, disclos, encrypt
주제 10	payment, mobil, phone, termin, internet, sevic, deposit, bank, onlin, account
주제 11	seller, buyer, bid, auction, negotiation, tax, , commerc, purchas, online, merchant
주제 12	risk, financi, determin, predict, select, polici, model, estimate, forecast, portfolio
주제 13	mobil, payment, devic, payor, send, sms, contactless, wallet, payee, wireless,
주제 14	card, credit, purchas, merchant, provid, account, servic, issu, cardhold, check
주제 15	financi, process, currency, verify, datum, insurance, plan, analysis, database, memory,
주제 16	request, online, platform, server, access, function, protocol, control, service, environment
주제 17	devic, wireless, mobil, magnet, card, rfid, tag, vehicl, reader
주제 18	payment, account, receiv, deposit, electron, paye, check, request, bank, withdrawal
주제 19	invest, asset, fund, investor, stock, option, properti, manag, bond, compani
주제 20	store, onlin, shop, sale, vendor, purchas, product, phone, custom, marketplace

출처 : 김태경, 2016

(연구동향 분석) LDA 분석은 기존의 문헌을 기반으로 텍스트 마이닝을 통하여 관련 주제를 추출하고, 연구에 대한 트렌드 분석과 핵심 키워드 파악을 통해 연구 동향에 대한 분석으로도 많이 활용되고 있다. 전통적으로 기존 연구 동향은 메타분석을 많이 이용하였으나 LDA 분석은 텍스트를 기반으로 데이터를 확보하는 것이 용이하기 때문에 다양한 분야에서 활용되고 있다. 그 중 정보시스템 분야의 연구 동향은 국제 논문을 기반으로 실시간 검색 데이터베이스를 구축하여 1980~2015년까지 35년간의 정보시스템 연구 데이터를 수집하여 LDA를 진행한 연구가 존재하며, 본 연구에서는 <표 2-8>과 같이 정보시스템 토픽을 분석하여 약 10개의 토픽을 도출하였으며, 각각의 유용한 키워드를 추출하였다(안정국 외, 2016).

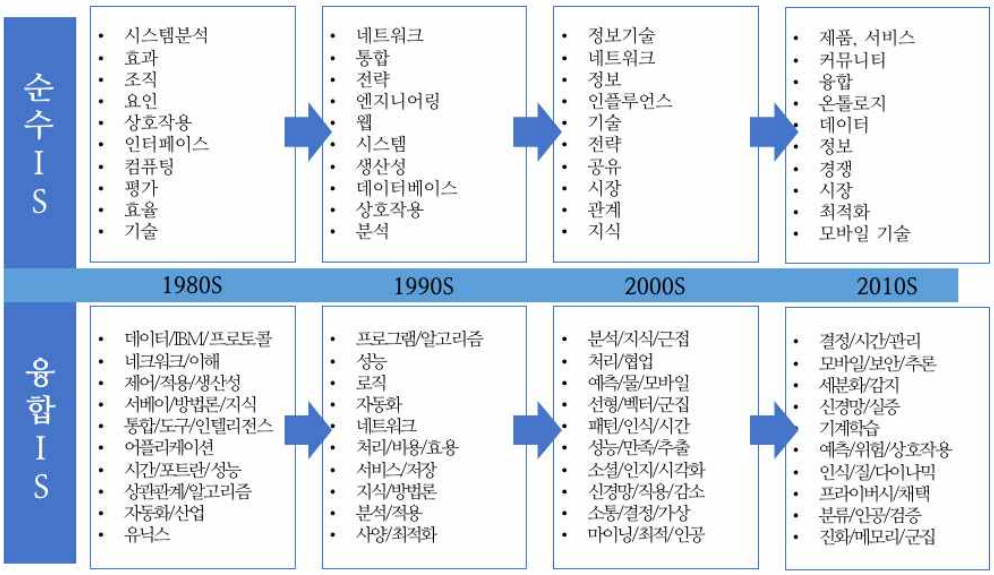
<표 2-8> 정보시스템 토픽(2010~2015) 샘플링

주제	라벨	키워드
02	커뮤니티	support, investigation, sharing, search, systems, governance, study, success, exploring, community, network, knowledge, framework, introduction, large, communities, risk
03	융합	multi, influence, issue, project, research, level, criteria, outsourcing, privacy, relational, structure, firm, projects, multiple, group, theoretical, dimensional, hybrid, perspective, high
04	온톨로지	based, decision, making, web, knowledge, content, role, time, cost, network, internet, ontology, real, usage, case, resource, event, distributed, effective, agent
05	데이터	data, model, system, online, theory, models, management, research, effects, performance
06	정보	information, empirical, strategy, efficient, perspective, adoption, approach, online,
07	경쟁	systems, approach, open, mining, management, supply, performance, digital, source,
08	시장	analysis, social, consumer, cross, firm, information, organizational, business, pricing,
09	최적화	information, systems, perspective, behavior, search, study, management, modeling,
10	모바일 기술	technology, user, effect, public, review, untitled, enterprise, virtual, impact, enabled, media, rfid, computer, trust

출처 : 안정국 외, 2016

35개년의 연구 내용을 분석한 결과에 따라 1980년대부터 2010년대까지의 동향은 [그림 2-14]와 같다. 1980년대의 경우 시스템 분석, 평가, 효율 등의 순수정보시스템의 내용과 네트워크/이해, 제어/적용/생산성 등에 대하여 주로 연구되었다. 2010년대에는 제품, 서비스, 커뮤니티, 최적화 등 순수정보시스템의 연구가 주로 진행되었으며, 융합정보시스템은 세분화/감지, 기계학습, 신경망/실증에 대한 연구가 주로 진행되었다(안정국 외, 2016).

[그림 2-14] 순수 정보시스템과 융합 정보시스템 토픽 모델링

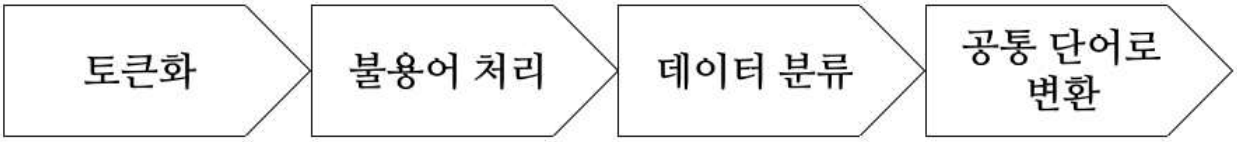


출처 : 안정국 외, 2016

상기 내용과 같이 수집된 텍스트 정보로부터 주요한 키워드를 추출하여 연도별 주요 이슈를 도출할 수 있으며, 이는 특정 분야의 핵심 주제를 파악하는데 유용하게 활용이 가능하다.

(산업공학) 국내 산업공학 연구동향 분석에서는 상위 25개 대학의 산업공학 관련학과의 전임 교수들이 논문을 가장 많이 게재하는 상위 5개의 국내 저널을 선정하여 총 3,003개의 논문을 데이터로 수집하여 분석을 실시하였다(김상겸, 2016). 일반적인 텍스트 분석의 데이터 전처리 과정은 [그림 2-15]와 같다.

[그림 2-15] LDA 데이터 전처리 과정



토큰화는 여러 문장에 존재하는 공백을 기준으로 분리하는 단계로 분석에 해당하지 않은 불용어를 제거한다. 그에 따라서 기본적인 데이터 셋을 <표 2-9>의 예시와 같이 구성할 수 있다. 명사 위주의 품사가 고려하여 의미적 정보를 가진 아젠다를 도출하고 동의어 및 다의어를 공통 단어와 다의어 처리를 통하여 데이터 전처리를 진행하였다. 빈도수가 1,000회 이상 나타난 키워드를 빈도 분석하여 정리한 결과 'system'이 가장 높은 빈도로 도출되었으며, 뒤로 'process'는 2,129회, 'service'는 1,950회의 순서로 도출되었다.

<표 2-9> 산업공학 분야 연구 동향 주요 키워드 빈도

단어	빈도	단어	빈도
System	3,114	Analysis	1,167
Process	2,129	Cost	1,126
Service	1,950	Company	1,117
Problem	1,750	Design	1,095
Performance	1,691	Customer	1,093
Information	1,621	Management	1,070
Time	1,416	User	1,041
Quality	1,403	Order	1,026
Product	1,366	Factor	1,003

출처 : 김상겸, 2016

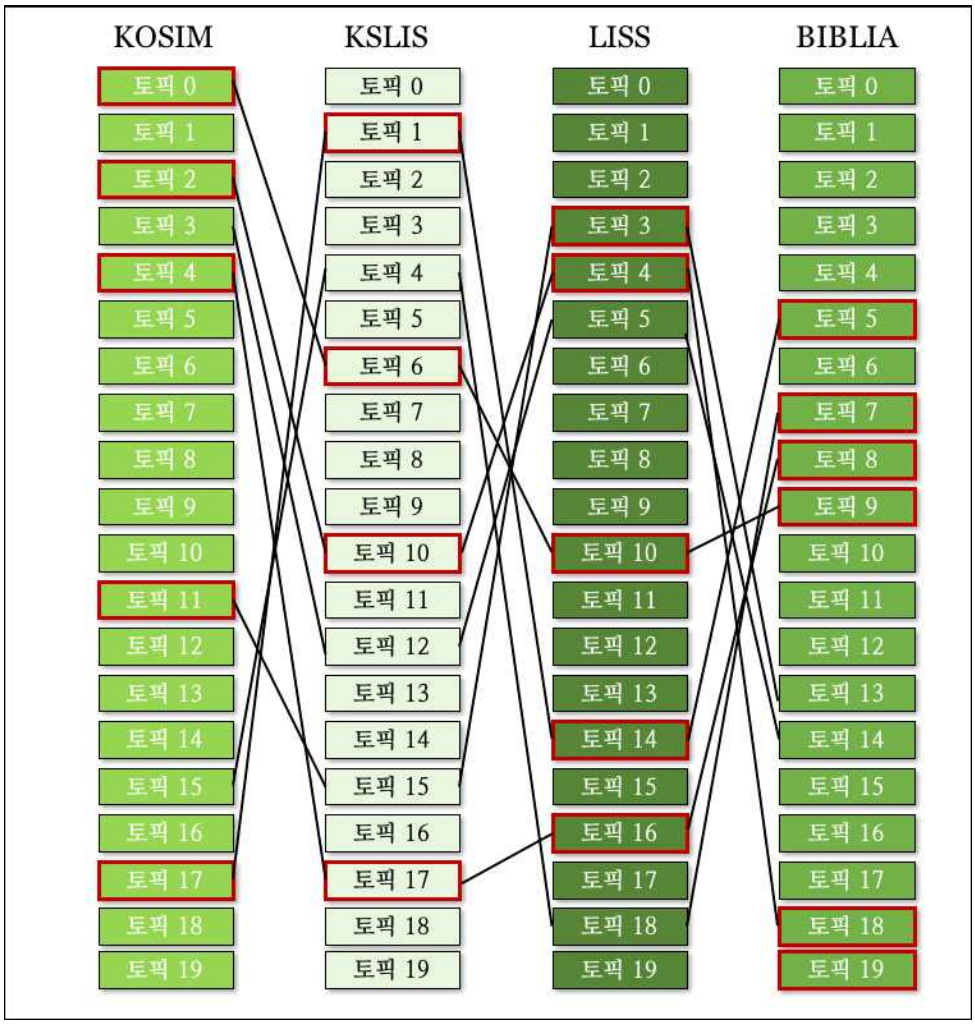
도출된 키워드를 기반으로 20개의 주제로 군집이 나눠졌다. 정의된 주제는 '공급망 관리', '공정관리', '시스템 모델링', '생산설비의 효율성', 'HW/SW 신뢰성 평가', '무선네트워크', '품질경영', '퍼지 이론', '물류네트워크', '기업성과 데이터마이닝', '컨테이너 터미널의 효율성',

‘고객 관계 관리’, ‘인간공학’, ‘최적화’, ‘교육 시스템’, ‘서비스 품질’, ‘금융공학’, ‘소셜 네트워크 분석’, ‘사용자 중심 추천 시스템’이 추출되었다. 또한, 도출된 토픽으로 추이를 볼 때, ‘품질 경영’이 2004년부터 2014년까지 가장 많은 점유율을 갖고 있음을 확인할 수 있다(김상겸, 2016).

(문헌정보학) 문헌정보학 분야에서도 LDA 분석은 연구 동향 분석으로 자주 사용된다. 기존의 선행연구들에서는 저자의 성향과 기준에 따라 분석하는 방식이나, 분류된 방법이 다르게 측정되어 신뢰성이 떨어졌다. 본 연구는 데이터를 기반으로 분석하는 LDA 분석을 실시함으로써 데이터의 신뢰성을 확보하고, 연구 동향을 파악하기 위한 확증적인 분석이라는 것을 증명하고 있다(박자현, 송민, 2013). 문헌정보학 연구에서는 논문 초록을 대상으로 텍스트 마이닝을 실시하였으며, 불용어(Stop word) 처리를 포함한 원 데이터의 전처리 과정을 거쳐 해당 키워드만 추출하도록 실시하였다. 추가적으로 문헌정보와 관련된 학술지별 논문 초록 텍스트 마이닝을 통하여 동시출현빈도 분석을 실시하여 어떤 키워드가 가장 높은 빈도가 있는 지를 확인하였으며, 총 논문 초록에서는 키워드를 대상으로 연도별 시계열 분석을 통하여 전체적인 연구동향을 파악하였다.

동시출현빈도 분석을 간단히 살펴보면 [그림 2-16]과 같다. 각 학술지에 기입된 초록의 키워드를 도출하여 분석한 결과 ‘도서관 서비스’, ‘정보학’, ‘자료조직’을 포함하는 해당하는 연구주제가 4개의 학술지에 공통적으로 도출된 것을 발견하였다. ‘도서관 서비스’ 영역의 도서관 유형별 서비스에 관한 연구주제는 ‘Library’, ‘Service’, ‘Public’, ‘University’ 등의 단어가 도출되어 정보관리학회지의 토픽2, 한국도서관·정보학회지 토픽4, 한국문헌정보학회지 토픽 10, 한국비블리아학회지 토픽18에 동시에 출현하였다. 이처럼 각 학술지에서 군집된 토픽들이 4개의 학술지에 동일하게 출현하는 것으로 보아 일정 수준 연관성이 있다는 것을 확인하였다. 따라서 선행연구를 통하여 4개 학술지에서 공통적으로 도출된 토픽의 경우 더 다양한 연구를 많이 수행하는 분야라는 것을 증명하고 있다.

[그림 2-16] 문헌정보학 관련 4개 학술지에서 도출된 토픽모델링 결과 비교 · 분석



출처 : 박자현, 송민, 2013

상기 언급된 다수의 연구 분야에서의 연구 동향을 기존의 메타 분석이나 다른 기준에 의거한 분석과는 다르게 순수하게 데이터를 기반으로 LDA 분석을 실시하였다. 따라서 LDA는 다양한 분야의 문헌들을 분석하는 분석도구로써 현재 많은 검증을 통하여 신뢰성 높은 분석으로 언급되고 있다.

(블로그 분석) 현재 하루에 약 12만 개의 블로그가 생성되고, 140만개의 블로그에서 새 글이 작성되고 있다. 이런 폐쇄형 소셜 네트워크의 발달로

소비자는 자신의 소비 행태나 상품 사용 후기, 여행 후기 및 여행 일정 공유 등 다양한 분야에 다양한 방법으로 블로그를 활용하고 있다. 따라서 블로그 내용에 대한 파악은 현재 소비자들이 어떤 이념을 갖고 있으며, 어떤 경향을 가지고 있는지를 파악하는데 용이하다. 나아가 소비자의 니즈를 파악함으로써 기업의 마케팅 전략으로 사용할 수 있는 유의미한 데이터를 도출해 낼 가능성도 매우 높다. 선행연구에서 블로그 데이터를 활용하여 LDA 분석을 통하여 사용자가 관심을 갖는 주제가 무엇인지 파악하고 해당되는 키워드를 도출하기 위한 시도가 지속적으로 이뤄지고 있다(조요한 외, 2009).

블로그 LDA 분석을 위한 기본 데이터는 University of Glasgow에서 블로그의 포스트 된 약 2천 8백만 개의 포스트 데이터를 수집하여 데이터셋을 구성하였다. 불용어 및 기준에 적합하지 않은 포스트를 제외하고 수정하여 구성된 데이터셋을 혼잡도(Perplexity) 계산을 통해 모델의 적합성을 측정하였다. 혼잡도는 모델을 평가하는 지표로 낮은 혼잡도가 나올수록 좋은 성능의 데이터이다. 혼잡도의 식은 다음과 같다.

$$Perplexity(D_{test}) = \exp\left\{-\frac{\sum_{d=1}^{M}\log p(w_d)}{\sum_{d=1}^{M}N_d}\right\}$$

분석결과는 [그림 2-17]과 같다. LDA 모델을 학습시켜 얻은 확률이 가장 높은 주제로 나타났으며, 주제 41은 은행, 16은 문학, 22는 인종과 관련된 키워드가 도출되었다.

[그림 2-17] 블로그 LDA 토픽 모델링 분석 결과 각 키워드별 확률 비중치

주제 41	
단어	확률
account	0.0601
review	0.0279
bank	0.0260
month	0.0251
cd	0.0203
rate	0.0188
check	0.0167
deposit	0.0165
credit	0.0155
save	0.153

주제 16	
단어	확률
book	0.0279
read	0.0266
write	0.0194
writer	0.0114
stori	0.0113
author	0.0103
raglin	0.0098
publish	0.0094
reader	0.0087
poem	0.0077

주제 22	
단어	확률
black	0.0184
white	0.0157
american	0.0140
wright	0.0140
king	0.0120
race	0.0120
america	0.0115
call	0.0098
racist	0.0097
flag	0.0089

출처 : 조요한 외, 2009

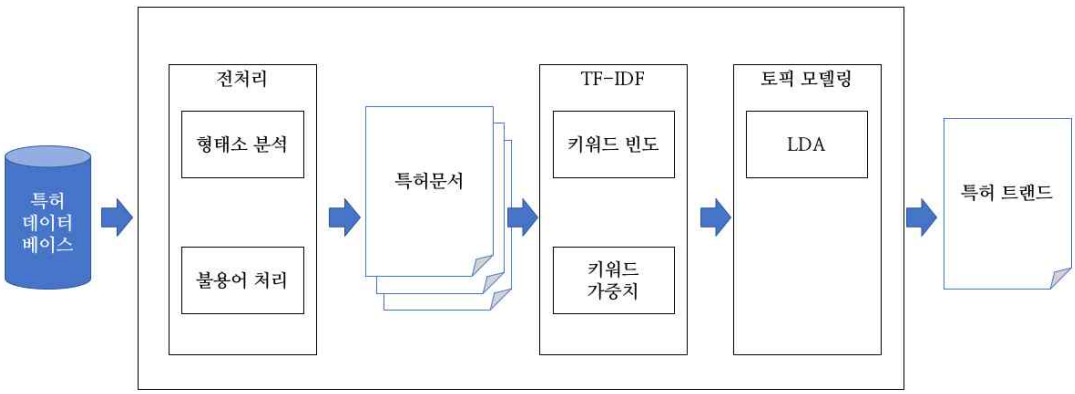
그러나 LDA 알고리즘의 편리성과 정확성에도 불구하고 블로그의 데이터의 경우 범위가 광범위 하고, 데이터 별로 절대적인 양의 차이가 존재하여 통제하기 쉽지 않아 데이터 수집의 속도와 분석의 제한이 존재한다. 따라서 선행연구에서는 차후의 연구에서 LDA 최적화에 대한 연구가 진행되어야 한다고 시사하고 있다(조요한 외, 2009).

특허 기반의 데이터에서도 LDA 분석을 적용하여 추세 분석을 진행하였다. 4차 산업으로 인한 정보 기술 발달의 속도가 더욱더 가속도가 붙으면서 정보 지식의 재산권에 대한 개념은 과거보다 더욱더 중요한 개념이 되었다. 따라서 많은 사용자들은 자신의 정보기술을 남에게 도용 및 남용되지 않도록 권리를 얻을 수 있는 방법인 '특허'를 이용하여 법적으로 보호를 받는다. 특허 출원량은 2012년에 2백만 건에서 2016년 2백5만 건으로 지속적으로 증가하고 있는 추세이다. 따라서 특허에 대한 빅데이터는 계속적으로 축적되고 특허에 대한 추세를 예측 할 수 있게 되었다. 선행연구에서 특허의 추세를 분석하여 향후 시장과 기술을 예측하고 유용한 정보를 얻게 된다(김동연, 2017).

특허 데이터의 LDA 분석 프로세스는 [그림 2-18]과 같다. 특허청 등에서 특허 데이터 베이스를 구축하고 전처리를 통하여 사전을 구축한다. 특허 문서에

적용된 TF-IDF 분석을 통하여 키워드 빈도를 산출하고 높은 빈도의 단어에 가중치를 부여한다. 이와 같은 분석 방식을 기반으로 토픽 모델링을 수행하여 특허에 어떤 주제의 트렌드가 있는지 예측할 수 있다. TF-IDF 알고리즘은 정보 검색 혹은 텍스트 기반 데이터 마이닝에서 주로 사용하는 알고리즘이다. TF-IDF를 이용하면 어휘들의 빈도를 측정할 수 있으며, 빈도에 따라 단어의 중요도까지 측정할 수 있어 매우 유용한 알고리즘이다. 기본적인 TF-IDF 수식은 <표 2-10>과 같다.

[그림 2-18] 특허 데이터 기반 LDA 분석 프로세스



출처 : 김동연, 2017

<표 2-10> TF-IDF의 기본 수식

수식 명	수식
TF	$tf(t,d) = 0.5\frac{0.5 \times f(t,d)}{\max f(w,d) : w \in d}$
IDF	$idf(t,D) = \log\frac{\lvert D\rvert}{\lvert d \in D : t \in d\rvert}$
TF-IDF	$tfidf(t,d,D) = tf(t,d) \times idf(t,D)$

출처 : 김동연, 2017

TF는 추출되는 단어의 빈도로 특정한 어휘가 문헌에서 얼마나 자주 등장하는지를 측정하는 것이다. 빈도가 높을수록 문헌에서 중요한 영향을

미치는 어휘가 된다. IDF는 문헌의 빈도수의 역수로 전체 문헌의 개수가 해당 어휘를 포함 한 문헌의 개수로 나눠주는 것을 말한다. TF-IDF 알고리즘은 특허 분석 등에 자주 사용되며, 단순히 어휘의 빈도만 확인 하는 것이 아니라 중요하게 자주 등장하는 어휘에 대해서 가중치를 부여하여 분석되는 것이기 때문에 중요도에 따라서 추후에 진행되는 LDA에 긍정적인 영향을 미치는 결과를 가져다준다(김동연, 2017). 즉, 가중치에 따라 분류된 주제는 기존의 총 어휘 빈도에 기반을 둔 분류에 비하여 보다 정확한 의미로 해석된다. TF-IDF을 활용하여 가중치를 부여한 키워드를 도출한 후에 K-means clustering과 엔트로피 값을 계산하여 보다 좋은 성능의 키워드 도출이 가능하다. 엔트로피 계산식은 다음을 참고하여 구성하여 분석에 응용하였다(Tsenf et al., 2007).

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

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

$$\text{상호 정보: } G_i = \max_2(\log(\frac{P(x_i, C_k)}{p(x_i)P(C_k)}))$$

$$G_i = 1$$

$$\text{전체적인 빈도 기간: IDF: } G_i = \frac{g_i}{d_i}$$

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

정보 수집:

$$G_i = -\sum_k P(C_k)\log P(C_k) + P(t_i)\sum_k P(C_k|t_i)\log P(C_k|t_i) + P(\overline{t_i})\sum^k P(C_k|\overline{t}_i)\log P(C_k|\overline{t_i})$$

a_{ij}: 문서 j에서 어휘 i가 나타낸 빈도

f_i: 전체 문헌에서 어휘 I가 나타난 빈도

n: 전체 문헌의 수

d_i: 용어 i가 나타난 문헌의 수

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

G_i: 용어 i의 어휘의 가중치

상기 내용을 기반으로 키워드를 도출하고, 해당하지 않은 불용어를 제거한 후 토픽모델링 알고리즘 분석 기법을 활용하여 결과 도출이 가능하다. 토픽모델링 알고리즘을 활용할 때 일반적인 LDA 알고리즘은 Blei et al. (2003)에 제안한 다음의 공식을 활용한다.

$$\prod_{i=1}^{K} p(\beta_i|\eta)\prod_{d=1}^{D} p(\theta_d|\alpha)(\prod_{n=1}^{N} p(Z_{d,n}|\theta_d)p(W_{d,n}|\beta_{1:k}, z_{d,n}))$$

상기 언급한 연구 동향 분석과는 다르게 오로지 동향을 알아보거나 탐색하는 정도에 그치는 것이 아닌, 온라인 웹상에 존재하는 데이터를 기반으로 실시간으로 토픽을 파악하고 숨어있는 의미를 찾는 것에 더 큰 의의가 존재한다. 토픽모델링 알고리즘을 통하여 도출된 토픽을 기반으로 보다 객관적이고 실증적인 정책적 제언 및 시사점을 얻게 된다. 기존의 설문조사 방식이나 인구통계학적인 방식보다는 보다 과학적, 통계적 접근의 데이터 분석을 통하여 실제적인 시사점을 얻을 수 있다는 점에서 토픽모델링

알고리즘은 다양한 연구로 진행되고 있다.

따라서 SW 산업 분야의 직종 및 직업 분류와 관련하여 수요 및 공급 예측을 위한 연구도 토픽모델링을 통해 수행이 가능하다는 결론을 확인할 수 있다. 현재 잡코리아 등의 채용공고에서 제공되는 직업 정보 및 직종 역량 정보는 실제적으로 반영되지 않은 표준적인 지표로 구성되어 있다. 급변하는 현재 인력 시장에서 증가되는 새로운 일자리에 대한 정보 부족과 자동화 및 4차 산업혁명의 일환으로 발생된 정적인 직종들의 감소 또한 인력 수요 공급에 있어 매우 중요한 사항이다. 토픽모델링 알고리즘을 기반으로 잡코리아 등의 채용공고 데이터 베이스를 구축하고, 분석 가능한 데이터로 전처리한 후에 분석을 함으로써 보다 객관적 지표의 SW 인력 토픽이 도출 될 것이다. 또한 본 연구를 통하여 새롭게 창출된 일자리를 확인하는데 실증적인 정보를 제공할 것이다. SW 인력 수요와 공급에 대한 연구는 많이 진행되지 않고 있다. 선행연구들은 주로 SW인력의 직업에 대한 만족도에 대한 연구(유종열 2014) 등의 성과에 대한 연구가 대부분이 이뤄지고 있다. 따라서 본 연구에서는 토픽 모델링을 활용한 SW인력 데이터 분석을 통하여 SW인력 수요를 예측하고 나아가 공급적 측면까지도 영향을 줄 수 있게 하는 것이 궁극적인 목표이다.

2. 다이내믹 토픽 모델링(DTM)

다이내믹 토픽 모델링(Dynamic Topic Modeling: 이하 DTM)은 대량의 문헌들에서 특정 주제를 분류하기 시작하는 LDA의 알고리즘 방식에 주제들의 시계열적인 요소를 반영한 확장된 알고리즘이다. 상기 설명했던 LDA 분석 기법의 방식은 시간의 변화에 따른 변화량인 시계열적인 요소를 고려하지 못한 값들을 기반으로 데이터 마이닝을 통해 분석결과를 도출하는 것이다. 반면에 DTM은 시계열 요소를 고려하여 사용자의 주요 관심 분야를 시간 변화에 따른 추이를 추적하여 보다 정확하게 미래를 예측할 수 있도록 하는 모델이다(Greene and Cross, 2017).

DTM은 Blei and Laffety (2006)의 연구에서 주장된 내용으로 국내에서는 많이 사용되지 않고 있으나 국제적으로 LDA의 확장된 개념으로 활용하려는 노력을 지속적으로 하고 있다. 전통적인 시계열 모델링은 연속 확률 분포에 중점을

두고 분석을 진행하지만 DTM은 이산 변수 즉, 토픽을 기반으로 설계되었다. 기존의 LDA와는 다르게 기본적인 토픽 다항식의 고유의 모수(Natrual parameter)에 대한 공간 모델뿐만 아니라, 로지스틱 정규 분포에 대한 고유 모수의 문헌 별 주제 비율을 활용하는데 의의가 있다.

첫째로, 기존의 정적 토픽 모델의 기본적인 통계적 가정을 검정한다. 기존의 LDA는 $\beta 1$: K를 K개의 주제로 지정하고, 각 화제는 고정된 어휘 사전을 구축하여, 사전을 기반으로 한 문서별 토픽을 추출하는 것이 특징이다.

1. 디리클레 할당과 같이 (K-1) 즉, 단일어에 대한 분포에서 토픽에 대한 비율 θ 선택

2. 각 단어 별:

(a) 주제 할당 선택 $Z \sim Mlt(\theta)$

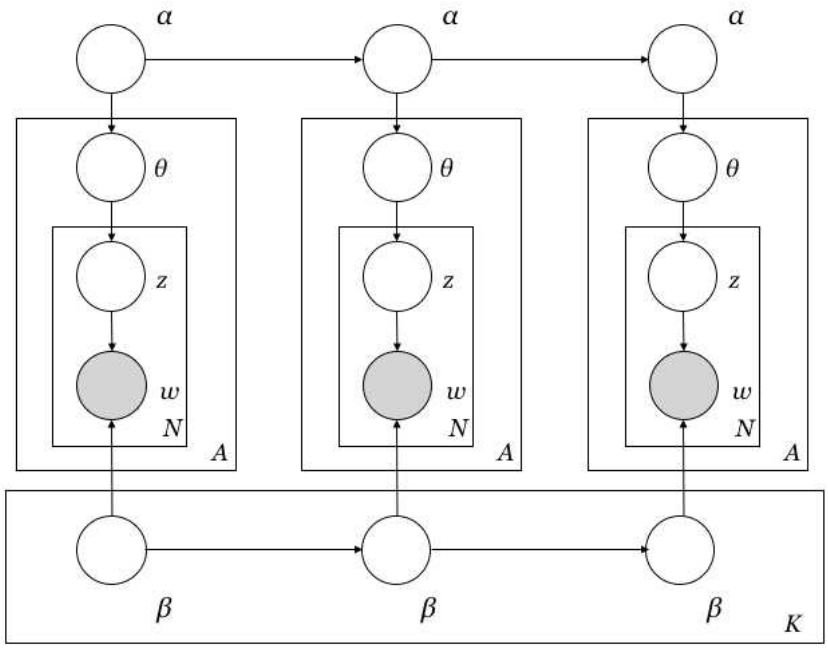
(b) 단어 선택 $W \sim Mlt(\beta_z)$

상기 언급된 수식의 경우는 문헌이 동일한 주제에서 교환 가능하게 형성되는 것을 가정한다. 그러나 많은 수집된 문헌의 경우 주제가 변경되거나 점차 진화하는 경향을 갖는다. 따라서 DTM에서는 데이터가 시간 단위로 t와 연관된 토픽이 t-1과 관련된 토픽으로부터 발전되는 “K-Component Topic Model” (K-컴포넌트 토픽 모델)로 문서를 모델링 한다.

즉, V항을 갖는 K-성분 모델에 대하여 $\beta_{t,k}$ 는 t 시간의 토픽 k에 대한 고유 모수의 V-벡터를 나타낸다. 다항식 분포의 일반적인 표현은 평균 매개의 변수화에 있다. V-차원 다항식의 평균 매개 변수를 π 나타낸다면, 고유 모수의 i번째 매핑은 $\beta_i = \log(\pi_i/\pi_V)$로 나타낼 수 있다. 기존의 LDA는 정적인 데이터에 의한 불확실성을 분석하는데 사용되지만, 순차적으로 진화하는 데이터에의 모델에는 적합하지 않다. 따라서 가우시안 잡음(Gaussian noise)를 통하여 변화하는 상태 공간 모델(State space model)에서 각 주제인 $\beta_{t,k}$의 고유 모수를 연결한다. DTM의 모델의 단순한 버전의 수식과 그래픽 모델은 [그림 2-19]와 같다.

$$\beta_{t,k}|\beta_{t-1,k} \sim N(\beta_{t-1,k}, \sigma^2 I)$$

[그림 2-19] DTM 그래픽 모델



출처 : Blei and Laffety, 2006

LDA 알고리즘은 문헌별 주제 비율 θ는 디리클레 분포에서 가져온 것이나, DTM에서 문헌별 주제 비율에 대한 불확실성을 표현하기 위해서는 로지스틱 평균 α를 사용한다(Greene and Cross, 2017). 모델 간의 순차적인 구조는 간단한 동적 모델로 표현된다. DTM은 토픽과 토픽 비율 분배를 연결함으로써 순차적으로 토픽 모델을 수집하여 순차적인 코퍼스(Corpus)를 생성하였으며, 생성의 과정은 다음과 같다(Blei and Lafferty, 2006).

1. 토픽 인출(Draw) $\beta_t|\beta_{t-1} \sim N(\alpha_{t-1}, \sigma^2 I)$

2. 인출 $\alpha_t|\alpha_{t-1} \sim N(\alpha_{t-1}, \delta^2 I)$

3. 각 문헌에 대하여 :

(a) 인출 $\eta \sim N(\alpha_t, \alpha^2 I)$

(b) 각 어휘에 대하여:

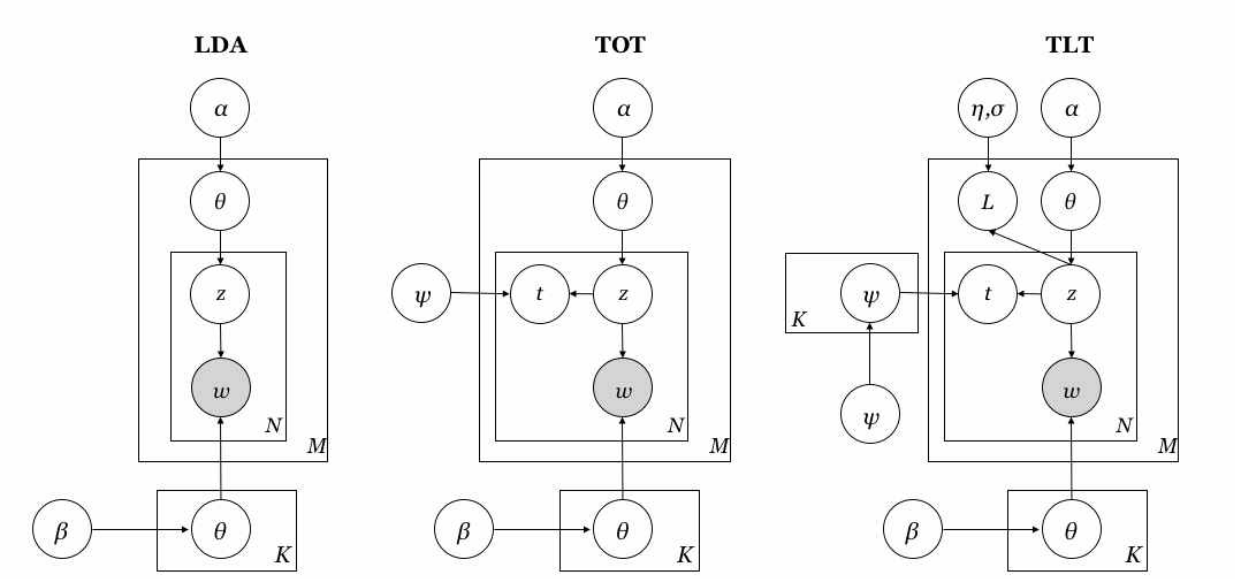
i. 인출 $Z \sim Mlt(\pi(\eta))$

ii. 인출 $W_{t,d,n} \sim Mlt(\pi(\beta_{t,z}))$

DTM 알고리즘이 LDA와 비교할 때 가지는 차이점이 있다면 그래픽 모델에서 t와 t-1간의 수평 화살표를 제거하면, 독립적인 3 시간 차원의 LDA 알고리즘으로 축소된다. 시간의 흐름에 따라 t와 t-1의 변화 및 관계성을 측정하고 k 번째 항목에서 진화한 것을 확인할 수 있다(Greene and Cross, 2017).

시간의 흐름에 따른 LDA에 대한 확장 연구는 다른 방식으로도 시도되었다. 시간 라벨 토픽 모델(Time Label Topic Model; 이하 TLT)의 형식으로 주장된 이 모델은 기존 분석 방식에 시간에 대한 변수를 추가하여 개발된 모형으로 동시 발생(co-occurance) 패턴에 영향을 주어 토픽 문제에 혼동을 줄 수 있다고 주장한다(Chen et al., 2015). 정적인 LDA 알고리즘을 개선하기 위한 노력이 지속되었다. 추후 다른 연구자에 의하여 TOT(topic over time)가 주장되면서 시간에 대한 요소가 추가되었지만, 각 레이블의 특성을 고려하지 못한 모델이여서 사용의 한계가 존재하였다. 따라서 시간의 변화에 따른 요소와 함께 응답 변수를 고려한 TLT를 주장하였으며, TLT 알고리즘의 내용은 [그림 2-20]과 같다(Chen et al., 2015).

[그림 2-20] 기존 LDA와 TOT, TLT의 그래프 모델 비교



출처: Chen et al., 2015)

K : 토픽의 수

M : 총 문헌의 수

V : 총 고유 어휘의 수

N_d : 문헌 d에 있는 어휘 토큰의 수

α_d : 문헌 d에 대한 사전 디리클레 K-벡터

β_k : 토픽 k에 대한 사전 디리클레 V-벡터

Y_k : ψk에 대한 사전 대칭 디리클레

θd : 문헌 d에 존재하는 주제의 다항 분포

ϕ_k : 토픽 k에 고유 어휘의 다항 분포

ψ_k : 시간 스탬프(Time Stamp)가 있는 주제 k에 고유 어휘의 다항 분포

σ_η : 응답 매개 변수 σ와 η

L_n : 문헌 d의 응답

Z_{di} : 토픽과 관련된 토큰

W_{di} : 문헌 d 안에서의 i번째 토큰

t_{di} : 문헌 d 안에서의 i번째 토큰에 관련된 타임 스탬프

TLT는 기존의 정적인 LDA에 시간 변화에 대한 예측률을 높이고 TOT에서 고려하지 못한 레이블의 특성까지 고려하여 다 고차원적인 데이터 마이닝 알고리즘을 구축하고 있다. TLT 알고리즘에서 문헌 d는 토픽에 대한 다항 분포인 θ_d와 연관이 있으며, 각 토픽은 어휘에 대한 다항 분포인 ϕ_z 및 해당 토픽에 대한 문헌의 각 어휘에 대한 타임 스탬프(Time stamp)가 있는 다항 분포인 ψ_z와 관련된다. L은 문헌의 응답 변수이며, σ와 η는 응답 매개 변수이다. 그래프 모델과 관련하여 TLT 알고리즘의 식은 다음과 같다.

$$P(Z,w,L,t|\alpha,\beta,\sigma,\eta,Y) = P(w,Z|\alpha,\beta)P(L|Z,\sigma,\eta)P(t|Z,Y)$$

TLT 알고리즘의 각 문헌과 응답은 다음과 같은 과정을 갖는다.

For each topic k ∈ {1, . . . ,K}:
Generate Øk=(Øk,1,...,Øk,v)^T ~Dir(.|β)
Generate Ψk=(Ψk,1,...,Ψk,T)^T~Dir(.|y)
For each document d:
Generate θd=(θd,1,...,θd,v) ~Dir(.|α)
For each word wdiθd)
Generate zdi ~Mult(.|θd)
Generate wdi ~Mult(.|Øzdi)
Generate twdi ~Mult(.|Ψzdi)
Generate response variable L|z,σ ,η~N(η ,σ).

상기 내용과 동일하게 기존 LDA의 정적인 알고리즘으로는 시간 변화에 따라 실시간적으로 변화하는 토픽이나 문헌의 내용을 따라가기엔 신뢰성의 부재와

정확성이 떨어지는 결과를 초래한다. 따라서 시간에 대한 고려는 무엇보다 실시간적인 토픽을 파악하는데 있어 정확성을 증가시키며, 나아가 분석의 신뢰성까지 증가시킨다.

국내의 선행연구 중에 DTM을 사용하는 연구는 매우 드물다. 최근 카쉐어링의 의견 교류와 관련하여 온라인상의 소비자의 리뷰를 대상으로 DTM 알고리즘을 활용한 연구가 진행되었다(안주언, 2017). 기존의 LDA의 방식을 활용하여도 전체적인 댓글의 트렌드 파악 및 니즈 파악 등이 가능하다. 그러나 실시간으로 댓글의 흐름과 내용은 변화되며, 보다 정밀한 정보를 얻기 위해서는 시간 변화에 따른 변수를 고려하여야 한다. 오피니언 마이닝도 이에 상응한다. 오피니언 마이닝은 온라인 웹상에서 사용자들이 자신들이 느낀 감정이나 생각을 표현하고 리뷰로 남김으로써 자기 표출(Self-disclosure) 혹은 정보 공유(Information Sharing)의 중요한 영향을 미친다. 따라서 현재 기업들은 이런 댓글, 후기 등에 민감하게 반응하며, 기업 성과에 직접적인 영향을 미치기도 한다. 오피니언 마이닝의 Elasticnet 목적함수는 다음과 같다.

$$\arg mim_{\beta}\sum_{i}(y_i-\beta\, x_i)^2+\lambda_1\sum_{k=1}^{K}\left|\beta_k\right|+\lambda_2\sum_{k=1}^{K}\beta_k^2$$

Lasso 회귀분석을 사용하여 DTM으로 얻어진 데이터를 감정 분석하여 소비자가 카쉐어링에 대한 생각이 무엇인지 파악하고자 하였다. 첫째, 구글 플레이와 네이버 blog에서 얻은 리뷰 및 정보를 토큰화 과정을 거쳐 POS 태깅으로 명사를 추출한다. 또한 불용어를 제거함으로써 데이터 전처리를 완료한다.

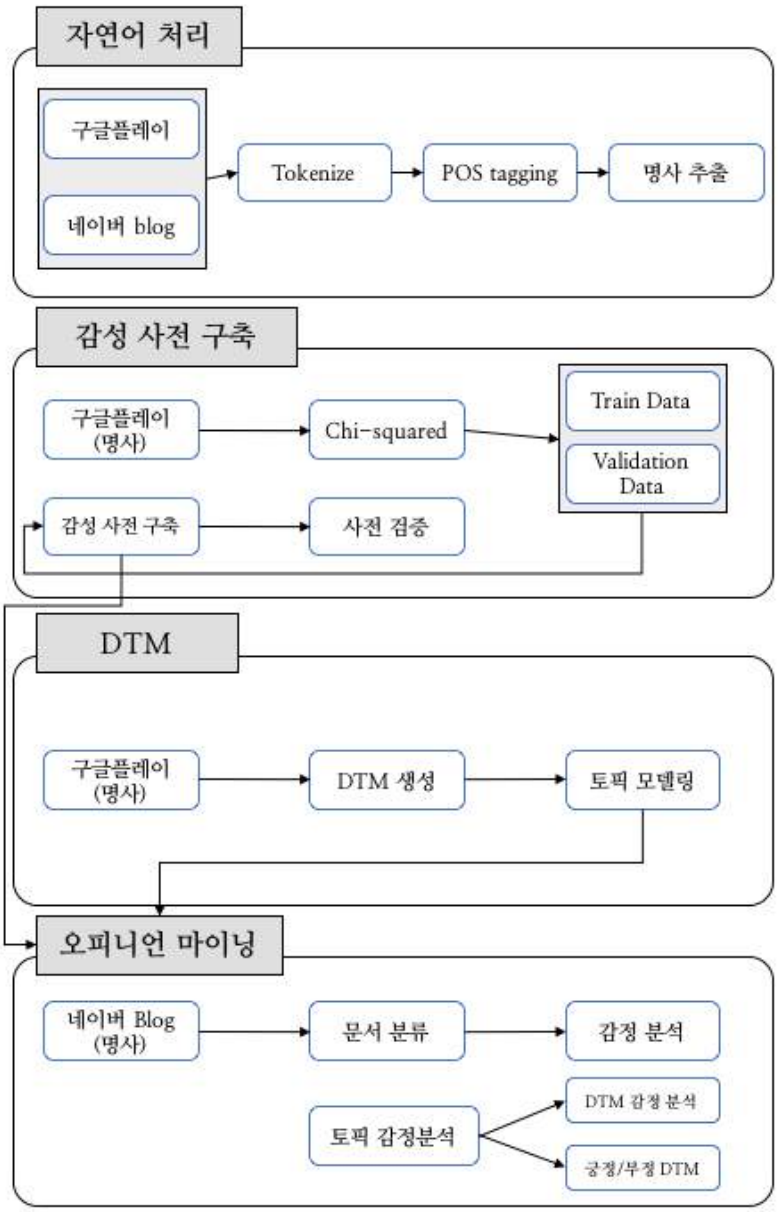
감정 분석을 위하여 감정 사전을 구축하기 위하여 구글 플레이의 명사를 카이제곱 검정을 통하여 훈련 데이터와 실제 사용할 데이터로 분류하고 감정 사전을 구축하여 검증까지 진행 한다. 감성 사전을 구축하기 위하여 나이즈 베이즈 분류기를 만들기 위한 산식은 다음과 같다. 나이즈 베이즈 분류기를 통하여 정의된 확률이 긍정적/부정적인 글인지 효과적으로 판단해준다.

$$P(X|y_i)=\prod_{i=1}^{m}P(x_i|y_i)$$

구글 플레이의 명사를 토대로 DTM 알고리즘을 사용하여 토픽 모델링 분석을 실시한다. [그림 2-21]과 같이 실시간으로 변화하는 변수를 반영함으로써 보다 정확한 토픽을 추출해 낼 수 있다(Blei and Lafferty, 2006).

[그림 2-21] 카쉐어링 DTM 알고리즘 분석 및 오피니언 마이닝

분석 절차



출처 : 안주언, 2017

기본적인 구조는 LDA와 동일하며, 주어진 데이터 내에서 어떠한 주제들이 분포해 있는지에 대한 확률 모형을 다음의 식과 같이 구성한다.

$$p(z_1,\ldots,z_N) = \int p(\theta)(\prod_{n=1}^{N} p(z_n|\theta))d\theta$$

$$p(w,z) = \int p(\theta(\prod_{n=1}^{N} p(z_n|\theta)p(w_n|z_n)$$

DTM은 LDA의 확장된 알고리즘으로 설정한 시간만큼 LDA를 실행하고 LDA 분석에서 모수 값을 송수신하면서 각 LDA를 연결한다. 즉, t 시간과 t-1 시간의 변화에 따른 LDA 분석 값을 송수신하여 시계열 적인 요소를 고려하는 것이다(Blei and Lafferty 2006). 오피니언 마이닝을 위하여 감성사전과 토픽 모델링 결과를 기반으로 문서를 분류하고 감정분석을 실시한다. 감정분석을 통하여 긍정적 어휘와 부정적 어휘를 추출하고 어휘에 따른 전망을 예측한다(안주언, 2017).

DTM 분석은 1년 단위로 진행하였으며, 시간에 따라 단어의 비중도의 차이가 변경되므로 시간에 따른 변수를 고려하여 10개 단위의 키워드로 구성하였다. 상기 사례와 같이 LDA는 정적인 데이터로 한정적 시간에 따른 분포된 어휘의 특징을 알 수 있었던 반면에 DTM은 시간별로 키워드의 변화량까지 파악할 수 있어 보다 정확한 동향 파악이나, 가중치 파악이 가능하다.

토픽은 시시각각 시간 단위에 따라 변화한다. 물론, 100년 동안 일관되게 사용되는 토픽도 존재 할 것이나 반대로 24시간을 넘기지 못하고 사라지는 토픽도 존재한다. 다른 국외 연구에서는 상기 언급한 시간에 대한 고려를 좀 더 고차원적으로 최적화하기 위하여 짧은 기간의 변화(Short-timescale)와 긴 기간(Long-timescale)의 변화에 대한 다차원적인 종속성을 고려해야한다는 주장을 하고 있다(Iwata et al. 2014). 예를 들어 정치 관련 뉴스 주제를 토픽 모델링 한다고 할 때, 의회 구성원의 직함 등의 토픽은 수십 년에 걸쳐 지속적으로 나타날 수 있으나 현재 의회에서 입법 절차를 진행 중인 아젠다의 경우 며칠 발생하고 사라질지 모른다. 따라서 여러 시간을 고려한 다중 규모의 시간 변화에 따른 DTM을 시도하여 정보손실을 줄이고 최적화된 다차원 DTM을 구성할 수 있다.

MDTM 알고리즘을 위하여 기존의 LDA에 각 토픽에 대한 단어에 대해 다중 시간 척도 집합을 고려하는 방법으로 시간 규모 속성을 통합한 산식은 다음과 같다.

$$\phi_{t,z} \sim Dirichlet(\sum_{s=0}^{S} \lambda_{t,z,s} \xi_{t-1,z}^{(s)})$$

$\phi_{t,z}$: t 시점에서의 z 토픽에 대한 어휘 다항분포

$\lambda_{t,z,s}$: 각 t,z,s에 대한 가중치

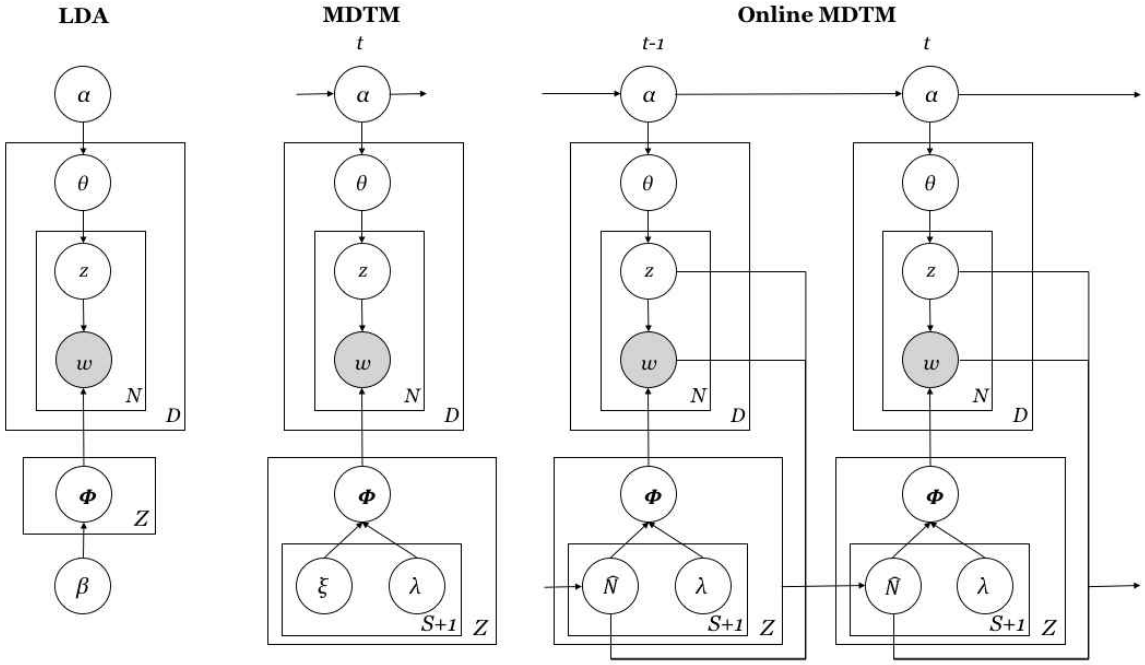
$\xi_{t-1,z}^{(s)}$: t-1 시점에 규모가 있는 z번째 항목의 토픽에 대한 다항분포

LDA에서 토픽 비율은 θ_t, d는 디리클레 분포로 부터의 표본 추출로 계산되며, MDTM으로 다차원적인 동적 자료를 측정하기 위하여 디리클레 매개 변수는 $\alpha_t = \{\alpha_{t,z}\}_{z=1}^{z}$로 가정한다. 특히, t 시점에서 토픽 z의 디리클레 모수는 감마(Gamma) 값을 사용하여 측정한다(Iwata et al., 2014).

$$\alpha_{t,z} \sim \Gamma(\gamma \alpha_{t-1,z,\gamma})$$

따라서 상기 확인된 식들을 기반으로 기존의 LDA와 다른 측면을 고려하여 그래프 모델을 확인하면 [그림 2-22]에서와 같이 시점 t와 t-1을 고려하는 점에서 가장 큰 차이를 보인다.

[그림 2-22] LDA, MDTM, Online DTM의 그래픽 모델 비교



출처 : Iwata et al., 2014

선행 연구들에서 확인 할 수 있듯이 DTM에서 시간에 대한 변수의 적용은 내용의 정확성과 정보의 질을 증가시키기에 중요한 요소 중 하나이다. 기본적인 LDA 알고리즘을 사용했을 때와는 다르게, DTM은 추출된 시간에 따른 흐름까지 파악할 수 있어 동시간대적인 시간 흐름에 따른 토픽의 변화와 키워드의 변화, 그리고 예측의 분석이 가능하다.

결과적으로 DTM은 LDA 알고리즘의 확장된 개념으로써 LDA 특성을 갖고 있다. 추가적인 시간에 대한 보완까지 이뤄져 있어 DTM이 LDA에 비해서 상대적으로 최적화된 알고리즘으로 볼 수 있다. 본 연구는 SW직종 및 직업의 변화에 관련된 분석을 수행하기 때문에 시계열에 대한 고려가 매우 중요하다.

SW분야와 관련하여 토픽모델링 알고리즘의 사용의 시도도 최근 진행되었다. 소프트웨어 직종의 성격과 대학 커리큘럼의 내용의 차이를 토픽모델링 분석을 활용하여 커리큘럼과 SW 직종에서 요구하는 능력의 차이를 증명하였다(최재원 외, 2017). 본 연구에서도 DTM을 활용하여 잡포털, 채용공고의 인력 수요 · 공급 매칭 데이터와 기존 직업 체계인 한국직업표준, 고용직업표준 등의 데이터를

수집하여 전처리를 하고 도출된 명사형의 데이터 셋으로 SW인력 수요를 예측할 것이다. 시간 변수를 추가하여 시간 별로 변화하는 일자리의 추이를 확인하고 나아가 숨겨져 있는 새로운 의미를 갖는 일자리를 도출하고 기존의 일자를 재정비하여 현재 직종 기준보다 발전된 직종 및 직업 분류를 제시하고자 한다.

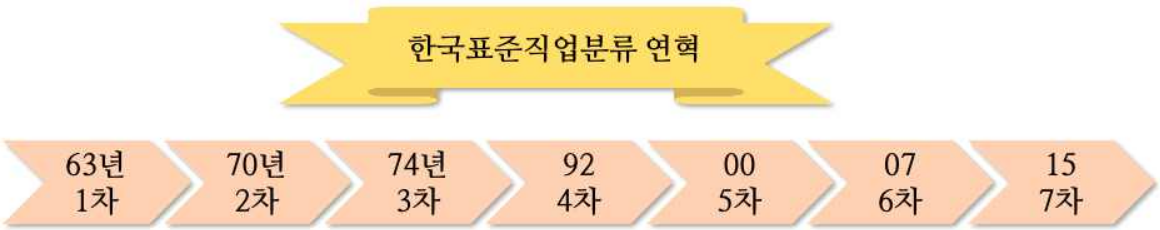
정보기술의 발전으로 점차 급변하는 사회에서 직업의 변화는 가속화되고 있다. 따라서 기존에 존재하는 직업 관련 표준은 새로운 일자리를 설명하는데 있어 부족하다. 4절에서는 기존의 직종 및 직업 분류체계에 대한 현황을 확인하여 각 직종 표준별 특징을 파악한 내용을 기술하였다. 나아가 DTM 토픽 모델링 알고리즘 분석을 통하여 도출된 토픽들에 대한 직종 재배치를 기존 분류체계와 비교 · 파악하여 정의하는 기초자료로 활용하였다.

제4절 SW산업 직종 및 직업 분류체계

1. 한국표준직업분류 (KCSO-07)

체계적인 직업 분류를 국내에서 규정하기 시작하게 된 것은 1960년 내무부 통계국 국세 조사에 사용된 것이 첫 시도였다. 1858년 국제노동기구(ILO)의 국제표준직업분류(ISCO)를 토대로 1963년 처음으로 한국표준직업분류가 제정되었다. 기존의 직업분류 개선 및 보수 작업을 진행하여 [그림 2-23]과 같이 1966년에 개정작업을 진행 2015년 기준 제 7차 개정작업을 추진하였다.

[그림 2-23] 한국표준직업분류 연혁



출처 : 통계청 (2017)

6차 개정인 2007년에는 국제표준직업분류(ISCO-08) 개정본을 반영하여 국제 정세 및 비교 가능하게 하고 전문가와 기술공의 대분류를 통합하여 국내 인력 시장을 통합하고 현장 적응력을 강화하였다. 전문가 및 관련 종사자를 세분화하고, 고용 인력규모를 파악하여 통합 · 축소하는 분류체계를 개정하였다.

제 7차 개정의 주요 내용은 다음과 같다. 첫째, 시간 경과에 따라 분류체계의 전반을 개정하는 것을 목표로 하였으며, 중분류 및 단위 분류 체계를 중심으로 개정하였다. 둘째, 국제표준직업분류의 기준을 적용하여, 구호 및 부호 체계 등 기존 체계를 유지하고, 특히 2007년 ISCO-08을 기준으로 반영하였다. 셋째, 국내 고용시장의 전반적인 직업 체계의 변화 특성을 반영하여 지식 정보화에 따른 사회 변화상을 반영하였다. 이 중 직종 세분화를 통하여 상대적으로 적게

분포되었던 사무 및 판매·서비스직 분류는 보다 세분화 하였고 기계화 및 자동화의 진전으로 기계 관련 기능직은 통합되었다. 마지막으로 직종 간 통합성을 제고하여, ISCO-08의 일관적 특징을 유지하기 위하여 세분류 수준에서 일대일로 연계 되도록 세분화하였다.

7차 개정의 특징으로는 [그림 2-24]와 같이 총 4가지 특징이 존재한다. 첫째, 전문 기술직의 직종 확장으로 지식 정보화 변화상을 반영하였다. 4차 산업 혁명과 관련하여 정보 기술의 신 성장 동력으로 형성된 직종인 데이터 분석가, 산업 특화 소프트웨어 프로그래머, 모바일 애플리케이션 프로그래머를 신설하여 변화상을 반영하였다. 또한 문화 콘텐츠의 확장과 기술 발전으로 문화·미디어 채널 콘텐츠의 생산과 유통 전반이 다변화하였다. 마찬가지로 미디어 콘텐츠 창작자, 공연·영화, 음반 기획자, 사용자 경험과 인터페이스 디자이너 등으로 직업도 세분화 되었다.

둘째, 사회 서비스와 관련된 일자리의 신설을 통하여 저출산·고령화에 대한 대비책을 마련하였다. 복지·돌봄 일자리에 대한 수요 증가에 따라 장애인과 노인 돌봄 서비스 종사원과 놀이 및 행동치료사를 신설 및 상담 전문가, 임상심리사 등 관련 직종을 조정하였다. 또한 여가생활의 증가와 관련된 문화 관광 및 숲과 자연환경 해설사, 개인 생활 서비스 종사원 등이 신설되었다.

셋째, 분류항목이 비교적 적었던 사무 관련 업무와 서비스·판매직 세분화를 통하여 지식 정보화 시대에 따른 서비스직의 세분화가 이루어졌다. 포괄적 직종으로 분류되어 있었던 사무직군 증권 사무원, 대학 행정 조교, 행정사, 기타 금융 사무원, 중개 사무원이 세분화되어 신설되었고, 서비스·판매직의 상점 경영과 일선 관리 종사원 등의 직업도 세분화 되었다.

넷째, 기계화·자동화에 따른 기능직 및 기계 직종 통합을 통하여 과실 및 채소 가공과 제조 관련 기능 종사원, 복합·다기능 기계에 맞춰서 세분화되었던 직종이 통합되었다.

[그림 2-24] 제 7차 한국표준직업분류 개정 특징



출처 : 통계청, 2017

주로 SW인력과 관련된 직종으로는 '전문가 및 관련 종사자' 부분이다. '전문가 및 관련 종사자' 는 7차 개정에서 4차 산업 등의 정보 기술 발견 이슈를 고려하여 포함되었다.

한국표준직업분류에서 직업의 정의 또한 중요한 요소이다. 국제표준직업분류에 기본 원칙에 따라 직종(Job)은 '자영업을 포함하는 한정된 고용주에게 개별 종사원들이 과업을 시행하거나 혹은 시행을 위한 일련의 과업 자체' 로 설정하고 있으며, 직업(Occupation)은 '비슷한 직종들의 집합' 으로 정의한다. 상기 언급된 '비슷한' 의 의미는 주어진 과업에 대하여 매우 유사한 성격을 갖는 것을 말한다. 추가적으로 직업은 과업의 계속성을 가져야 하며 다음의 조건에 충족하여야 한다.

(1) 매월, 매주, 매일 등의 주기적 행태가 존재함

(2) 계절적 성격을 갖고 있음

(3) 뚜렷한 주기는 없으나 지속적으로 실행됨

(4) 현재 시행하고 있는 과업을 지속적으로 행할 의미 및 가능성 존재

제 7차 개정을 진행함으로써 <표 2-11> 과 같이 대분류, 중분류, 소분류, 세분류, 세세분류 등으로 분류단계를 설정하고 중분류 52개 항목, 소분류 156개

항목, 세분류 450개 항목, 세세분류 1,231개 항목으로 나누어 총 1,231개의 직업을 분류하였다.

<표 2-11> 제7차 개정 후 분류단계별 항목 수

대분류	중분류	소분류	세분류	세세분류
전체	52	156	450	1,231
1. 관리자	5	16	24	82
2. 전문가 및 관련 종사자	8	44	165	463
3. 사무 종사자	4	9	29	63
4. 서비스 종사자	4	10	36	80
5. 판매 종사자	3	5	15	43
6. 농림 · 어업 숙련 종사자	3	5	12	29
7. 기능원 및 관련 기능 종사자	9	21	76	198
8. 장치 · 기계 조작 및 조립 종사자	9	31	65	220
9. 단순 노무 종사자	6	12	24	49
A. 군인	1	3	4	4

출처 : 통계청, 2017

SW인력의 한국표준직업분류 항목표는 <표 2-12>와 같다. <표 2-11>과 같이 기존의 직종에서 ‘정보시스템 개발 전문가’를 ‘데이터 및 네트워크 관련 전문가’와 ‘컴퓨터 시스템 및 소프트웨어 전문가’로 소분류 되었다. 또한 ‘웹 운영자’ 소분류 위치 변화(→ ‘정보 시스템 및 웹 운영자’) 등으로 위치변경과 함께 ‘산업 특화 소프트웨어 프로그래머’ 등이 신설 · 세분되었다. 추가적으로 ‘데이터 분석가’는 기존의 정형 데이터 분석뿐만 아니라 머신러닝 및 뉴럴 네트워크 등의 비정형 데이터 분석까지 포함하도록 업무 및 직종의 범위가 조정되었다(통계청2017).

<표 2-12> 한국표준직업분류에서 SW 관련 직종 분포

2. 전문가 및 관련 종사자				
직업 코드				직업 명
대분류	중분류	소분류	세분류	
22				정보 통신 전문가 및 기술직
	221			컴퓨터 하드웨어 및 통신공학 전문가
		2211		컴퓨터 하드웨어 기술자 및 연구원
			22110	컴퓨터 하드웨어 기술자 및 연구원
		2212		통신공학 기술자 및 연구원
			22121	통신기기 기술자 및 연구원
			22122	통신장비 기술자 및 연구원
			22123	통신기술 기술자 및 연구원
			22124	통신망 운영 기술자 및 연구원
			22129	그 외 통신공학 기술자 및 연구원
	222			컴퓨터 시스템 및 소프트웨어 전문가
		2221		컴퓨터 시스템 전문가
			22211	정보 통신 컨설턴트
			22212	컴퓨터 시스템 감리 전문가
			22213	컴퓨터 시스템 설계 및 분석가
		2222		시스템 소프트웨어 개발자
			22221	시스템 소프트웨어 설계 및 분석가
			22222	시스템 소프트웨어 프로그래머
		2223		응용 소프트웨어 개발자
			22231	범용 소프트웨어 프로그래머
			22232	산업 특화 소프트웨어 프로그래머
			22233	모바일 애플리케이션 프로그래머
			22234	게임 프로그래머
			22235	네트워크 프로그래머
			22239	그 외 응용 소프트웨어 프로그래머
		2224		웹 개발자
			22241	웹 프로그래머
			22242	웹 기획자
		2229		기타 컴퓨터 시스템 및 소프트웨어 전문가
			22290	그 외 컴퓨터 시스템 및 소프트웨어 전문가
	223			데이터 및 네트워크 관련 전문가
		2231		데이터 전문가
			22311	데이터 설계 및 프로그래머
			22312	데이터 분석가
			22313	데이터 관리 및 운영자
		2232		네트워크 시스템 개발자
			22320	네트워크 시스템 개발자
		2233		정보 보안 전문가

			22330	정보 보안 전문가
		2239		기타 데이터 및 네트워크 관련 전문가
			22390	그 외 데이터 및 네트워크 관련 전문가
	224			정보 시스템 및 웹 운영자
		2241		정보 시스템 운영자
			22410	정보 시스템 운영자
		2242		웹 운영자
			22420	웹 운영자
	225			통신 및 방송 송출 장비 기사
		2250		통신 및 방송 송출 장비 기사
			22501	통신장비 기사
			22502	방송 송출 장비 기사

출처 : 통계청, 2017

2. ICSO-88 (국제표준직업분류)

국제표준직업분류는 2007년 2월에 개정된 국제 표준 직업 체계를 표시한 것이다. 국제 노동 기구(ILO: International Labour Organization)의 기술 전문가 그룹의 자문으로 전문가 회의를 통하여 최종 업데이트 되었다.

ICSO는 모든 국가 및 관련 이해 당사자의 의견을 수렴하기 위하여 ILO에 설문항목을 게재하여 최종적으로 직무를 분류할 수 있는 체계를 조정하였으며, 2007년 전문가 회의에서 논의되어 최종 2006년 12월 수용, 공고되었다. 해당 분류 체계는 국제적인 지표를 포괄하여야하기 때문에 일반적으로 요구되는 공식적인 자격, 과업, 동일하거나 유사한 일의 성격을 필요로 하는 모든 직업이 분류되어야 한다는 대전제 하에 구성되었다.

4차 산업 등의 정보 기술 발전으로 인하여 정보기술(Information Technology)의 발전은 기존 분류를 부적합하게 만들었으며, 기존 분류체계가 트렌드를 반영하지 못하여 IT 관련 직종을 재정립하였다. ISCO-88에서는 이전의 분류 방식과는 다르게 좀 더 세분화되고 통합되는 전체적인 변경을 통하여 일자리의 변화에 대한 점진적 대처를 진행하였다. 따라서 ISCO-88의 일부 범주는 노동 시장의 직업 및 기술 변화를 반영하거나 기존 문제점을 보완하기 위하여 통합 및 세분화와 직종 간 하위 항목 이동 등을 진행하였다.

세부항목별 변화 양상을 살펴보면 <표 2-13>과 같다.

<표 2-13> 각 단계별 그룹의 수 ISCO-88

Major groups	Sub-major groups	Minor groups	Unit groups
1. Managers, Senior officals and legislator	4(3)	11(8)	28(33)
2. Professionals	6(4)	24(18)	89(55)
3. Technicians and associate professionals	5(4)	20(21)	87(73)
4. Clerks	2	9(7)	28(23)
5. Service and sales workers	2	12(9)	40(23)
6. Skilled agricultural, fishery, and forestry workers	2	6(6)	18(17)
7. Craft and related trades workers	5(4)	16	66(70)
8. Plant and machine operators, and assemblers	3	13(20)	42(70)
9. Elementary occupations	6(3)	11(10)	33(25)
0. Armed forces occupations	3(1)	3(1)	3(1)
Total ISCO-08	38(28)	125(116)	434(390)

출처 : ILO, 2007

분류 체계 단계별 그룹 수를 확인한 결과 공장 및 기계 운영자의 단위 그룹 수는 다른 직종에 비하여 축소된 것을 확인 할 수 있다. 또한 하위 항목에 대한 주제별 직종 표준을 개발이 추가되어야 한다는 결론을 도출하였다.

(1) Health

(2) Teaching

(3) Agriculture

(4) Information and communications technology

(5) Tourism occupations

SW인력과 관련된 그룹은 25번 Information and Communications Technology Professionals, 35번 Information and communication technicians이다. 25번과 35번의 단계별 항목은 <표 2-14>와 같이 구성된다.

<표 2-14> IT 관련 직업 분류 체계

Sub-Major Groups	Miner Groups	Unit Groups	Occupations
25			Information and communications technology(ICT)
	251		Software and applications developers and analysts
		2511	Systems anaylsts
		2512	Software developers
		2513	Web and multimedia developers
		2514	Applications programmers
		2519	Software and multimedia developers and analysts not elsewhere classfied
	252		Database specialists and systems administrators
		2521	Database designers and administrators
		2522	System administrators
	253		ICT network and hardware professionals
		2531	Computer network professionals
		2532	Telecommunications engineering professionals
		2539	ICT network and hardware professionals not elsewhere classified
35			Information and communications technicians
	351		ICT operations and user support technicians
		3511	ICT operations technicians
		3512	ICT user support technicians
		3513	Systems testing technicians
	352		Web technicians
		3520	Web technicians
	353		Communications technicians

		3531	Broadcasting and sound and vision recording technicians
		3532	Telecommunications engineering technicians

출처 : ILO, 2007

25번 그룹은 정보 및 통신 기술 전문가에 대한 내용으로써 특정 응용 프로그램에 대한 정보 기술 시스템, 하드웨어, 소프트웨어 및 관련 개념을 계획, 설계, 작성, 테스트를 통하여 기술하고 있다. 최적의 성능 및 데이터 무결성 및 보안을 보장하는 데이터베이스 및 기타 정보 시스템을 설계, 개발, 제어, 유지하는 등의 직종들을 정의하고 재정립하였다.

하위 주요 그룹의 근로자가 수행하는 과업에는 대개 비즈니스 기능에서 정보 기술 사용연구, 컴퓨터 미디어 시스템을 포함한 특정 응용 프로그램의 하드웨어 또는 소프트웨어 구성 평가, 계획 및 컴퓨터 설계 및 프로그램의 설계, 작성, 시험 및 유지 등의 과업들을 실행하는 그룹을 단위 그룹으로 분류하였다. 25번의 하위 그룹의 직업은 다음과 같은 단위 그룹으로 나눠진다.

25번 그룹의 하위 항목으로 251과 252가 존재하며 251에 대한 정의는 다음과 같다. Miner group 251은 Software and Applications Developers and Analysts이다. 소프트웨어 및 응용 프로그램 개발자 및 분석자는 하드웨어, 소프트웨어 및 기타 응용프로그램과 같은 정보 기술 시스템을 계획, 설계, 작성, 테스트, 조언 및 개선하여 특정 요구 사항을 충족하는 기술을 하는 직업을 말한다. 수행되는 업무는 비즈니스 기능에서 정보 기술사용을 조사하고 효율성과 효율성을 극대화하기 위한 영역을 파악하여 컴퓨터의 이론적 측면과 운영방법에 대한 연구를 통한 다양한 컴퓨팅 업무를 수행하는 것을 말한다. 해당하는 직무를 간단하게 살펴보면 다음과 같이 세분류된다.

2511 System Analysts

2512 Software Developers

2513 Web and Multimedia Developers

2514 Applications Programers

2519 Software and Applications Developers and Analysts Not Elsewhere Classified

2511 System Analysts를 간단하게 살펴보면 고객의 정보 기술 요구 사항, 절차 또는 문제점을 조사, 분석 및 평가하고 현재 또는 미래의 정보 시스템을 개선하기 위한 제안, 권장 사항 및 계획을 개발 및 구현하는 업무로 정의하고 있다. 각 유닛 단위의 직업에는 과업의 종류 및 이 직종에 분류되는 이름의 직업들, 다른 곳에서 분류된 몇 가지 직업들을 표기하여 해당 업무와 비교하도록 하였다.

2511 System Analysts에 할당된 과업 :

(a) 사용자가 문서 등의 요구 사항을 확인하고 시스템에 대한 원리를 측정하고 조율하기 위하여 경영진과의 합의가 필요하다.

(b) 비즈니스 프로세스, 절차 및 업무를 식별하고 분석한다.

(c) 과업간 비효율 측면을 구별 및 평가하고 최적의 비즈니스 프로세스 및 시스템 기능을 추천하는 일을 한다.

(d) 시스템 테스트 계획의 작성, 구현 등과 같은 기능 솔루션을 배포하는 것에 대한 책임을 진다.

(e) 시스템 개발자가 사용할 수 있도록 직무 명세서를 개발한다

(f) 워크 플로우를 개선하거나 새로운 목적을 위해 시스템을 확장하거나 수정한다.

(g) 업무의 호환성을 높이기 위해 조직 내의 컴퓨터 시스템을 조정하고 연결한다.

해당 직종에 분류되는 직업들 :

1. Business analysts (IT)
2. Computer Scientist
3. Information system analyst
4. Systems consultant
5. Systems designer

다른 연관된 직업 분류에서의 직업들:

1. System administer (computers) - 2522
2. Communication analyst (compusters) - 2523
3. Network analysts - 2523

상기 언급된 내용과 같이 직종에 대한 과업들을 재정리 하고 유닛 단위의 직업에 대한 세부적인 내용을 제시함으로써 국제적인 표준 직업을 재분류 하고 있는 것을 볼 수 있다. 이를 기반으로 KCSO 등의 기초자료로 활용하고 있으며, 해당 직종의 교과서와 같은 역할을 하고 있다(ILO, 2007).

그럼에도 불구하고 국제적 표준에 준하는 기준을 맞추고 있어야하기 때문에 급변하고 있는 현재 일자리의 전체적인 설명은 부족한 부분이 존재한다. 일반적인 분류에 대해서는 표준 분류체계로써 큰 역할을 담당하고 있으나, 드론, AR 등의 신기술 혹은 혁신 기술에 대한 직업들을 정의하기에는 부족한 부분이 존재한다.

3. 한국고용직업분류 (KECO-09)

한국고용직업분류(Korean Employment Classification of Occupations)는 국내 고용시장의 현상과 수요, 공급의 현실적인 구조를 반영하여 직종 체계를 적절하게 다시 분류한 것으로 각종 고용정보를 연계하여 활용하는데 그 의의가 존재한다. 즉, KECO는 고용 동향 데이터 수집, 취업알선과 훈련 · 자격정보에 대하여 개발 및 정보 제공을 목적으로 국내 노동시간에 활용 가능한 정보를 쉽게 제공 받고 효과적으로 이용하기 위하여 개발되었다.

한국고용정보원에서 개발된 한국고용직업분류 체계는 2001년에 '산업 · 직업별' 고용관련 구조조사를 위하여 노동연구원에서 연구된 직업조사를 일부 조정하여 사용하였으나, 조사결과에 대한 코딩 문제로 인하여 어려움을 겪었다. 따라서 국내 고용시장의 상황에 맞는 직능 유형(Skill specialization) 중심의 새로운 분류에 대한 필요성이 부각되었다. 의미 있는 통계정보를 제공하여 직업정보의 활용도를 높여 일반인의 시각에서 쉽게 이해하고 활용 할 수 있는 직업분류체계를 개발할 필요성이 지속적으로 발생하였다.

상기 배경을 기반으로 2002년에 직능 유형을 중심으로 하는 한국고용직업분류(KECO)가 개발되었다(한국고용정보원, 2015). 또한 KECO를 기반으로 하는 직업사전을 개발하여 취업 알선 및 교육 훈련, 필요한 역량 등을 한 눈에 확인할 수 있는 체제를 구축하여 사용자가 보다 직관적으로 직종을 파악하고 수요에 맞는 적합한 공급을 할 수 있는 창구로써 활용되어 왔다. 2007년에는 기존에 개발된 직종에 대하여 직업 세분류(4-단계) 단위를 제정하기 위한 통계자료를 수집 및 활용하여 직종 내용의 훼손을 막고 직종 간의 연계성을 강화시켜 국가통계의 활용에 더 많은 기여를 가져다주었다. KECO의 분류 체계는 <표 2-15>와 같다.

<표 2-15> KECO의 구조

대분류	중분류	소분류	세분류
1. 관리직	01. 관리직	9	24
2. 경영재무직	02. 경영 · 회계 · 사무 관련직	9	31
	03. 금융 · 보험 관련직	3	11
3. 사회 서비스직	04. 교육 및 자연과학 · 사회과학 연구 관련직	8	22
	05. 법률 · 경찰 · 소방 · 교도 관련직	3	8
	06. 보건 · 의료 관련직	8	23
	07. 사회복지 및 종교 관련직	3	9
	08. 문화 · 예술 · 디자인 · 방송 관련직	8	30
4. 판매 및 개인 서비스직	09. 운전 및 운송 관련직	5	15
	10. 영업 및 판매 관련직	5	19
	11. 경비 및 청소 관련직	5	16
	12. 미용 · 숙박 · 여행 · 오락 · 스포츠 관련직	7	23
	13. 음식서비스 관려직	2	11
5. 건설 · 생산직	14. 건설 관련직	7	29
	15. 기계 관련직	9	23
	16. 재료 관련직	7	22
	17. 화학 관련직	3	9
	18. 섬유 및 의복 관련직	6	18
	19. 전기 · 전자 관련직	7	13
	20. 정보통신 관련직	6	15
	21. 식품가공 관련직	4	14
	22. 환경 · 인쇄 · 목재 · 가구 · 공예 및 생산단순직	9	27
6. 농림어업직	23. 농림어업 관련직	5	13
7. 군인	24. 군인	1	3
계	24	139	429

출처 : 한국고용정보원, 2015

1~4 단계로 구성되며 대분류는 7개로, 중분류는 24개, 소분류는 139개, 세분류는 429개로 직업을 구성하였다. 대분류는 사실상 분류 코드에서 사용되지 않아 중분류 중심의 체제라 할 수 있다.

KECO의 분류 원칙으로는 크게 직능유형 우선, 중분류 중심체계, 포괄성의 원칙, 베타성의 원칙, 연계성의 유지의 원칙, 최소 고용과 노동시장 우선을 기준으로 분류 체계를 설정하였다.

첫째, 직능유형 우선은 일하는 작업자가 과업을 수행하는 일이 갖는 여러 가지 측면의 성격을 말하는 것으로, 업무 수행에 따른 결과가 생산되는 최종 생산물과 과업을 수행하는 방법 및 프로세스, 과업을 수행하는데 필요한 업무 지식, 주요 활동 도구 및 장비 등이 포함된다. 관련된 업무를 수행하는데 있어 필요한 능력의 높낮이를 말하는 직능수준(Skill level)과는 상반되는 개념이다. 직능 유형에서 제일 주요한 내용은 수행결과물이다. 최종 결과물의 도출에 따라 직종의 성격이 달라지며, 같은 사무직이라고 해도 운송 서비스에 대한 업무 내용과 제조업에 대한 내용은 다르기 때문이다.

둘째, 중분류 중심체계는 인위적인 10진법의 체계를 사용하지 않으며, 영역이 다른 분야를 다르게 분류하고, 하나의 분류 각각이 독립적인 성격을 갖기 위하여 인위적이지 않은 틀을 사용하여 구성하였다. 물론 기존 분류체계에서 10진법을 사용하는 것에 있어서 장점은 존재한다. 중분류 중심체계는 하위분류가 10개 이하로 구성되며 보다 쉽게 설명하고 직관적으로 볼 수 있다는 장점을 갖는다. 그러나 단점은 매우 치명적인데, 분류된 항목을 10개로 제한하기 때문에 관련성이 낮은 항목임에도 불구하고 하나로 묶이게 되는 경우가 종종 발생한다. 중분류 중심의 한국고용직업분류는 중분류 중심으로 이뤄져 구성되며, 4자리 직업코드 분류체계와도 별개의 개념적으로 활용되고 있어, 데이터의 활용성을 증대시키려는 노력이라 할 수 있다.

셋째, 포괄성의 법칙은 국내에 존재하는 모든 직종과 직업을 분류체계 내에서 모두 포괄해야한다는 원칙이다. 유사한 직업이라고 하더라고 사업체 규모에 따라 직종 범위에 대한 차이는 매우 다르게 나타날 수 있기 때문에 아래의 순서와 같게 처리한다.

1. 주요한 직종의 우선 원칙 : 과업 2개 이상을 수행하는 경우에 수행하고 있는 직종 내용 및 분류 항목에 대하여 직종 내용을 비교 · 평가하여 직종 내용의 상관성이 높은 항목으로 분류

2. 최상급의 직능수준의 우선 원칙: 수행하고 있는 직종이 상이한 수준의 훈련과 경험을 통하여 얻어지는 업무 능력을 사용하고자 한다면, 가장 높은 등급의 직종 능력을 요구하는 일로 분류
3. 생산업무에 대한 우선 원칙 : 제품의 생산과 공급이 같이 수행되는 경우는 생산 단계에서 발생된 관련된 업무를 우선적으로 분류하는 원칙

넷째, 베타성의 원칙이란 한 개의 직종이 동일한 직업 단위의 레벨에서 2개 이상의 직업으로 분류되지 못하는 것을 말한다.

다섯째, 연계성 유지란 KECO는 직업의 분류 그 자체로 사용되기도 하며, 직업과 관련되는 교육 훈련, 취업 알선, 자격, 대학교의 학과 등의 다양한 분야에서 연계성을 갖기 위하여 직관적으로 쉽게 분류하며, 항상 그 중심에는 KECO가 있도록 하는 것을 원칙으로 한다. 해당하는 모든 정보가 웹을 통하여 제공되는 현재 시장에서 직무들에 대한 연계성은 매우 중요한 요소 중 하나이다.

여섯째, 최소의 고용 노력과 노동시장 우선이란 단위 직업은 최소의 고용하려는 인원이 있는 직무단위로 구성되었고, 너무 세분화된 직무 단위를 사용하지 않아 직업정보로 유의미하게 사용이 가능하다. 따라서 세분류 단위를 1,000명의 종업원 수를 기준으로 하여 설정하고, 고용인원이 과다한 경우에 직무성격에 따라 직무 단위로 세분화 하였다. 단, 고용시장에서 직업 행태 별로 해당되는 제도 혹은 관습이 존재하는 경우에 있어서는 고용 인원에 상관없이 단위 직업의 규모로 분류하였다.

SW 인력에 해당하는 분류체계는 20인 정보통신 관련직으로 해당 직무의 정의는 정보를 생산하거나 전달하는 일련에 처리 과정에 관련된 직업들이 이에 해당하며, 정보시스템 및 컴퓨터 관련직으로 핵심 IT 활용 직무뿐만 아니라 방송과 통신장비 설치에 대한 직무까지 고려하여 포괄적인 IT분야를 정의하였다. 정보 기술과 관련된 직업은 산업의 특성상 서로 다른 명칭으로 사용되고 있으며, 서로 다른 과업을 실시함에도 불구하고 동명의 직업명을 사용하는 경우도 빈번하게 나타난다. 예를 들면, 웹 개발자, 웹마스터, 웹프로그래머 등이 그러한 오류를 범하기 쉽다.

KECO의 직무 코드 2011인 ‘컴퓨터 하드웨어 기술자 및 연구원’을 분류체계의 표시 사례로 들면 다음과 같다. 2011은 가정, 군사, 산업 등의 과학용 컴퓨터 혹은 컴퓨터와 관련된 장비를 연구하고 설계하여 개발하는 전문가를 말하며, 컴퓨터 관련 장비의 구성요소에 대한 설치 및 제조를 감독하고 테스트하기도 한다. 컴퓨터 관련 사업체, 교육 및 연구기관 등의 다양한 분야에서 고용된다.

2011의 주요 업무는 컴퓨터와 관련 장비들의 개발을 설계 및 계획하고 조정하며, 부품과 회로 및 제어시스템을 테스트하기 위하여 테스트 방식을 개발하기도 한다. 중앙처리 장치와 주변장치, 출력장치와 입력장치 및 보조기억장치들의 성능의 향상을 위하여 컴퓨터시스템을 설계하고 개발하여 운영 자문 등의 지도를 수행한다. 관련된 직업 예시는 컴퓨터 하드웨어 설계기술자, 콘트롤러개발원, 하드디스크 개발원, 컴퓨터 제어시스템개발원 등에서 업무한다. 한국표준직업분류 코드에서 매칭할 경우에는 2211의 컴퓨터 하드웨어 기술자 및 연구원과 같은 맥락의 직무라 볼 수 있다. 이와 같이 각 직무별 주요 업무, 직업 예시, 다른 직무기준과의 연계성을 고려하여 정보를 활용 가능하다. KECO 09에서 분류된 SW 인력 관련 직무 체계는 <2-16>과 같다.

<표 2-16> KECO SW 인력관련 분류체계표

중코드	중분류	소코드	소분류	세코드	세분류	세세코드	세세분류
20	정보통신 관련 직	201	컴퓨터하드웨어 및 통신공학 기술자 및 연구원	2011	컴퓨터하드웨어 기술자 및 연구원	20110	컴퓨터하드웨어 기술자 및 연구원
						20111	컴퓨터하드웨어 기술자 및 연구원
						20112	컴퓨터 부품 및 주변기기 개발기술자 및 연구원
				2012	통신공학 기술자 및 연구원	20120	통신공학 기술자 및 연구원
						20121	통신기기 및 장비(설비) 개발자 및 연구원
						20122	통신기술개발자 및 통신망운영기술자
20	정보통신 관련 직	202	컴퓨터시스템 설계 및 분석가	2021	컴퓨터시스템 설계 및 분석가	20210	컴퓨터시스템 설계 및 분석가
						20211	IT컨설턴트
						20212	컴퓨터시스템 설계 및 분석가
						20213	컴퓨터시스템감리기술자
				2022	네트워크시스템 분석가 및 개발자	20221	네트워크시스템 분석가 및 개발자(운영자 제외)
				2023	컴퓨터시스템 및 네트워크보안 전문가	20231	컴퓨터시스템 및 네트워크 보안전문가
		203	소프트웨어개발전문가	2031	시스템소프트웨어개발자	20310	시스템소프트웨어개발자
						20311	시스템소프트웨어 및 운영체제 프로그래머
						20312	임베디드 및 펌웨어 프로그래머
				2032	응용소프트웨어개발자(응용프로그래머)	20320	응용소프트웨어개발자(응용프로그래머)
						20321	응용소프트웨어프로그래머
						20322	네트워크프로그래머
						20323	컴퓨터 및 모바일 게임프로그래머
		204	웹전문가	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
20	정보통신 관련 직	201	컴퓨터하드웨어 및 통신공학 기술자 및 연구원	2012	통신공학 기술자 및 연구원	20642	방송 및 인터넷 케이블 설치 및 수리원

출처 : 한국고용직업표준, 2009

4. 국가직무능력표준 (NCS)

국가직무능력표준(National Competency Standard: NCS)은 산업별 현장의 직무를 수행하기 위하여 요구하는 기술, 지식, 역량에 대한 내용을 국가 차원에서 산업 부분별 및 수준별로 분류체계를 만든 것으로 과업과 직업교육훈련 및 자격을 연계하여 직업교육훈련과 직업에 대한 자격체계를 직무수행의 능력 중심으로 전환하여 인재개발의 실효성을 획득하는데 목적이 있다. 교육부와 고용노동부를 중심으로 NCS 교육과정과 자격, 인사관리, 직업훈련 체계의 개편을 이뤄 국가의 인재개발 정책이 능력중심사회라는 목표를 지향 할 수 있도록 하는 수단적 도구로써 중요한 역할을 담당하고 있다. 즉, 교육훈련-자격-일의 연결하는 고리 역할을 수행하며, 수요 공급 지표의 핵심 가치로써 <표 2-17>과 같이 기능한다.

첫째, NCS를 활용하여 교육훈련기관에서는 훈련과정과 직업능력의 개발, 훈련기준에 맞는 교재 개발 등에 NCS를 활용하여 산업 수요에 맞는 인재를 양성하는 것에 의의가 있다.

둘째, 기업은 NCS를 활용하고 종업원의 경력 개발의 도구, 직업 능력의 자가 진단 도구로써의 가치, 승진 · 배치 등에 관련 역량을 체크하는 직무기술서로서 활용하도록 권고하고 있다.

셋째, 자격 관련 시험 기관(산업인력공단 등)에서 NCS를 활용한 출제 기준을 설정하고 직업별 훈련기준을 제 · 개정하는 지표로 활용하도록 기회를 제공하고 있다. 구체적 추진 방향으로는 일학습병행제도(대학교와 기업이 연계한 현장실무 교육의 교육 인정) 운영과 NCS 기반 새로운 직업 자격의 개편, NCS 기반 성과중심 임금과 직무중심 임금체계의 개편, NCS 기반 인사관리 전반(채용, 승진 등)에 대한 체계 구축에 있어서 NCS의 역할은 매우 주요한 영향을 주고 있다.

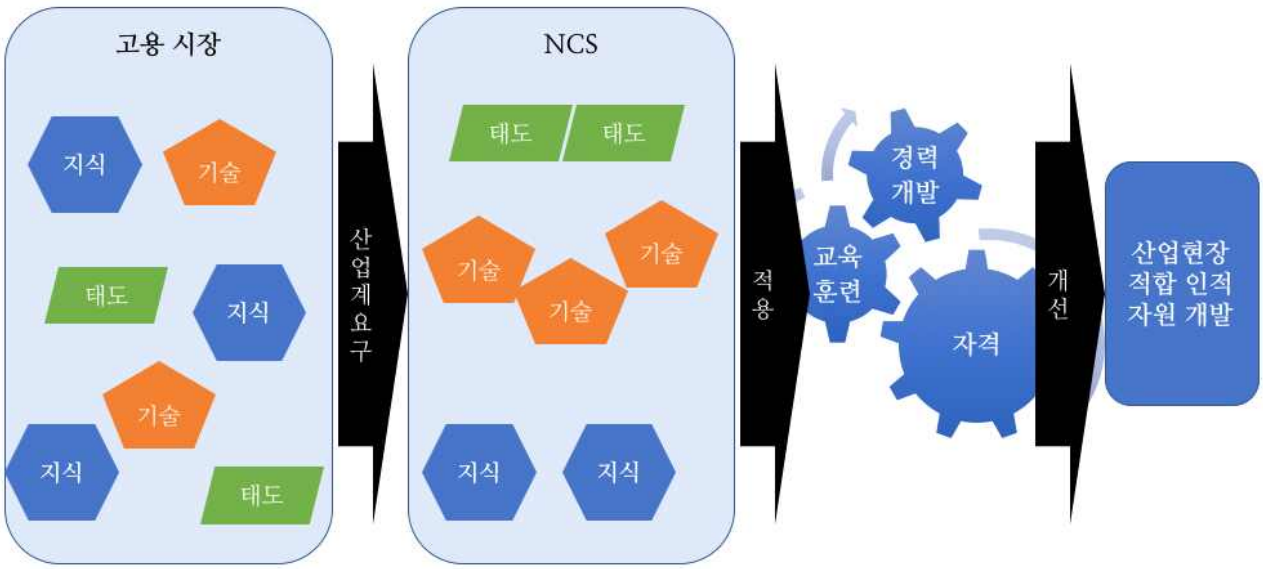
<표 2-17> NCS의 활용 영역

구분		활용 콘텐츠
산업 현장	근로자	직업능력자가진단, 평생경력개발
	기업	직무기술서, 채용 · 승진 · 배치 체크리스트
교육훈련기관		훈련기준, 훈련교재, 교육훈련과정
자격시험기관		시험문항, 시험방법, 출제기준, 자격종목 설계

출처 : 한국고용정보원, 2015

NCS의 개념도는 [그림 2-25]와 같다. 산업 현장에서 얻을 수 있는 지식, 기술, 태도 등에 대한 정보를 토대로 산업계가 요구하는 요구사항을 정리하고, 국가직무능력표준에 각각 직무에 대한 핵심적인 지식, 기술, 태도를 재정립하여 실제 현장에 접목시켜 관련된 교육훈련을 개발하고, 자격 제도를 만들고, 해당 내용을 토대로 경력을 쌓을 수 있는 연수 환경을 조성하여 기존의 체계적이지 못했던 고용 시장에서의 필요한 역량을 확보하고, 일반 종업원에게는 공급을 원하는 기업을 쉽게 매칭 받고 기업에게는 수요에 따른 공급을 원활하게 진행할 수 있는 하나의 매개적 지표를 설정하는데 의의가 있다.

[그림 2-25] 국가직무능력표준 개념도



출처 : 한국고용정보원, 2015

NCS에 대한 특성은 [그림 2-26]과 같다. 첫째, NCS는 한 명의 종업원이 해당 직업 안에서 요구되는 과업을 성공적으로 이행하기 위해 요구되는 수행능력을 말한다.

둘째, NCS는 당면한 과업을 수행하기 위하여 모든 방법의 수행능력을 포괄적으로 설명하고 있다. 즉, 과업에 대한 작업능력(특정한 과업을 수행하기 위하여 부여되는 능력), 작업에 대한 관리 능력(다채로운 직업들을 계획하고 조직하는 능력), 돌발 상황 발생 시 대처 능력, 과업의 미래 지향적인 능력(해당 과업의 기술적 환경을 분석하고 시기적절하게 조정해 나가는 능력)이 모두 포함된 개념이다.

셋째, NCS는 모듈(Module)의 형태로 구성되어 있으며, 하나의 직업 안에서 종업원이 수행해야하는 역량들의 직무능력을 유닛화하여 개발하며, NCS 세분류 직무 역량은 다수의 능력단위의 집합으로 구성한다.

넷째, NCS는 산업 분야의 단체들이 주도적, 능동적으로 참여하여 자신의 분야에 대한 전문적인 지식과 현장의 실제적으로 적용 가능한 직무에 대한 기술하는 것을 원칙으로 하며, 현장에서 매우 뛰어난 종업원 및 전문가를 초빙하여 개발한다.

다섯째, NCS 개발 능력단위 요소는 8개 단계의 분류체계를 원칙으로 하며, 수준에 따라서 제시되었다. 수준체계는 현장에서 알 수 있는 현장에서의 업무별 난이도를 중심으로 NCS를 최종적으로 구성하며, '자격-교육훈련-산업현장'의 연계를 기반으로 평생학습과, 자격의 수준체계를 구성하는 것에 활용 지표로 사용되고 있다.

[그림 2-26] NCS분류체계 구성도(예시)



출처 : 한국고용정보원, 2015

NCS의 분류체계 구조는 <표 2-18>과 같다. 분류체계 구성도를 보면 여타 분류체계랑 유사하게 대분류 안에 해당 중분류가 중분류 안에 소분류가, 소분류 안에 세분류가 존재한다. NCS는 4가지의 단계로 구분되어 있으며 대분류에는 총 24개, 중분류 80개, 소분류 238, 세분류 887개로 구성되어 있다.

<표 2-18> NCS 분류체계 구조

대분류명	대분류	중분류	소분류	세분류
계	24	80	238	887
01. 사업관리	1	1	2	5
02. 경영 · 회계 · 사무	1	4	11	27
03. 금융 · 보험	1	2	9	35
04. 교육 · 자연 · 사회과학	1	3	5	13
05. 법률 · 경찰 · 소방 · 교도 · 국방	1	2	4	15
06. 보건 · 의료	1	2	7	34
07. 사회복지 · 종교	1	3	6	16
08. 문화 · 예술 · 디자인 · 방송	1	3	9	63
09. 운전 · 운송	1	4	7	26
10. 영업판매	1	3	7	17
11. 경비 · 청소	1	2	3	6
12. 이용 · 숙박 · 여행 · 오락 · 스포츠	1	4	12	42
13. 음식서비스	1	1	3	9
14. 건설	1	8	26	109
15. 기계	1	10	29	115
16. 재료	1	2	7	34
17. 화학	1	4	11	32
18. 섬유 · 의복	1	2	7	23
19. 전기 · 전자	1	3	24	72
20. 정보통신	1	3	11	58
21. 식품가공	1	2	4	20
22. 인쇄 · 목재 · 가구 · 공예	1	2	4	23
23. 환경 · 에너지 · 안전	1	6	18	49
24. 농림어업	1	4	12	44

출처 : 한국소프트웨어산업협회, 2017

NCS의 대분류에는 사업관리부터 농림어업까지 존재하며, SW인력과 관련된 직무는 20. 정보통신이며 정보통신은 중분류인 01. 정보기술, 02. 통신기술 03. 방송기술로 나눠진다. 정보기술의 소분류, 세분류는 <표 2-19>와 같다.

<표 2-19> 20. 정보통신-01. 정보기술 SW인력 관련 직무

NO.	대분류	중분류	소분류	세분류	능력단위수	Lv. 5 미만
	총계				51	22
1	20. 정보통신	01. 정보기술	01. 정보기술 전략·계획	01. 정보기술전략	10	0
2				02. 정보기술컨설팅	11	1
3				03. 정보기술기획	10	6
4				04. SW제품기획	10	3
5				05. 빅데이터분석	9	1
6				06. IOT융합서비스계획	11	1
7			02. 정보기술 개발	01. SW아키텍쳐	11	1
8				02. 응용SW엔지니어링	25	4
9				03. 임베디드SW엔지니어링	18	5
10				04. DB엔지니어링	12	2
11				05. NW엔지니어링	11	1
12				06. 보안엔지니어링	13	2
13				07. UI/UX엔지니어링	10	4
14				08. 시스템SW엔지니어링	9	0
15			03. 정보기술 운영	01. IT시스템 관리	11	1
16				02. IT기술 교육	11	2
17				03. IT기술 지원	12	3
18			04. 정보기술 관리	01. IT프로젝트 관리	14	0
19				02. IT품질보증	11	0
20				03. IT테스트	17	1
21				04. IT감리	11	0
22			05. 정보기술 영업	01. IT기술영업	10	2
23				02. IT마케팅	10	0
24			06. 정보보호	01. 정보보호관리·운영	15	1
25				02. 정보보호진단·분석	10	0
26				03. 보안사고분석대응	8	1

출처 : 소프트웨어산업협회, 2017

5. IT 분야 역량인정체계 (ITSQF)

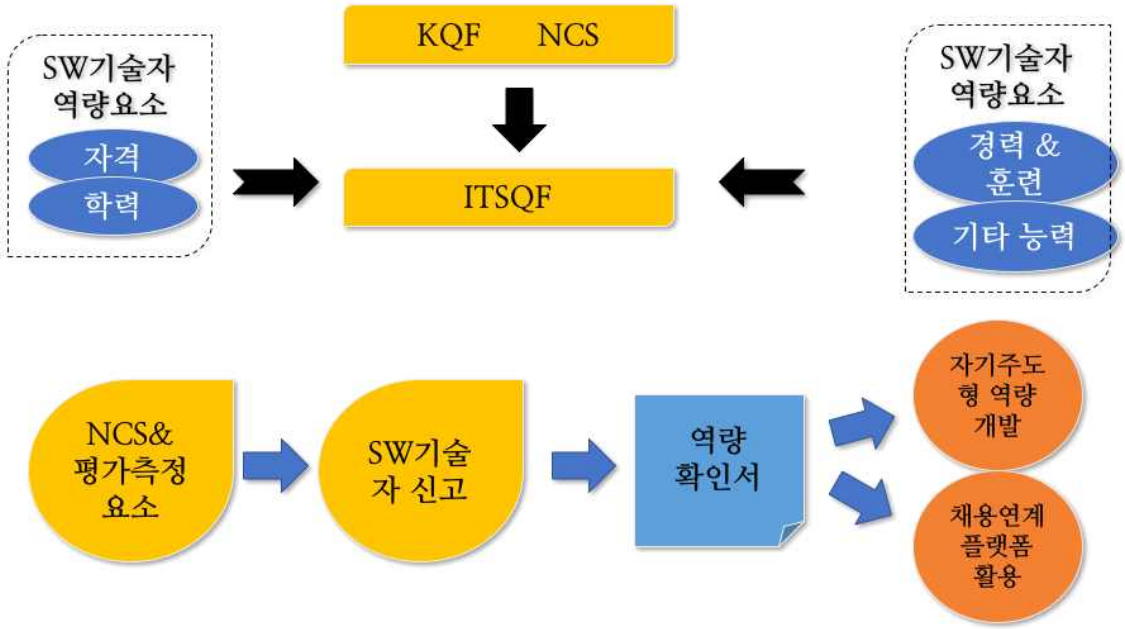
4차 산업의 발전으로 소프트웨어 기술자의 역량과 시장의 변화가 가장 크게 변화하고 있다. 그러나 산업 고용 현장에서는 여전히 기존 지표를 활용한 경력과 학력 중심의 채용이 일어나고 있는 추세이다. 특히, IT 산업에서 활용되고 있는 SW기술자간의 등급 체계가 한 가지 방법으로 체계화 되어 있어 다양한 직무를 반영하기 어려우며, 객관적 역량측정들의 요소가 부족하다. 이를 보안하기 위하여 IT 분야의 역량중심의 직무체계 ITSQF(IT Sectoral Qualifications Framework) 구축과 산업 현장으로의 확산이 필요하다.

산업별 역량인정체계(SQF)란 능력중심사회를 구현하기 위하여 학력중심의 역량체계에서 다양한 요소들로 구성된 역량 체계이다. 기존 학력 외에 개개인의 직무능력을 다양한 방식으로 인정받을 수 있는 산업현장 수준에서의 통합적이고 구체적인 자격에 대한 상호인정 틀을 기반으로 체계가 구축되었다.

SQF는 특정 산업분야에서 자격, 교육, 훈련, 현장경력 등을 연계하는 일련의 역량평가의 인정의 틀로 볼 수 있으며, 수직적-수평적 능력인정체계를 말한다. 특정 직무에서 요구되는 능력은 국가직무능력표준(NCS)로 표현하여 기술, 지식, 자율성과 책임성으로 구분하여 정의한다. NCS는 훈련에서는 훈련 목표가, 자격에서는 평가기준이, 고용시장과 교육훈련 및 자격을 연계하는 매개적 역할을 도모한다.

ITSQF 구축 목적은 기존 분류체계와는 다르게 다양한 능력의 반영으로 SW기술자 역량을 공정하게 인정받을 수 있는 생태계를 조성하는 것이다. 특히 SW기술 우수인재 발굴, 객관적인 역량 평가, 부족 역량 진단 등의 목적으로 사용된다. 그에 따라 KQF의 자격-학력-경력 등의 상호 인정 체계 및 연계 기준 등의 KQF의 원리와 정신을 반영하여 역량체계를 구성하고 NCS의 표준화된 능력단위 및 요소 등을 활용하여 산업 직무역량체계를 구성한다.

[그림 2-27] IT분야 역량체계의 개념



출처 : 한국소프트웨어산업협회, 2017

2017년까지의 추진 상황을 보면 직무체계를 새롭게 검토하고 직무기술서를 완성하였다. 또한 역량 인정을 위한 준비를 바탕으로 ITSQF 현장에 적용하고 있다.

직무기술서는 직무체계를 면밀히 검토하고 직무기술서 작성을 위하여 기존의 산업계 전문가 회의를 실시하여 고용 현장에서 필요한 실제적인 직무에 대한 검토가 이루어졌다. 이 중 ITSQF의 실행력 강화를 위한 TF팀(과학기술정보통신부, 한국IT서비스 산업협회, 대기업/중견기업의 담당자, 대학교수 등으로 구성)을 구성하여 직무체계에 대한 검토를 통하여 직무기술서를 구축하였다.

상기 언급한 NCS를 기반으로 하여 특정 직무를 특정 수준으로 구성하여 능력단위를 다양한 방법으로 이수를 하면 그에 상응하는 역량수준을 인정받는 한국형 국가역량체계의 기본원리를 반영하여 <표 2-20>가 설계되었다. ITSQF에서 SW기술자는 9개의 역량 수준으로 구분되고, 학교 졸업과 자격 취득으로 인하여 SW 관련 업무에 대한 직업적 성장의 체계적 흐름을 나타낸다.

<표 2-20> ITSQF 수준표

역량 수준			1. 현장경력 & 교육훈련	2. 학력	3. 자격
ITSQF 수준		NCS 수준	수준상향조건	학력	NCS기반자격/국가기술자격
ME	마스터	L8	EE+8년		
EE	특급		A8+7년+50시간 교육		
A8	고급		I7+3년+50시간 교육	유관 박사	IT관련기술사
I7	중급	L7	I6+3년		
I6	중급	L6	J5 +3년	유관 석사	NCS기반자격_L6
J5	초급	L5	J4+1년	유관 학사	NCS기반자격_L5 / IT관련기사
J4	초급	L4	J3+1년	유관 3년제 전문학사	
J3	초급	L3	B+2년	유관 2년제 전문학사	NCS기반자격_L3 / IT관련산업기사
B	기능사	L2		고교 졸업	IT관련 기능사

출처 : 한국소프트웨어산업협회, 2016

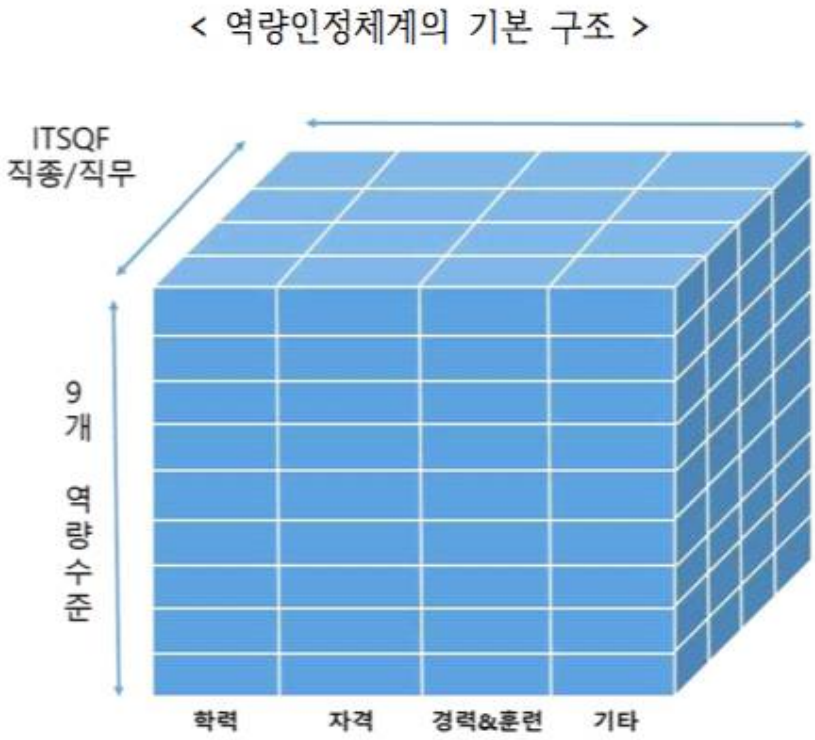
IT 분야의 역량인정체계의 특징으로는 SW기술자의 다양한 능력들을 평가하고 역량으로 인정 할 수 있는 조건과 환경을 만들기 위하여 설계되었으며, 기술 · 능력 · 학력, 자격 등과 관련하여 현장 경험 및 훈련과 대회 수상 경력 등 IT 관련 역량을 증진시킬 수 있는 활동의 총체적 측정에 의해 수행된다.

그에 따라 ITSQF는 정보기술 분야 전반에 직무에 대한 특성을 반영하고 있다. 정보기술 분야는 특성상 직종/직무 간의 연계성이 높고, 직종/직무별 역할에 대한 대체가 쉽게 일어나는 특징을 갖는다. 따라서 직종/직무 간 필요한 역량 차이가 반드시 존재하며, 직무를 전환하여 상위 수준으로 올라가는 수순을 주로 따라가는 형태를 취한다.

[그림 2-28]과 같이 ITSQF의 역량 인정체계의 기본적인 원리는 직종과 직무 분류체계와 역량 수준이라는 2가지 차원의 축으로 구성되어 있다. 이 중 기술자 역량은 교육 · 훈련 · 자격 등의 다양한 방법으로 충족된다. ITSQF에서는 자격,

현장 경력, 학력 및 교육훈련의 4가지 차원과 대회수상과 현장 경력 등의 내용을 추가하여 기술자간 차별화된 역량에 더하여, KQF의 상호연계성의 특성을 부여하여 종합적으로 합리적인 직무 역량을 파악하는 방식으로 구성되었다(한국소프트웨어산업협회 2016).

[그림 2-28] ITSQF 역량의 기본적 구조



출처 : 한국소프트웨어산업협회, 2016

따라서 상기 ITSQF는 기본적 구조를 토대로 직종과 직무 분야별로 자격, 현장경력, 학력, 교육훈련과 더불어 대외실적 등의 다양한 지표를 추가하여 명확한 규칙을 형성하고 소프트웨어 기술자의 역량 체계를 NCS 분류체계와 연계하여 구성했다고 판단할 수 있다.

[그림 2-29] ITSQF 직무군별 소프트웨어기술자 신고현황



[그림 2-29]는 ITSQF 체계에 대하여 직무별로 SW기술자에 대한 현장 기술자의 신고현황이다. 내용을 보면 SW엔지니어의 비중이 약 50%에 육박할 정도로 높다. 그 외에도 기술자 역량 수준을 직무와 직종별로 파악하기 위하여 NCS를 기반으로 직종을 정의하였고 직종에 따라 파생되는 직무가 설정되어 있다. 그에 따라서 <표 2-21>과 같이 KQF와 NCS를 연계하여 총 12개의 직종으로 구성되었다.

<표 2-21> ITSQF 직무역량체계

ITSQF 직종	직종 정의	직무
IT컨설턴트	조직의 최종 경영목표를 달성하기 위하여 정보기술(IT)을 효율적으로 사용 할 수 있는 각종 전략들을 파악하고 수행하기 위한 구체적인 모형과 운영 정책을 모색한다.	정보기술기획자 정보기술컨설턴트 정보보호컨설턴트
IT프로젝트 관리자	IT관련 프로젝트를 진행함에 있어 해당 프로젝트의 산출물의 납기를 준수하며, 프로젝트를 기획하고 관련된 의사소통, 정보시스템, 보안, 성과, 조달, 변경, 품질, 인적자원, 위험 등을 총체적으로 관리한다.	IT프로젝트 관리자
SW아키텍트	소프트웨어의 목적을 성취하기 위하여 각종 업무에 대한 시나리오를 설계하고 요구 분석을 통하여 관련	SW아키텍트 빅데이터분석

	정보를 취득하고 자료를 시각화를 하기위한 UI/UX의 구조 설계 및 기획한다.	가 UX/UI기획자 업무분석가
SW엔지니어	소프트웨어 프로그래밍 언어로 제각각의 소프트웨어의 목적성에 맞춘 소프트웨어 기능을 설계 및 구현하고, 정보에 대한 시각화를 위한 UX/UI를 구축하며, 구현이 되고 나서 단위당 테스트를 통하여 실제적으로 사용자가 사용할 수 있도록 배포하며, 소프트웨어의 버전을 관리하고 각종 기술지원을 통하여 소프트웨어의 제품의 성능을 향상시킨다.	빅데이터개발자 UX/UX퍼블리셔 응용SW개발자 시스템SW엔지니어
임베디드SW 엔지니어	하드웨어 등의 플랫폼에 대한 이해를 바탕으로 플랫폼별 운영체제를 이식하고 펌웨어와 디바이스 드라이버의 관리와 함께 어플리케이션 등의 임베디스 소프트웨어를 구축하고, 하드웨어 플랫폼에 구현하여 최적화시킨다.	임베디드SW 엔지니어
DB엔지니어	데이터베이스를 설계 및 구축할 업무를 파악하고 데이터베이스에 대한 효과 및 보안 등의 안정성을 확보하여 데이터 자체의 물리구조를 설계하고 데이터 규모를 산정하여 최적화 된 데이터베이스 배치를 이뤄낸다. 데이터 베이스 및 온라인상에 클라우드와 같은 기술을 파악하여 트렌드를 분석하고 소프트웨어를 변경 및 운영을 관리 및 통제한다. 데이터베이스의 문제발생시 수정 및 보안, 백업 등의 업무를 진행한다.	DB엔지니어 데이터아키텍트
NW엔지니어	네트워크의 전반적인 시스템의 구조를 파악 및 분석하여 문제점과 보안 사항을 찾아 네트워크 설계에 반영하고, 사용자가 요구하는 사항에 알맞은 네트워크의 기본적 시스템을 구축하여 안정성, 확장성, 기능성, 변경 가능성, 관리 용이성을 확보할 수 있는 네트워크를 설계하고 구현한다. 구현이 완료된 네트워크 시스템에 대해서 운영 기간 동안은 보장된 서비스를 제공하며, 네트워크의 성능을 향상시켜 사용자가 용이하게 사용될 수 있도록 준비한다. 또한 네트워크와 관련되어 클라우드 서비스 등의 현재 추세를 파악하여 추세에 맞는 서비스를 제공한다.	NW엔지니어
시스템 엔지니어	시스템 자체에 대한 연구 · 분석 · 설계를 하는 전문가로써 컴퓨터의 설치 및 폐지와 운영 전반에 대한 기술적인 상담과 엔지니어링을 담당한다. 다양한 전산처리 시스템을 운영하며, 서버 컴퓨터에 대한 운영체제에 대한 지식을 통하여 프로그래밍 언어를 활용하여 하드웨어, 데이터베이스 등을 운영 관리하고 전체적인	IT시스템관리자 IT지원기술자

	전산지식으로 일반 업무에 시스템 적으로 활용 할 수 있는 전반적인 시스템을 구축 운영한다. 컴퓨터 업계에 대한 시장 동향을 파악하여 시스템에 대한 지속적이고 효율적인 운영을 할 수 있도록 지원한다.	
IT마케터	정보기술 및 마케팅에 대한 전반적인 지식을 바탕으로 고객을 관리하고 정보기술과 관련된 마케팅 기술을 확보하여 영업기회를 만들어내고 고객의 요구에 적합하게 솔루션을 제안하고 계약, 판매, 협상, 판매 및 사후관리를 수행한다. 또한 새로운 소프트웨어의 개발을 위하여 내부 외부의 환경적 요구, 시장성 등을 파악하여 소프트웨어 제품의 로드맵 및 사용을 수립하고 소프트웨어 제품의 개발, 판매, 지원, 마케팅 계획을 수립하며, 개발된 솔루션을 사용자에게 제공하기 위한 전략을 구축하여 가격, 제품, 판촉, 유통, 고객을 관리한다.	SW제품기획자 IT기술영업 IT서비스기획자
IT보안 엔지니어	IT에 관련한 보안 이론을 토대로 실무능력을 가져 정보 자산을 보호하고 지키기 위한 물리적이며 논리적인 보안 계획을 형성하고 위험을 평가하며, 요구된 사항에 따라 보안의 체계를 구축 및 운영하여 기업의 IT 시스템 및 자료 정보를 내부 외부로부터 지키는 역할을 하며, 반출, 불법적인 접근 및 노출을 방지한다.	정보보안전문가 침해사고대응 전문가
IT품질관리자	정보시스템의 효율을 증가시키기 위하여 안정성을 확보하고 정보시스템의 구현 및 운영을 통하여 종합적이고 체계적으로 점검하고 문제점을 개선 할 수 있도록 하는 감리의 업무를 수행하며, 품질 보증을 위하여 소프트웨어의 각종 테스트를 담당하여, 이에 관련한 필요한 기획 및 진단 컨설팅, 분석, 설계, 계획, 환경조성 등의 실행을 통하여 자동화를 수행하고 관리한다.	IT품질관리자 IT테스터 IT감리원
IT교육 강사	IT 분야의 기술과 관련된 교육을 체계화하고 재정립하여 효율적으로 훈련을 받을 수 있는 환경을 구성하고 교육 훈련을 통하여 소프트웨어와 정보기술 관련 분야에 필요한 인재를 육성하는 전략을 수립, 운영, 기획하는 역할을 도모하여 교육 체계를 확립하고 그에 맞는 프로그램을 기획, 개발하여 운영하고 필요한 교수자를 수급함으로써 소프트웨어 교육환경의 질을 증가시키는 역할을 담당한다.	IT교육강사

출처 : 한국소프트웨어산업협회, 2017

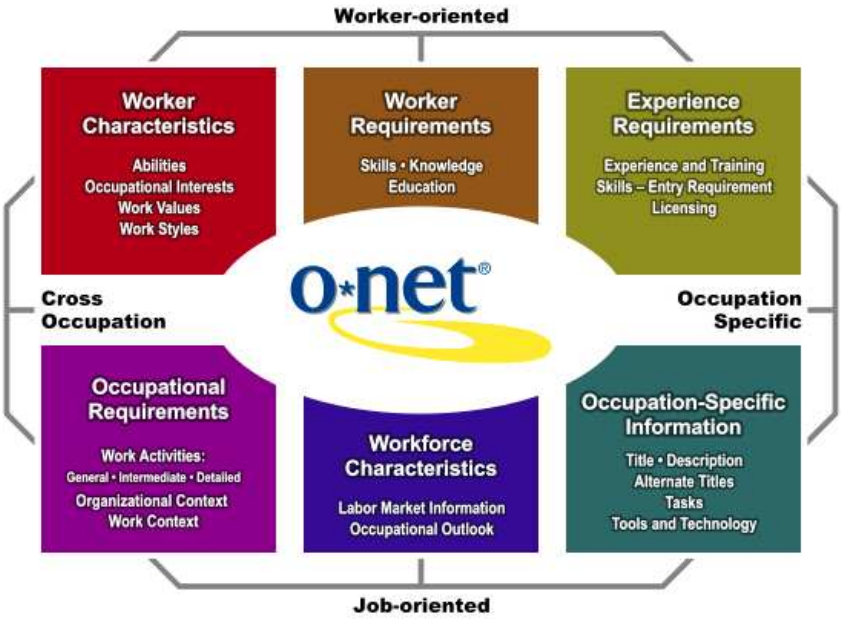
6. O*net (미국직업정보체계)

미국직업정보체계인 O*net 프로그램은 국가 차원에서 직종의 이해를 돕고 사용자들이 활용하기 편하게 하도록 미국 내에서 일련의 체계를 만든 것을 말한다. 해당 데이터는 급변하는 직종의 성격과 노동의 방식, 비중에 대하여 미국 경제의 전반에 미치는 영향을 이해하기 위한 중요한 기초 자료로 사용되고 있으며, 해당 정보를 바탕으로 고용시장에서 숙련된 SW인력을 개발 및 유지를 용이하게 활용하기 위한 응용 프로그램으로 개발되었다.

해당 프로젝트의 중심에는 인력 데이터베이스가 있으며, 미국 내의 전체 경제를 망라하는 약 1,000개의 직종에 대하여 표준화하고 직업별로 상세한 설명이 포함되어 있다. 또한 무료로 웹 사이트에서 제공하고 있어 고용 시장을 파악하는데 용이하며, 현업에 있는 각 직종의 광범위한 종업원들이 지속적으로 자료를 업데이트 하고 있다.

O*NET의 정보는 O*net online, My network move 및 기타 공공기관과 민간 개발 응용프로그램을 활용하여 매년 수백만 명 이상이 활용하는 프로그램이다. 제공되는 데이터는 사용자가 필요한 교육 및 일자리를 찾도록 직관적으로 정보를 제공하고 있으며, 기업의 고용주 역시 시장에서 경쟁력을 갖출 수 있는 핵심 조건 중 하나인 역량이 풍부한 종업원을 찾는 데 중요한 징검다리 역할을 하고 있다. 특히 O*NET에서 제공하고 있는 콘텐츠 모형은 개념적 토대가 되며 직업에 대한 중요한 유형의 정보를 식별하고 이론적으로나 경험적으로 튼튼한 프레임워크를 구축하고 있다고 판단할 수 있다.

[그림 2-30] O*net Online에서 제공되고 있는 O*net 콘텐츠 구성도



출처 : O*net Online website[1)]

[그림 2-30]과 같이 O*net의 콘텐츠 프레임워크로 제시하고 있는 콘텐츠 구성도는 직업 및 조직 분석에 대한 연구를 활용하여 개발되었다. 직업 관련 전문가 회의를 통하여 직업의 성격, 분류, 근로자 중심의 기술서 등의 보다 현실적인 성격을 반영하여 프로세스를 구현하였다. 콘텐츠 프레임워크는 직업 정보가 직업, 업종, 직종 간에 적용될 수 있도록 구현하여 현장에 적용될 수 있게 한다. 그에 따라서 프레임워크는 총 6가지의 주요 영역으로 구성되어 있어 정보를 사용하는 사용자는 종업원의 직종별 주요 특성을 파악하는 정보를 제공하고 있다.

1. 근로자의 성격(Worker Characteristics) - 효과적인 업무 수행에 필요한 지식과 기술을 습득 할 수 있는 역량과 역량에 영향을 줄 수 있는 지속적인 특성
2. 작업자 요구사항(Worker Requirements) - 직종과 관련된 경험 및 교육을 통하여 획득되거나 개발된 작업 관련된 속성을 설명하는 부문
3. 직업별 정보(Occuapation-Specific Information) - 선택된 직종이나 특정

1) O*net Online 홈페이지 : https://www.onetonline.org/

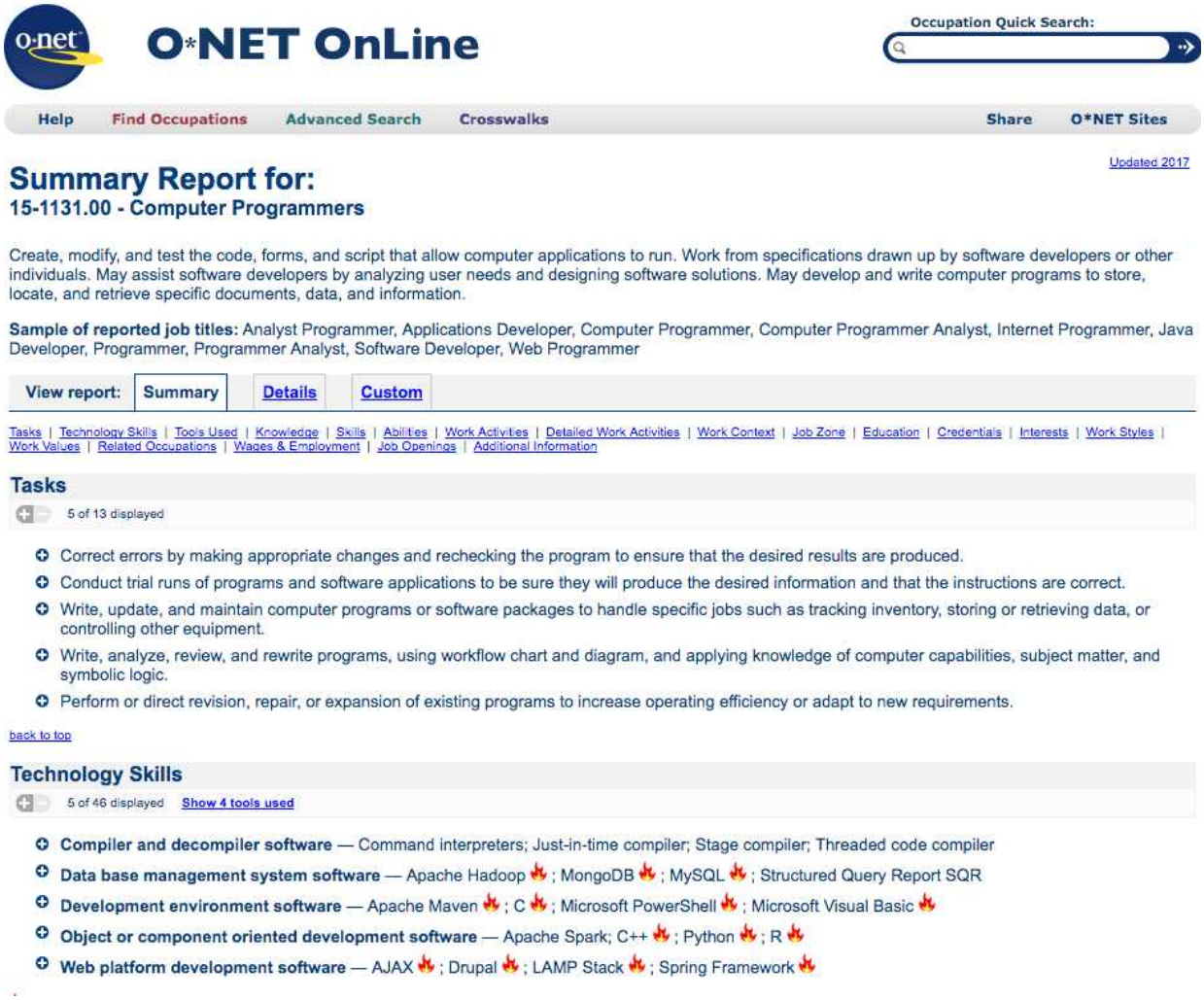
직종의 변수 또는 기타 콘텐츠 모델 요소

4. 노동력 특성(Workforce Characteristics) - 직업과 관련된 요구 사항에 영향을 줄 수 있는 직업의 일반적인 특성을 정의하고 설명하는 요소
5. 직업 요구(Occupational Requirements) - 다양한 직업에 필요한 요소들을 설명하는 변수 혹은 세부 요소들의 포괄적인 집합

O*net 온라인에서는 상기 언급된 모델을 기반으로 사용자에게 정보를 제공하고 있으며 제공되는 정보는 다음과 같다.

1. 35개의 기술(Skills)
   (1) 기본 기술 (수학, 작문, 독해력 등)
   (2) 교차 가능 기술 (장비 선택, 품질 관리 분석 등)
2. 일반적인 업무 활동 - 여러 작업 상에 발생하는 41가지의 일반적인 유형의 행동(업무 조직, 계획 및 우선순위 지정, 컴퓨터와의 상호 작용 등)
3. 흥미(Interests) - 어떠한 직종에서 가장 성취가 잘되는가를 파악하기 위하여 종업원의 개인적 차원의 이익과 관련된 6가지 흥미 유형
4. 업무 스타일(Work Style) - 구상, 지속성, 협력 등 직종상 가치를 발전시키거나 종업원에게 중요한 역량을 연결시키는 16가지 작업 스타일
5. 업무 맥락(Work context) - 물리적 및 구조적 직종의 특성과 같은 업무의 본질에 해당하는 57가지 물리적 및 사회 구조적 요인
6. 경험과 훈련(Experience and Training) - 직업과 관련된 교육 및 훈련의 수준을 구별하는 5개의 경험 영역

[그림 2-31] O*net Online에서 제공되고 있는 직업 정보 예시(Computer Programmers)



SW인력에 맞춘 정보는 Information Support and Service, Network System, Programming and Software Development, Web and Digital Communcations 등의 업무 맥락에 따라 직관적으로 분류되어 있다. 각 직종은 분류별 코드 부여와 세분류 항목이 제공되고, 상기 언급한 콘텐츠 프레임워크를 활용하여 각 세분류 항목별 직종이 설명되고 있다.

<표 2-22> O*net Information Technology 직종 분류

대분류	직업 코드	세분류
Information Support and Service	15-1199.08	Business Intelligence Analysts
	15-1199.00	Computer Occupations. All other
	15-1199.02	Computer User Support Specialists
	15-1151.00	Data Warehousing Specialists
	15-1199.07	Database Architects
	15-1199.06	Document Management Specialists
	15-1199.12	Geographic Information Scientist and Technologists
	15-1199.05	Geospatial Informations Scientists and Technologists
	15-1199.04	Information Technology Project Managers
	15-1199.09	Search Marketing Strategists
	15-1199.10	Software Quality Assurance Engineers and Tseters
	15-1199.11	Video Game Designers
	15-1199.03	Web Administrators
Network Systems	15-1143.00	Computer Network Architects
	15-1152.00	Computer Network Support Specialists
Programming and Software Development	15-1141.00	Computer Programmers
	15-1122.00	Computer Systems Analysts
	15-1142.00	Informatics Nurse Specialists
	15-1143.01	Software Developers, Applicatoins
	15-1131.00	Software Developers, Systems Software
Web and Digital Communication	15-1134.00	Web Developers

출처 : O*net Online

7. SFIA (The Skills Framework for the Information Age)

SFIA는 2000년에 설립되어 정기적으로 업데이트되어 현재 7레벨까지 능력 표준을 구비해 놓은 직업 표준 중 하나의 지표이다. SFIA에서 활용하고 있는 협업 개발 기술 스타일(Collaborative development Style)은 기업 등의 현장과 교육 훈련 환경에서 역량 관리의 실질적인 경험을 갖고 있는 전문가들로부터 열린 회의를 통하여 의견을 수렴하고 국제적인 표준을 제정하였다.

SFIA는 다른 이론적 방식과는 차별적인 부분이 존재하는데 위와 같이 현업에서 인사(승진, 채용), 교육 훈련 등의 전문가에 대한 담화를 통하여 발견된 합의적 직종이기 때문에 이미 약 2000개의 국가에서 SFIA 표준을 준수하고 있다.

SFIA는 2000년대 영국의 이니셔티브(UK Initiative) 기업을 선두로 시작되었다. IT 산업을 위한 역량과 관련한 능력의 결정적인 프레임을 제공하기 위하여 공동 프로젝트를 실시 및 구축하였다. 그 결과 현재는 몇 가지 사용자 개인의 기술 역량의 전략적 요소로 자리매김 하였다.

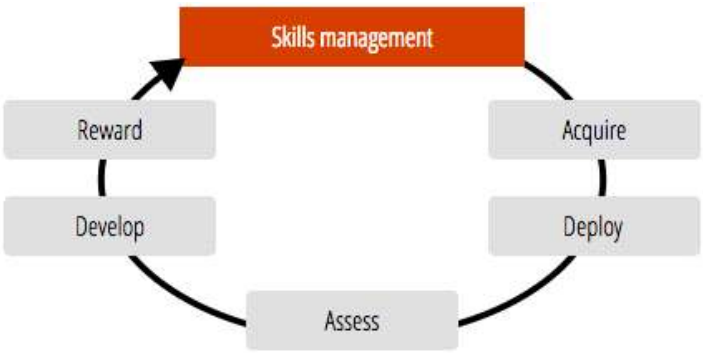
2000년 이래 SFIA는 글로벌 IT 기술 프레임 워크의 표준이 되었다. 그 이유는 기업 조직 및 사용자 개인이 역량 및 갖고 있는 관련된 기술을 간단하고 포괄적으로 작성하였기 때문이며 광범위한 글로벌 네트워크 기반으로 지속적으로 역량이 갱신되고 있다. SFIA 프레임워크는 사용자를 지원하는 자금까지 구축된 에코 시스템으로 독립적이고 자립적인 역량 정보를 제공하고 있다. SFIA의 역량 정보는 사용자의 요구와 정보 수명이 짧아지는 인력시장에서 더 빠른 주기로 갱신된다.

SFIA 역량 정보는 개인과 조직에게 기술, 능력 및 전문 지식을 일관된 방식으로 정의하고 기술 전문 용어 및 약어를 사용하지 않고 누구나 알 수 있는 쉬운 언어로 해석해 놓음으로써 누구든지 인적 자원에 대한 이해부터 전문적인 개발전문가들이 해당 내용을 수용할 수 있다. 따라서 영문 기반으로 되어있지만 번역 문제로 인한 오역 등에 문제를 최소화 가능하다. 따라서 SFIA는 기업의 비즈니스의 요구 사항을 설명하고 요구 사항을 충족시킬 수 있는 역량 있는

인재를 적재적소에 뽑기 위해 인력을 평가하는데 있어 핵심 지표로 사용이 가능하다.

[그림 2-32]와 같이 SFIA에서 제공되는 역량 정보는 교육 기관, 대학교에서도 수준 있을 기술을 보유하고 있는 인재양성에 도움을 준다. 추가적으로 실제 현장에서 요구하는 역량에 대한 정보를 파악할 수 있기 때문에 SFIA는 개인이 '나의 직업이 아닐 것이라고 생각하는 경우'에 있어서도 업무를 수행하는데 올바른 의사결정과 기술을 함양할 수 있게 직관적인 가이드를 제공한다. 따라서 교육 기관에서 SFIA를 활용하는 것은 학생들에게 올바른 수준의 지식을 전달 할 수 있는 단계적 지표가 된다는 점에서 활용성이 매우 높은 국제 표준이다.

[그림 2-32] SFIA 역량 관리 프레임워크



출처 : sfia-online.org[2)]

SFIA 역량관리 프레임워크는 역량의 통합 관리를 위하여 모든 시점에서 일관적인 평가 방식을 채택한다. SFIA의 역량관리에는 총 7단계의 난이도(Level)에 [그림 2-33]과 같이 따라서 각 위치에 따른 역량을 구분하였다.

2) SFIA SW역량 소개 홈페이지 : sfia-online.org

[그림 2-33] 역량 난이도별 전문 역할 및 조직 역할

Professional role	SFIA Levels of responsibility	Organizational role
Practice Lead	L7. set strategy, inspire, mobilize	Senior Management
	L6. initiate, influence	
Subject Specialist	L5. ensure, advise	First Line Management
	L4. enable	Team Leader
Skilled Professional	L3. apply	Experienced Team Member
On learning curve	L2. assist	On learning curve
	L1. follow	

출처 : sfia-online.org

Level 7. 가장 높은 조직 수준에 있는 역량으로 기업 전체의 정책 수립 및 적용을 포함하여 중요한 업무 영형의 모든 측면에 대하여 권한이 부여된다. 책임의 범위가 넓기 때문에 대부분의 의사결정을 진행하여 의사결정에 책임이 매우 크다.

Level 6. 기술, 재무 및 품질 관리를 포함하는 중요한 업무 영역에서 하위 종업원들의 행동 및 결정에 대한 권한 및 책임을 갖고 있으며 조직의 중장기적 목표를 수립하고 책임을 하위 종업원들에게 할당하는 역할을 한다.

Level 5. 기업의 업무의 대부분의 분야에서 일을 창출하며, 프로젝트 생성 및 해당 프로젝트의 감독과 프로젝트 목표에 대한 성과관리를 책임진다. 이정표를 수립하여 하위 종업원에게 과제 및 역할을 부여하는 역할을 한다.

Level 4. 책임이라는 명확한 틀 안에서 일반적인 업무를 진행하며, 실질적이며 개인적인 책임과 자율권을 행사할 수 있다. 계획은 상급 종사원에게 받은 목표와 프로세스를 충족하기 위하여 과업을 진행하는 역할을 한다.

Level 3. 복잡한 문제를 해결하고, 과업을 수행하며, 구체적인 사항을 상급자를 통하여 지시 받아 과업을 수행한다. 수행 일정에 대한 검토를 지속적으로 하여, 해당 과업에 대한 재량권을 갖고

업무를 수행한다.

Level 2. 일반적인 문제와 과업을 수행하며, 제한된 재량권을 사용하여 과업을 진행한다. 다른 사람과의 협업보다는 자신에게 주어진 일에 전념하는 특징을 갖는다.

Level 1. 상급자의 감독 하에 과업을 진행한다. 매우 작은 재량권을 행사할 수 있으며, 불특정한 상황에서는 상급자의 지시없이 행동하지 못하는 특징을 갖는다.

SFIA는 각 역량 단위에 따라 7가지의 난이도가 부여되며, 기업 내에서의 위치에 따라 행사할 수 있는 재량권이 다르고 업무의 범위도 다르다. 따라서 기업에서는 SFIA를 통하여 자신의 위치(Position)에 따른 역할을 숙지하는데 기초자료로 활용이 가능하며, 대학 등의 교육기관에서는 난이도에 따른 역량을 교육하는데 기초적인 자료로 활용 가능하다.

SFIA는 컴퓨터와 관련된 SW 전문가들의 역량 정보를 수록한 내용이다. 아직 Level 7 단계는 개발 중에 있다. 현재까지의 Level 7까지 수록된 역량 리스트는 <표2-23>과 같다. 또한 모든 역량에 1~7의 난이도 존재하는 것은 아니며 각 역량 정보에 따라서 약간의 차이를 보이고 있다. 예를 들어 <표 2-23>와 같이 'Requirements definition and management'의 역량의 경우 Level 6, 5, 4, 3, 2로 구성되어 있으며, 'IT governance'의 경우 Level 7, 6, 5로 구성되어 있어 역량별 할 수 있는 범위의 차이에 따라 차이가 존재한다.

<표 2-23> SFIA 역량 명 리스트

SFIA 역량 명	
Analytics	Network planning
Animation development	Network support
Application support	Organisation design and implementation
Asset management	Penetration testing
Availability management	Performance management
Benefits management	Portfolio management
Business analysis	Portfolio, programme and project support

Business modelling	Porting/software configuration
Business process improvement	Problem management
Business process testing	Product management
Business risk management	Professional development
Capacity management	Programme management
Change implementation planning and management	Programming/software development
Change management	Project management
Configuration management	Quality assurance
Conformance review	Quality management
Consultancy	Quality standards
Continuity management	Radio frequency engineering
Contract management	Relationship management
Customer service support	Release and deployment
Data analysis	Requirements definition and management
Data management	Research
Database administration	Resourcing
Database design	Safety assessment
Digital forensics	Safety engineering
Digital marketing	Sales support
Emerging technology monitoring	Security administration
Enterprise and business architecture	Selling
Facilities management	Service acceptance
Financial management	Service level management
Hardware design	Solution architecture
Incident management	Sourcing
Information assurance	Storage management
Information content authoring	Sustainability assessment
Information content publishing	Sustainability engineering
Information management	Sustainability management
Information security	Sustainability strategy
Information systems coordination	System software
Innovation	Systems design
IT governance	Systems development management

IT Infrastructure	Systems installation/decommissioning
IT management	Systems integration
IT strategy and planning	Teaching and subject formation
Learning and development management	Technical specialism
Learning assessment and evaluation	Testing
Learning delivery	User experience analysis
Learning design and development	User experience design
Methods and tools	User experience evaluation
Network design	

출처 : sfia-online.org

그에 따라 SFIA는 총 95개의 역량 명칭을 통하여 소프트웨어 분야에 대한 직업들의 난이도, 능력의 종류 등을 누구나 알기 쉬운 언어로 풀이해 놓고 있어 많은 국가에서 직업 표준의 기초자료로 사용하고 있다(SFIA foundation, 2011).

본 장에서는 각 역량 지표들의 특성과 비교를 바탕으로 역량의 정의 및 수준별 구성의 중요성과 SW분야에서의 필요성을 설명하였다. 다음 3장에서는 SW산업에 대한 DTM을 이용한 분석 결과를 바탕으로 직종 및 직업별 분류를 구분한다. 또한 DTM분석을 바탕으로 도출된 직업 분류와 직업의 변화상을 확인하고 전문가 기반의 직종 및 직업 도출을 통해 직업 및 직종의 변화를 예측하였다.

제3장 분석 결과

제1절 텍스트마이닝 분석

1. SW인력 수요 정보 데이터

연구의 목적에 따른 분석 결과를 도출하기 위하여 본 연구는 SW인력 관련 고용 시장 수요 데이터를 확보하여 텍스트 마이닝을 실시하였다. SW인력 수요 데이터는 워크넷에 게재된 데이터, 워크넷 잡 매칭 데이터, 그리고 인력공급 데이터를 수집하여 활용하였다.

워크넷은 고용노동부와 한국고용정보원이 공식적으로 운영하는 직업 정보 및 구직 · 구인정보를 제공하는 취업정보사이트이다. 한국직업분류표준에 따라 한국직업 사전을 개발하게끔 사용자에게 필요한 정보를 직관적으로 제공하고 있다. 따라서 워크넷의 구인 정보의 활용성 및 신뢰성은 선행연구를 통하여 확인 가능하다. 한국고용정보원에서는 [그림 3-1]과 같이 워크넷에서 제공되는 구인 · 구직 데이터베이스를 통하여 임금 동향을 분석하여 공시하고 있다 (한국고용정보원, 2015).

[그림 3-1] 임금충족률 상위 5개 직종



(임금충족률 = (평균 제시임금/평균 희망임금) x 100 출처: 한국고용정보원, 2015

직종 간 만족도는 [그림 3-1]의 차트와 같이 구인 기업의 지불임금과 구직하는 사용자들의 희망임금의 비율로 임금충족률을 분석하여 측정할 수 있다. 즉, 워크넷에서 제시 되고 있는 자료의 신뢰성은 높은 편이며, 기업 및 사용자 모두 구인 · 구직의 부분에서 많이 사용하는 창구로 판단된다.

본 연구에서도 [그림 3-2]와 같이 실제 SW인력의 현장에서 요구하는 수요에 따른 데이터 기반 수요 예측과 새로운 직종을 분류하기 위하여 구인 기업의 구인 정보를 워크넷 구인정보에서 수집하여 사용하였다. 특히 워크넷의 채용정보 중 IT · 정보통신 · 웹과 관련된 분야의 정보를 검색하여 도출된 기업들의 모집 내용을 데이터베이스화 하였다.

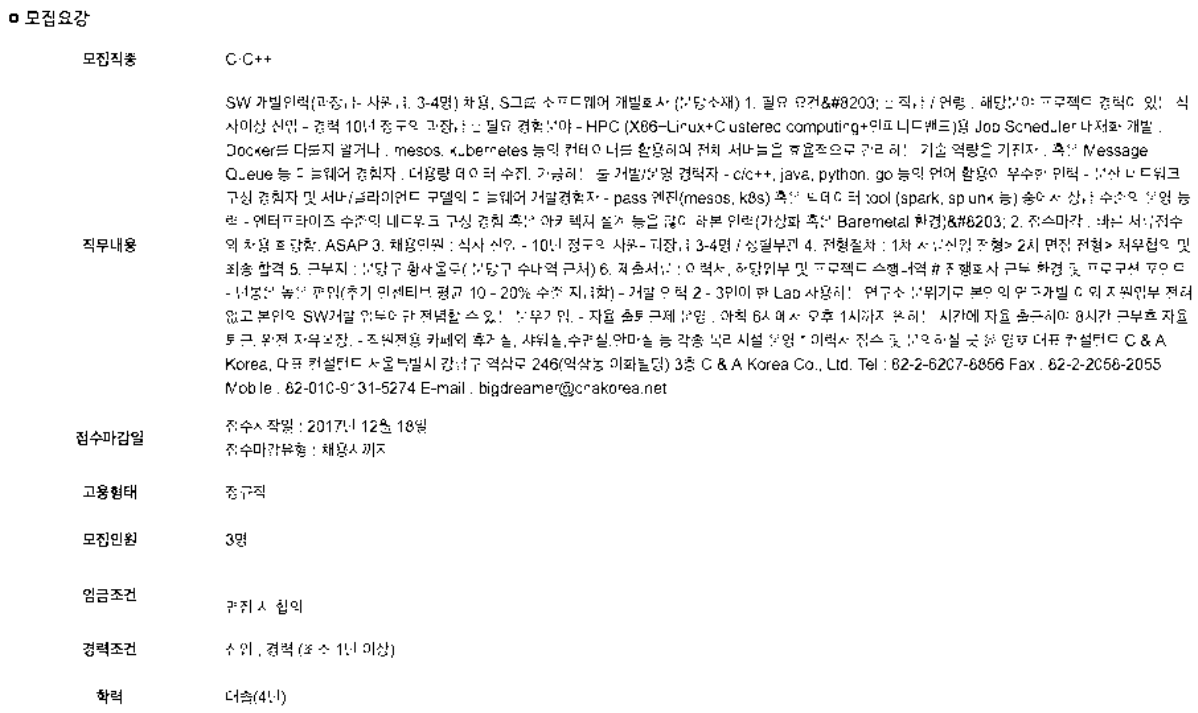
[그림 3-2] 워크넷 채용정보 직종별 'IT · 정보통신 · 웹' 검색 결과



출처 : 워크넷.or.kr

또한, [그림 3-3]과 같이 직종별 분류에서 각 모집 내용을 확인하여 실제 SW 인력 현장에서 모집 직종, 업무내용, 접수마감일, 모집 인원, 임금 등의 기본적인 정보를 수집하여 본 연구의 수요 데이터로 활용하였다.

[그림 3-3] 워크넷 채용정보 세부항목 수집 예시



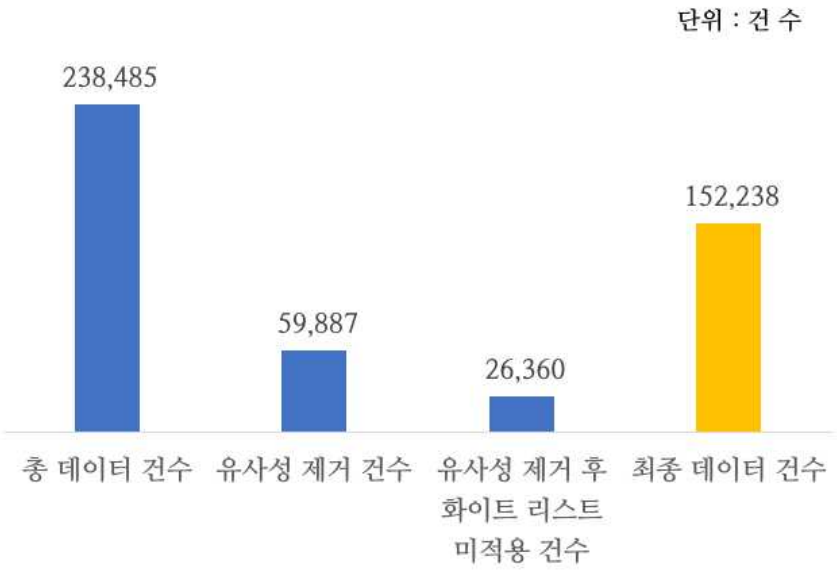
▫ 모집요강

모집직종	C·C++
직무내용	SW 개발인력(과장급 - 사원급, 총 3-4명) 채용, S그룹 소프트웨어 개발회사 (협력업체) 1. 필요 요건 소프트웨어 개발 . 해당분야 프로젝트 경력자 있는 석사이상 신입 - 경력 10년 정도의 과장급 필요 경력년수 - HPC (X86-Linux+C usterec computing+인피니밴드)용 Job Scheduler 나저층 개발 . Docker등 다운지 않거나 . mesos. kubernetes 등의 컨테이너를 활용하여 전체 서버들을 효율적으로 관리 하는 기술 역량을 가진자 . 혹은 Message Queue 등 분산 플랫폼 경험자 . 대용량 데이터 수집, 가공하는 솔루션 개발/운영 경력자 - c/c++, java, python, go 등의 언어 활용이 우수한 인력 - 분산 노드 워크 구성 경험자 및 서버/클라이언트 모델의 소프트웨어 개발경험자 - pass 엔진(mesos, k8s) 혹은 노드 터 tool (spark, splunk 등) 숙련자 상기 수준의 운영 능력 - 엔터프라이즈 수준의 네트워크 구성 경험 혹은 하드웨어 설계 등을 갖춘 하부 인력(가상화 혹은 Baremetal 환경) 2. 접수마감 . 빠른 서류전수의 처리 요망. ASAP 3. 채용인원 : 석사 신입 - 10년 정도의 사원~ 과장급 3-4명 / 경력부관 4. 전형절차 : 1차 서류심사 전형> 2차 면접 전형> 처우협의 및 최종 합격 5. 근무지 : 판교 창사옥로(판교역 수내역 근처) 6. 제출서류 : 이력서, 해당업무 및 프로젝트 수행내역 #주행회사 근무 환경 및 프로그램 유의 - 너무도 놀라운 혜택(추가 인센티브 평균 10 - 20% 수준 지급) - 거절 이력 2 - 3번이 한 Lab 사용하는 연구소 내부기로 복잡 업무 경력을 C 이지 찍힌부 제외 없고 본인의 SW개발 업무에 관 전념할 수 있는 근무 환경 - 자율 출퇴 근제 운영 . 매칙 6시 에서 오후 1시까지 운영 시간에 자율 출근하여 8시간 근무효 자율 퇴근 유연 자율근장. - 조식제공 카페의 확장 술, 샤워실,수면실,안마술 등 각종 복리 시설 운영 . 이력서 접수 및 면접 시 궁금 점은 언제든지 C & A Korea, 대표 컨설턴트 서울특별시 강남구 역삼로 246(역삼동 이화빌딩) 3층 C & A Korea Co., Ltd. Tel : 82-2-6207-8856 Fax . 82-2-2056-2055 Mobile . 82-010-9131-5274 E-mail . bigdreamer@cnakorea.net
접수마감일	접수시 작성 : 2017년 12월 18일 접수마감유형 : 채용시 까지
고용형태	정규직
모집인원	3명
임금조건	회사 내규
경력조건	신입 , 경력 (또는 1년 이상)
학력	대졸(4년)

출처 : 워크넷.or.kr

상기 방법을 토대로 [그림 3-4]와 같이 텍스트마이닝을 위한 기초 데이터를 수집하였다. 기간은 2017년 1월 1일부터 2017년 9월 28일까지의 데이터를 수집하였으며, 총 수집된 데이터는 238,485개의 기업의 데이터를 수집하였고 유사한 데이터 제거 건수로 59,887건, 유사한 데이터 제거 후에 어휘 사전에 적용되지 않은 26,360건을 제외한 최종 152,238건을 기반으로 분석을 실시하였다.

[그림 3-4] SW 인력 수요 데이터베이스



데이터 구성은 [그림 3-5]와 같이 '시간', '기업명', '채용 정보 제목', '채용 정보 세부 내용', '키워드' 의 순으로 정리하였다.

[그림 3-5] SW인력 수요데이터 예시

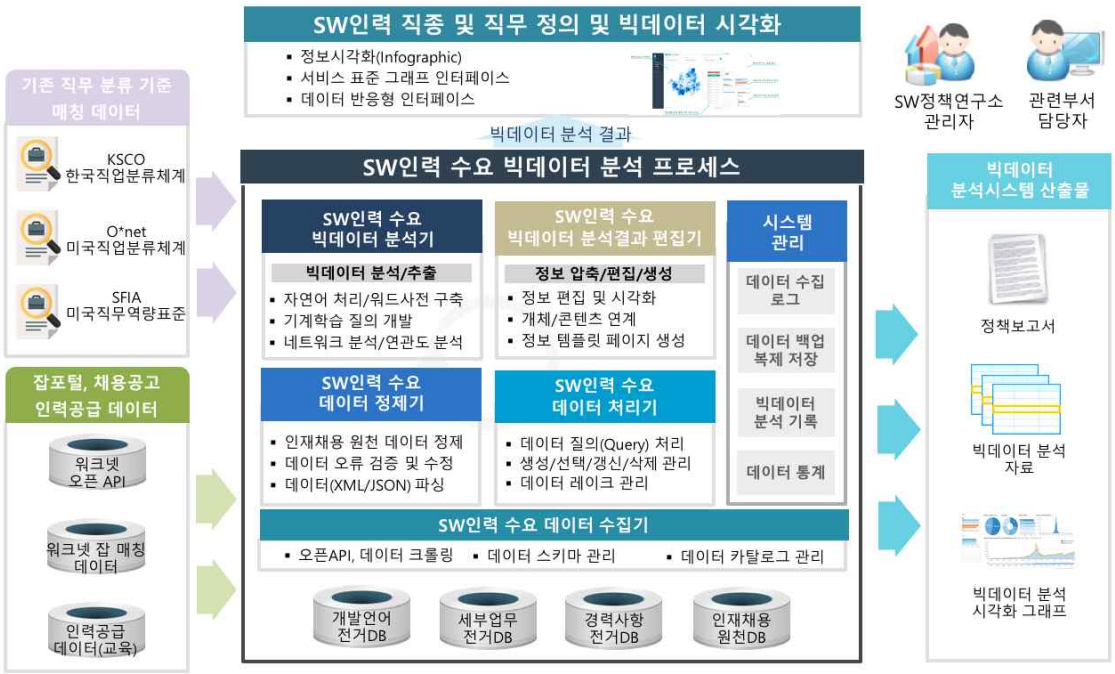
1/2/2017	에이티씨	소프트웨어(C/S C# Mysql) 개발자 모집	Windnow CE6 /C# 기반의 C/S 프로그램 개발 C#/JAVA 기반의 socket Programming 가능자 C#/JAVA 기반의 Serial Programming 가능자 JAVA 웹 프로그램 개발가능자	시스템프로그래머,운영체계소
1/2/2017	동도이엠아이	[IT기술지원] 건설현장 출입안전시스템 설치공사 및 기술지원	-[IT기술지원] 건설현장 출입안전시스템 설치공사 및 기술지원 -다수 건설현장 출입안전시스템 IT기술 지원업무, 건설현장 출입시스템 설치공사(가설) - 학력무관 - 운전가능자 - 직무는 건설현장에 적용되는 안전시스템을 구축 하는 작업입니다. - 배우고자하는 의지만 있으면 배우시면서 직무에 적용 하실 수 있습니다. - 주5일 근무이며 출근 9시 퇴근6시 야근은 거의없습니다. - PC사용능력 중급 (Ms office MS Sql,네트워크환경 및	-,
1/2/2017	NYinternational	화성에서 열심히 일하고 보람을 느낄 가족 같은 분을 모십니다<웹디자	아르헨티나에서 와인을 수입판매하는 회사로 경리 및 영업관리 사무업무 보조 혹은 웹디자인 및 광고디자인 홍보 업무 중 선택 **자가운전자 유류대 지급/ 운전불가능자인 경우 출근차량픽업가능**	영업관리사무원,영업실적정리/ 업결산사무원,판매관리사무원, 보조,영업관리,고객응대,
1/2/2017	오에스컴퍼니	웹디자이너 채용	포토샵, 일러스트, 디자인 편집, 웹디자인, 기타 사무 업무	웹디자이너,-,
1/2/2017	유니드	참신 하고 성실한 인재을 찾습니다.-웹디자이너 구인	안녕 하세요. 본사는 인터넷 판매을 기획 디자인 판매 할수 있는 성실한 인재을 찾고 있 습니다..	웹디자이너,웹 UI디자이너,포트

본 연구에서는 수집된 내용을 토대로 빅데이터 기반 분석 프로세스를 적용하여 다이나믹 토픽 모델링 분석을 실시하고, 24개의 토픽으로 분류하여 현재 상황을 반영한 데이터 기반 직종 분류체계를 재정립하였다.

2. 빅데이터 기반 분석 프로세스

앞에서 서술한 것과 같이 수집된 SW인력 수요 데이터베이스를 기반으로 빅데이터 기반 분석 프로세스를 [그림 3-6]과 같이 구성하였다.

[그림 3-6] SW인력 수요 빅데이터 분석 프로세스



워크넷 및 잡(Job) 채용 사이트에서 획득된 인력공급 수요 데이터와 상기 언급한 한국직업분류체계(KSCO), 미국직업분류체계(O*net), 국제직종역량표준(SFIA) 등 7개의 기존 직종 및 직종 분류체계를 기반으로 SW인력에 대한 유효 데이터진단 및 프로세스에 따른 데이터 분석을 실시하였다.

수요 데이터를 SW인력 수요 데이터 수집기에 적용하여, 워크넷 오픈 API 및 데이터 크롤링을 활용하여 웹페이지에 있는 데이터를 수집하여 데이터베이스화 하였다. 각 데이터 분석 프로세스 간 진행되는 과업은 다음과 같다

1. SW인력 수요 데이터 처리기 : 데이터가 데이터 분석을 위한 원시데이터

처리 과정을 실시한다. 데이터 분석 전에 텍스트 데이터베이스의 질의에 명시된 데이터를 생성/선택/갱신/삭제 등의 데이터 처리를 통하여 분석할 수 있는 데이터도 전환한다. 데이터 레이크 작업을 통하여 분석하고자 하는 모든 원시데이터를 저장한다.

2, SW인력 수요 데이터 정제기 : 수집된 데이터베이스 즉, 원시 데이터를 분석이 용이하게 만들기 위하여 전처리(Preprocessing) 과정을 실시한다. 실시하는 과정에는 데이터 오류 검증 및 수정, 웹 기반 데이터(XML/JSON) 파싱(Parsing)작업이 소요된다.

3. SW인력 수요 데이터 분석기 : 데이터 정제기를 통하여 분석 가능한 데이터로 만들어진 다음에는 텍스트 기반의 자연어 처리 및 어휘 사전을 구축하여 토픽 모델링 분석을 할 수 있는 사전준비를 실시한다. 또한 기계학습을 할 수 있는 질의를 개발하여 실시간으로 지속적인 훈련과 테스트를 동반하여 빅데이터 분석을 실시하고, 네트워크 분석과 키워드별 연관도 분석을 통하여 유의한 분석 결과를 도출한다.

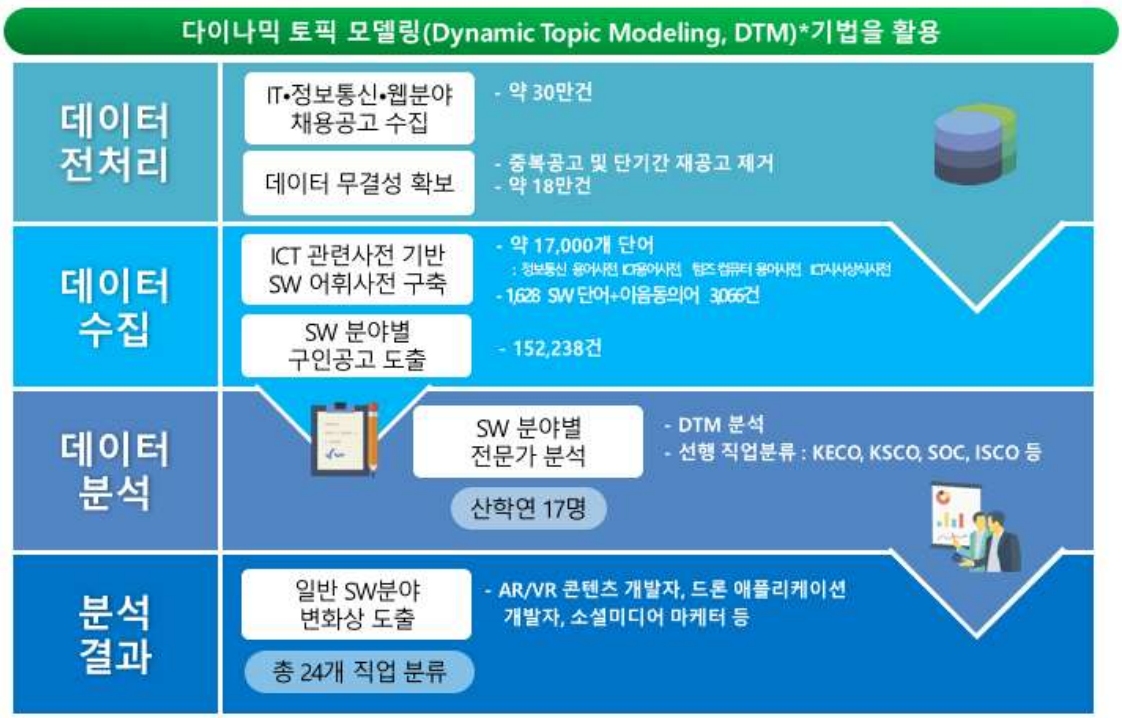
4. SW인력 수요 데이터 분석결과 편집기 : 데이터 분석기에서 분석된 결과에 대한 정보를 편집 및 시각화하여 직관적으로 분석결과를 확인 할 수 있도록 하며, 정보 템플릿 페이지를 생성하여, 해당 분석결과를 실시간으로 보여줄 수 있는 작업을 진행한다.

이후 빅데이터 분석 프로세스를 기반으로 하여 분석된 결과에 대하여 정보시각화 및 서비스 표준 그래프 인터페이스와 데이터 반응형 인터페이스를 구축하여 사용자에게 보다 인지적으로 용이한 정보를 제공함으로써 SW인력의 수요 공급을 연결시키는 역할을 도모하였다.

2장에서 기술한 DTM 토픽 모델링은 기존의 텍스트 데이터를 전 처리하여 자연어 기반으로 텍스트를 분석 가능한 단위(예를 들어, 명사 등)로 파싱하고 분석하여 유의미한 결과를 창출해 내는데 의의가 있다. 본 연구는 DTM 토픽 모델링을 기반으로 데이터 분석을 실시하여 유의미한 SW인력 수요 예측 및 새로운 직종 및 직종 분류체계를 제안하는데 의의가 있다.

본 연구의 분석을 위한 프로세스는 [그림 3-7]과 같이 '데이터 전처리' - '데이터 수집' - '데이터 분석' - '분석 결과'의 순으로 시행되었다.

[그림 3-7] 빅데이터 기반 세부 프로세스



1. 데이터 전처리 : 워크넷에서 IT · 정보통신 · 웹 분야 키워드 채용공고에 대한 구직 데이터를 수집하여 총 30만 건의 데이터를 확보하였다. [데이터 무결성 확보] 30만 건 데이터에서 중복공고 및 단기간 재공고를 제거하였으며, 약 18만 건의 데이터를 확보하였다.

2. 데이터 수집 : 전처리 한 데이터를 확보하여 토픽 모델링을 할 수 있도록 ICT관련 사전 기반 SW 어휘사전을 구축한다. ICT관련 사전 즉, 정보통신 용어사전, ICT용어사전, 텀즈 컴퓨터 용어사전, ICT 시사상식사전에서 얻은 단어를 취합하여 Black List로 만들고 ICT와 관련 없는 용어 및 중복되는 단어를 삭제 및 조정하여 총 17,000개의 어휘 사전을 구축하였다.

3. 데이터 분석 : SW 분야별 전문가 17인을 통하여 다이나믹 토픽 모델링(DTM)을 실시하여 토픽 모델링을 진행하였다. 선행 직업 분류로 상기 언급한 한국직업표준분류 등 7가지 선행 분류와 비교 파악하여 새로운 직종 및 직종을 분류하였다.

4. 분석 결과 ; 데이터 분석을 통하여 도출된 직종을 통하여 기존에 존재하지 않았던 직종들은 새롭게 정의하고 현재 있던 직종 또한 재정립 및 재분류를 통하여 SW인력 수요에 맞는 인재상을 새롭게 제시하여 총 24개의 직업 분류를 도출하였다.

추가적으로 [그림 3-8]과 같이 SW인력 수요 데이터 수집은 웹 크롤러를 개발하여 분석에 알맞은 데이터를 수집하였으며 파이썬 2.7 프로그래밍 언어를 활용, scrapy/.xpath를 통하여 정보 데이터를 수집하였다.

[그림 3-8] SW인력 수요 데이터 수집 방법 (Python 2.7 Scrapy)



<표 3-1> SW인력 수요 데이터 수집 방법

구분	요소	기능
개발요소	LW crawler	• HTTP GET/POST Method를 통해 웹 페이지(HTML, XML, JSON) 수집 • Python Scrapy Import
	LW_parser	• Xpath 기반 객체 식별 기능 제공 • JSON 객체 식별 기능 제공 • 데이터 객체 리스트 추출 기능 제공 • 데이터 객체 자료형 변환 기능 제공
	LW_crawlUtils	• RoundTime 설정 기능 제공 • 수집이력관리 / 에러관리 / 페이지 스케쥴링

토픽 모델링을 위하여 SW 어휘 사전을 구축하기 위하여 <표 3-1>과 같은 방법을 활용하여 전처리 작업을 실시하였다. 우선적으로 기존 ICT관련 사전에 데이터를 모두 추출한 다음, SW와 관련된 어휘를 도출하기 위하여 SW관련 직종에 종사하는 전문가 30인에게 3차례에 걸쳐서 사전 검토를 완료하여 White List 17,000개의 어휘를 확보하였다. 이후 [그림 3-9]와 같이 1차적으로 White List를 확보한 후에 부족한 어휘가 존재할 수 있어 기존에 모아 두었던 Black List를 3차례 재검토를 통하여 해당 단어를 추출하였다.

추출된 단어의 특징을 살펴보면 동의어 처리를 할 경우 한 줄에 해당 단어의 영어 표기, 한글표기, 약자 등의 내용을 표기하여 알고리즘 적용 시 '동류 적용'으로 할 수 있도록 데이터 셋을 구성하였다.

[그림 3-9] SW인력 텍스트 마이닝을 위한 사전 구성 및 추출 단어 파싱

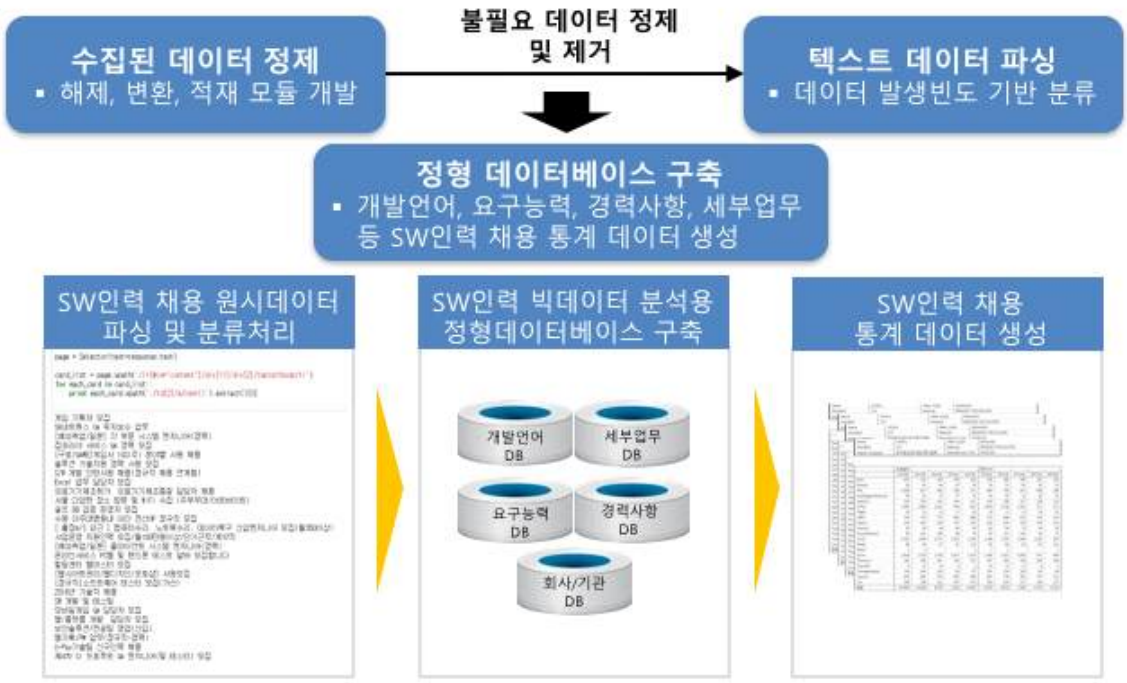


<표 3-2> SW직종 토픽모델링분석을 사용하기 위한 SW 어휘사전 구성

구분	요소	내용
사전 구성요소	출현단어 전체 List	• SW 관련된 단어 (정보통신 용어사전,ICT 용어사전 등 추출)
	White List	• 실제 SW 관련된 유의한 단어 추출 • SW 관련 직종의 전문가에 30인에게 3차례에 걸쳐 활용 어휘 검수
	Black List	• 전체 단어에서 White List를 뺀 나머지 단어 • 추후 부족한 부분을 추가적으로 추가하기 위하여 3차에 걸쳐 어휘 검수

상기 내용을 토대로 [그림 3-10]과 같이 SW인력 채용의 원시데이터를 분석 가능한 단위의 데이터로 전처리를 하여 분석 가능한 정형데이터로 변경하고 데이터베이스를 구축하였다. 텍스트 데이터 파싱은 데이터 발생빈도를 기반으로 분류하였다.

[그림 3-10] SW직종 분석용 데이터 생성 과정



3. 분석 결과

<표 3-3> DTM 토픽모델링 분석결과

분류	1	2	3	4	5	6	7	8
키워드 1	프로그래머(35.0%)	IT (34.4%)	관리(76.9%)	일러스트(15.5%)	photoshop(28.8%)	방송(36.9%)	상품(40.1%)	시스템(31.6%)
키워드 2	Development (17.3%)	HTML (14.5%)	home page (18.5%)	3D(15%)	userinterface(25.1%)	document(13.5%)	제작(37.4%)	엔지니어(16.2%)
키워드 3	PHP (16%)	사진(13%)	거래(0.4%)	photoshop(14%)	computer(15%)	data(8%)	영상(16%)	네트워크(15%)
키워드 4	웹개발자(7.9%)	Coding (8.2%)	업데이트(0.4%)	그래픽디자인(11.7%)	UX(12.3%)	영상(6.9%)	시각디자인(1.1%)	서버(12.9%)
키워드 5	ASP.NET (4.6%)	Licence(7.1%)	스케줄링 (0.2%)	CG(5.4%)	OA(6.1%)	multimedia(5.9%)	CG(0.7%)	보안(6.3%)
키워드 6	Mysql(1.9%)	CSS (5.6%)	스크립트 (0.2%)	character(5%)	QA(3.5%)	검색(5%)	웹페이지(0.5%)	리눅스(3.2%)
키워드 7	JSP (1.7%)	Certificate (2.7%)	웹호스팅(0.1%)	그래픽(4.4%)	GUI(1.6%)	agency(3.8%)	pop-up(0.2%)	테스트(2.7%)
키워드 8	윈도우(1.7%)	Javascript(2.3%)	QC(0.1%)	2D(4.1%)	QC(0.3%)	WAS(2.3%)	CF(0.2%)	유닉스(1.2%)
키워드 9	웹페이지(1.4%)	웹표준(2.1%)	계정(0.1%)	모델링(4%)	모션그래픽(0.1%)	스크립트(1.6%)	섬네일(0.1%)	firewall(0.9%)
키워드 10	리눅스(1.1%)	Jquery(0.9%)	intellectual property (0.1%)	모션그래픽(3.6%)	웹페이지(0.1%)	encoding(1.1%)	콘텐츠관리(0%)	윈도우(0.8%)
키워드 11	PYTHON (0.9%)	ip address (0.9%)	P2P(0.0%)	animation(3.2%)	인터페이스(0.1%)	시뮬레이션(0.9%)	모션그래픽(0%)	visualize(0.7%)
키워드 12	Javascript(0.9%)	웹접근성(0.5%)	pop-up(0.0%)	flash(1.8%)	레이아웃(0%)	미들웨어(0.9%)	source(0%)	백업(0.7%)

분류	9	10	11	12	13	14	15	16
키워드 1	보안(13.7%)	구축(18%)	전산(52.5%)	기획(70.1%)	openmarket(28.6%)	MD(47.1%)	컨텐츠(41.2%)	운영체제(74.4%)
키워드 2	ERP(13.4%)	프로그래밍(10.1%)	블로그(21%)	정보(12.9%)	관리(13.5%)	online(21%)	SNS (24.8%)	관리(15.3%)
키워드 3	solution(9%)	네트워크(6%)	바이럴(14%)	교육(7%)	상품(14%)	검사(9%)	Digital.(13%)	모니터링(3%)
키워드 4	정보보안(5.8%)	데이터베이스(5.1%)	cafe(9.5%)	direct digital marketing(3.9%)	소셜커머스(12.7%)	AMD(6.6%)	포털사이트(3%)	웹서비스(0.5%)
키워드 5	SI(5.3%)	frame(5.1%)	컴퓨터공학(0.3%)	웹서비스(1.2%)	photoshop(7.8%)	DBA(3.9%)	VR. (2.4%)	클라이언트(0.5%)
키워드 6	consulting(4.8%)	장애(4.9%)	백업(0.2%)	financial(0.7%)	그래픽(7.1%)	가공(2.2%)	통합인증(2.4%)	모델링(0.4%)
키워드 7	MES(4.5%)	ORACLE(4.8%)	원격(0.1%)	정보처리(0.3%)	수정(4.3%)	검수(1.7%)	Editor. (2%)	모형(0.4%)
키워드 8	offline(1.9%)	SQL(4.7%)	community(0.1%)	계정(0.1%)	e-commerce(3.2%)	FAE(1.7%)	BI. (1.6.%)	community(0.4%)
키워드 9	취약점(1.2%)	라이브러리(4.5%)	component(0%)	agent(0.1%)	제품디자인(2.6%)	ISP(0.7%)	Creative.(1.5%)	DRONE(0.4%)
키워드 10	해킹(0.9%)	서버(3.8%)	CPA(0%)	community(0.1%)	SMS(0.4%)	리테일(0.7%)	CI (1.1%)	operator(0.1%)
키워드 11	암호(0.9%)	전산(3.7%)	CPM(0%)	portal(0%)	계정(0.2%)	Amazon0.4(%)	AR (1%)	원격(0.1%)
키워드 12	개인정보(0.7%)	solution(3.1%)	CPS(0%)	PPL(0%)	social media(0.1%)	VOC(0.3%)	Developer(0.3%)	CDN(0.1%)

분류	17	18	19	20	21	22	23	24
키워드 1	development(32.4%)	모바일(28.7%)	development(45.3%)	설계(39.1%)	online(72.8%)	인터넷(39.9%)	웹사이트(59.9%)	유지(45.6%)
키워드 2	JAVA(24.8%)	development(10.1%)	C(12.1%)	analysis(15.4%)	소프트웨어(18.2%)	CS(18%)	마켓(12.8%)	보수(42.4%)
키워드 3	안드로이드(8%)	어플리케이션(15%)	program(8%)	데이터(8%)	로봇(4%)	반도체(12%)	리뉴얼(9%)	정보기기(0.1%)
키워드 4	JSP(5.6%)	solution(15.1%)	소프트웨어(6.9%)	자동화(5.4%)	offline(1.6%)	연구(12%)	communication(6.3%)	RFID(0.1%)
키워드 5	IOS(5.2%)	platform(7.7%)	응용프로그램(4.6%)	consulting(5.3%)	build(0.4%)	R(6.9%)	APP(3.8%)	원격(0.1%)
키워드 6	어플리케이션(3.6%)	클라이언트(2.4%)	embeded(3.3%)	autocad(5%)	CORE(0.2%)	클라우드(6.2%)	CRM(2.9%)	circuit(0.1%)
키워드 7	서버(1.7%)	유니티(1%)	펌웨어(3.2%)	시각디자인(3%)	paging(0.1%)	intelligent(0.1%)	sharing(1.9%)	EDI(0.1%)
키워드 8	SPRING(1.6%)	스마트폰(0.8%)	.NET(1.3%)	빅데이터(2.3%)	DSP(0%)	컴퓨터공학(0.1%)	업데이트(0.9%)	ip address(0.1%)
키워드 9	APP(1.5%)	게임개발(0.7%)	사물인터넷(1.2%)	검증(2%)	encoding(0%)	프레임(0.1%)	리소스(0.2%)	source(0%)
키워드 10	backend(1.4%)	웹서비스(0.3%)	리눅스(1%)	검토(1%)	fpga(0%)	SOC(0.1%)	backend(0.1%)	최적화(0%)
키워드 11	frontend(1.3%)	API(0.2%)	프로그래밍(0.9%)	반도체(1%)	RUBY(0%)	어셈블리(0.1%)	editor(0%)	QC(0%)
키워드 12	응용프로그램(0.9%)	open source(0.2%)	윈도우(0.9%)	process(0.7%)	source code(0%)	.NET(0%)	IRC(0%)	정보(0%)

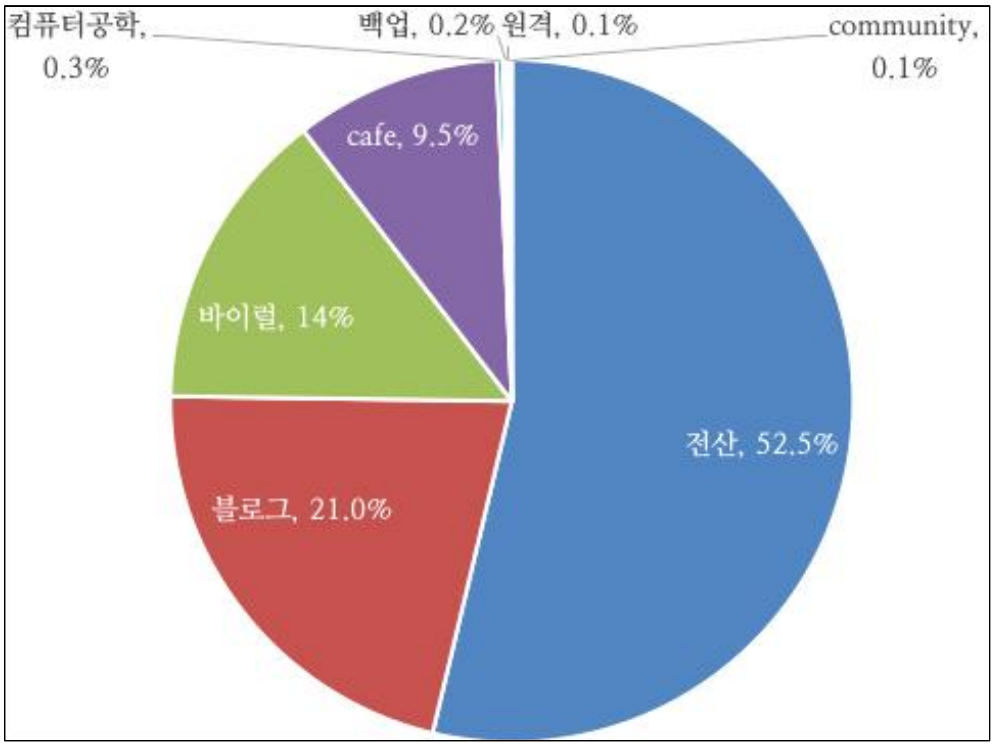
상기 언급한 분석 프로세스에 따라 <표 3-3>과 같이 DTM 알고리즘을 이용한 텍스트 마이닝을 수행한 결과 총 24개의 직종으로 분류되었다. 상기 확보된 데이터의 시간의 흐름에 따라 토픽들이 가장 두드러지게 나오는 군집을 확인한 결과 24개로 정리되었으며, 각 12개의 핵심단어로 비중을 주어 24개의 직업으로 분류하였다. 어휘 사전에 존재하는 어휘를 기반을 두어 수집된 데이터의 의미를 토픽을 시간의 흐름에 따라 변화 값까지 계산하여 분석하였다. 기존 2장에서 DTM에 언급했듯이 다이나믹 토픽 모델링의 기본이 되는 수식을 활용하여 분석하였다.

$$\beta_{t,k}|\beta_{t-1,k} \sim N(\beta_{t-1,k}, \sigma^2 I)$$ (DTM 알고리즘 산식)

각 직종별 특성을 확인하면 다음과 같다.

1. 소셜미디어 마케터

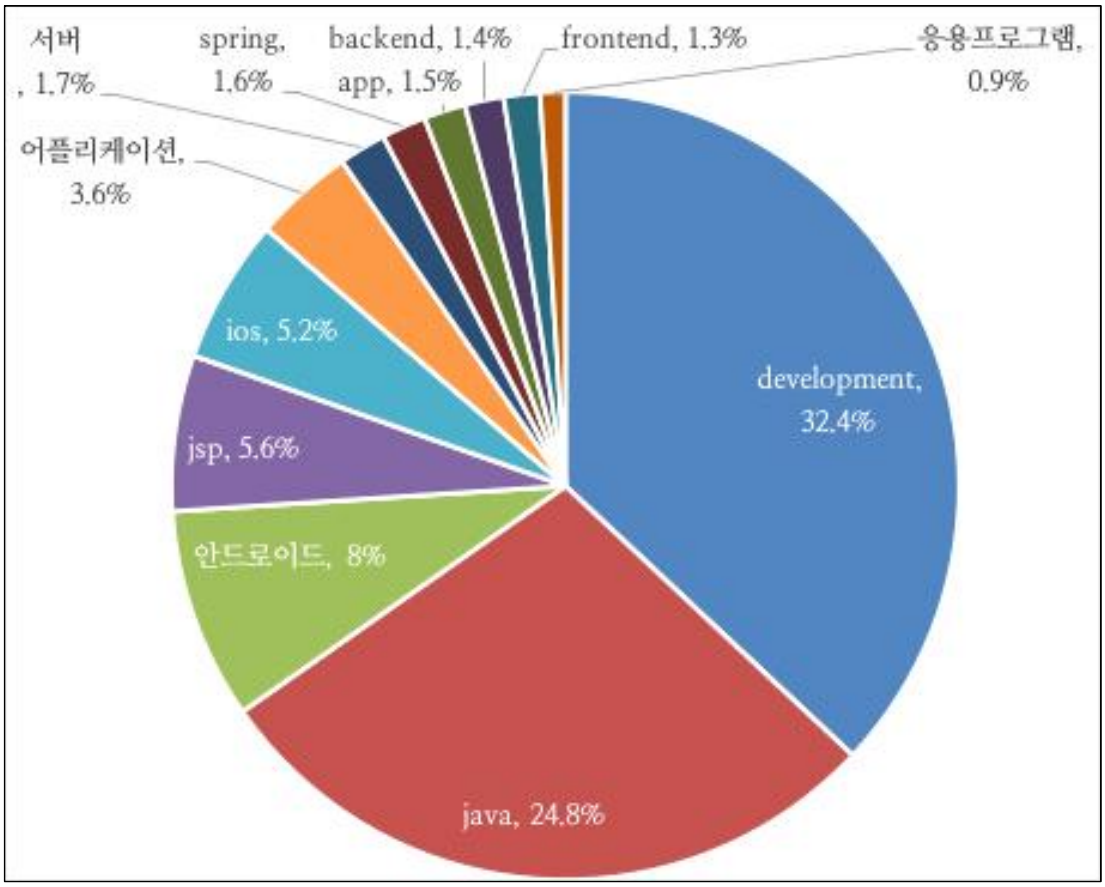
[그림 3-11] 소셜미디어 마케터 주요 키워드 분포



24개의 직종 및 직업에 첫 번째의 직업의 키워드 분포는 [그림 3-11]과 같이 '전산'이 가장 높은 52.5%이며, '블로그'가 21.0%, '바이럴' 14%, 'cafe' 9.5%로 분포하고 그 뒤를 '컴퓨터 공학', '백업', '원격', 'Community'로 분포된다. 기본적은 컴퓨터 공학적 지식을 가지고 있으면서, 블로그나 cafe 등의 키워드와 바이럴이라는 키워드로 유추해 볼 때 웹상의 커뮤니티를 관리하는 관리자 혹은 소셜 네트워크 기반의 마케터와 같은 직종으로 분석된다.

2. 모바일 어플리케이션 프로그래머

[그림 3-12] 모바일 어플리케이션 프로그래머 주요 키워드 분포

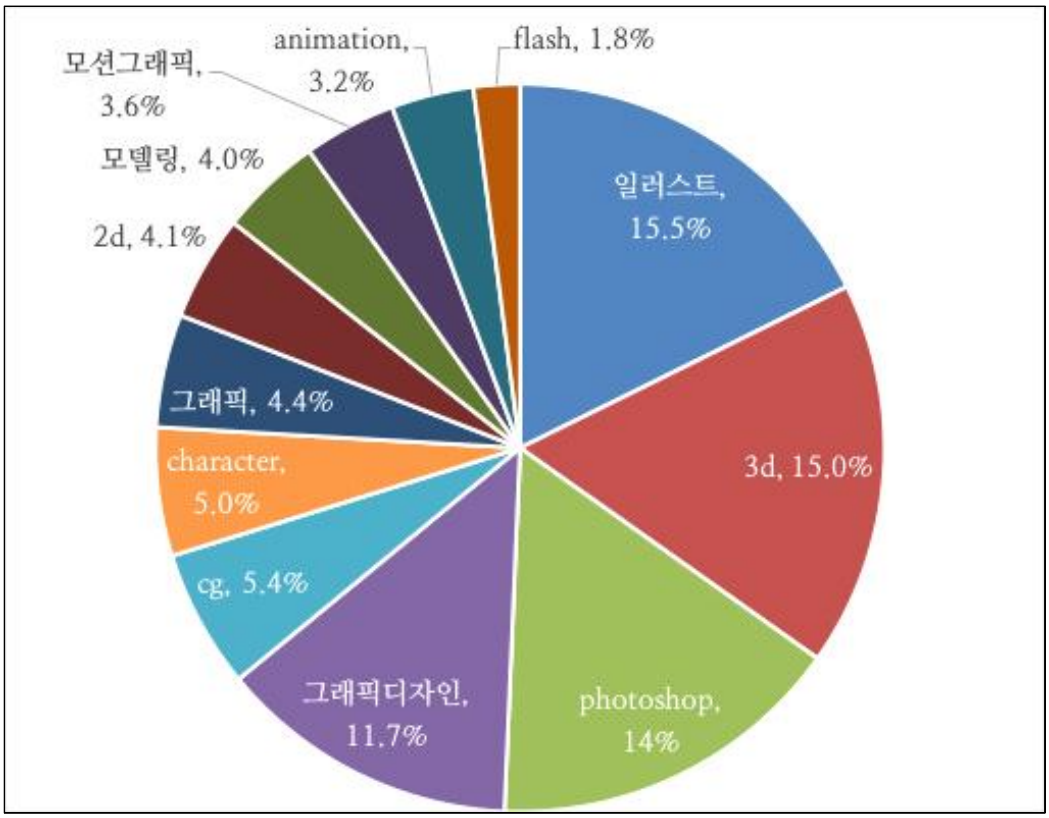


[그림 3-12]와 같이 2번째 직업의 키워드는 'development'가 32.4%로 가장 높게 나타났으며, 그 뒤로 'java'(24.8%), '안드로이드'(8%)와 'jsp', 'ios', '어플리케이션' 등의 순서로 분포되었다.

java와 jsp의 프로그래밍 언어를 이용한 개발을 하며, ios와 어플리케이션이 등장하는 것으로 유추해 볼 때 모바일 관련 개발자와 같은 직종으로 분석된다.

3. 멀티미디어 디자이너

[그림 3-13] 멀티미디어 디자이너 주요 키워드 분포

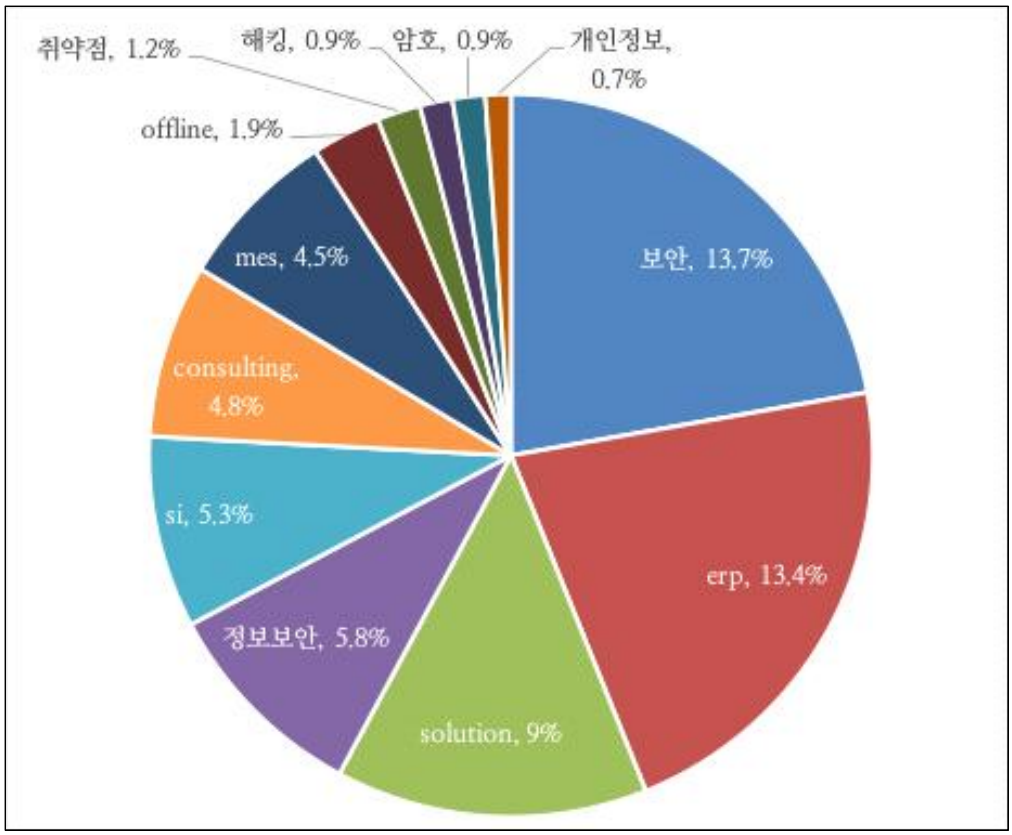


[그림 3-13]과 같이 3번째 직업의 키워드는 '일러스트'가 15.5%로 가장 높게 나타났으며, 그 뒤로 '3d'(15%), 'phtoshop'(14%), '그래픽디자인'(11.7%), 'cg'(5.4%) 등의 순서로 분포되었다.

일러스트 및 photoshop 등의 디자인 관련 프로그램의 도출과 그래픽 디자인, 3d 등의 디자인, cg과 같은 키워드가 등장함으로써 멀티미디어 관련 디자인을 설계하거나 직접 디자인하는 것과 같은 직종으로 분석된다.

4. 정보보안 전문가

[그림 3-14] 정보보안 전문가 주요 키워드 분포

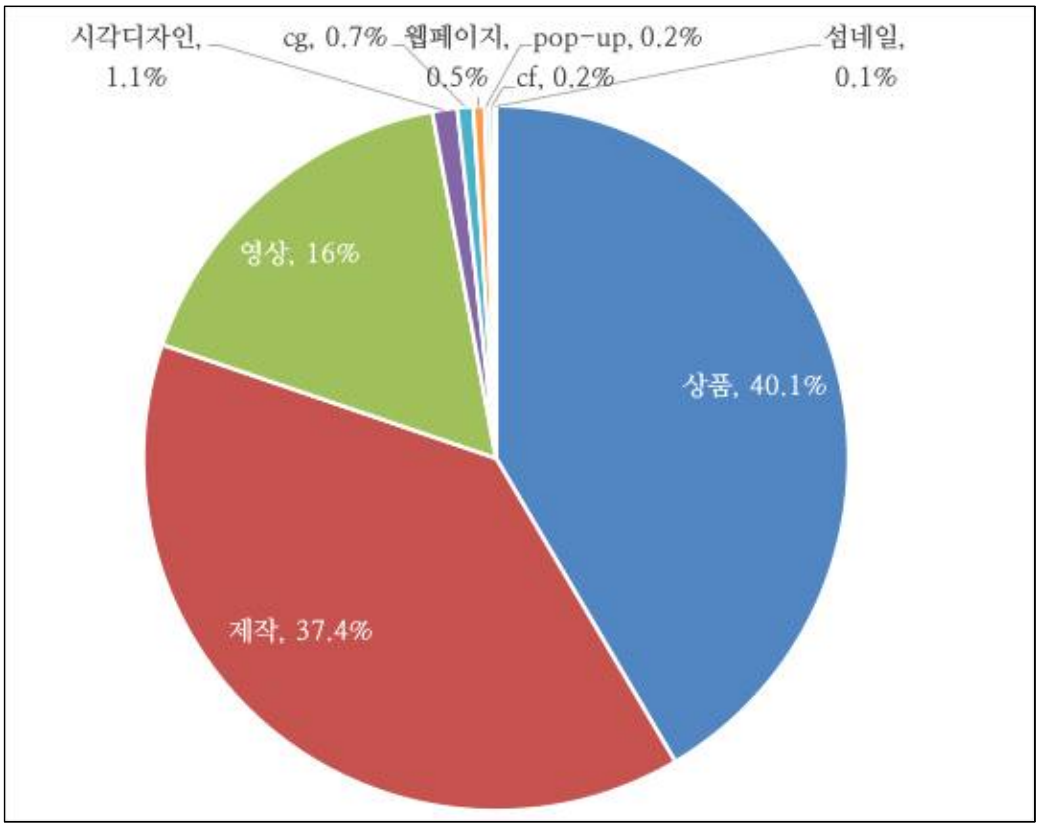


[그림 3-14]와 같이 4번째 직업의 키워드는 '보안'이 13.7%로 가장 많은 분포를 차지하였으며, 그 뒤로 'erp'(13.4%), 'solution'(9%), '정보보안'(5.8%), 'si'(5.3%) 등의 분포를 보였다.

보안 및 solution, 정보보안으로 비추어 보아 정보 보안과 관련된 업무를 시행하며, 보안과 관련된 솔루션을 제공하고 컨설팅하는 전문가 같은 직종으로 분석된다.

5. 멀티미디어 콘텐츠 관리자

[그림 3-15] 멀티미디어 콘텐츠 관리자 주요 키워드 분포

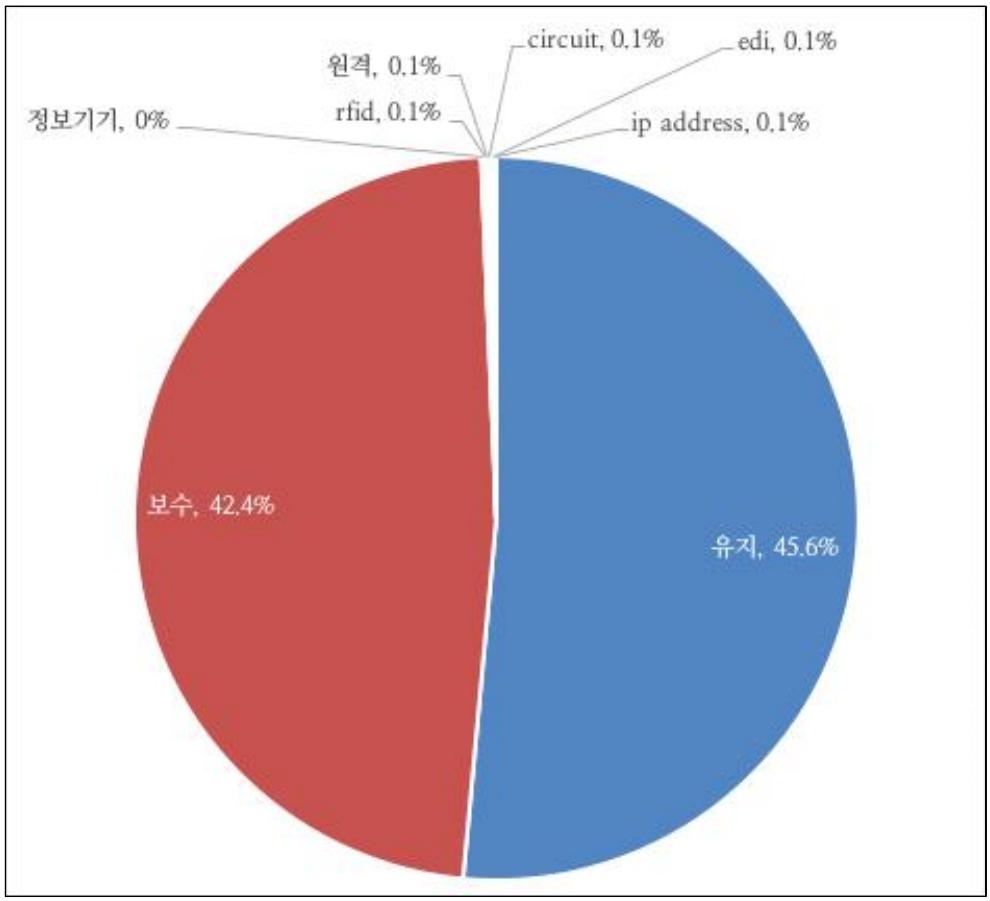


[그림 3-15]와 같이 5번째 직업의 키워드는 '상품'이 40.1%로 가장 많은 분포를 하였으며, 그 뒤로 '제작'(37.4%), '영상'(16%) 등이 분포 되었다.

상품에 대한 웹이나 기타 다른 매체에서 소비자에게 보이는 제작이나 영상 등의 키워드로 유추해 볼 때 멀티미디어와 관련한 콘텐츠를 제작하거나 영상 촬영한 콘텐츠를 관리하는 관리자와 같은 직종으로 분석된다.

6. 정보시스템 운영자

[그림 3-16] 정보시스템 운영자 주요 키워드 분포

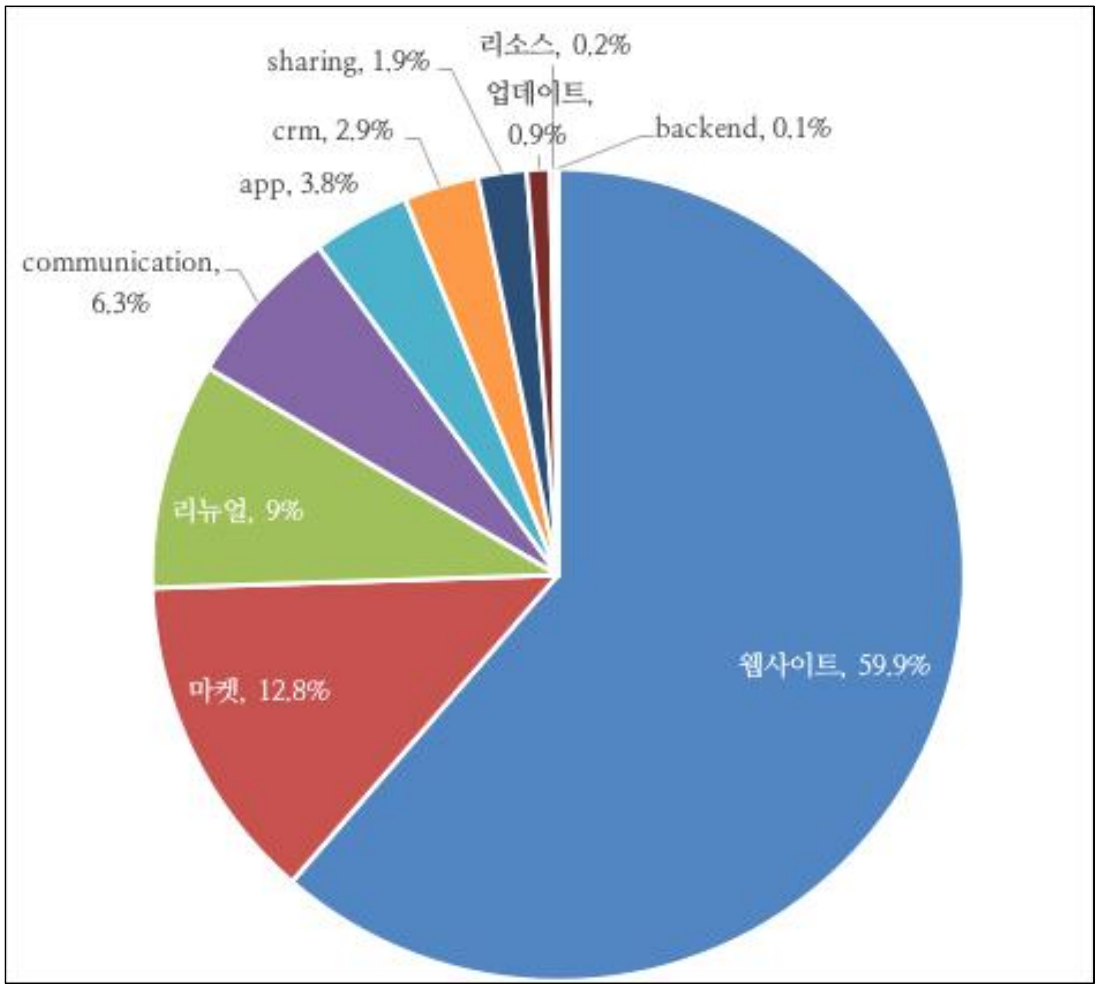


[그림 3-16]과 같이 6번째 직업의 키워드는 '유지'가 45.6%로 매우 높은 분포를 이루고 있으며 그 뒤로 '보수'(42.4%), 'rfid'(0.1%), '원격'(0.1%) 등이 분포되었다.

따라서 유지와 보수를 가장 중요하게 간주하는 직종의 일종으로 원격 등의 키워드도 나오기 때문에 정보 시스템에 대한 유지보수를 관리 및 실행하는 운영자와 같은 직종으로 분석된다.

7. 웹 운영자

[그림 3-17] 웹 운영자 주요 키워드 분포

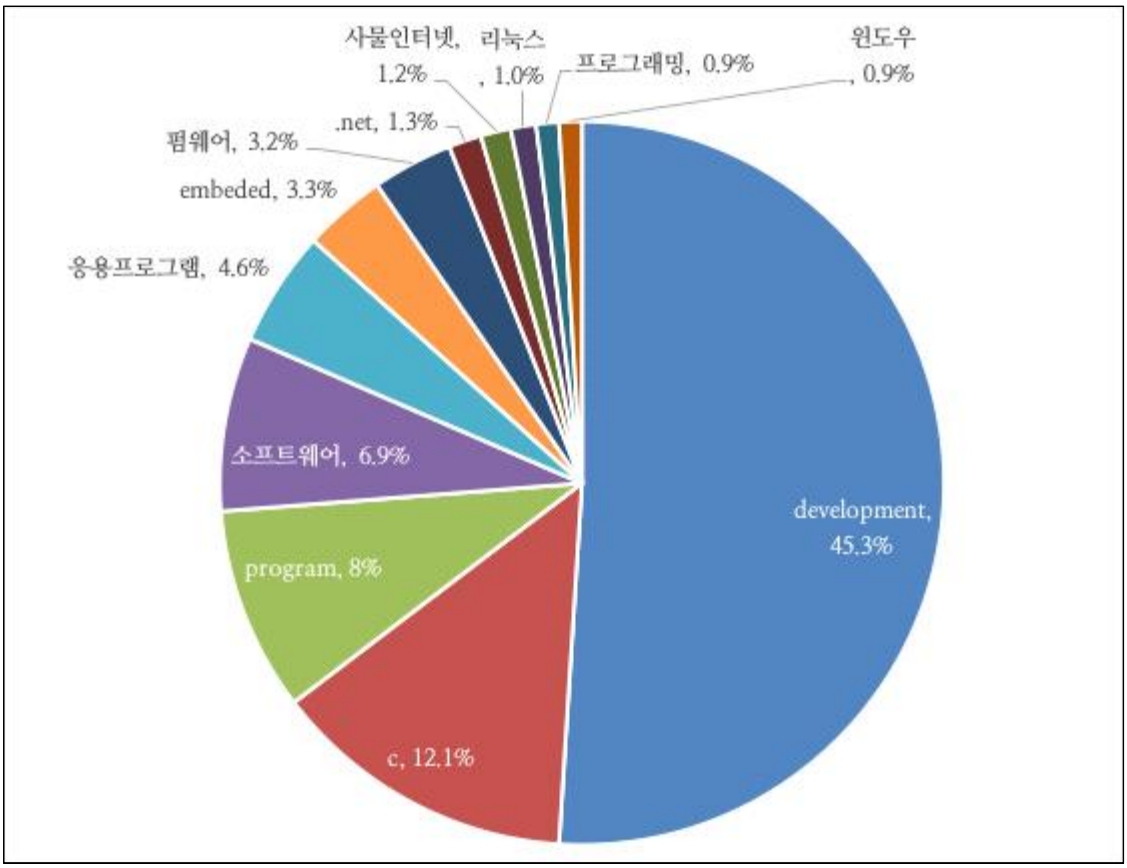


[그림 3-17]과 같이 7번째 직업의 키워드는 '웹사이트'가 59.9%로 매우 높은 분포를 이루고 있으며 그 뒤로 '마켓'(12.8%), '리뉴얼'(9%), 'communication'(6.3%) 등이 분포되었다.

웹사이트에 대한 마켓과 리뉴얼 등의 운영과 관련된 키워드가 도출된 것으로 유추해 볼 때 웹사이트를 관리하고, 사용자와 커뮤니케이션을 하면서 운영하는 운영자와 같은 직종으로 분석된다.

8. 사물인터넷 소프트웨어 개발자

[그림 3-18] 사물인터넷 소프트웨어 개발자 주요 키워드 분포

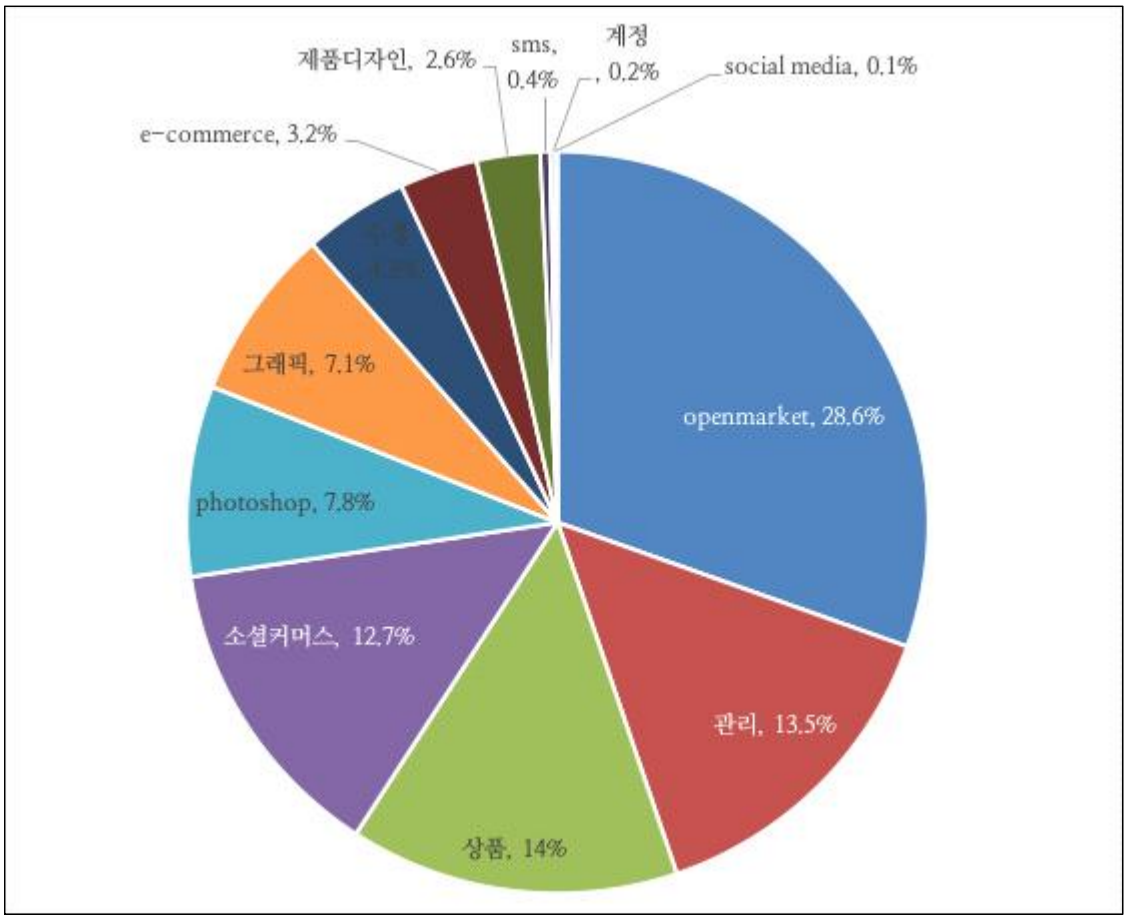


[그림 3-18]과 같이 8번째 직업의 키워드는 'development'가 45.3%로 매우 높은 분포를 이루고 있으며 그 뒤로 'c'(12.1%), 'program'(8%), '소프트웨어'(6.9%) 등이 분포되었다.

c언어의 키워드 분포 및 응용프로그램의 특성인 임베디드와 펌웨어의 작동여부에 대한 프로그램 구성 등의 키워드를 유추해 볼 때 사물인터넷에서 사물과 연결하는 펌웨어 및 임베디드와 관련된 프로그램을 개발하는 개발자와 같은 직종으로 분석된다.

9. E-커머스 관리자

[그림 3-19] E-커머스 관리자 주요 키워드 분포

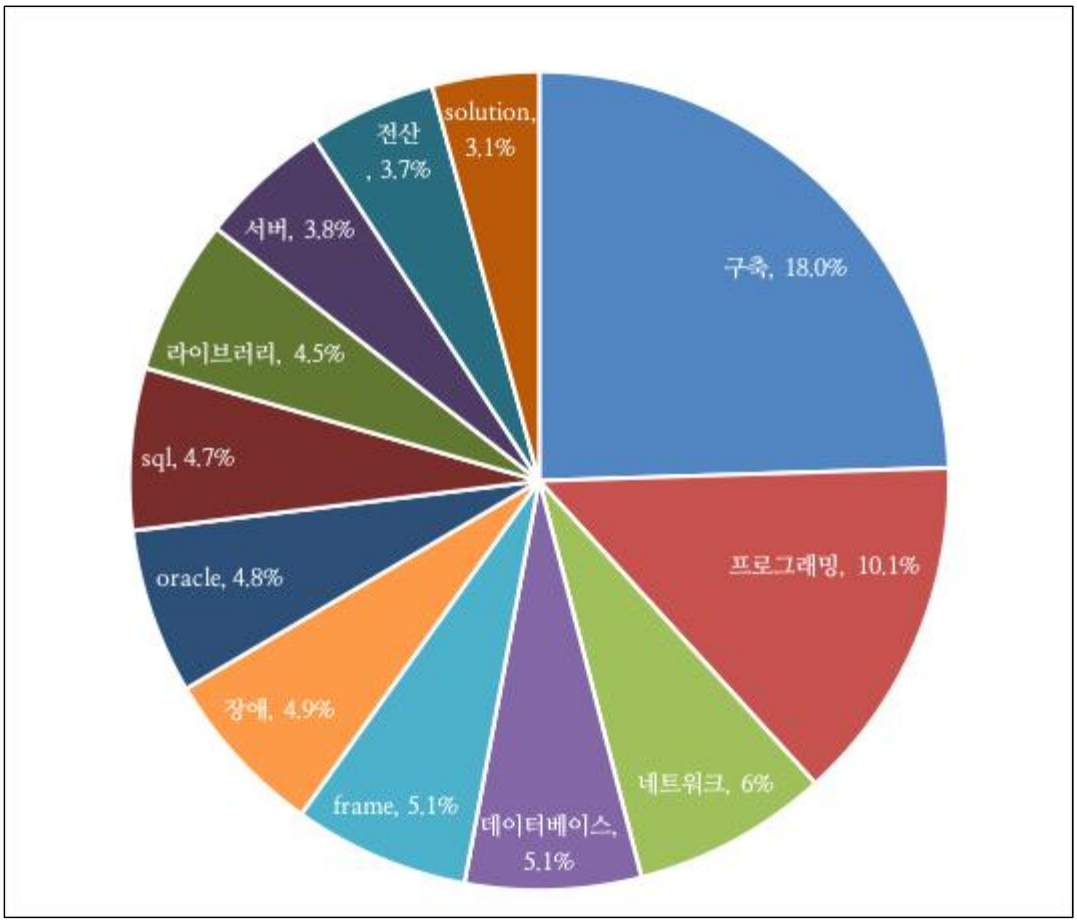


[그림 3-19]와 같이 9번째 직업의 키워드는 'openmarket'이 28.6%로 가장 높은 분포를 이루고 있으며 그 뒤로 '관리'(13.5%), '상품'(14%), '소셜커머스'(12.7%) 등이 분포되었다.

오픈마켓 및 소셜커머스, e-커머스 등의 키워드 도출과 상품, 관리 등의 키워드를 종합하여 유추해 볼 때 e-커머스 사이트 운영 및 전체적인 상품을 게재하고 올릴 수 있는 플랫폼을 관리하는 관리자와 같은 직종으로 분석된다.

10. 서버 개발자

[그림 3-20] 서버 개발자 주요 키워드 분포

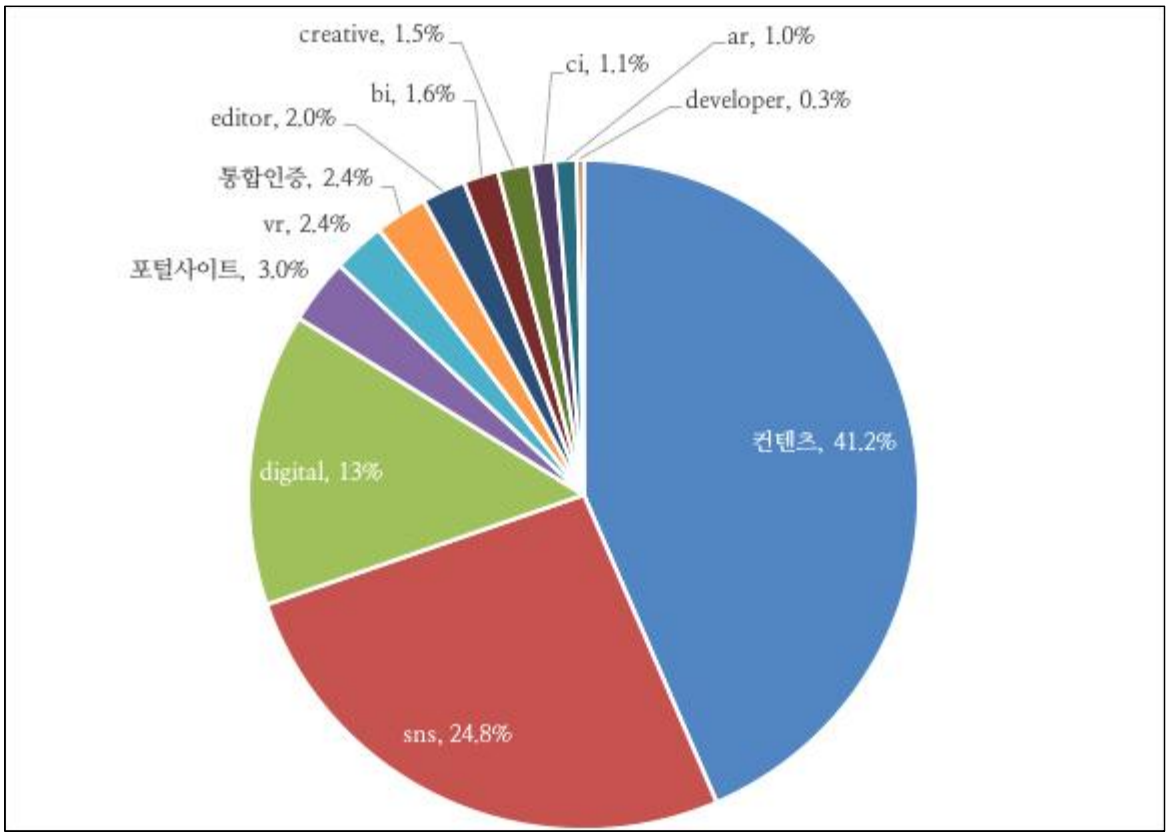


[그림 3-20]과 같이 10번째 직업의 키워드는 '구축'이 18.0%로 가장 높은 분포를 이루고 있으며 그 뒤로 '프로그래밍'(10.1%), '네트워크'(6%), '데이터베이스'(5.1%) 등이 분포되었다.

네트워크의 구축, 서버 장애 등에 대한 프로그래밍 기술과 데이터베이스 서버 등에 대한 라이브러리 관리 및 개발의 키워드로 유추해 볼 때 네트워크 및 시스템 서버에 대한 개발자와 같은 직종으로 분석된다.

11. AR/VR 콘텐츠 개발자

[그림 3-21] AR/VR 콘텐츠 관리자 주요 키워드 분포



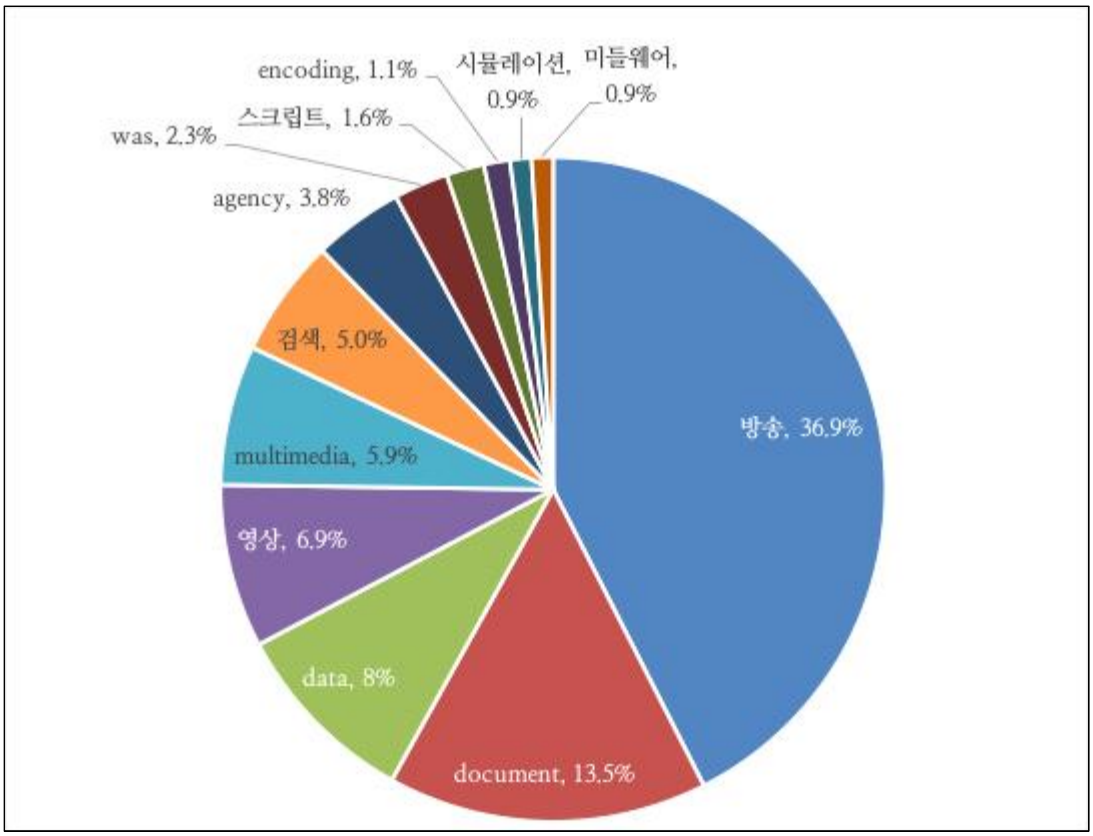
[그림 3-21]과 같이 11번째 직업의 키워드는 '컨텐츠'가 41.2%로 가장 높은 분포를 이루고 있으며 그 뒤로 'sns'(24.8%), 'digital'(13%), '포털사이트'(3.0%) 등이 분포되었다.

상대적으로 낮은 분포지만 'AR'(1.0%), 'VR'(2.4%)가 등장하고 콘텐츠의 비중이 매우 높으며, AR/VR의 콘텐츠를 sns 상에서 유통하거나 게시하는 등의 시스템을 관리하고 포털사이트를 관리하는 콘텐츠 관리자와 같은 직종으로 분석된다.

또한 기존의 직종에는 없는 AR/VR이 등장하는 것으로 보아 4차 산업 혁명의 시작으로 산업 전반의 변화에 따라서 생성된 새로운 직업이라고 판단된다.

12. 웹 방송 엔지니어

[그림 3-22] 웹 방송 엔지니어 주요 키워드 분포



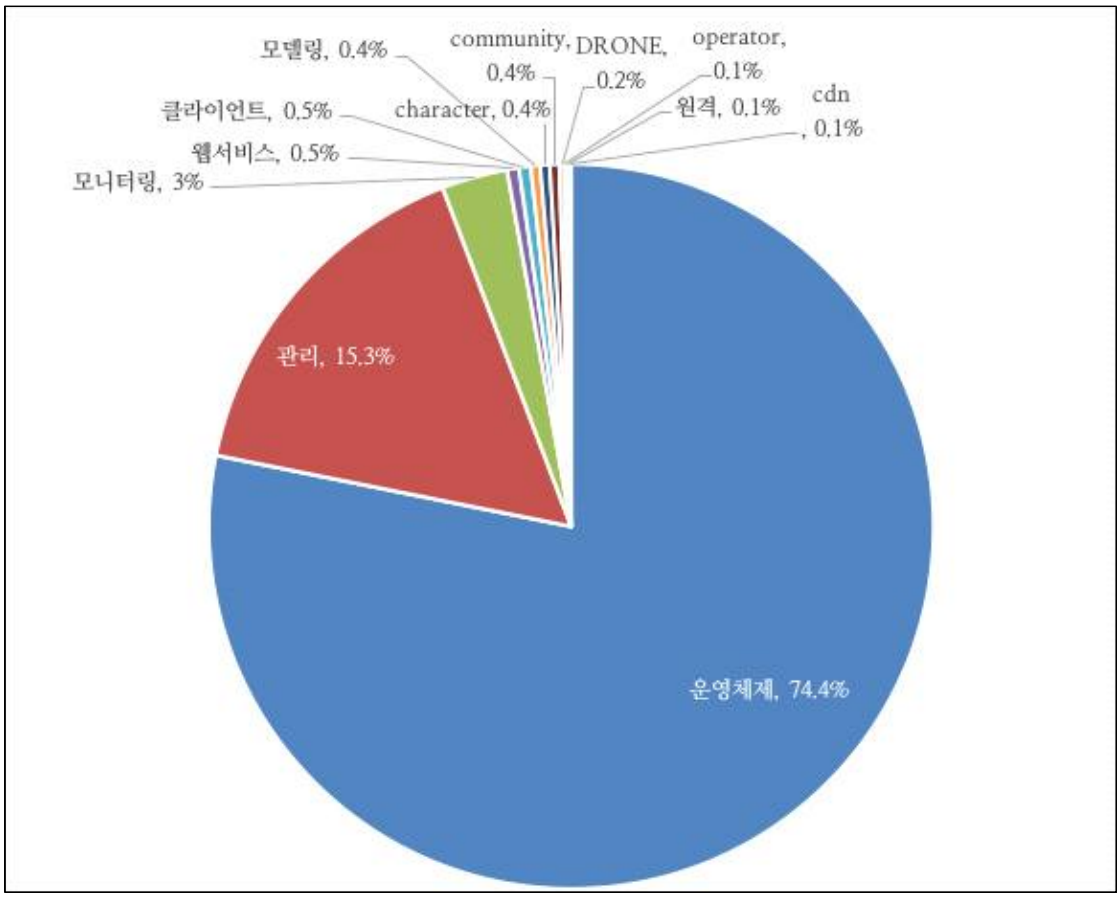
[그림 3-12]와 같이 12번째 직업의 키워드는 '방송'이 36.9%로 가장 높은 분포를 이루고 있으며 그 뒤로 'document'(13.5%), 'data'(8%), '영상'(6.9%) 등이 분포되었다.

가장 많은 비중을 차지하는 웹상에서 방송을 제작하거나 영상을 만들고 에이전시 등의 키워드를 통하여 유추할 때 방송과 관련된 시스템을 구축하고 기획하는 엔지니어와 같은 직종으로 분석된다.

웹 방송 엔지니어는 기존의 멀티미디어 방송 엔지니어 등과 유사한 직종이 존재하였으나, 최근 유투브나 개인 방송의 활용의 증가로 인하여 영상의 트래픽 문제 로딩 문제들을 해결하고 최적화를 하는 역량을 갖은 엔지니어의 수요가 급격하게 늘어 새롭게 향상된 직종으로 판단된다.

13. 드론 애플리케이션 개발자

[그림 3-23] 드론 애플리케이션 개발자 주요 키워드 분포



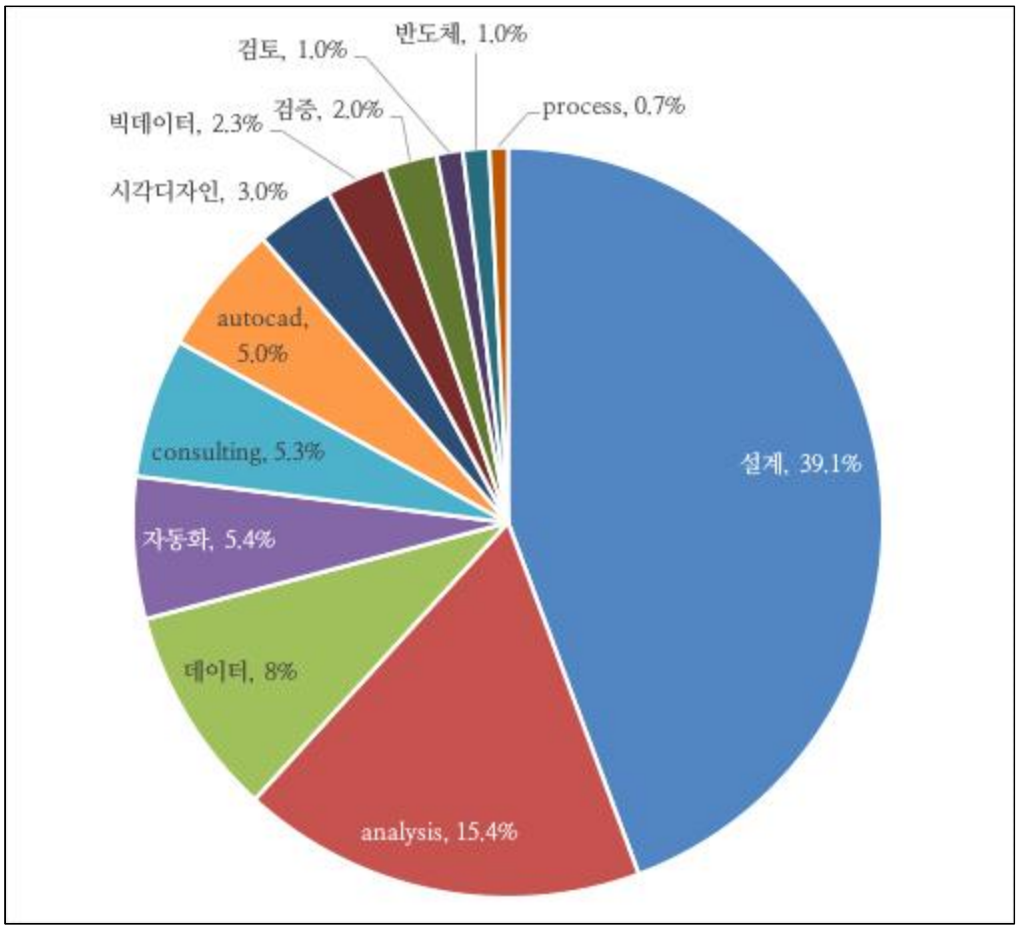
[그림 3-23]과 같이 13번째 직업의 키워드는 '운영체제'가 74.4%로 가장 높은 분포를 이루고 있으며 그 뒤로 '관리'(15.3%), '모니터링'(3%), '클라이언트'(0.5%), '드론'(0.2%) 등이 분포되었다.

드론이라는 키워드의 등장과 운영체제, 원격, 클라이언트 등의 내용을 유추해 볼 때 드론 하드웨어를 활용하여 운영체제를 개발하고 관리하며, 원격으로 시스템을 운영 및 조정해주는 역할을 하는 개발자의 직종과 같이 분석된다.

최근 무인항공에 대한 수요와 공급으로 인하여 많은 사용자가 드론을 활용한 비즈니스와 취미 등에 활발하게 확인되고 있다. 기존에는 없었던 새로운 직종의 생성이라고 판단된다.

14. 자동화 설계자

[그림 3-24] 자동화 설계자 주요 키워드 분류

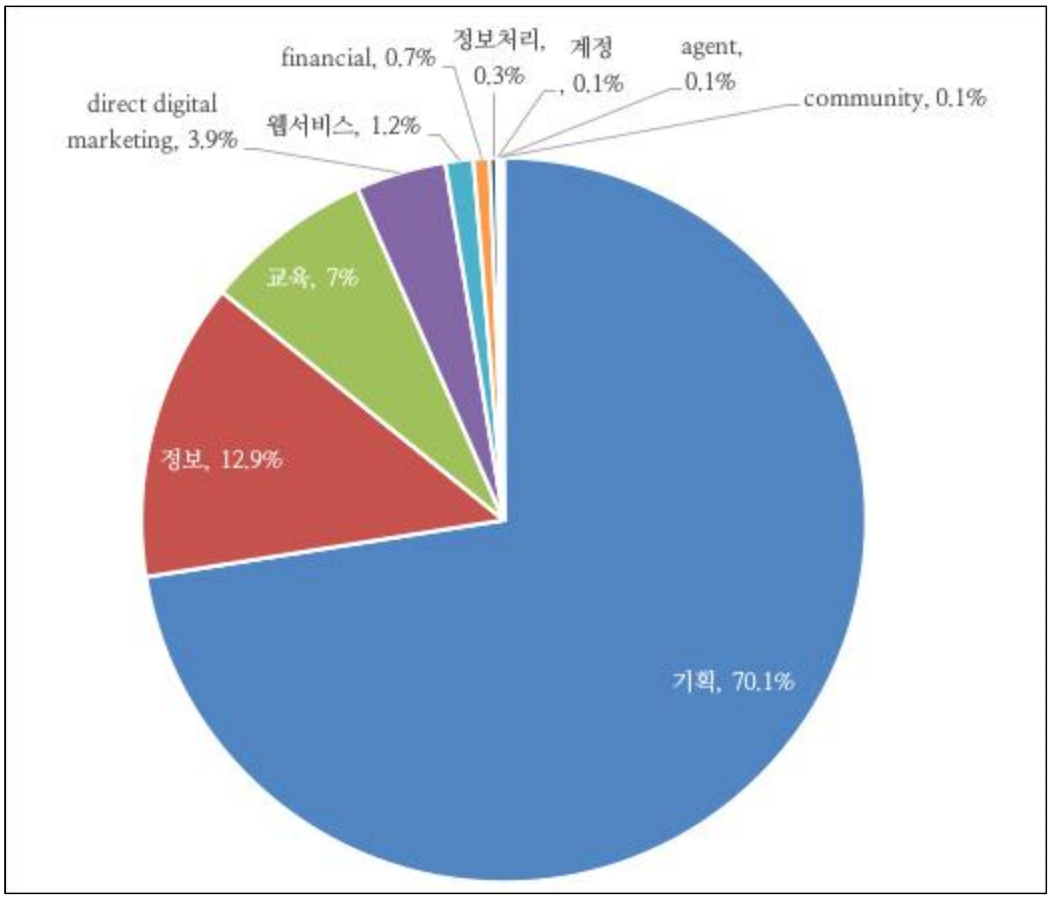


[그림 3-24]와 같이 14번째 직업의 키워드는 '설계'가 39.1%로 가장 높은 분포를 이루고 있으며 그 뒤로 'analysis'(15.4%), '데이터'(8%), '자동화'(5.4%), 'consulting'(5.3%) 등이 분포되었다.

설계 및 자동화 그리고 컨설팅, 반도체 등의 키워드로 유추해 볼 때 공정 및 반도체 등의 자동화를 설계하고 분석하며, 공정이 원활하게 운영될 수 있도록 컨설팅 해주는 직종으로 분석된다.

15. 디지털 마케터

[그림 3-25] 디지털 마케터

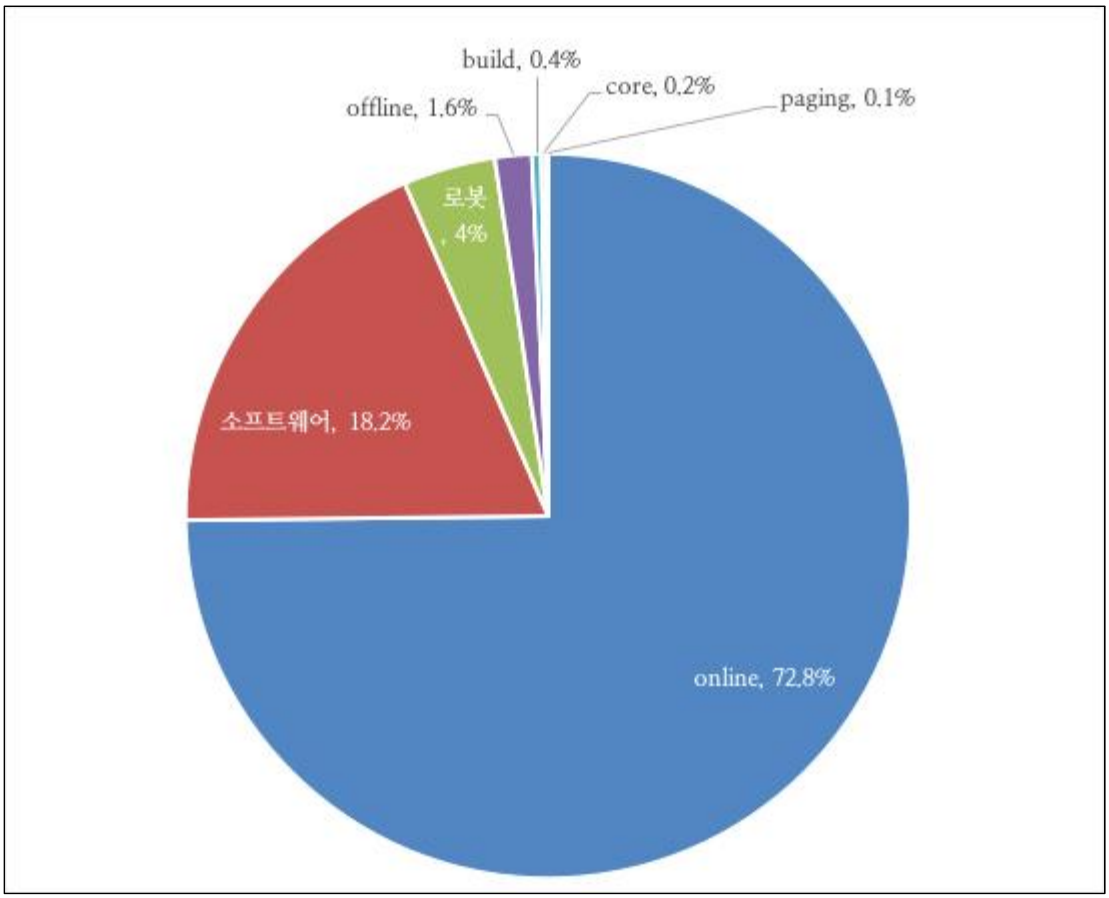


[그림 3-25]와 같이 15번째 직업의 키워드는 '기획'이 70.1%로 가장 높은 분포를 이루고 있으며 그 뒤로 '정보'(12.9%), '교육'(7%), 'direct digital marketing'(3.9%), '웹서비스'(1.2%) 등이 분포되었다.

정보와 기획 교육 등 그리고 direct digital marketing의 키워드를 조합하여 유추해 볼 때 웹 혹은 디지털 상에서 이루어지는 정보 교환과 교육, 마케팅과 관련된 부분에 기획을 하는 마케터의 직종으로 분석된다.

16. 로보틱스 시스템 개발자

[그림 3-26] 로보틱스 시스템 개발자 주요 키워드 분포

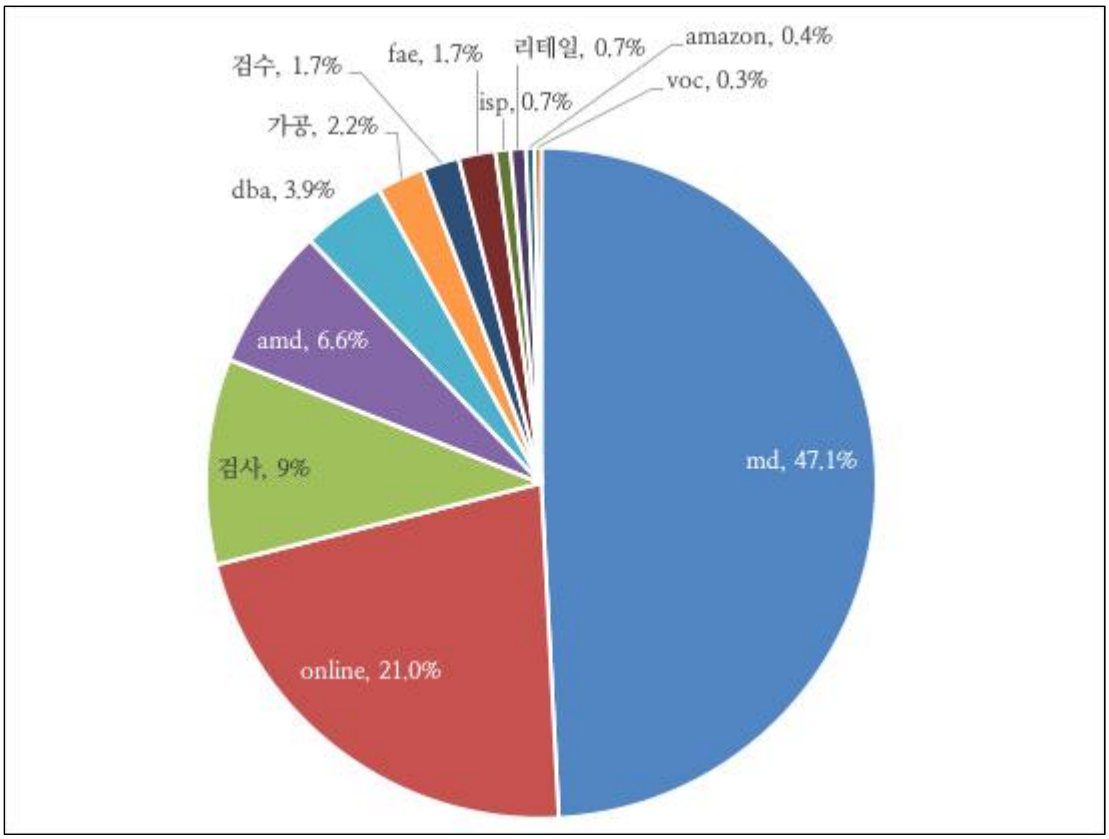


[그림 3-26]과 같이 16번째 직업의 키워드는 'online'이 72.8%로 가장 높은 분포를 이루고 있으며 그 뒤로 '소프트웨어'(18.2%), '로봇'(4%), 'offline'(1.6%), 'build'(0.4%) 등이 분포되었다.

'online', 'offline', '로봇'과 '소프트웨어' 등의 키워드를 보고 유추해 볼 때 로봇을 활용하여 온라인/오프라인의 소프트웨어를 구축하고 로봇과 관련한 소프트웨어를 개발하는 개발자와 같은 직종으로 분석된다.

17. 온라인 머천다이저

[그림 3-27] 온라인 머천다이저 주요 키워드 분포

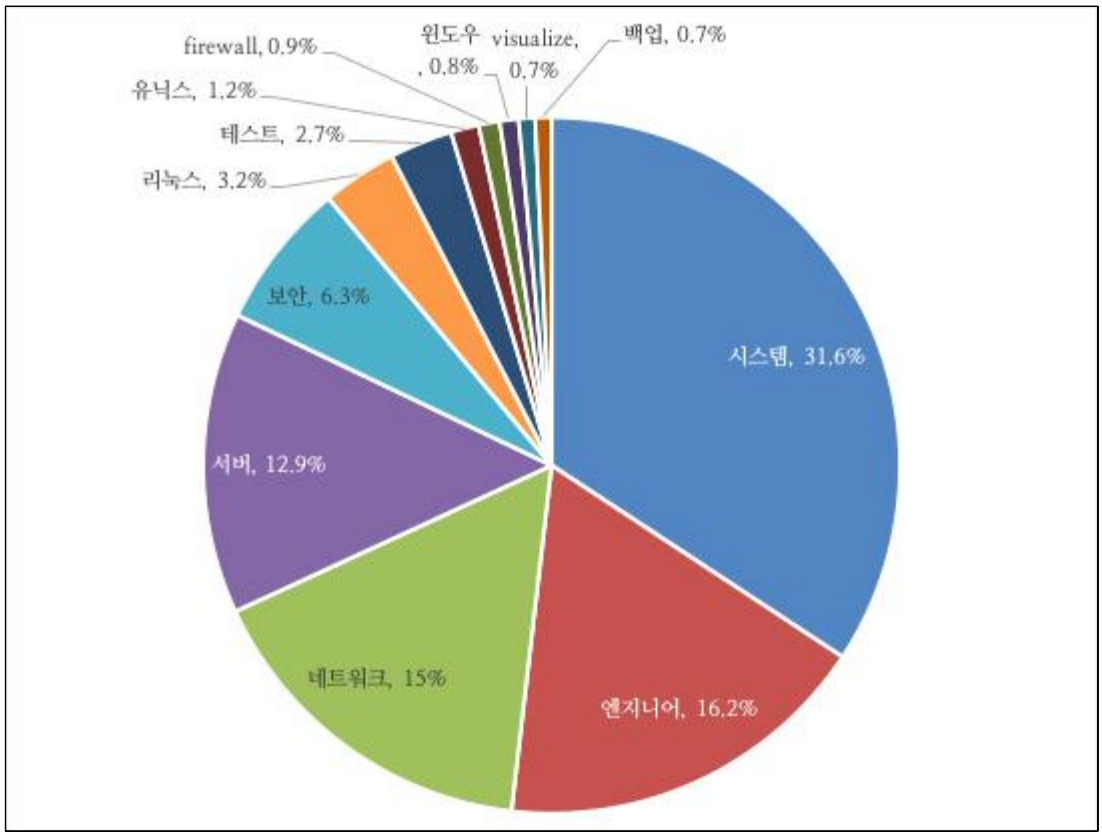


[그림 3-27]과 같이 17번째 직업의 키워드는 'md'가 47.1%로 가장 높은 분포를 이루고 있으며 그 뒤로 'online'(21.0%), '검사'(9%), 'amd'(6.6%), 'dba'(3.9%) 등이 분포되었다.

가장 비중이 많은 md와 두 번째 online 그리고 amd의 소프트웨어 및 dba 등의 데이터베이스 관리자가 등장하는 것으로 보아 온라인상에서 소프트웨어와 관련된 상품이나 제품을 거래하기 유의하게 데이터를 최적화 하고 관련 데이터베이스를 관리하는 관리자와 같은 직종으로 분석된다.

18. 시스템 소프트웨어 개발자

[그림 3-28] 시스템 소프트웨어 개발자 주요 키워드 분포

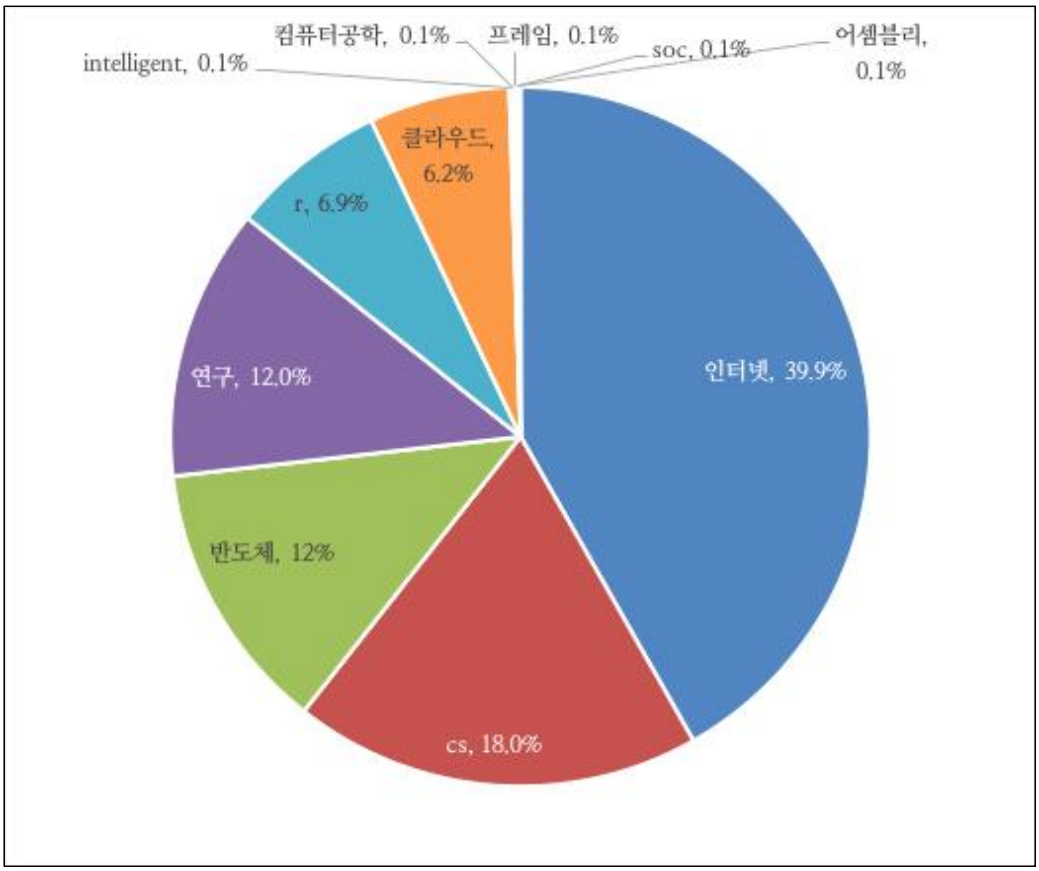


[그림 3-28]과 같이 18번째 직업의 키워드는 '시스템'이 31.6%로 가장 높은 분포를 이루고 있으며 그 뒤로 '엔지니어'(16.2%), '네트워크'(15%), '서버'(12.9%), '보안'(6.3%) 등이 분포되었다.

'시스템', '엔지니어' 및 '서버 네트워크', '보안'에 대한 키워드와 '테스트', '유닉스' 등의 키워드로 유추해 볼 때 시스템 전반의 소프트웨어를 운영 및 관리하고 서버와 네트워크를 개발하여 시스템에 문제가 없이 테스트와 백업을 할 수 있는 소프트웨어 개발자와 같은 직종으로 분석된다.

19. 임베디드 소프트웨어 개발자

[그림 3-29] 임베디드 소프트웨어 개발자

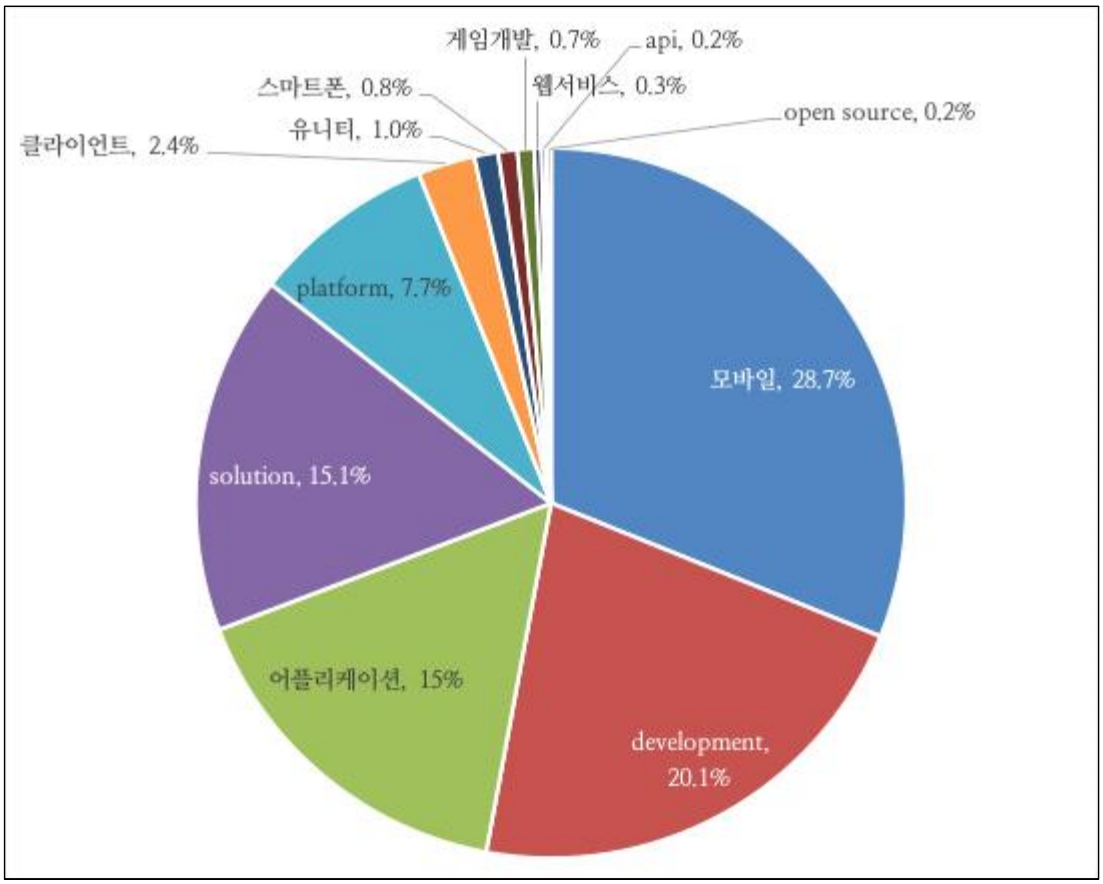


[그림 3-29]와 같이 19번째 직업의 키워드는 '인터넷'이 39.9%로 가장 높은 분포를 이루고 있으며 그 뒤로 'cs'(18.0%), '반도체'(12%), '연구'(12.0%), 'r'(6.9%) 등이 분포되었다.

'인터넷'과 '반도체'와 'cs 업무', '클라우드'와 '어셈블리' 등의 키워드로 유추해 볼 때 반도체와 관련된 소프트웨어 즉, cs(computer system) 업무를 하며 어셈블리를 통하여 반도체 공정을 하는데 도움을 주는 소프트웨어를 개발하는 개발자와 같은 직종으로 분석된다.

20. 모바일 게임 개발자

[그림 3-30] 모바일 게임 개발자 주요 키워드 분포

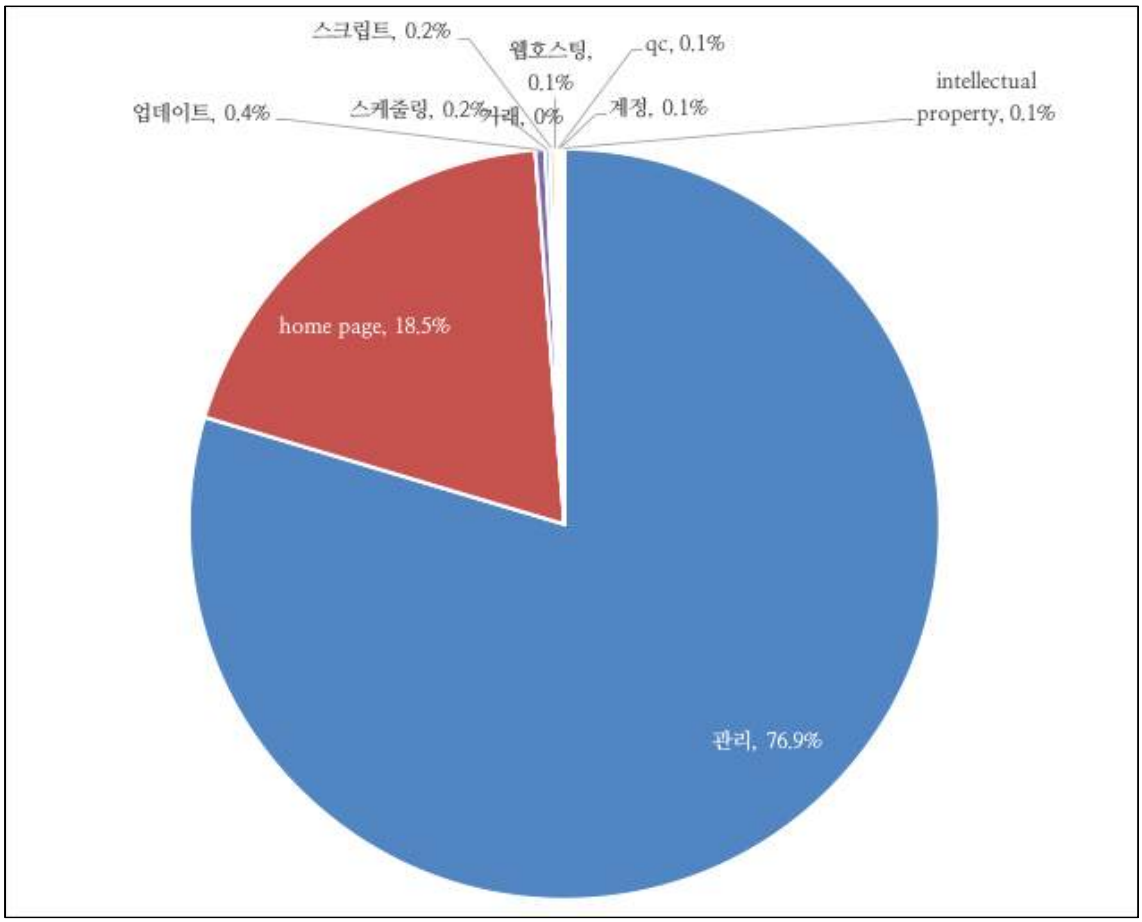


[그림 3-30]과 같이 20번째 직업의 키워드는 '모바일'이 28.7%로 가장 높은 분포를 이루고 있으며 그 뒤로 'development'(20.1%), '어플리케이션'(15%), 'solution'(15.1%), 'platform'(7.7%) 등이 분포되었다.

상기 언급한 모바일 어플리케이션 개발자와는 다르게 '모바일', '어플리케이션' 및 '플랫폼' 그리고 '클라이언트'와 '게임개발' 등의 키워드를 유추해 볼 때 모바일 플랫폼 기반에서 게임개발을 하는 개발자와 같은 직종으로 분석된다.

21. 웹 호스팅 관리자

[그림 3-31] 웹 호스팅 관리자

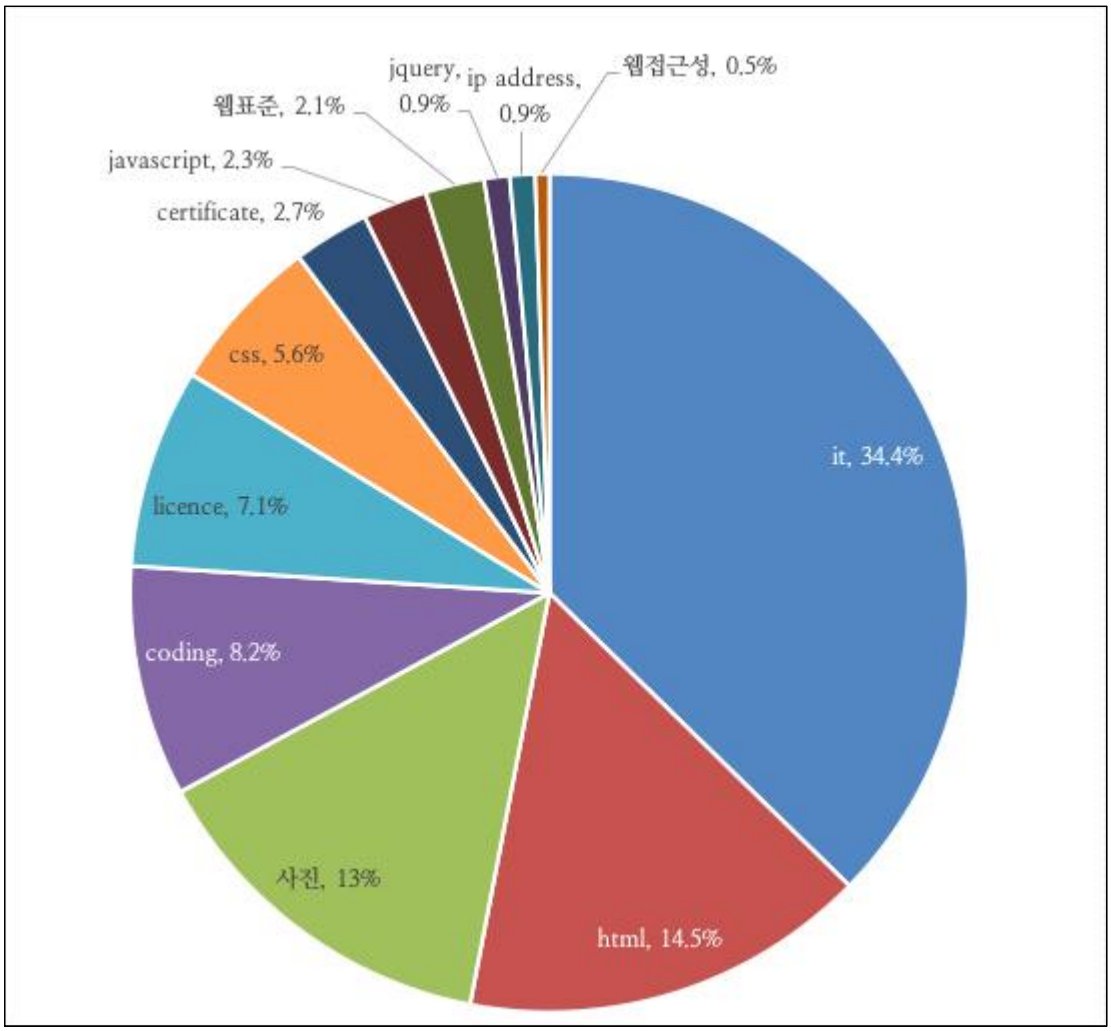


[그림 3-31]과 같이 21번째 직업의 키워드는 '관리'가 76.9%로 가장 높은 분포를 이루고 있으며 그 뒤로 'home page'(18.5%), '업데이트'(0.4%), '스케줄링'(0.2%), '스크립트'(0.2%) 등이 분포되었다.

'home page', '관리', '업데이트', '웹호스팅', '계정'과 같은 키워드로 유추해 볼 때 웹 호스팅과 관련하여 계정 관리 및 품질 관리와 홈페이지를 들어 갈 수 있도록 도와주는 관리자와 같은 직종으로 분석된다.

22. 웹 표준 전문가

[그림 3-32] 웹 표준 전문가 주요 키워드 분포

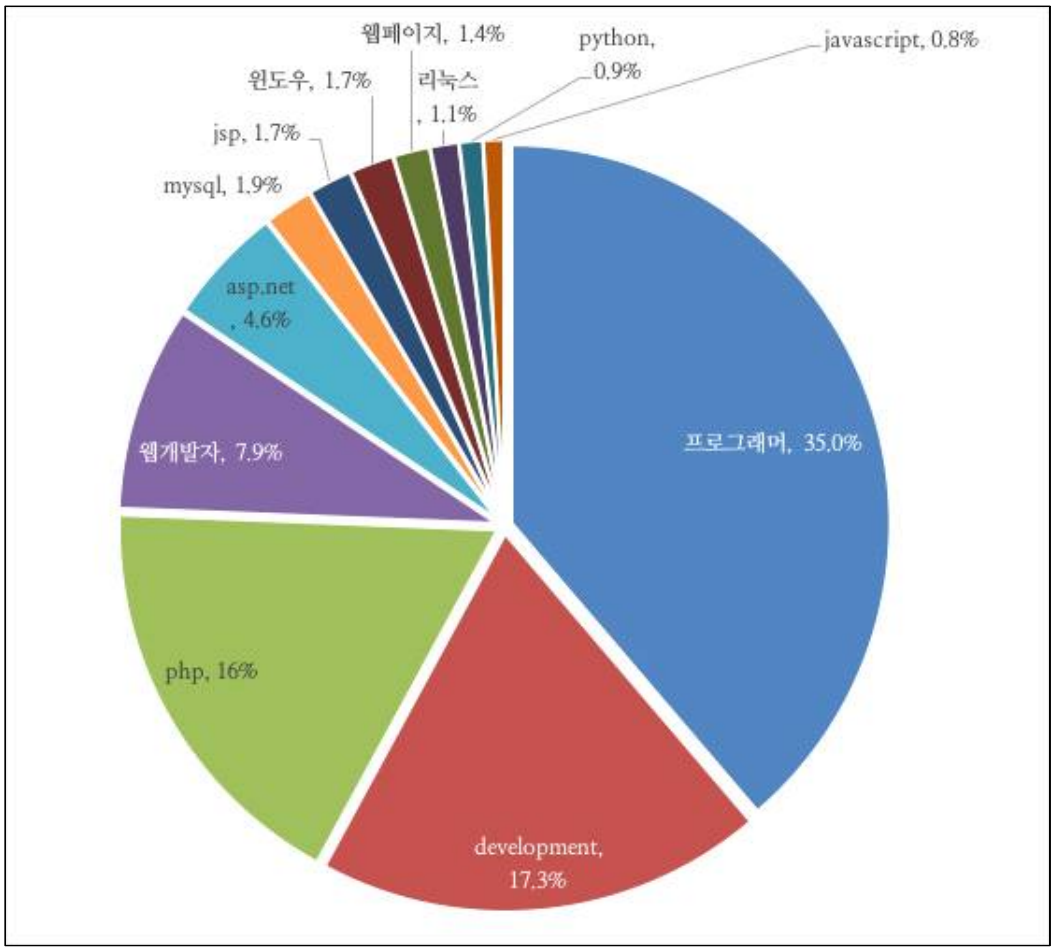


[그림 3-32]와 같이 22번째 직업의 키워드는 ‘it’가 34.4%로 가장 높은 분포를 이루고 있으며 그 뒤로 ‘html’(14.5%), ‘사진’(13%), ‘coding’(8.2%), ‘licence’(7.1%) 등이 분포되었다.

‘html’, ‘licence’ 등의 기술 관련 용어와 ‘웹 표준’, ‘웹 접근성’의 키워드로 유추해 볼 때 웹에 대한 분화된 직종으로 볼 수 있으며, 웹 표준과 접근성을 향상시켜주는 전문 직종으로 분석된다.

23. 웹 프로그래머

[그림 3-33] 웹 프로그래머 주요 키워드 분포

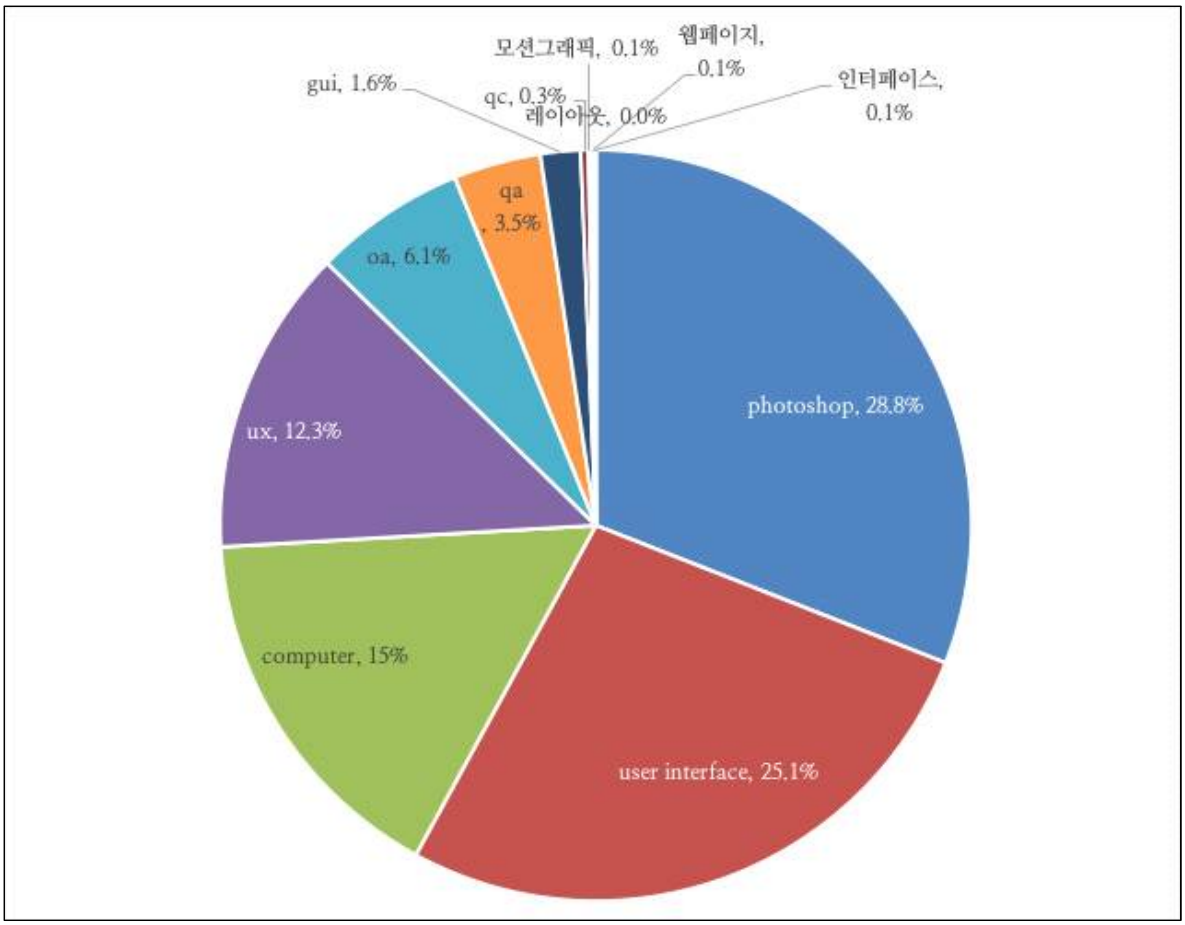


[그림 3-33]과 같이 23번째 직업의 키워드는 '프로그래머'가 35.0%로 가장 높은 분포를 이루고 있으며 그 뒤로 'development'(17.3%), 'php'(16%), '웹개발자'(7.9%), 'asp.net'(4.6%) 등이 분포되었다.

'프로그래머', 'development' 등 개발자의 역량이 필요하며, 'php', 'asp.net', 'jsp' 등 웹에 대한 개발 관련 영역을 도출하는 것을 확인 할 수 있다. 따라서 본 그룹의 직종은 웹 개발자와 관련된 것으로 확인할 수 있다.

24. UX/UI 디자이너

[그림 3-34] UX/UI 디자이너 주요 키워드 분포



[그림 3-34]와 같이 24번째 직업의 키워드는 ‘photoshop’이 28.8%로 가장 높은 분포를 이루고 있으며 그 뒤로 ‘user interface’(25.1%), ‘computer’(15%), ‘ux’(12.3%), ‘oa’(6.1%) 등이 분포되었다. ‘user interface’, ‘ux’ 등의 사용자 인터페이스 관련 키워드 및 ‘photoshop’, ‘모션그래픽’, ‘인터페이스’, ‘gui’ 등의 디자인 업무 관련 요소의 도출을 바탕으로 본 직종 그룹은 UX/UI에 대한 디자인을 하는 디자이너로 분류된다.

따라서 본 연구에서 분석된 24개의 직종 및 직업 분류체계는 <표 3-4>와 같다.

<표 3-4> 24개 직업분류를 통한 직업 이름 정의

No.	소셜미디어 마케터	모바일 어플리케이션 프로그래머	멀티미디어 디자이너	정보 보안 전문가	멀티미디어 콘텐츠 관리자	정보시스템 운영자	웹 운영자	사물인터넷 소프트웨어 개발자	E-커머스 관리자	서버 개발자	AR/VR 콘텐츠 개발자	웹 방송 엔지니어
1	전산	development	일러스트	보안	상품	유지	웹사이트	development	openmarket	구축	컨텐츠	방송
2	블로그	java	3d	erp	제작	보수	마켓	c	관리	프로그래밍	sns	document
3	바이럴	안드로이드	photoshop	solution	영상	정보기기	리뉴얼	program	상품	네트워크	digital	data
4	cafe	jsp	그래픽디자인	정보보안	시각디자인	rfid	communication	소프트웨어	소셜커머스	데이터베이스	포털사이트	영상
5	컴퓨터공학	ios	cg	si	cg	원격	app	응용프로그램	photoshop	frame	vr	multimedia
6	백업	어플리케이션	character	consulting	웹페이지	circuit	crm	embeded	그래픽	장애	통합인증	검색
7	원격	서버	그래픽	mes	pop-up	edi	sharing	펌웨어	수정	oracle	editor	agency
8	community	spring	2d	offline	cf	ip address	업데이트	.net	e-commerce	sql	bi	was
9	component	app	모델링	취약점	섬네일	source	리소스	사물인터넷	제품디자인	라이브러리	creative	스크립트
10	cpa	backend	모션그래픽	해킹	콘텐츠관리	최적화	backend	리눅스	sms	서버	ci	encoding
11	cpm	frontend	animation	암호	모션그래픽	qc	editor	프로그래밍	계정	전산	ar	시뮬레이션
12	cps	응용프로그램	flash	개인정보	source	정보	irc	윈도우	social media	solution	developer	미들웨어

No.	드론 애플리케이션 개발자	자동화 설계자	디지털 마케터	로보틱스 시스템 개발자	온라인 머천다이저	시스템 소프트웨어 개발자	임베디드 소프트웨어 개발자	모바일 게임 개발자	웹 호스팅 관리자	웹 표준 전문가	웹 프로그래머	UX/UI 디자이너
1	운영체제	설계	기획	online	md	시스템	인터넷	모바일	관리	it	프로그래머	photoshop
2	관리	analysis	정보	소프트웨어	online	엔지니어	cs	development	home page	html	development	user interface
3	모니터링	데이터	교육	로봇	검사	네트워크	반도체	어플리케이션	거래	사진	php	computer
4	웹서비스	자동화	direct digital marketing	offline	amd	서버	연구	solution	업데이트	coding	웹개발자	ux
5	클라이언트	consulting	웹서비스	build	dba	보안	r	platform	스케줄링	licence	asp.net	oa
6	모델링	autocad	financial	core	가공	리눅스	클라우드	클라이언트	스크립트	css	mysql	qa
7	character	시각디자인	정보처리	paging	검수	테스트	intelligent	유니티	웹호스팅	certificate	jsp	gui
8	community	빅데이터	계정	dsp	fae	유닉스	컴퓨터공학	스마트폰	qc	javascript	윈도우	qc
9	DRONE	검증	agent	encoding	isp	firewall	프레임	게임개발	계정	웹표준	웹페이지	모션그래픽
10	operator	검토	community	fpga	리테일	윈도우	soc	웹서비스	intellectual property	jquery	리눅스	웹페이지
11	원격	반도체	portal	ruby	amazon	visualize	어셈블리	api	p2p	ip address	python	인터페이스
12	cdn	process	ppl	source code	voc	백업	.net	open source	pop-up	웹접근성	javascript	레이아웃

4. 분석 후 직종/직업 재배치

기존 분류 체계인 한국표준직업분류(KCSO)의 경우 총 22개의 직종으로 SW인력 분류체계를 형성하고 있다. 또한 ITSQF의 경우는 총 28개의 역량으로 SW인력 분류체계를 형성하고 있다. 상기 언급한 7개의 직업분류체계와 본 연구에서 DTM을 활용한 데이터 기반 분류체계의 유사성을 확인하기 위하여 기존 분류체계의 코드와 매칭을 실시하였다.

24개의 대부분의 분류가 기존에 있었던 분류를 대변하고 있는 것을 확인할 수 있다. 그러나 ‘AR/VR 콘텐츠 개발자’와 ‘드론 애플리케이션 개발자’, ‘웹 방송엔지니어’, ‘멀티미디어 콘텐츠 관리자’ 등 기존 분류체계에서도 분류되지 않은 새로운 일자리에 대하여 코드 매칭이 비교적 잘 이뤄지지 않고 있다. 즉, 4차 산업 혁명 및 정보 기술의 고도화로 인하여 기존의 분류 체계의 시간에 따른 ‘정시성(Punctuality)’가 고려되지 않아 현재 상황에 맞는 일자리를 정의해주지 못한다는 것을 반증한다.

본 연구는 시간 흐름에 따른 추이를 고려를 할 수 있도록 기간을 설정하여 기간의 흐름별 해당되는 주요한 키워드가 무엇인지를 도출하기 위하여 DTM 알고리즘을 사용하였고, 분석을 적용하기 위한 수요 데이터로 실제 현장에서 요구하고 있는 모집 공고, 채용 공고의 역량, 필요 요구 사항, 해당 분야의 전문성 측정 등의 다양한 정보를 수집 및 정제하여 분석에 사용하였다.

또한 기존 분류체계의 경우 많은 사용자의 협업과 전문가의 회의를 통하여 발견된 직종이기 때문에 분류 자체가 문제가 있고 문제를 진단하려는 차원의 본 연구의 목적이 아니다. 또한 분석결과가 보여주고 있는 다른 시사점으로는 기존 분류체계를 반영하여 분류체계를 보다 정교화하고 최신화하여 계승한다는 점에서 해당 연구의 가치가 있다.

따라서 본 연구를 통하여 실시간으로 변화하는 일자리에 대한 진단과 함께 기존 분류체계들의 면밀한 검토를 통하여 보다 개선되고 계승된 데이터 기반의 분류체계를 분류 했다는 점에서 의의가 있다.

<표 3-5> 기존 분류체계 코드 매칭

명칭	ISCO-88	KSCO-07	KECO-09	NCS	ITSQF	O*net	SFIA
미디어 콘텐츠 디자이너							
멀티미디어 디자이너	2513	-	2042 20420	20-03-02-05	-	-	Animation development
UX/UI 디자이너	3532	22242	20411 20421	20-01-02-07	ITSQF11	15-1199.03 15-1134.00	User experience analysis User experience design User experience evaluation
시스템 소프트웨어 개발자							
시스템 소프트웨어 개발자	2411	2222 22221 22222	20310 20212	20-01-02-08	ITSQF13	15-1121.00	System design System installation/decommissioning System development management System integration
정보 보안 전문가							
정보 보안 전문가	2522	22212 22330	2023 20231	20-01-06-01 20-01-06-02 20-01-06-03 20-01-06-04 20-01-06-05	ITSQF03 ITSQF26 ITSQF27	15-1122.00	Digital forensics Information security Security administration Safety engineering Safety assessment
서버 개발자							
서버 개발자	3520	22241	20521	20-01-03-01	ITSQF17	15-1199.03	Facilities management Quality assurance Quality management Quality standards
미디어 콘텐츠 전문가							

멀티미디어 콘텐츠 관리자	2519	-	2042 20422	20-03-02-05	-	-	Animiationdevlopment
웹 방송 엔지니어	3531	22420 2250	20422	20-03-03-02 20-03-03-04 20-03-03-05	-	-	Methods and tools Performance management Quality assurance Quality standards
디지털 마케팅 전문가							
소셜 미디어 마케터	3532	22410	20421 20321	20-01-05-02	ITSQF19 ITSQF21	15-1199.1 0	Digital marketing
디지털 마케터	3532	13503	-	20-01-05-01 20-01-05-02	ITSQF19	15-1199.1 0	Business anaylsis IT strategy and planning Selling Sales support
E-커머스 관리자	3520	22420	20520 20420	20-01-03-01	ITSQF17	15-1199.0 3	Customer service support Quality management Information systems coordination
온라인 머천다이저	-	22420	20420	20-01-05-02	ITSQF20	15-1199.1 0	Sales support Selling
AR/VR 콘텐츠 관리자							
AR/VR 콘텐츠 개발자	-	22239	-	20-01-02-03	ITSQF14 ITSQF12	15-1132.0 0	Innovation Methods and tools Sustainability management
드론 애플리케이션 개발자							
드론 애플리케이션 개발자	-	22232	-	20-01-02-02	ITSQF14 ITSQF12	15-1132.0 0	Innovation Methods and tools

기존 분류체계 코드 매칭을 통하여 해당 직종의 특성을 보다 면밀하게 파악하고 [그림 3-35]와 같이 12개의 직종, 24개의 직업을 재배치하였다.

[그림 3-35] SW산업 직종/직업 재배치



변화된 일자리는 4개 직종인 ‘모바일 개발자, 서버 개발자, 웹 전문가, 임베디드 소프트웨어 전문가’ 등의 변화가 있었으며, 총 9개의 직업이 변화하였다.

발견된 일자리로는 4개 직종인 ‘미디어 콘텐츠 전문가, 디지털 마케팅 전문가, 드론 애플리케이션 개발자, AR/VR 콘텐츠 개발자’ 등을 발견하였으며, 총 9개의 직업을 발견하였다.

해당 직종에 대한 기존 코드 매칭을 통하여 면밀히 분석 후 <표 3-6>과 같이 총 12개의 직종과 24개의 직업에 대한 정의를 하였다.

<표 3-6> 신규 일자리 정의 (12개 직종, 24개 직업)

직종	직업	정의
모바일 개발자		프로그래밍 언어 활용 능력을 이용하여 모바일 환경에서의 소프트웨어 설계 및 개발을 수행한다. 모바일 어플리케이션 개발 및 구현을 위한 조직의 요구 분석 수행 및 방법론 설정 등 개발 관련한 전반적인 업무를 담당한다.
	모바일 애플리케이션 프로그래머	스마트폰이나 태블릿 PC 등 휴대용 단말기를 위해 개발되는 응용 소프트웨어(애플리케이션)를 설계, 작성하고, 유지 및 개선하는 자를 말한다.
	모바일 게임 개발자	모바일에 기반하여 시나리오의 순서에 따라 배경과 상황 등을 각종 언어로 프로그래밍하며 컴퓨터 게임의 운용을 위한 프로그램을 제작하는 자를 말한다.
임베디드 소프트웨어 전문가		하드웨어와 소프트웨어의 지식을 활용하여 산업기기 및 스마트 시스템 등에 내장되는 임베디드 소프트웨어의 아키텍처를 설계하고 구체화하는 전반적인 업무를 담당한다.
	사물인터넷 전문가	펌웨어와 임베디드 소프트웨어 개발 능력을 기반으로 사용자의 목적에 맞는 사물인터넷 소프트웨어, 펌웨어 등을 개발, 설계, 작성하고, 유지 및 개선하는 자를 말한다.
	자동화 설계자	생산 공정 등의 자동화 시스템을 구축 및 운영하여 공정의 효율성을 향상시키며, 자동화에 따라 파생되는 데이터를 수집한다. 데이터 분석을 통해 설비의 문제점 등을 검증 및 검토하여 최적화 된 자동화를 지원하는 설계 및 컨설팅을 하는 업무를 담당한다.
	로보틱스 시스템 개발자	로봇의 구성요소를 연구하고, 인공지능 및 센서기술, 영상처리기술을 이용하여 소프트웨어를 개발하고, 장치를 설비하여 단일체로 로봇을 조립, 제작하여 운용하는 업무를 담당한다.
	임베디드 소프트웨 개발자	반도체 등과 관련된 하드웨어와 소프트웨어 문제를 해결하여 장비를 설치하고 엔지니어들이 장비를 사용할 수 있도록 환경을 조성하는 업무를 담당한다.
정보시스템 및 웹 운영자		시스템 사용자들에게 기술적인 지원 및 훈련을 시키고, 시스템 전반의 관리와 운영을 담당하며, 장애 발생 시 문제를 처리한다. 웹 사이트나 인터넷 홈페이

	지의 운영 전반에 걸쳐 실무적인 책임을 진다.
정보시스템 운영자	시스템 사용자들에게 시스템 최적화, 품질 관리 등의 기술적인 지원 및 시스템 전반의 관리와 운영을 담당하며, 장애 발생 시 문제를 처리한다. 웹 사이트나 인터넷 홈페이지 및 관련 정보기기의 운영 전반에 걸쳐 실무적인 책임을 진다.
웹 운영자	웹 사이트나 온라인 마켓 운영 전반에 걸쳐 실무적인 책임을 지는 자를 말한다. 웹 사이트 구축목적에 맞게 서비스를 리뉴얼하거나 운영 시스템의 리소스를 관리한다. 웹 사이트 활성화를 위한 부가서비스를 운영한다. 웹 사이트 업데이트를 위한 적절한 관리를 담당하여 고객들의 요구사항, 불만사항에 대한 신속한 처리를 한다.
웹 전문가	웹서버 구축 및 웹 호스팅 운영에 대한 기술적인 책임을 지며 웹의 신기술을 습득하고 상품화하며, 시험하는 업무를 수행한다. 기본적인 프로그래밍 능력 및 웹 표준과 규정에 관한 이해를 기반으로 웹에서 비즈니스를 생산하기 위해 필요한 고객의 요구사항을 관철하는 업무를 담당한다.
웹 호스팅 관리자	웹 호스트를 고객에게 제공하며 관리하는 업무를 수행한다. 고객에게 발생하는 기술적인 장애를 해결하며, 고객의 지적재산권 및 호스팅 품질 향상을 위한 주기적인 시스템 업데이트와 계약사항 문서화, 신규 호스팅 상품을 기획하고 개발하는 자를 말한다.
웹 표준 전문가	웹 기반의 콘텐츠를 이용하기 쉽게 혹은 웹 서비스개발 및 관리를 위하여 웹을 표준화하여 사용자들의 사용성과 접근성을 올릴 수 있도록 설계하는 업무를 담당한다.
웹 프로그래머	프로그래밍 언어 능력을 활용하여 웹상에서 디자이너가 기획한대로 포맷을 잡으면 이를 그대로 보여주기 위해 필요한 프로그램을 제작한다.
미디어 콘텐츠 디자이너	컴퓨터 그래픽 등을 활용하여 웹페이지, 동영상, 게임, 컴퓨터 및 모바일 애플리케이션과 같은 미디어 콘텐츠에 필요한 아이콘, 이미지, 영상 효과 등의 시각적 구성요소 및 사용자 환경을 디자인 한다.
멀티미디어 디자이너	디자인 소프트웨어 활용 및 모션그래픽 편집 능력을 기반으로 인터넷 혹은 방송, 영화 등에 활용하기 위한 자막, 이미지, 영상효과 등의 그래픽을 디자인하는 자를 말한다.
UX/UI 디자이너I	인지공학, 인체공학, 심리학, 산업디자인 등의 융복합

	전문 지식을 활용하여, 웹페이지, 애플리케이션, 컴퓨터 및 모바일 기기 등의 사용자 환경(UI, User Interface)을 사용자 경험(UX, User Experience)에 입각하여 편의성과 활용성을 극대화할 수 있도록 연구 · 개발하고, 제품이나 서비스에 그 결과를 반영하여 디자인하는 자를 말한다.
시스템 소프트웨어 개발자	컴퓨터 시스템의 자체기능 수행명령체계인 시스템 소프트웨어를 연구 및 개발하고 설계하며, 이와 관련한 프로그램을 작성하는 업무를 수행한다.
시스템 소프트웨어 개발자	상동
정보 보안 전문가	해커의 해킹이나 개인정보 유출 및 바이러스 발생에 대비하여 온라인, 오프라인 전산망의 보안을 유지하기 위하여 필요한 보안프로그램을 개발하고, 보안 상태를 점검하며 보안 취약점을 위한 다각적인 해결책을 제시하는 자를 말한다. 정보 시스템 상 개인정보의 유출을 막고 관련 보안 체계를 개발하는 전문가도 이 분류에 포함한다.
정보 보안 전문가	상동
서버 개발자	데이터베이스 및 네트워크 서버에 관한 지식을 이용하여 서버내 최적화 된 데이터베이스를 설계 및 구축하고 서버간 네트워크망을 개발, 기획하는 업무를 수행한다. 구축 된 서버망에 발생하는 네트워크 장애현상에 대한 해결책을 제시하고 데이터베이스 수집 자료의 효용성을 높이기 위해 데이터베이스 프로그램을 만들고 관리 및 백업을 수행하는 자를 말한다.
서버 개발자	상동
미디어 콘텐츠 전문가	모바일 및 웹 콘텐츠를 제공하기 위한 전략을 수립하고 실행하며, 시각디자인, 모션그래픽 등의 멀티미디어 편집 기술을 활용하여 콘텐츠화 하는 전반적인 업무를 담당한다.
멀티미디어 콘텐츠 관리자	모바일 및 웹 콘텐츠를 제공하기위 한 전략을 수립하고 실행하며, 시각디자인, 모션그래픽 등의 멀티미디어 편집 기술을 활용하여 요구 사항을 충족하는 콘텐츠 솔루션을 설계하고 관리하는 자를 말한다.
웹 방송 엔지니어	디지털 장비를 이용해 촬영 혹은 녹음한다. 촬영 혹은 녹음 된 파일을 방송에 적합한 형태로 편집한다. 효과음과 배경음을 컴퓨터로 믹싱한다. 편집된 파일을 배포 가능 형태로 인코딩한다. 신규 방송용 장비를 도입하고 관리한다. 업계의 신기술 동향을 파악하

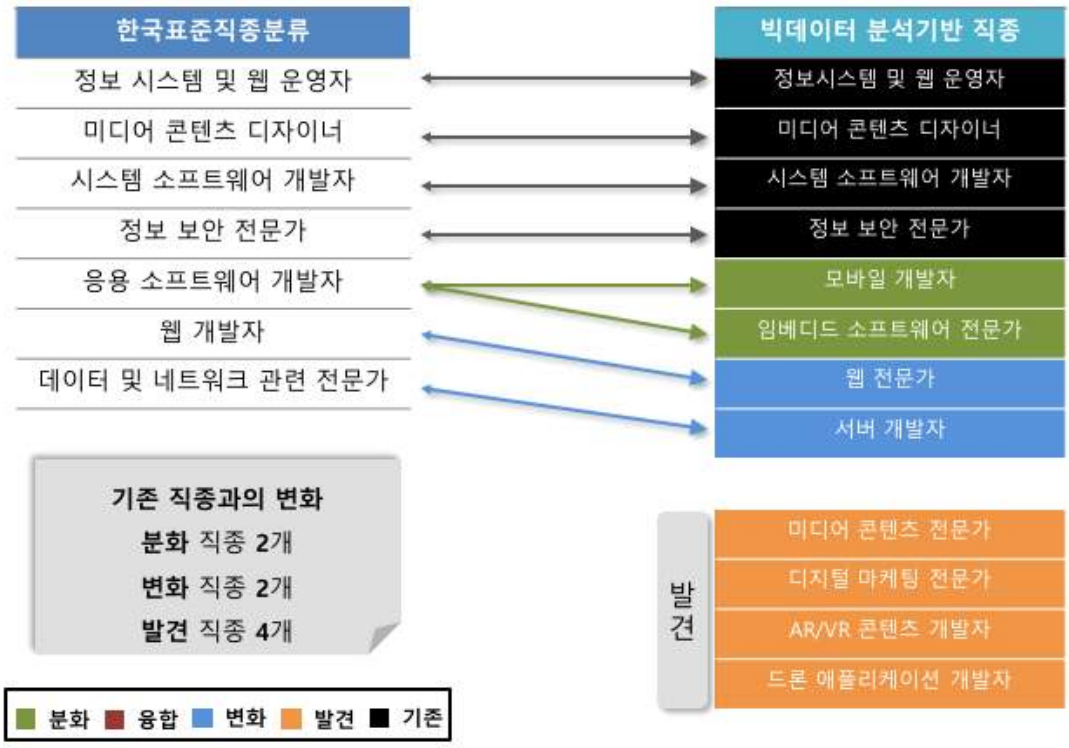
	고 연구한다.
디지털 마케팅 전문가	소셜미디어, 웹 환경을 바탕으로 조직의 목표에 부합하는 디지털 마케팅을 기획하고 실행하며 내외부 이해 관계자를 관리한다. 현지화 된 디지털 및 전통적인 미디어에 대한 이해를 바탕으로 디지털 구매 판매 전략에 관한 미디어 영향을 전반적으로 분석하는 업무를 담당한다.
소셜 미디어 마케터	컴퓨터공학과 마케팅 지식을 기반으로 소셜 미디어를 통해 고객과의 상호 작용, 매력적인 콘텐츠 홍보, 수익 창출 기회 확대 등을 위한 기술 및 영업 전반의 업무를 담당한다
디지털 마케터	소비자와의 긴밀한 접촉을 통해 직접 제품을 판매하고 프로젝트 목표 설정, 브리핑, 결과 분석 등을 수행한다. 제품 공급업체를 관리하고 주요 프로젝트의 달성 목표 관리 및 일정 수립, 모니터링 등의 전반적인 관리를 수행하는 자를 말한다.
E-커머스 관리자	소셜커머스, 오픈마켓 등 전자상거래를 통해 상품 판매를 지원하며, 상품 공급업체를 관리하여 상품의 갱신 및 변경점을 추적하여 업데이트를 수행한다. 업계 소식, 기술, 동향, 및 수익의 기회를 파악하고 전자상거래 관련 기술을 지원하는 자를 말한다.
온라인 머천다이저	오픈마켓 또는 전자상거래 상품과 관련해 모인 데이터를 최적화하며 제품 경쟁력을 비교 분석한다. 고객의 의도를 파악하고 그에 맞는 상품 콘텐츠 및 보완 제품을 기획하고 홍보하는 마켓 운영의 전반적인 업무를 수행하는 자를 말한다.
AR/VR 콘텐츠 개발자	증강현실 및 가상현실 기술을 활용하여 소셜미디어에 최적화 된 실감형 콘텐츠를 개발한다. 소셜미디어 및 커뮤니티의 의견을 수집하여 창의적인 AR/VR콘텐츠를 기획하며 성능 최적화를 수행하는 업무 담당하는 자를 의미한다.
AR/VR 콘텐츠 개발자	상동
드론 애플리케이션 개발자	드론의 이동, 통신 및 분석 방법을 개선하기 위한 새로운 기술을 연구한다. 무인 항공 시스템에 적용 할 드론 모형을 제작하고 이를 위한 어플리케이션을 개발하고 시뮬레이션 하는 업무를 담당한다
드론 애플리케이션 개발자	상동

5. 일자리 변화상(직종/직업)

다이나믹 토픽 모델링의 결과 총 12개의 직종과 24개의 직업을 도출하였다. 도출된 직종 및 직업은 한국표준직업분류(KSCO)를 기준으로 기존의 직종과의 변화 양상을 파악하였다.

한국표준직업분류(KSCO)에서 SW인력에 대한 분류는 대분류 5개, 중분류 14개, 세분류 30개로 분류되어 있다. 기존의 30개의 세분류로 된 직업에 대하여 24개로 재정립하고 변화에 이슈를 분석한 결과는 [그림 3-36]과 같다.

[그림 3-36] 일자리 직종 변화

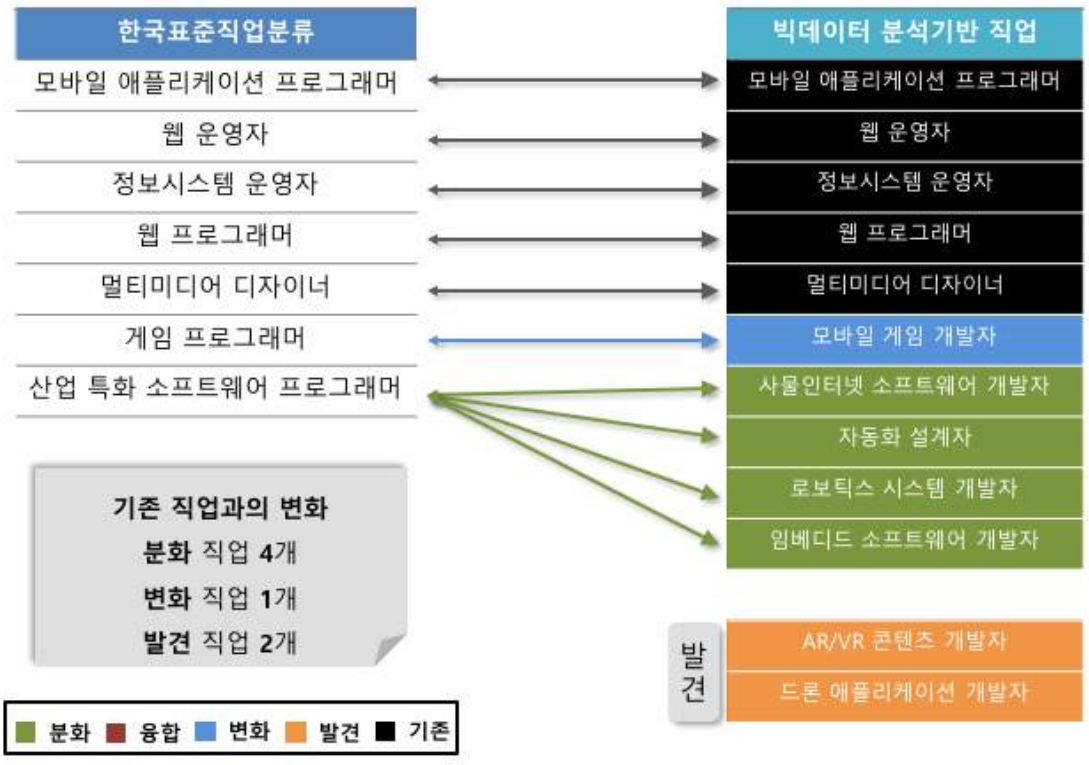


기존의 분류체계와 새로운 분류체계의 변화상은 기존의 하나의 직업 및 직종에서 기술적 고도화와 수요의 급증으로 인하여 분화된 직종인 '응용 소프트웨어 개발자(KSCO)'가 존재하며, 새로운 분류 체계에서 '모바일 개발자' 직종과 '임베디드 소프트웨어 전문가' 직종으로 분류되었다.

기존 직종에서 새롭게 변화된 직종은 '웹 개발자'와 '데이터 및 네트워크 관련 전문가'에서 '웹 전문가'와 '서버 개발자'로 변화하였다. 기존의 직종의 정의와 신규 발생 직종의 정의를 비교하였을 때 기존의 웹 사용량과 현재의 웹 사용량의 트래픽 수는 기하급수적으로 차이가 존재한다. 따라서 단순히 웹을 운영하는 것이 아닌 웹에 대한 전문적인 지식을 겸비하여 운영, 관리 때로는 개발까지 할 수 있는 역량이 필요하다. 따라서 기존의 협의적인 내용의 웹 운영자에서 광의적으로 역량을 포괄할 수 있는 웹 전문가로 변화하였다.

데이터 및 네트워크에 전반적인 내용을 다루던 직종에서 그 내용들의 기반이 되는 서버를 개발하는 개발자로 변화하면서 기존의 하드웨어나 소프트웨어 체계의 서버 구축과는 다르게 클라우딩 기술의 발전과 가상 머신의 발전으로 기존과 비교 할 수 없는 데이터의 증가와 사용자의 네트워크 기술 인지도의 증가로 인하여 서버는 가장 중요한 핵심 분야 중 하나로 도약하였다.

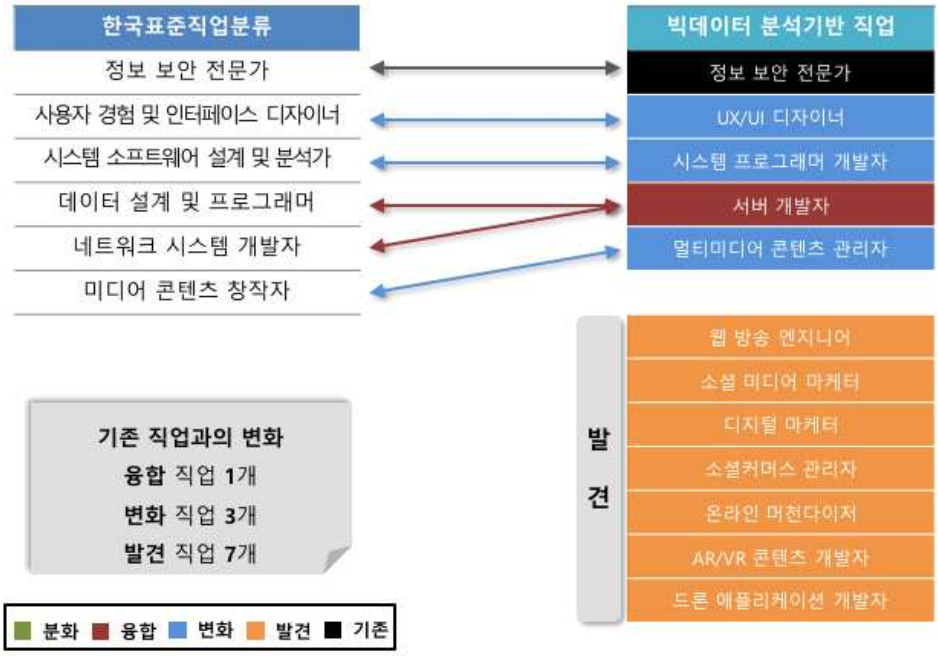
[그림 3-37] 일자리 직업 변화



[그림 3-37]과 같이 기존의 분류체계에서 직업적 변화는 ‘산업 특화 소프트웨어 프로그래머’의 분화가 가장 두드러진다. 기존의 분류에서는 산업에 특화된 소프트웨어를 프로그래밍 한다는 포괄적인 정의로 직종을 정의하고 있으나 현재 산업에서 정보기술의 접목 속도는 매우 빠르기 때문에 도출된 토픽에 의거하여 4가지로 분야인 ‘사물인터넷 소프트웨어 개발자’, ‘자동화 설계자’, ‘로보틱스 시스템 개발자’, ‘임베디드 소프트웨어 개발자’로 분화하였다.

기존의 게임 프로그래머의 경우 최근 PC 게임보다 모바일의 게임 산업이 두드러지게 커지는 현상을 반영하여 ‘모바일 게임 개발자’로 변화하였다. 또한 AR/VR 기술의 발전과 스마트폰의 발전은 새로운 분야를 창출하게 하였으며, AR/VR로 즐길 수 있는 콘텐츠를 개발하는 ‘AR/VR 콘텐츠 개발자’ 직업을 발견하였다. 더불어 무인항공의 이슈가 점점 증가하여 드론의 개인 소유가 점차 증가하고 산업 측면에서도 드론을 활용하여 유통 등의 다양한 방법의 적용 시도가 있었다. 따라서 드론에 대한 소프트웨어의 개발이 필요하게 되었으며, 어플리케이션 구현 또한 중요한 산업 중 하나로 자리매김하고 있다. 이에 따라 ‘드론 애플리케이션 개발자’를 발견하였다.

[그림 3-38] 일자리 직업 변화



추가적으로 [그림 3-38]과 같이 데이터 설계와 네트워크 시스템을 개발하는 직종은 온라인 클라우딩 서비스 및 스토리지 소프트웨어의 기술 발전과 강화로 서버에 대한 이슈가 더욱더 중요한 부분이 되었다. 따라서 각각의 데이터 설계와 네트워크를 따로 분별해서 보는 것이 아닌 서버 상에서 이뤄지는 모든 개발 업무를 담당하는 직종으로 '서버 개발자'로 융합되었다.

온라인 마켓의 발전과 사용자간의 커뮤니티의 형성을 통한 정보 교류 플랫폼의 증가로 오프라인을 이용하지 않고 온라인에서 물건을 보다 쉽고, 싸게 구매 할 수 있는 환경이 조성되었다. 따라서 해당 현상을 관리하고 개발하고 구현하는 직종의 필요성이 증가하였다. 또한 1인 미디어의 등장으로 많은 사람들이 유투브 등을 이용하면서 영상에 대한 마케팅과 엔지니어링이 필요해졌다. 따라서 '웹 방송 엔지니어, 소셜 미디어 마케터, 디지털 마케터, 소셜커머스 관리자, 온라인 머천다이저' 등의 웹과 관련되며 커뮤니티에 관련된 개발 및 구현을 할 수 있는 역량이 단계적으로 증가하였다.

이와 같은 결과를 바탕으로 다음 2절에서는 2장 2절에서 제시한 PEST분석을 바탕으로 SW직종에 대한 매칭을 통해 환경 분석을 통한 직종에 대한 특성을 도출하였다.

제2절 거시 환경 분석과 일치성 검토

1. 거시 환경 분석(PEST)과 SW직종 재배치의 일치성 검토

상기 분석된 자료를 기반으로 거시 환경을 분석한 PEST 분석에 맞춰 현재 시장에 알맞은 분석이 되었는지 일치성을 검토하고자 한다. 기존 분석된 PEST를 정리하면 [그림 3-39]와 같다.

[그림 3-39] SW인력 거시 환경분석



거시 환경의 정치적(Political)을 간단히 정리하면 다음 3가지 현황으로 파악된다.

(가상현실 전문 펀드 조성) 2016년부터 2017년까지 가상현실에 대한 산업을 지원하기 위하여 정부차원으로 총 400억 원의 투자비용을 충당할 수 있는 전문 펀드를 조성하였다. 정보 기술의 발달로 현재 스마트폰과 같은 스마트 기기는 지속적인 성장 추세에 있으며, 미래 성장 동력으로 가상현실을 주목하고 있다(미래창조과학부 외 3개 부처 합동, 2016). 삼성, 구글 , 페이스북, 소니 등도 가상현실에 대한 기술을 활용하는 시장의 현황을 유심하게 지켜보고 있으며, 초기 생태계의 선점을 위한 노력을 지속적으로 진행 중이다.

주요 국가의 정책적인 방향으로 첫째, 중국은 '가상현실산업발전백서 5.0'을 발표하여 독자적 기술개발의 표준안을 마련하고, 미국은 '혼합현실(MR:Mixed Reality)' 명칭으로 하여 10대 미래 정보 기술로 지정하였으며, 유럽 권에서는 VR로 구현된 가상 투어를 상용화하여 유럽 문화유산을 소개 할 수 있는 '아키오가이드(Archeoguide)' 프로젝트를 수행하고 있다. 국내 시장에서는 아직 플랫폼과 콘텐츠의 경쟁력이 부족하다. 따라서 콘텐츠와 관련한 투자 펀드를 확보하고 안정적으로 투자를 모색 할 수 있는 환경을 조성할 수 있는 '가상현실 전문 펀드'를 조성하여 신 시장 및 플랫폼을 선점하고, 글로벌 기업으로 부상 할 수 있도록 기회를 제공하고 있다(미래창조과학부 외 3개 부처 합동, 2016).

(디지털콘텐츠코리아 펀드 운영) IT 기술의 발전에 따라 디지털 상에서 활용되는 콘텐츠의 발전도 종속 관계로 같이 발전하고 있다. 그러나 디지털 콘텐츠 영역은 아직도 총 콘텐츠의 매출에 30%에 미치지 못한다는 실정이다. 2013년 기준으로 디지털 콘텐츠 시장의 비중은 지식정보와 게임의 비중에 전체의 66%에 육박하여 매우 비평준화된 시장을 구성하고 있다는 것을 확인하였다. 따라서 정부는 해당 디지털 콘텐츠의 발전과 지속적인 경쟁력을 확보하기 위하여 디지털콘텐츠코리아 펀드를 14년도부터 운영하고 있다(주기환, 2014).

(SW분할발주 및 SW품질성능평가시험) 국내의 산업체가 실질적으로 활용할 수 있도록 SW품질성능평가시험의 기준을 만들어 SW에 대한 품질을 평가하고 서비스의 경쟁력을 확보 할 수 있는 여건을 조성해주어 글로벌 시장에서 경쟁력을 갖게 하는 것에 의의가 있다. 기능성, 효율성, 신뢰성, 유지보수성, 사용성, 이식성의 6가지의 항목에 대한 측정을 더욱 강화할 방침이다 (한국정보통신기술협회, 2010).

경제적 관점에서 SW 시장을 확인하면, 먼저 기존에는 없었던 새로운 소비 패턴이 등장하였다.

(소셜커머스) 인터넷을 활용하는 소비자는 과거와는 다르게 정보 수용의 속도와 양, 다양성에 있어서 변화의 폭이 매우 크다. 오프라인에서 직접 물건을 보고 구매하는 것과는 다르게 오프라인에서 물건을 확인하고 온라인에서 구매를 진행하는 사람들도 많이 증가하였다. 커뮤니티 안에서의 공동구매 등의 구매의 패턴이 변화함에 따라서 구매를 할 수 있게 만들어주는 플랫폼인 '소셜커머스' 및 'O2O 비즈니스 모델'의 수요가 급증하고 있다. 'Z세대(Z generation)'의 등장으로 95년생 이후의 출생한 사용자의 경우 온라인에서 접하는 정보는 보다 신뢰하는 경향을 가지고 있으며, 온라인에서 구매 활동이 왕성하다는 것을 알 수 있어 기존의 오프라인에서만 한정적으로 제품과 서비스를 구매하는 것과는 다른 소비패턴을 보여주고 있다.

(공유경제) 소셜 커머스의 발전과 함께 협력적 구매 즉, 공동구매로 규모의 경제를 활성화시켜 비용을 감소하는 것을 넘어서 직접 소유하지 않고 '공유'하는 개념의 경제적 환경 변화가 일어나고 있다. 자동차를 공유하거나, 고가의 유모차를 직접 사지 않고 한시적으로 이용 할 수 있도록 공유하는 플랫폼의 산업이 발전하고 있다. 에어 비앤비(Air BnB)의 경우 빈집을 대여하여 여행자에서 보다 경쟁적인 비용을 지불하고 사용할 수 있게 하여 사용자는 비용 절감을 운영자는 수익 창출을 할 수 있는 새로운 비즈니스 모델까지 제시하고 있다.

(인더스트리 4.0) 4차 산업의 일환으로 기존의 제조업의 기술과 접목 할 수 있는 정보 기술을 개발하여 현재보다 효율을 증가시키고 비용을 절감시켜 고수익을 얻을 수 있도록 인더스트리 4.0을 정부에서 공시하였다. 3D 프린터의 기술과, 사물인터넷(IoT), 로봇 산업, 빅데이터를 활용한 제조업 시스템 최적화, 인공지능(AI) 등의 기술을 접목하여 사양되고 있는 제조업을 다시 재도약 시킬 수 있는 환경을 조성하고 있다. 제조업 기반의 성장 동력을 갖고 있던 우리나라의 경제력에서 인더스트리 4.0과 스마트 팩토리 (Smart Factory)의 발전은 지속적으로 우리나의 경제 성장 동력 중 중요한 부분을 차지할 것이다.

사회문화적 요소로는 소비자의 니즈에 정확하게 맞는 맞춤형 서비스를 받는 소비자의 선호가 증가하고 있다는 점이다.

(커스터마이징) 최근 기업들이 소비자가 원하는 것을 맞춰서 제공하는 이른바 '커스터마이징' 전략을 사용하면서 소비자는 남들과 똑같은 것이 아닌 남들과는 차이가 있어 개성적이고 최신의 것을 제공 받는데 상당히 만족감을 갖는다. 스타벅스의 '사이렌 오더' 등의 정보 기술의 확산으로 소비자는 언제 어디서든 자신이 원하는 자신만의 제품을 제공 받을 수 있다.

(모바일 온리족) 국내 스마트폰 보급률이 90%가 넘을 정도로 현재 국내 시장에서의 스마트폰의 입지는 매우 중요하고 모바일 시장의 파이도 점점 커지고 있다. 기존에는 전화 및 메시지 등의 기능으로 한정적으로 사용했던 모바일을 현재는 영화부터 게임, 비즈니스, 투자, 은행 업무, 독서 등 다양한 분야에서 다양한 내용들을 한꺼번에 하나의 모바일 디바이스에서 활용하고 있다. 또한 태블릿 등의 다양한 크기의 이동 기기의 발달로 인하여 중장년층에서는 립프로깅(Leapfrogging) 현상이 나타나고 있으며, 20-30대에서는 스마트폰이 없으면 활동을 못하게 되는 '노모포비아(Nomophobia)' 현상이 발생하고 있다.

(1인 미디어의 출현) 과거에도 존재하고 있던 미디어가 4차 산업 등 정보기술의 발전으로 인하여 연예인뿐만 아니라 일반인도 자유롭게 미디어를

제작하고 유포할 수 있는 환경이 조성 되었다. 대표적으로 유튜브와 국내에서는 아프리카 tv의 플랫폼이 대표적인 예시이며, 소비자는 이런 미디어 창작자의 콘텐츠를 보면서, 대리 만족과 간접 경험 등의 만족감을 얻는다. 기존의 공영 방송에서 진행하는 미디어와 다르게 남녀노소 누구든지 만들 수 있는 구조로 콘텐츠를 직접 운영할 수 있으며, 운영을 위한 영상 편집 지식들이 요구되고 있다.

(콘텐츠 소비 패턴의 변화) 2차 산업혁명을 통하여 제품의 대량생산이 가능해져 공산품의 재화 소비가 많았던 20세기와는 다르게 21세기에는 개인용 컴퓨터의 보급과 스마트폰의 발달 등의 기술력의 발전으로 소비자는 보다 자신의 개성에 부합한 콘텐츠를 추구하는 성향을 갖게 되었다. 타인이 소유하고 있는 공산품이 아닌 자신 만의 콘텐츠를 갖고 스토리를 갖고 소비를 함으로써 자신의 존재의 이유를 모색하고 있다.

기술적 시장은 모바일, AR/VR, 드론, 로봇, 사물인터넷 등의 개발과 발전이다.

(모바일) 상기 언급한 것과 마찬가지로 개인용 컴퓨터에서 이제는 모바일로 사용자의 사용 패턴의 변화가 생겼다. 전화와 메시지만 가능하였던 이동전화에서 국제적인 전산망의 구축과 반도체의 발달로 인하여 스마트폰이 개발되었다. 기존의 버튼식의 휴대폰이 아닌 터치식으로 변경함으로써 사용 용이성을 증가시켰으며, 데이터 활용이 가능해짐에 따라 개인용 컴퓨터에서 인터넷을 연결한 것과 다르게 언제 어디서든 인터넷을 사용하고 정보를 제공 받을 수 있는 ‘N 스크린’ 시대가 되었다.

(AR/VR) 정치적인 거시 환경에서도 언급했듯이 과거 가상현실의 연구가 지속적으로 일어나고 모바일화가 진행되면서, 영상 기술도 함께 개선되었다. 시장 규모도 2020년에는 약 5조원을 예측하고 있으며, 2025년에는 약 13조원을 예측하고 있다. 최근의 기술동향으로 ‘포켓몬go’의 증강현실을 활용한 게임이 개발되면서 전 세계적으로 파급력 있는 파워 콘텐츠로 급부상하였다. 최근에는 가상현실을 체험할 수 있는 ‘VR방’이 증가하고 있으며, 영상 기술이

고도화해짐에 따라 보다 좋은 품질의 가상증강현실을 제공하고 있다.

(드론) 기존의 군사 목적으로 개발되었던 무인항공기인 '드론'은 상용화가 진행되면서 개인 소비자에게 유통되고 있다. 배터리 기술의 발전과 영상 기술, 드론에 들어가는 항법 및 고도 측정 등의 고도화된 항공기술의 발전으로 드론의 활용도는 지속적으로 증가하고 있다. 아직은 항공과 관련한 법적 규제 및 상용적 활용이 엄격히 제한되고 있지만 기술 계량화와 보안 등의 기술 확산으로 상용화도 가까운 미래에 가능하다고 예측되고 있다.

(로봇) 제조업에서 주로 자동화 공정으로 많이 활용되고 있는 로봇은 이제 가정의 청소까지 할 수 있을 정도로 기술이 발전하였고 더 많이 고도화 되었다. 인공지능을 활용하여 사람과 대화할 수 있는 로봇 개발이 이뤄지고 있으며, 지속적인 기술의 발전을 보이고 있다.

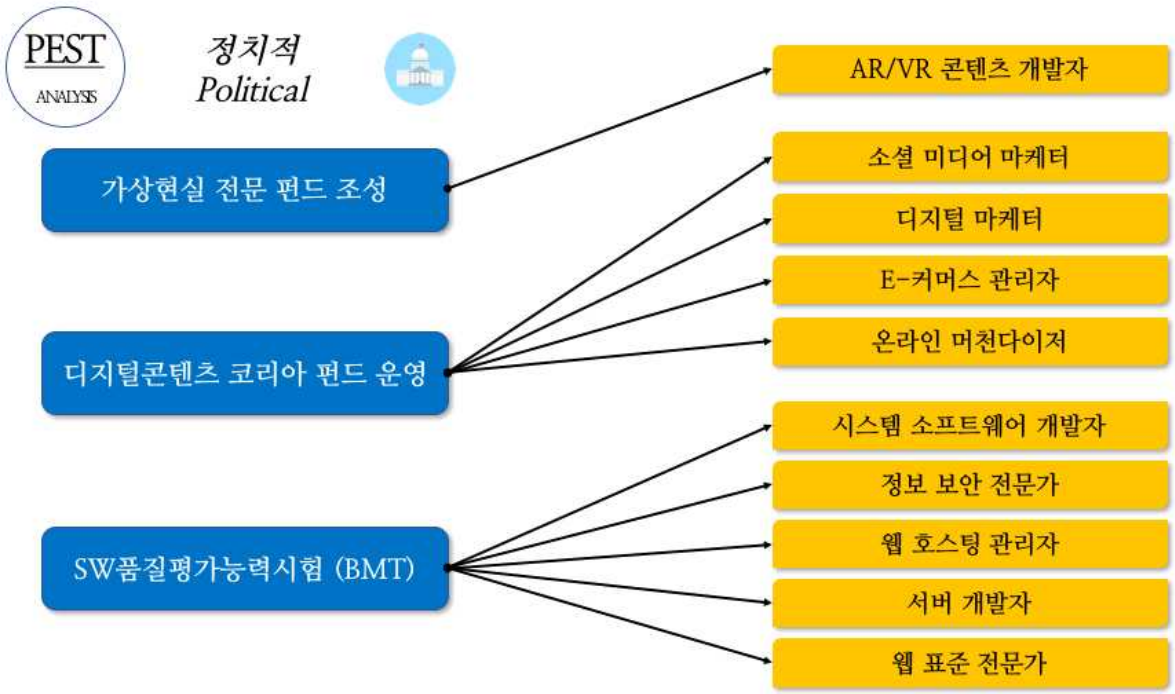
(사물인터넷) 기존의 제품들에 인터넷을 연결하는 개념으로써 사물인터넷은 소비자의 생활면에서 보다 편리한 편익을 제공하는 역할을 하고 있다. 전등을 스마트폰으로 점등하는 등의 기술 등의 발전을 통하여 보다 많은 산업에서 기회를 얻고 있다.

2. 결과 비교

거시 환경과 DTM으로 분석된 12개의 직종 및 24개의 직업과의 일치성 검토 결과는 거시 환경의 4가지 차원에 해당하는 직업을 매칭하여 결과 비교를 진행하였다.

1) 정치적 차원에서의 직업 일치성

[그림 3-40] 정치적 차원의 직업 일치성



[그림 3-40]과 같이 첫 번째로 도출되었던 가상현실 전문 펀드 조성은 지속적으로 가상/증강 현실 시장을 확장시키고 있다. 따라서 직종분류의 'AV/VR 콘텐츠 개발자' 직업과 관련이 있다. 본 연구에서 직업의 정의로 증강현실 및 가상현실 기술을 활용하여 소셜 미디어에 최적화 된 콘텐츠를 개발하는 직종으로 정의되어 있으며, 기존 분류체계에서는 존재하지 않았던 AR/VR 콘텐츠 개발자를 발견함으로써 현재 정치적 차원에서의 환경에

일맥상통하고 있다는 것을 증명하고 있다. 즉, 시간의 반영으로 인하여 보다 정확한 거시 환경을 반영하고 있는 분석 결과라고 할 수 있다.

두 번째로 도출되었던 디지털콘텐츠코리아 펀드 운영은 직종분류에서 총 4가지 직업에 반영되어있다. ‘소셜 미디어 마케터’는 컴퓨터 공학의 지식과 마케팅 지식을 기반으로 소셜 미디어에서의 고객과의 상호 작용을 통한 콘텐츠 홍보 등의 역할을 수행하는 직종으로써 아직 콘텐츠 산업의 비율이 30% 정도 밖에 유지하지 못하는 디지털 콘텐츠 산업에서 관련 역량을 통하여 홍보 등의 역할을 함으로써 디지털콘텐츠 활성화에 이바지 할 수 있는 직업이다.

‘디지털 마케터’는 소비자와의 접촉을 통한 제품을 판매하고 프로젝트 수행 등의 과업을 수행하는 직업으로써 소비자에게 필요한 디지털 콘텐츠를 제공함으로써 과거보다 디지털 콘텐츠를 소비자들이 자신의 니즈에 맞게 잘 활용 할 수 있도록 도움을 주는 직업이다.

‘E-커머스 관리자’는 소셜커머스 등의 전자상거래를 통하여 상품 판매를 지원하고 있으며, 상품 판매 간에 발생하는 디지털 콘텐츠에 대한 간접적인 영향을 미칠 수 있는 직업으로써 디지털콘텐츠의 발전에 도움을 줄 것으로 판단된다.

‘온라인 머천다이저’는 오픈마켓 또는 전자상거래 상품과 관련한 데이터를 최적화하여 제품 경쟁력을 비교분석하며, 상품 콘텐츠 및 보완제품을 기획하고 홍보하는 역량을 갖춘 직업이다. 사용자가 원하는 상품 콘텐츠와 보완제품을 기획하기 때문에 디지털 콘텐츠에 간접적인 영향을 주는 직업이며, 디지털 콘텐츠의 발전으로 다채로운 콘텐츠를 분석하고 운영할 수 있는 위치에 있는 직업이다.

세 번째로 도출되었던 정치적 환경인 SW품질평가능력시험의 의무화는 사용자에게 제공되는 SW의 품질을 진단하고 보다 개선된 품질을 산업 전반에 활용 할 수 있는 여건을 만들어 줄 수 있도록 법제화, 규제화를 통하여 Benchmark Test(BMT)를 의무화 하고 있다.

관련하여 해당되는 직업으로는 시스템 소프트웨어 개발자, 정보보안전문가, 웹호스팅 관리자, 서버 개발자, 웹 표준 전문가가 있다.

‘시스템 소프트웨어 개발자’는 컴퓨터의 시스템의 수행명령체계인 시스템

소프트웨어를 연구 및 개발하고 설계하는 직종으로써 SW품질에 많은 영향을 주는 직종이다. 기존에도 존재하던 직업으로써 컴퓨터 시스템을 유지보수하며, 시스템 최적화를 위하여 시스템 연구를 함으로써 BMT에서 제시하고 있는 기능성, 효율성, 유지보수성, 사용성의 평가를 수행하고 관리 운영 가능하다.

‘정보 보안 전문가’는 해커의 해킹이나 개인정보 유출 및 바이러스 발생에 대비하여 온/오프라인 전산망의 보안을 유지하거나 보안프로그램을 개발하고, 보안 상태를 점검하여 보안 취약점을 위한 다각적인 해결책을 제시하는 직종으로써, 컴퓨터 및 네트워크에서 침해되는 정보에 대한 해결책을 제시함으로써 품질강화를 진행하는 직업이다. 품질 강화를 위하여 기능성, 신뢰성, 유지보수성, 사용성 등의 역량을 평가 받고 운영 관리함으로써 현재 시장에 부합한 직업이다.

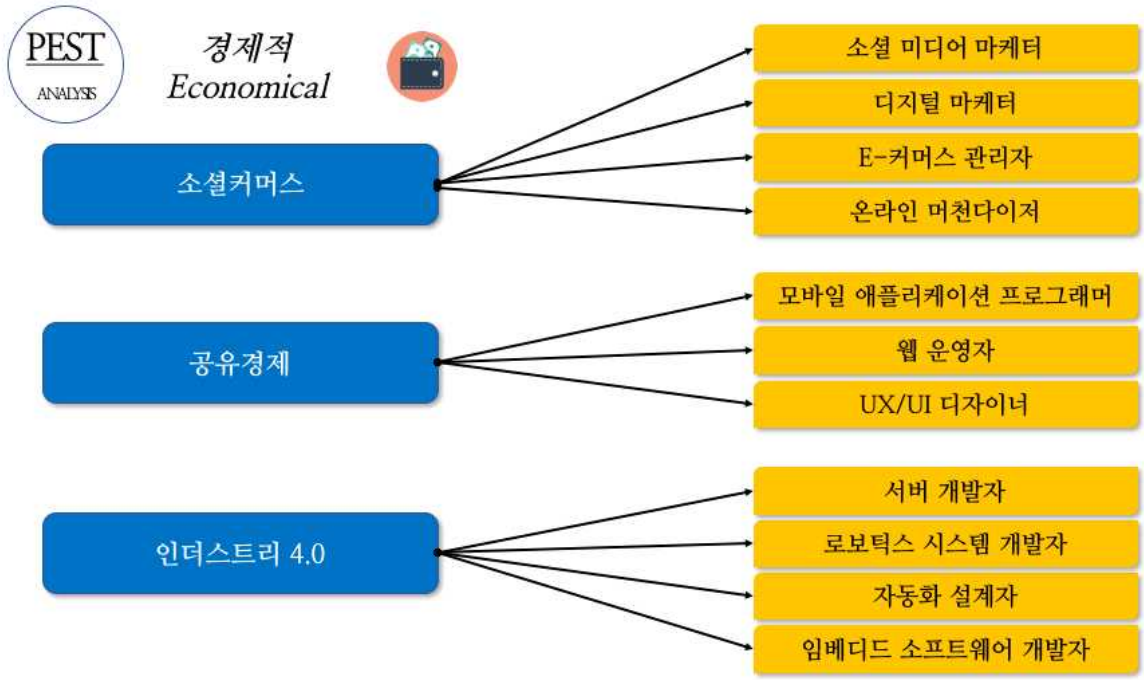
‘웹 호스팅 관리자’는 웹 호스트를 고객에게 제공하여 관리하는 업무를 수행하며, 고객에게 발생하는 기술적인 장애를 해결하고 고객의 지적재산권 및 호스팅 품질 향상을 위하여 시스템적으로 도움을 줄 수 있는 직업으로 사용자의 시스템 및 컴퓨터 사용으로써 발생하는 기술적 장애를 해결하여 품질 평가 요소인 기능성, 효율성, 신뢰성, 유지보수성, 사용성 등의 품질을 향상시키는 직종을 수행해야하기 때문에 현재 시장에 필요한 직업이다.

‘서버 개발자’는 데이터베이스 및 네트워크 서버에 관한 지식을 이용하여, 서버 내 최적화 된 데이터베이스를 설계 및 구축하고 서버 간 네트워크 망을 개발, 기획하는 직업으로써 서버에서 발생 할 수 있는 서버간의 데이터의 신뢰와 서버가 붕괴 되었을 때의 유지보수를 신속하게 진행하고, 서버가 문제가 있을 때 다른 서버 및 다른 체제에 이식 할 수 있어야 하며, 서버와 관련한 다양한 이슈를 해결 할 수 있는 자로써 SW품질향상을 위한 기능성, 효율성, 신뢰성, 유지보수성, 사용성, 이식성을 최적화 및 적절하게 수행해야하기 때문에 현재 시장에 필요한 직업이다.

해당 직업들에게 BMT의 평가 요소인 기능성, 효율성, 신뢰성, 유지보수성, 사용성, 이식성을 적용하여 SW 품질 향상을 시켜 사용자가 보다 쉽고 편리하게 이용할 수 있는 환경을 제도적 차원에서 정부에서 의무 시행하고 있다.

2) 경제적 차원에서의 직업 일치성

[그림 3-41] 경제적 차원의 직업 일치성



[그림 3-41]과 같이 첫 번째로 도출된 소셜 커머스의 발전으로 인하여 소비자의 비용절감과 소비 시간의 감소에 영향 받는 직업으로는 소셜 미디어 마케터, 디지털 마케터, E-커머스 관리자, 온라인 머천다이저 등이 있다.

‘소셜 미디어 마케터’는 상기 설명과 동일하게 소셜 미디어 관련 콘텐츠를 홍보하는 직업으로써 소셜 커머스의 발달로 인하여 새로운 소비 패턴을 보이고 있는 소비자에게 매력적인 콘텐츠를 홍보함으로써 보다 쉽게 수익 창출을 도모할 수 있게 도움을 주는 직종이기 때문에 시장에 필요한 직업이다.

‘디지털 마케터’는 소비자와 접촉을 하는 직업으로써 현재 오프라인에서 온라인으로 변경된 소비 패턴에 있어서 더 중요한 직종이라 할 수 있다. 오프라인과는 다른 특별함을 제공하여 소비자의 이목을 집중시켜야하기 때문에 현재 시장에서 필요한 직업이다.

‘E-커머스 관리자’는 소셜 커머스에 있어서 매우 중요한 직업이다. 공동

구매 등의 플랫폼 상에서 이뤄지는 소비 행동에 대하여 관리하고 상품의 갱신 및 변경을 소비자의 구매 패턴에 맞추어 운영 관리해야하는 직종이다. 따라서 본 직업은 소셜커머스로 변경되고 있는 경제적 환경에 매우 필요한 직업이다.

‘온라인 머천다이저’는 온라인 전자상거래에서 필요한 전반적인 업무를 담당하고 있어 E-커머스 관리자와 동일하게 소셜 커머스에서 매우 중요한 요소이다 . 또한 온라인뿐만 아닌 O2O 시자에서 소비 패턴 파악은 필수 불가결한 부분이기 때문에 본 직업 또한 소셜 커머스 발달에 따른 필요한 직업이다.

두 번째, 공유 경제의 소비 경향이 발달함에 따라 영향을 받는 직업으로는 모바일 애플리케이션 개발자, 웹 운영자, UX/UI 디자이너의 직업이 도출되었다.

‘모바일 애플리케이션 개발자’는 스마트폰이나 태블릿 PC 등 휴대용 단말기를 위해 개발되는 응용 소프트웨어를 설계하고 유지 보수하는 직업을 말하며, 공유경제를 통하여 플랫폼의 중요성은 더욱더 중요해지고 있다. 서로간의 물건을 공유하고 협력적으로 구매를 위하여 활동이 일어날 공간이 필요하며, 모바일의 발전으로 모바일에서 구매하려는 소비경향도 같이 발전하여 플랫폼을 담당하는 어플리케이션의 개발이 필요해졌다.

‘웹 운영자’는 웹 사이트나 온라인 마켓 운영 전반에 걸쳐 실무적인 책임을 지는 자를 말하며, 웹 사이트 구축목적에 맞게 서비스를 리뉴얼하거나 운영 시스템의 리소스를 관리하는 직업이다. 공유 목적으로 형성된 웹사이트에 대한 관리 및 운영 방법은 다른 비즈니스 모델을 가지고 있는 사업과 다르기 때문에 해당 특성에 맞는 운영이 필요하다. 따라서 공유 경제 시스템을 운영할 수 있는 웹 운영자의 역량이 필요하다.

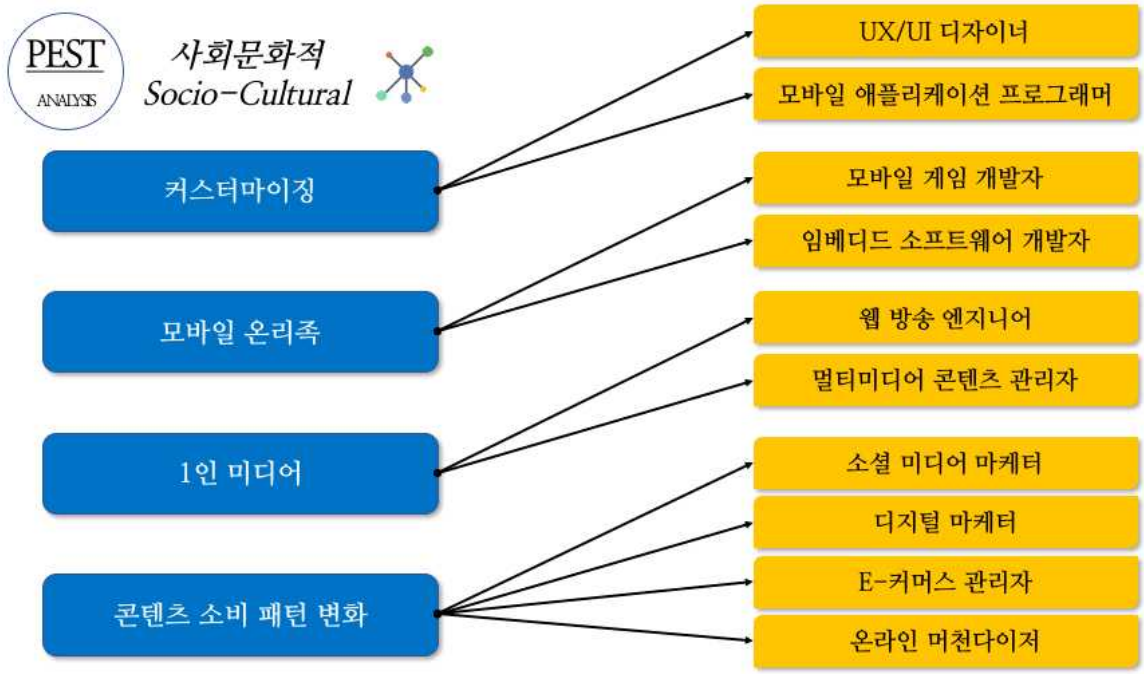
‘UX/UI 디자이너’는 웹페이지, 애플리케이션, 컴퓨터 및 모바일 기기 등의 사용자 환경을 사용자 경험에 입각하여 편의성과 활용성을 극대화 할 수 있도록 연구, 개발 하고 제품이나 서비스에 그 결과를 반영하여 디자인을 하는 직업으로 공유경제의 특성에 맞는 사용자 경험과 그에 따른 인터페이스를 구축함으로써 아직은 접근하기 어려운 공유경제 시스템을 보다 쉽게 접촉할 수 있도록 도움을 줄 수 있는 직업이다.

공정 기술의 발전으로 공장은 점차 자동화가 진행되고 있으며, 4차 산업의

발전으로 인더스트리 4.0을 통하여 비용절감 등의 경제적 발전을 이루고 있다. 영향을 받는 직업으로는 '자동화 설계자', '로보틱스 시스템 개발자', '임베디드 소프트웨어 개발자', '서버 개발자' 등이 있다.

3) 사회문화적 차원에서의 직업 일치성

[그림 3-42] 사회문화적 차원의 직업 일치성



[그림 3-42]와 같이 첫 번째로 도출된 커스터마이징은 사용자의 니즈에 맞춤형으로 제공하는 것을 선호하는 방식을 말하며, 커스터마이징에 영향을 받는 직업으로는 UX/UI 디자이너, 모바일 애플리케이션 프로그래머가 도출되었다.

'UX/UI 디자이너'는 사용자의 경험과 관련된 업무를 진행하며, 사용자가 사용할 때 편리한 구조, 디자인 그리고 사용자의 움직임을 분석하고, 커서의 이동, 클릭 및 사용자의 시선 등을 파악하여 보다 사용자에게 편의를 줄 수 있는 환경을 만드는 역할을 하는 직업이다. 커스터마이징은 공공재 및 누구나

똑같은 환경의 똑같은 제품을 제공 받는 것이 아니라 소비자 각자의 개성과 소비 패턴, 성격과 성향 등을 고려하고 주요 고객들의 클릭 위치, 클릭 수, 접속 패턴 등을 파악하여 보다 편리한 서비스를 제공 할 수 있는 기회를 제공하려고 하고 있다. 커스터마이징이 되어있는 사이트 및 구매에서는 소비자는 자신이 원하는 데로 구성되어 있기 때문에 사이트 및 제품에 대한 만족도 높아지며, 기업의 브랜드 태도 또한 좋아지기 때문에 많은 기업에서 찾고 현재 찾고 있는 직업이다.

‘모바일 어플리케이션 개발자’는 모바일 상에서 사용할 수 있는 어플리케이션을 개발하는 업무를 하는 직업으로써 개인의 취향 및 패턴에 맞는 어플리케이션을 개발함으로써 소비자는 어플리케이션을 더 많이 사용할 것이다. 또한 소비자가 제공하는 댓글 등의 정보를 기반으로 소비자들이 원하는 어플리케이션에 요소를 개발함으로써 만족도를 높일 수 있으며, 많은 소비자가 어플리케이션을 사용할 것이다.

두 번째로 사회문화적 환경의 요소는 모바일 온리족이 도출되었다. 모바일을 기반으로 모든 행동을 하려는 집단을 말하며, 영향을 받는 직업으로는 모바일 게임 개발자, 임베디드 소프트웨어 개발자 등이 있다.

‘모바일 게임 개발자’는 모바일 기반에서 시나리오의 순서에 따라 배경과 상황 등을 각종 언어로 프로그래밍하여 컴퓨터 게임의 운용을 위한 프로그램을 제작하는 직업이며, 모바일로 모든 것을 하는 모바일 온리족은 이제 컴퓨터 게임에서 벗어나 모바일 게임을 실시간으로 어떤 장소에 구애 받지 않고 실시한다. 따라서 모바일 게임의 시장은 일반 컴퓨터 게임의 시장보다 크게 성장하고 있으며, 고용 시장에서도 많은 개발자를 요구하고 있다. 한국콘텐츠진흥원에서 발표된 ‘2016년 게임이용자 실태조사’ 보고서에 따르면, 게임 시장에서 가장 큰 분포를 차지하고 있는 종류는 2016년 기준 ‘모바일 게임’으로써 전체 시장의 60.2%로 국민의 절반 이상이 모바일 게임을 즐기고 있다는 것을 확인하였다. 모바일 온리족으로 가장 많은 시간을 보내는 콘텐츠 또한 게임으로 모바일 게임에 대한 개발은 현재 시장에서 가장 큰 시장 성장률을 가지고 있으면서, 가장 경쟁력 있는 산업으로 부상하고 있다. 따라서 사회문화적인 요소로 모바일 온리족에게 있어 모바일 게임을 개발한다는 것은 필수 불가결적인 요소임에 틀림없다.

‘임베디드 소프트웨어 개발자’는 반도체 등과 관련된 하드웨어와 소프트웨어 문제를 해결하여 장비를 설치하고 엔지니어들이 장비를 사용할 수 있도록 환경을 조성하는 직업으로 모바일을 구성하는 하드웨어 요소인 반도에 대한 이슈와 모바일 기술의 발전 등에 대하여 지속적으로 관심을 갖는다. 2007년부터 시작된 스마트폰의 발전은 현재의 모습을 갖추기 위해서 지속적으로 발전하고 있다. 카메라 기능, 전화 품질 향상, 배터리 증가, 무선 충전기 등의 다양한 모바일 하드웨어는 지속적으로 개발되고 있으며, 모바일 온리족에게 모바일을 사용할 수 있도록 가장 기본적인 환경인 하드웨어와 하드웨어를 움직일 수 있는 소프트웨어의 개발 및 관리는 매우 중요한 요소 중 하나이다.

세 번째로 사회문화적 시장 환경의 요소는 1인 미디어이다. 과거와는 다르게 온라인 플랫폼의 발전으로 소비자가 정보 제공자가 되는 1인 미디어의 발달로 영향을 받는 직업으로는 웹 방송 엔지니어, 멀티미디어 콘텐츠 관리자 등이 있다.

‘웹 방송 엔지니어’는 디지털 장비를 이용해 촬영 혹은 녹음 된 파일을 방소에 적합한 형태로 편집하고, 효과음과 배경음을 컴퓨터로 믹싱(Mixing)하여, 편집된 파일을 배포 가능한 형태로 인코딩을 하면서, 신규 방송용 장비를 도입하고 관리하는 직업으로 온라인 플랫폼의 발달로 인하여 개인 방송의 송출 능력과 영상 편집 능력, 컴퓨터 운용 능력 및 오디오, 비디오 등의 하드웨어 제품의 관리 역량 등이 필요해지고 있다. 1인 미디어를 위해서는 각종 플랫폼에 업로드 할 수 있는 확장자의 이해뿐만 아니라 콘텐츠에 대한 이해와 함께 웹 문서의 이해까지 다양한 지식을 필요로 한다. 따라서 웹 방송 엔지니어는 1인 미디어를 진행하는 콘텐츠 리더(Content reader)가 컴퓨터 공학적으로 이해하지 못하거나 접근하지 못하는 부분에 대한 지원과 동영상 및 사진 편집 능력 등을 지원함으로써 보다 원활하게 소비자들에게 1인 미디어 콘텐츠를 제공할 수 있는 환경을 구축한다.

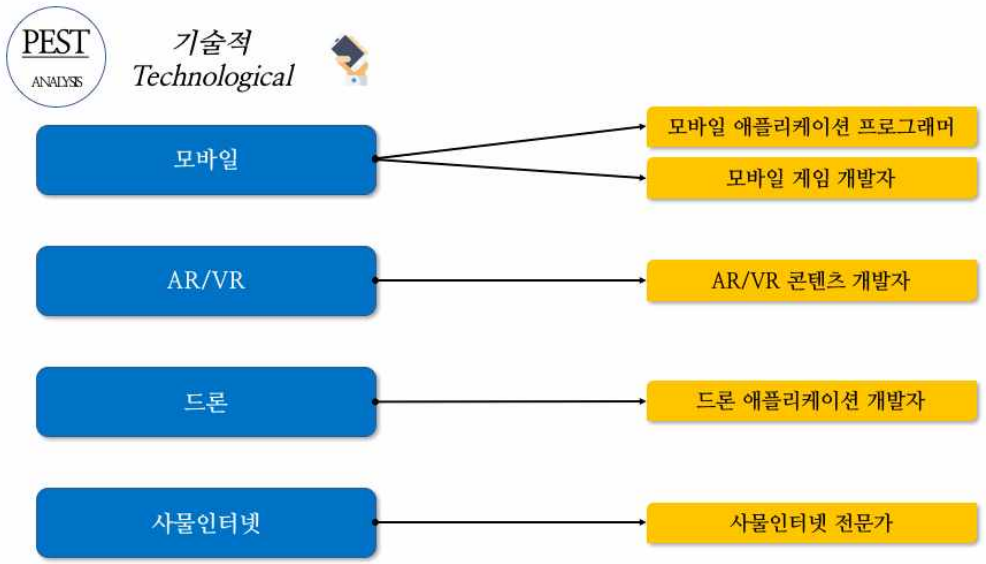
‘멀티미디어 콘텐츠 관리자’는 모바일 및 웹 콘텐츠를 제공하기 위한 전략을 수립하고 실행하며, 시각디자인, 모션 그래픽 등의 멀티미디어 편집 기술을 활용하여 요구사항을 충족하는 콘텐츠 솔루션을 설계하고 관리하는 직업으로 1인 미디어의 콘텐츠를 검토하고 관련 디자인 기획을 협업하여 보다 원활하게 방송될 수 있도록 도와줌으로써 현재 사회문화적인 1인 미디어의

형태를 보완한다.

네 번째는 콘텐츠 소비 패턴의 변화로 기존의 공공재를 소비했던 기성세대와는 다르게 자신이 원하는 니즈를 개성 있게 소비하는 현재는 다양한 콘텐츠의 개인화가 이루어지고 있다. 따라서 온라인상에서 제품을 팔거나 온라인에서 제공되는 영상 및 사진, 글 등의 콘텐츠에 있어서도 기존과는 다른 패턴의 소비가 발생하고 따라서 디지털 마케팅을 전문적으로 하는 전문가 직종의 4가지 직업인 소셜 미디어 마케터, 디지털 마케터, E-커머스 관리자, 온라인 머천다이저 등의 직업이 중요하게 작용된다.

4) 기술적 차원에서의 직업 일치성

[그림 3-43] 기술적 차원의 직업 일치성



[그림 3-43]과 같이 마지막 거시 환경인 기술적 요소에는 모바일, AR/VR과 드론 및 사물인터넷의 기술적 발전이 분류체계와 어떤 일치성이 있는지 비교 분석하였다.

첫 번째 모바일의 발달로 인하여 소비자는 언제 어디서든 인터넷을 검색하고

게임을 하며, 노래를 듣고, 영화를 볼 수 있게 되었다. 해당 키워드와 관련하여 영향을 받는 직종은 모바일 어플리케이션 개발자와 모바일 게임 개발자가 도출되었다.

‘모바일 어플리케이션 개발자’는 스마트폰의 발달로 인하여 어플리케이션이 하나의 시장으로 성장하고 있다. 현재의 어플리케이션의 개수는 셀 수 없을 만큼 많이 존재하고 있으며, 많은 어플리케이션에서 살아남기는 매우 어려운 일이라 할 수 있다. 따라서 어플리케이션 개발자는 넘쳐나는 어플리케이션에서 보다 소비자에게 사용하기 쉽고 접근하기 쉬운 방식으로 어플리케이션을 개발하여 다른 많은 어플리케이션과의 차별성을 가져야 한다. 따라서 모바일 기술의 발전이 점차 고도화 될 수 있는 이유도 SW 직군에 모바일 어플리케이션 개발자가 지속적으로 개발을 하고 있기 때문에 가능한 일이다.

‘모바일 게임 개발자’는 상기 언급한 모바일 어플리케이션 개발자에 포함되는 개념으로 기존 분류체계에서는 분류하고 있으나, 상기 언급과 같이 현재 모든 게임 산업에서 가장 큰 시장을 차지하고 있는 분야가 바로 모바일이기 때문에 모바일 게임은 다른 성격의 어플리케이션 보다는 더욱더 중요하다. 또한 쾌락재인 ‘게임’의 특성상 빠른 피드백과 보상 등의 다른 어플리케이션과의 차별점이 존재하기 때문에 다른 차원의 직종이라 보는 것이 옳다. 따라서 분류 체계에서도 모바일 어플리케이션 개발자와는 독립된 직업으로 분류하였으며, 가장 많은 애플리케이션 이용률을 보이는 범주가 게임이기 때문에 지속적인 모바일 게임 개발은 중요하다.

두 번째 AR/VR은 현재 정부에서도 적극 지원하고 있는 산업군 중 하나로 새로운 신 성장 동력 사업이다. 따라서 AR/VR 콘텐츠 개발자는 중요하다.

증강현실은 현실에 적용하여 가상의 콘텐츠를 형상화 하는 기술이며, 가상현실은 관련된 디바이스를 착용하여 전체적인 시각을 지배하여 현실과는 다른 가상의 세상을 투영하는 기술을 말한다. AR/VR을 활용한 콘텐츠 개발은 지속적으로 이뤄지고 있으며, 교육, 건설, 게임 등 다양한 분야에 적용되고 있다. 따라서 계속적인 기술발전은 필수 불가결적인 요소이며, 지속적인 지원이 필요하다.

세 번째 드론은 현재 가장 뜨거운 감자 중 하나에 속하는 콘텐츠이며, 군용 항법 개발에서 상용화 시도를 통하여 현재는 많은 소비자들이 취미 및 다양한

분야에서 활용을 시도하고 있다. ‘드론 애플리케이션 개발자’는 드론의 이동, 통신 및 분석 방법을 개선하기 위한 새로운 기술을 연구하며, 무인 항공 시스템에 적용할 드론 모형을 제작하고 이를 위한 어플리케이션을 제작하는 직업이다. 기존 분류체계에는 없었던 드론은 현재 상용화 기술을 거쳐 산업 전반에 활용까지 시도되고 있다.

네 번째 사물 인터넷(IoT)는 사물에 인터넷이 되는 센서를 부착하여 네트워크로 연결하는 기술로 ‘사물인터넷 전문가’는 펌웨어와 임베디드 소프트웨어 개발 능력을 기반으로 사용자의 목적에 맞는 사물인터넷을 개발하는 직업으로써 지속적인 사용자의 욕구에 맞춰 사물인터넷의 개발이 시도되고 있다.

3. 거시 환경 분석과 재배치 검토의 시사점

본 연구의 2장 2절에서 언급한 거시 환경 분석과 3장에서 분석한 12개의 직종과 24개의 직업과의 일치성을 비교 분석하였다. 분석결과를 요약하면 기존에 분류체계에는 없었지만 거시 환경 분석을 한 결과 가상현실 콘텐츠 강화를 위한 펀드 운영과 가상/증강현실의 기술적 진보로 도출된 AR/VR과 관련된 직업이 두드러지게 나타났다. AR/VR은 신성장 동력으로써 많은 국가에서 투자하고 있는 산업이다. 현재 콘텐츠의 부족으로 많은 사용자가 두드러지게 이용하고 있지는 않지만 콘텐츠의 강화와 기술의 성장으로 VR방과 AR 모델하우스 등의 게임뿐만 아니라 건설업에서도 제조업의 시뮬레이션과 가상 교육 등 다채롭게 활용되고 있다.

디지털 콘텐츠의 강화는 실제 SW인력의 소구에도 부합하였다. 디지털 콘텐츠의 비중이 계속적으로 상승하고 있으며, 이와 관련하여 콘텐츠를 운영 및 관리, 유지 보수하는 직업의 필요성은 점진적으로 증가하고 있으며, 전자상거래 또한 활발히 이용되고 있다. 따라서 디지털 마케팅 전문가의 4가지 직종인 소셜 미디어 마케터, 디지털 마케터, E-커머스 관리자, 온라인 머천다이저 등은 유망 직종으로 지속적으로 발전할 것이라 추측된다.

4차 산업 혁명은 정보기술 산업인 SW인력에 많은 영향을 주었다. 비단 컴퓨터 산업뿐만이 아닌 제조업 등의 컴퓨터를 활용 및 응용하는 분야 또한 기존의 산업 동향과는 다른 양상을 보이고 있다. 컴퓨터 하드웨어에 꼭 필요한 반도체 산업의 경우 SW인력에서 중요한 측면 중 하나로 부각되고 있었다. 뿐만 아니라 공정에 대한 자동화 설계, 로봇을 이용한 공정 등의 인더스트리 4.0을 대변하는 직업들이 현재 SW인력 시장에서 중요한 비중으로 요구되고 있다.

1인 미디어의 발전은 기존의 콘텐츠와는 다른 방식의 콘텐츠를 구현하고 환경을 조성하였다. 기존에는 연예인이 위주인 공영 방송을 즐겨보는 텔레비전의 무조건적인 수용으로만 전달되었던 미디어가 최근 온라인 기술과 영상 기술의 발전으로 쌍방 소통이 가능한 형태로 변화하였다. 또한 비용의 절감과 함께 남녀노소 원하는 누구나 영상을 제작하고 미디어를 배포 할 수 있는 환경을 조성함으로써 방송과 관련된 엔지니어의 필요성은 점차 증가하고

있다.

이동전화의 기능이 기존과는 비교도 안될 만큼 변화하면서 컴퓨터를 활용하지 않고 모바일만 사용하는 '모바일 온리족' 이라는 신조어까지 등장하였다. 시장의 환경은 모바일에 특화된 기술력이 있는 인력을 요구하고 있다. 특히 모바일 시장에서 가장 큰 비중을 갖고 있는 모바일 게임 시장의 경우 전체 게임 시장의 60.2%를 보유 할 만큼 중요한 시장이다. 모바일 게임과 관련된 개발자는 그만큼 더 많은 수요를 갖고 있으며, 게임 이외의 어플리케이션 시장의 전망도 지속적으로 좋아 관련 직업의 수요는 지속적으로 증가할 것이 예측된다.

군사용도로 사용되었던 '드론' 은 이제 하나의 취미생활로 변화하였다. 따라서 드론과 관련된 소프트웨어 기술 및 반도체, 하드웨어 공법 등의 다양한 기술의 활용도가 증가하고 있으며, 드론 전문가 등의 자격증도 탄생하게 되었다. 법적 규제에만 한정되어 있던 드론의 활용이 점차 산업의 공정에 활용이 되어가고 있으며, 인간의 공정 능력을 한층 업그레이드 시켜줄 수 있는 하나의 기술로써 도약하고 있다. 기존 분류 체계에는 존재하지 않았던 직업을 발견함으로써 본 연구의 목적성을 확인할 수 있었으며, 현재 어떤 주제가 SW 인력에 중요한 요소인지 반증하고 있다.

사물인터넷은 어떤 사물이든 인터넷 네트워크로 연결이 된다는 개념을 통해 다양한 분야에서의 확장성을 가지고 있다. 즉, 모든 제품에 인터넷이 연결되어 관리, 운영을 할 수 있으며, 원격조종을 하여 사용자가 집에 없어도 보일러 온도를 조정하거나 혹은 불을 끄거나 모든 것을 관리하는 디바이스에서 활용 할 수 있다. 사물인터넷은 최근의 정보 기술의 발전과 모바일의 1인 보급, 그리고 모바일에서의 기능강화 등을 통하여 관련 어플리케이션과 사물을 연동할 수 있는 칩 등의 반도체 연구의 산실이며, 보다 사용자에게 편리함을 주고 있다. 기존 제품에도 붙일 수 있으며, 다양한 원격을 조종할 수 있기 때문에 악용 될 수 있다. 그렇기 때문에 전문적으로 사물인터넷 기술을 활용을 할 수 있도록 하는 전문 인력이 현재 시장에서 필요하다.

정부 또한 다양한 노력을 기울이고 있다. 콘텐츠를 형성하기 위하여 펀드를 조성하는 사업을 지원하고 있으며, SW품질평가능력시험을 의무화하여 기존보다 좀 더 경쟁적인 시장을 구축하기 위한 품질 강화와 기회 창출에 많은

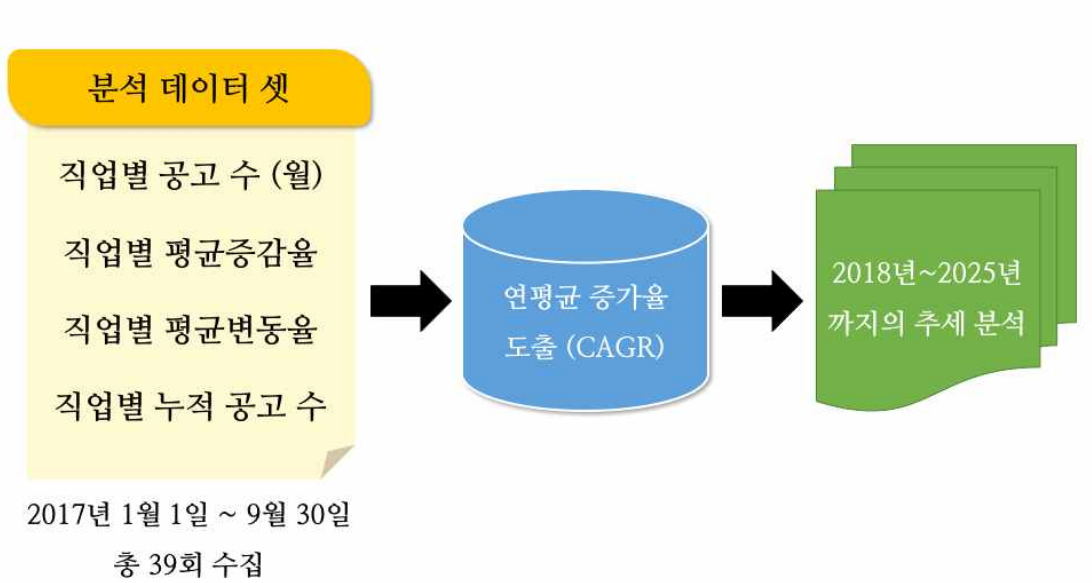
노력을 기울이고 있다. 기존의 분류체계에서는 존재하지 않던 몇 가지의 새로운 직종의 발견은 현재 거시 환경을 적용하는데 있어 매우 중요하게 재배치의 시사점을 가져다준다. 상기 언급된 분석 결과 등을 종합하여 3절에서는 분석결과를 바탕으로 각 직종별로 연평균복합성장률(CAGR)을 분석하였다. 그에 따라 각 직종별 향후 전망을 예측하였다.

제3절 SW직종 수요 예측을 위한 연평균복합성장률(CAGR) 분석

1. 연평균종합성장률 (CAGR) 분석 프로세스

본 절에서는 연구를 통하여 발견한 새로운 분류 체계 및 직업에 대한 정의를 실시하고 난 후에 2017년 1월 1일부터 9월 30일까지의 직업별 공고 수 및 평균 증감률, 평균 변동률, 누적 공고 수를 확인하였다.

[그림 3-44] 연평균 종합성장률 분석 프로세스



4가지 분석 데이터 셋을 확보하고 연평균 수요 증가율을 분석하기 위하여 연평균종합성장률(Compound Annual Growth Rate: CAGR)을 도출하여 지정된 기간 동안의 데이터의 연평균 성장률을 측정하고자 한다. CAGR은 매년 일정한 성장률이 존재한다고 가정하고 이에 따른 가중 평균의 성장률을 환산한 값이다. 본 연구과제는 2017년 데이터를 기준으로 2025년까지 향후 8개년의 추세를 도출하여 각 12개의 직종 및 24개의 직업의 전망을 확인함으로써 사양되는 직업과 유망 직업을 도출 가능하다(이동소, 2014).

평균종합성장률은 다양한 분야의 선행연구에서 도출되었다. 한국고용정보원에서

발표된 ‘일반기계산업 인력수요전망’ 보고서에 따르면 무역 실적을 2001~2012년까지 수출에 대한 CAGR은 한국, 중국, 미국, 독일, 일본 5개국 총 9.6% 평균 성장률을 보였으며, 나라별 평균 성장률을 확인한 결과 중국이 25.6%로 가장 높은 성장률을 보였으며, 한국이 16.9%로 그 뒤를 이었다(권우현 외, 2014). SW 인력 예측을 위한 분석에서도 산업의 연평균 성장률에 대한 연구도 시행되었다. 창조경제 시대의 일자리 창출을 위한 인터넷 경제의 효과 분석과 관련된 보고서에서 인터넷 경제에 대한 일자리 추계에 대한 결과를 CAGR을 통하여 도출하였다(오동환 외, 2014). 마찬가지로 미래 유망 신산업에 대한 시장 연구에서 인력수요에 따른 각 산업별 CAGR을 파악하여 산업별 동향을 파악하였다. 전기자동차와 관련된 산업의 인력 수요의 경우 전기차가 84.4%의 평균 성장률을 보이고 있으며, 항공의 경우 4.0%의 성장률을 보일 것으로 전망하고 있다(한국산업기술진흥원, 2016).

상기 선행연구들을 통하여 확인 한 결과 산업 동향 및 미래의 시장과 일자리 등의 수요에 대한 예측하기 위하여 연평균성장률 CAGR을 다양한 분야에서 사용하고 있다. 기존의 데이터를 통하여 년간 성장률의 지표에 기하평균을 적용하여 얻어낸 시간에 따른 평균 성장률을 도출하여, 평균적으로 해당 데이터인 산업이나 일자리 등의 분야의 미래 수요를 예측하는데 기초자료로 활용되고 있다.

본 연구에서도 상기 도출된 12개의 직종과 24개의 직업에 대한 분석 결과에 따라 2017년에 수집한 공고 수 및 증감률 등의 자료를 토대로 상기 연구들과 같이 2025년까지의 수요를 예측하고자 한다.

수요예측을 위한 CAGR의 공식은 다음과 같다.

$$CAGR = Rate(\text{기간}, 0, -\text{첫기간의 측정값}, \text{마지막 기간의 측정값})$$

2. 재배치된 SW 직종별 CAGR 결과

CAGR 분석 프로세스에 의거하여 각 직업별 CAGR을 분석한 결과는 다음과 같다. 각 4가지의 요소의 특성을 요약하면 평균증감률은 2017년 1월 1일에 수집된 데이터를 기준으로 100%로 기준 비율로 설정하고 기준 비율에 따라 증가 했는지 감소했는지에 대한 비율을 %(퍼센트)로 표기한 것을 말하고, 평균변동률도 증감률과 동일하게 2017년 1월 1일 기준을 100%로 설정하고 비율에 따라 변동이 얼마나 많이 되었는지를 비율 치로 표기한 것을 말한다. 월별 공고 수는 수집된 39번의 측정 일자에 따른 공고 수를 일주일 단위로 수집한 데이터 값을 말하며, 누적 월별 공고 수는 월별 공고 수를 누적하여 합산한 값을 말한다.

해당 값들의 특성을 기반으로 하여, 39회의 평균값을 구하여 2017년의 실측치로 구성하고 2018년부터 2025년까지의 예상된 평균 성장률을 상기 언급한 CAGR의 계산식에 맞추어 계산하였다. 2018년부터의 예상 평균값은 전년도에서 CARG의 비율을 계상하여 비율 값으로 증감하는 것으로 계산한다.

$$E = \text{전년도 도출 값} * (1 + CAGR\text{값})$$

E = CAGR을 반영한 평균 연평균 성장 예측 값

1) 소셜 미디어 마케터 (▲)

<표 3-7> 소셜 미디어 마케터 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	176	179	181	184	187	190	193	195	198	1.49%
평균 변동률	195	197	200	202	204	207	209	211	214	1.14%
월별 공고수	939	899	862	826	791	758	726	696	667	-4.00 %
누적월별 공고수	22929	24823	26873	29092	31494	34095	36911	39959	43258	8.00%

소셜 미디어 마케터의 CAGR 분석결과 <표 3-7>과 같이 평균 증감률을 1.49%로 증가하고 있으며, 평균 변동률은 1.14%로 증가하고 있다. 월별 공고 수는 -4.00%로 감소하고 있으며, 누적 월별 공고 수는 8.00%로 증가하고 있다. 소셜 미디어 마케터는 최근 소셜 미디어의 수요 증가로 인하여 유일하게 평균 증감률이 증가하고 있는 직업이다. 2025년까지 꾸준한 증가 추세를 보이며, 지속적으로 수요가 증가할 직업으로 판단할 수 있다.

2) 모바일 어플리케이션 개발자 (▲)

<표 3-8> 모바일 어플리케이션 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	98	98	98	98	98	97	97	97	97	-0.17 %
평균 변동률	122	123	124	124	125	125	126	127	127	0.50%
월별 공고수	212	203	194	186	179	171	164	157	151	-4.15 %
누적월별 공고수	5,211	5,636	6,096	6,593	7,131	7,713	8,342	9,022	9,758	8.16%

모바일 어플리케이션 개발자는 <표 3-8>과 같이 평균 증감률이 -0.17%, 평균 변동률이 +0.50%, 월별 공고 수가 -4.15%, 누적 월별 공고 수가 8.16%로 도출되었다. 평균 변동률을 확인한 결과 +0.50%의 상승이 있었으며, 직업에 대한 월별 공고 수는 적은 수준의 성장률에 머물지만, 모바일의 어플리케이션의 활용 범위의 증가와 '모바일 온리족' 등의 등장으로 어플리케이션과 관련된 직업은 지속적으로 증가추세로 판단된다.

3) 멀티미디어 디자이너 (▼)

<표 3-9> 멀티미디어 디자이너 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	68	67	66	65	63	62	61	60	59	-1.78 %
평균 변동률	91	90	90	89	88	88	87	86	86	-0.79 %
월별 공고수	110	106	103	100	96	93	90	87	84	-3.24 %
누적월별 공고수	2,657	2,887	3,136	3,407	3,702	4,022	4,370	4,747	5,158	8.64%

멀티미디어 디자이너는 <표 3-9>와 같이 평균증감률의 CAGR이 -1.78%, 평균변동률 CAGR이 -0.79%, 월별 공고 수의 CAGR이 -3.24%, 누적된 월별 공고 수의 CAGR이 8.64%로 도출되었다. 미세하게 증감률에서 감소 추세를 보이고 있으며, 멀티미디어와 관련된 직종이 전체적으로 감소하는 양상을 보이고 있다.

4) 정보 보안 전문가 (▲)

<표 3-10> 정보 보안 전문가 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	55	54	53	52	51	50	49	48	47	-1.80%
평균 변동률	78	77	77	76	75	75	74	74	73	-0.75%
월별 공고수	63	61	59	57	55	53	51	49	48	-3.37%
누적월별 공고수	1678	1825	1985	2159	2348	2553	2777	3020	3284	8.76%

정보 보안 전문가는 <표 3-10>과 같이 평균증감률의 CAGR이 -1.80%, 평균변동률 CAGR이 -0.75%, 월별 공고 수의 CAGR이 -3.37%, 누적된 월별 공고 수의 CAGR이 8.76%로 도출되었다. 정보 보안 전문가의 경우도 상대적으로 감소 추세를 보이고 있으나, 누적 월별 공고 수는 8.76%로 높은 성장률을 보이고 있어, 향후 프라이버시와 관련된 이슈를 해결하거나 방화벽 등의 보안과 관련된 업무가 필요하다는 것을 확인할 수 있었다.

5) 멀티미디어 콘텐츠 관리자 (▼)

<표 3-11> 멀티미디어 콘텐츠 관리자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	50	49	48	47	46	45	44	43	42	-2.33%
평균 변동률	72	71	70	70	69	68	67	67	66	-1.11%
월별 공고수	143	139	134	129	125	121	117	113	109	-3.37%
누적월별 공고수	3,318	3,601	3,907	4,240	4,601	4,993	5,418	5,879	6,379	8.51%

멀티미디어 콘텐츠 관리자는 <표 3-11>과 같이 평균증감률의 CAGR이 -2.33%, 평균변동률 CAGR이 -1.11%, 월별 공고 수의 CAGR이 -3.37%, 누적된 월별 공고 수의 CAGR이 8.51%로 도출되었다. 상기 언급된 멀티미디어 관련된 직종은 상대적으로 감소 추세를 보이고 있다. 평균증감률의 경우 -2.33%로 상대적으로 높은 감소율을 보이고 있어 멀티미디어의 콘텐츠 시장이 어느 정도 성숙기에 접어 들었다는 것을 증명하고 있다.

6) 정보 시스템 운영자 (▲)

<표 3-12> 정보 시스템 운영자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	91	90	89	89	88	88	87	86	86	-0.71%
평균 변동률	111	111	111	111	111	111	111	111	111	0.01%
월별 공고수	137	132	127	122	118	114	109	105	102	-3.68%
누적월별 공고수	3108	3365	3642	3943	4269	4621	5003	5416	5863	8.26%

정보 시스템 운영자는 <표 3-12>와 같이 평균증감률의 CAGR이 -0.71%, 평균변동률 CAGR이 0.01%, 월별 공고 수의 CAGR이 -3.68%, 누적된 월별 공고

수의 CAGR이 8.26%로 도출되었다. 정보 시스템 운영자는 시스템 전반에 대한 설계 구축 및 운영을 하는 직업으로써 평균변동률도 대부분의 직업이 감소율을 보이고 있을 때 증가율을 보이고 있다. 향후에도 지속적으로 유지를 하는 직업으로 판단된다. 정보 시스템에 대한 전반적인 지식을 갖은 전문가는 정보 고도화가 되어가고 있는 사회에서 중요한 역할하기 때문이다.

7) 웹 운영자 (▲)

<표 3-13> 웹 운영자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	59	59	58	57	56	56	55	54	53	-1.33 %
평균 변동률	78	77	77	76	76	76	75	75	75	-0.50 %
월별 공고수	24	23	22	22	21	20	20	19	18	-3.00 %
누적월별 공고수	584	634	688	747	811	880	955	1037	1125	9.00%

웹 운영자는 <표 3-13>과 같이 평균증감률의 CAGR이 -1.33%, 평균변동률 CAGR이 -0.50%, 월별 공고 수의 CAGR이 -3.00%, 누적된 월별 공고 수의 CAGR이 9.00%로 도출되었다. 웹 2.0의 출현으로 인하여 정보 기술의 대부분이 인터넷과 관련된 기술들이 주로 개발되고 운영되어 지고 있다. 웹 기반으로 할 수 있는 많은 콘텐츠와 웹을 운영하는 직업은 주요한 직업이다. 누적별 월별 공고 수의 CAGR이 9.00%나 나와 상대적으로 높은 성장률을 보이고 있어, 미래에도 많은 영향을 미칠 것으로 예측된다.

8) 사물인터넷 소프트웨어 개발자 (-)

<표 3-14> 사물인터넷 소프트웨어 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	62	61	59	58	57	56	55	54	53	-1.79%
평균 변동률	85	84	83	83	82	81	81	80	80	-0.75%
월별 공고수	367	354	341	329	317	306	295	285	275	-3.55%
누적월별 공고수	8390	9086	9840	10656	11541	12498	13535	14659	15875	8.30%

사물인터넷 소프트웨어 개발자는 <표 3-14>와 같이 평균증감률의 CAGR이 -1.79%, 평균변동률 CAGR이 -0.75%, 월별 공고 수의 CAGR이 -3.55%, 누적된 월별 공고 수의 CAGR이 8.30%로 도출되었다. 사물인터넷은 최근의 반도체 기술의 발전과 스마트폰의 보급으로 인한 이동성(Mobility)이 증가하면서 각광받는 분야 중 하나이다. 그러나 SW인력 수요로 파악한 결과는 평이한 수준의 증감률을 보이고 있으며, 기반이 되는 웹과 임베디드 등과 관련된 직업이 상대적으로 높은 성장률을 보이고 있는 것을 도출하였다. 따라서 사물인터넷 소프트웨어 개발자는 평이하게 유지 및 감소하는 형태의 성장률을 보이고 있다.

9) E-커머스 관리자 (-)

<표 3-15> E-커머스 관리자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	69	68	67	66	65	64	63	62	61	-1.46%
평균 변동률	92	91	91	90	90	89	89	88	87	-0.59%
월별 공고수	108	104	100	96	93	90	86	83	80	-3.65%
누적월별 공고수	2662	2890	3137	3406	3698	4014	4358	4731	5136	8.56%

E-커머스 관리자는 <표 3-15>와 같이 평균증감률의 CAGR이 -1.46%, 평균변동률 CAGR이 -0.59%, 월별 공고 수의 CAGR이 -3.65%, 누적된 월별 공고 수의 CAGR이 8.56%로 도출되었다. 온라인상에서 사용자간에 일어나는 거래를 관리하는 E-커머스 관리자는 상대적으로 감소의 폭이 좁아 평이한 수준의 직업이라 할 수 있다. SW 관련 지식과 마케팅적인 지식을 겸비하는 E-커머스 관리자 및 디지털 마케터의 경우도 상대적으로 평이한 수준의 증감형태를 보이고 있다.

10) 서버 개발자 (▲)

<표 3-16> 서버 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	64	63	62	61	60	59	58	57	57	-1.44%
평균 변동률	85	84	84	83	83	82	81	81	80	-0.64%
월별 공고수	50	48	46	44	42	41	39	37	36	-4.05%
누적월별 공고수	1506	1640	1787	1946	2119	2308	2513	2737	2981	8.91%

서버 개발자는 <표 3-16>과 같이 평균증감률의 CAGR이 -1.44%, 평균변동률 CAGR이 -0.64%, 월별 공고 수의 CAGR이 -4.05%, 누적된 월별 공고 수의 CAGR이 8.91%로 도출되었다. 상대적으로 작은 폭의 감소율과 누적 월별 공고 수의 높은 성장률로 향후 장래에 상승하는 직업으로 서버 개발자가 있으며, 클라우드 서비스 및 빅데이터로 인한 서버 관리가 중요한 책무 중 하나로 자리매김한 이유로 유망한 직업 중 하나이다.

11) AR/VR 콘텐츠 개발자

<표 3-17> AR/VR 콘텐츠 개발자 CAGR 분석결과 (▲)

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균증감률	78	77	76	76	75	74	73	73	72	-1.00%
평균변동률	103	103	102	102	102	102	102	102	102	-0.15%
월별공고수	72	69	67	65	62	60	58	56	54	-3.43%
누적월별 공고수	1842	2002	2175	2364	2568	2791	3033	3296	3582	8.67%

AR/VR 콘텐츠 개발자는 <표 3-17>과 같이 평균증감률의 CAGR이 -1.00%, 평균변동률 CAGR이 -0.15%, 월별 공고 수의 CAGR이 -3.43%, 누적된 월별 공고 수의 CAGR이 8.67%로 도출되었다. CAGR의 감소폭이 상당히 작으며, 현재 정부 차원으로 콘텐츠 향상을 위한 투자 및 지원을 계속적으로 하고 있어, 현재 보다 더 많은 수준의 인력이 필요할 것으로 판단된다. 특히 2025년의 경우 2017년 대비 약 2배의 누적 월별 공고 수의 수요를 창출 할 것으로 예측되기 때문에 전망이 좋은 직업이다.

12) 웹 방송엔지니어 (▲)

<표 3-18> 웹 방송엔지니어 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	63	62	61	60	59	59	58	57	56	-1.31%
평균 변동률	83	83	82	82	82	81	81	81	80	-0.40%
월별 공고수	37	36	35	34	33	32	31	30	29	-3.00%
누적월별 공고수	816	886	963	1045	1135	1233	1339	1454	1579	9.00%

웹 방송엔지니어는 <표 3-18>과 같이 평균증감률의 CAGR이 -1.30%, 평균변동률 CAGR이 -0.40%, 월별 공고 수의 CAGR이 -3.00%, 누적된 월별 공고 수의 CAGR이 9.00%로 도출되었다. 전체적으로 감소율의 폭이 적으며, 누적

월별 공고 수가 상당히 높게 측정되었다. 즉, 1인 미디어의 발전으로 인하여 웹 방송과 관련한 콘텐츠의 증가 추세에 있어 방송관련 엔지니어링 기술은 지속적으로 필요하다고 예측된다. 따라서 방송과 관련된 인력의 수요는 장래에도 지속적으로 성장 할 것이라는 것을 입증한다.

13) 드론 소프트웨어 개발자 (▲)

<표 3-19> 드론 소프트웨어 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	82	82	81	81	80	79	79	78	78	-0.71%
평균 변동률	109	110	110	110	110	110	110	110	110	0.05%
월별 공고수	116	111	106	101	96	92	88	84	80	-5.00%
누적월별 공고수	3355	3640	3949	4284	4648	5042	5470	5934	6437	8.00%

웹 방송엔지니어는 <표 3-19>와 같이 평균증감률의 CAGR이 -0.71%, 평균변동률 CAGR이 0.05%, 월별 공고 수의 CAGR이 -5.00%, 누적된 월별 공고 수의 CAGR이 8.00%로 도출되었다. 다른 직업들의 경우 대부분이 감소율을 보이는 평균변동률에서 드론 소프트웨어 개발자는 미세하지만 상승률을 보이고 있다. 초기 드론의 활용은 취미의 일부분으로써 여가를 즐기는 용도로 활용되었으나 최근의 법적 규제 완화 및 기술의 발전으로 항법 설계 등의 전문적인 기술이 요구되기 때문에 전문가의 수요가 필요하다. 따라서 드론 소프트웨어 개발자의 경우 장래에 지속적인 상승을 보일 직업 중 하나이다.

14) 자동화 설계자 (▲)

<표 3-20> 자동화 설계자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	87	86	85	85	84	83	83	82	81	-0.80 %
평균 변동률	110	110	110	110	110	110	110	110	110	-0.02 %
월별 공고수	121	116	112	108	104	100	97	93	90	-3.64 %
누적월별 공고수	2915	3162	3429	3720	4035	4377	4748	5150	5586	8.47%

웹 방송엔지니어는 <표 3-20>과 같이 평균증감률의 CAGR이 -0.80%, 평균변동률 CAGR이 -0.02%, 월별 공고 수의 CAGR이 -3.64%, 누적된 월별 공고 수의 CAGR이 8.47%로 도출되었다. 최근 반도체의 발달과 공정에 따른 자동화를 통하여 비용을 절감하고 보다 수익을 극대화 할 수 있는 방법들이 고안되고 있다. 또한 2차 산업의 발란으로 인더스트리 4.0과 스마트 팩토리 등의 공정과정과 정보기술의 접목은 기술력 확보와 더불어 경쟁적인 생산성을 가져다 줄 것이며, 장래에 더욱더 많은 인력 수요가 필요할 것으로 예측된다.

15) 디지털 마케터 (-)

<표 3-21> 디지털 마케터 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	90	89	88	87	86	85	84	83	82	-1.23%
평균 변동률	112	112	112	111	111	110	110	110	109	-0.36%
월별 공고수	175	169	163	157	151	145	140	135	130	-3.71%
누적월별 공고수	4187	4536	4914	5323	5767	6247	6768	7332	7943	8.33%

디지털 마케터는 <표 3-21>과 같이 평균증감률의 CAGR이 -1.23%, 평균변동률 CAGR이 -0.36%, 월별 공고 수의 CAGR이 -3.71%, 누적된 월별 공고 수의 CAGR이 8.33%로 도출되었다. 디지털 마케터의 경우 상기 언급된 E-커머스 관리자와 마케팅 관련 직업과 동일하게 평이한 수준에서 증가율을 보이고 있으며, 장래에 지속적인 유지 형태의 수요가 있을 것으로 예측된다.

16) 로보틱스 시스템 개발자 (-)

<표 3-22> 로보틱스 시스템 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	68	67	67	66	65	65	64	63	63	-1.06 %
평균 변동률	90	90	89	89	89	89	89	88	88	-0.25 %
월별 공고수	119	115	111	107	104	100	97	93	90	-3.45 %
누적월별 공고수	2659	2877	3112	3367	3643	3942	4264	4614	4991	8.19%

로보틱스 시스템 개발자는 <표 3-22>와 같이 평균증감률의 CAGR이 -1.06%, 평균변동률 CAGR이 -0.25%, 월별 공고 수의 CAGR이 -3.45%, 누적된 월별 공고 수의 CAGR이 8.19%로 도출되었다. 로보틱스 시스템 개발자는 자동화 설계자와 동일하게 반도체 등의 공정 작업에서 필요한 로봇을 설계하고 테스트하고 구현하는 업무를 진행한다. 평이한 평균 성장률을 보이고 있으며, 장래에 지속적으로 꾸준한 성장을 보일 것으로 전망된다.

17) 온라인 머천다이저 (-)

<표 3-23> 온라인 머천다이저 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	66	65	64	63	62	61	61	60	59	-1.28 %
평균 변동률	86	85	85	85	84	84	84	83	83	-0.41 %
월별 공고수	89	86	82	79	76	72	69	67	64	-4.12 %
누적월별 공고수	2250	2435	2637	2854	3090	3345	3621	3921	4244	8.26%

온라인 머천다이저는 <표 3-23>과 같이 평균증감률의 CAGR이 -1.28%, 평균변동률 CAGR이 -0.41%, 월별 공고 수의 CAGR이 -4.12%, 누적된 월별 공고 수의 CAGR이 8.26%로 도출되었다. 온라인 머천다이저의 경우도 E-커머스 관리자, 디지털 마케터와 동일하게 평이한 수준의 성장률을 보이고 있다. 웹상에서 이뤄지는 온라인 쇼핑을 즐겨하면서 소비 패턴의 변화하였고, 앞으로도 아마존, 알리바바 등의 온라인 기반 쇼핑몰이 시장의 크기를 확장해 나갈 것이다. 따라서 관련 직종인 온라인 머천다이저 또한 소폭 감소는 있지만 지속적으로 평이하게 유지되는 성장률을 보일 것으로 예측된다.

18) 시스템 소프트웨어 개발자 (▲)

<표 3-24> 시스템 소프트웨어 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	87	86	86	85	85	84	84	83	83	-0.62 %
평균 변동률	110	111	111	111	111	111	111	111	112	0.14%
월별 공고수	154	149	144	139	134	130	125	121	117	-3.37 %
누적월별 공고수	3561	3860	4184	4535	4916	5328	5775	6260	6785	8.39%

시스템 소프트웨어 개발자는 <표 3-24>와 같이 평균증감률의 CAGR이 -0.62%, 평균변동률 CAGR이 0.14%, 월별 공고 수의 CAGR이 -3.37%, 누적된 월별 공고 수의 CAGR이 8.39%로 도출되었다. 시스템의 전반적인 소프트웨어를 구축하는 시스템 소프트웨어 개발자는 다양한 시스템 구축 및 설계, 구현 등의 역량을 겸비하고 있는 전문가이며, 정보 시스템의 경우 모든 산업에서 다양한 방식으로 사용되고 있기 때문에 더욱더 많은 수요가 생성될 것이라고 전망된다. 평균변동률의 경우에도 상승률을 보이고 있기 때문에 추후 지속적인 상승세를 보이는 직업으로 전망된다.

19) 임베디드 소프트웨어 개발자 (-)

<표 3-25> 임베디드 소프트웨어 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	59	58	57	56	55	54	53	52	52	-1.60%
평균 변동률	80	80	79	79	78	78	77	77	76	-0.64%
월별 공고수	75	72	69	66	63	61	58	56	54	-3.97%
누적월별 공고수	1706	1844	1992	2153	2327	2515	2718	2937	3174	8.07%

임베디드 소프트웨어 개발자는 <표 3-25>와 같이 평균증감률의 CAGR이 -1.60%, 평균변동률 CAGR이 -0.64%, 월별 공고 수의 CAGR이 -3.97%, 누적된 월별 공고 수의 CAGR이 8.07%로 도출되었다. 기존의 임베디드의 기술은 PC를 제외한 각종 전자기기 등에 설치되어 지정된 특정 기능을 수행하는 소프트웨어를 말하며, 사물인터넷, 모바일 등의 기본적인 하드웨어 등이 포함되어 있었다. 그러나 더 세분화되어지고 각자의 직종별 수요가 높아짐에 따라 분화되었으며, 임베디드는 평이한 성장률을 보이고 있다.

20) 모바일 게임 개발자 (▲)

<표 3-26> 모바일 게임 개발자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균증감률	99	98	97	97	96	96	95	95	94	-0.57%
평균변동률	119	119	119	119	119	119	119	119	120	0.07%
월별공고수	115	111	107	103	99	95	92	89	85	-3.68%
누적월별공고수	2618	2833	3066	3318	3591	3885	4205	4550	4924	8.21%

모바일 게임 개발자는 <표 3-26>과 같이 평균증감률의 CAGR이 -0.67%, 평균변동률 CAGR이 0.07%, 월별 공고 수의 CAGR이 -3.68%, 누적된 월별 공고 수의 CAGR이 8.21%로 도출되었다. 모바일 소프트웨어에서 분리된 모바일 게임 개발자는 최근의 모바일에서 사용자들의 수요와 사용이 가장 많은 게임을 집중적으로 개발하고 설계하는 역할을 하는 전문가로 정의되었다. 평균변동률의 경우도 증가 추세를 보이고 있으며, 전반적으로 2025년에도 유망 있는 직종 중 하나로 자리매김 가능한 직업으로 해석된다. 시간이 부족한 현대인에게 이동하면서 쾌락적 요소를 즐길 수 있는 수단으로써 게임 개발은 유망한 분야임에 틀림없다고 예측된다.

21) 웹 호스팅 관리자 (▲)

<표 3-27> 웹 호스팅 관리자 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	93	92	91	90	90	89	88	87	87	-0.86%
평균 변동률	117	117	117	117	116	116	116	116	116	-0.10%
월별 공고수	245	237	230	223	217	210	204	198	192	-2.99%
누적월별 공고수	6445	7028	7665	8359	9115	9941	10841	11823	12893	9.05%

웹 호스팅 관리자는 <표 3-27>과 같이 평균증감률의 CAGR이 -0.86%, 평균변동률 CAGR이 -0.10%, 월별 공고 수의 CAGR이 -2.99%, 누적된 월별 공고 수의 CAGR이 9.05%로 도출되었다. 웹에서 원격 및 호스팅과 관련된 연결 등을 원활하게 해주는 직종을 시행하는 웹 호스팅 관리자의 경우, 프라이버시 및 보안의 위험과 클라우드 서비스 등의 웹 기반으로 시행되는 서비스가 증가하고 있으므로 중요한 직업 중 하나이다. 기존에는 하드웨어 기반의 저장소 및 운영체제를 활용했던 것과 달리 정보 기술의 발전으로 실시간으로 웹상에서 호스팅을 이용하여 자료를 보관하고 운용할 수 있는 기술 및 서비스가 발전하였다. 따라서 지속적인 호스팅 관리는 매우 중요하며, 비즈니스 및 개인사용에서도 호스트 관리는 중요하다. 따라서 상대적으로 미세한 감소폭으로 분석되었으나, 누적 월별 공고 수는 24개의 직업 중 가장 높은 비율로 분석되어 수요가 지속적으로 증가하는 양상으로 예측된다.

22) 웹 표준 전문가 (▼)

<표 3-28> 웹 표준 전문가 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	48	47	46	45	44	43	42	41	40	-2.12%
평균 변동률	70	70	69	69	68	67	67	66	66	-0.89%
월별 공고수	73	70	67	65	62	60	58	56	54	-3.73%
누적월별 공고수	1714	1855	2008	2174	2354	2548	2759	2987	3234	8.26%

웹 표준 전문가는 <표 3-28>과 같이 평균증감률의 CAGR이 -2.12%, 평균변동률 CAGR이 -0.89%, 월별 공고 수의 CAGR이 -3.73%, 누적된 월별 공고 수의 CAGR이 8.26%로 도출되었다. 웹 표준은 WWW(World Wide Web) 등 국제적으로 통용되는 웹으로 공식 표준 및 규격을 맞추는 역할을 하는 직업을 말하며, 기존의 웹이 지속적으로 발전함에 따라서 웹의 활용도가 높아져 웹

표준은 중요한 요소였다. 그러나 최근에는 웹에서 어플리케이션으로 옮겨가는 추세를 갖고 있으며, 웹 표준의 기본적인 구축은 이제 성숙 단계에 도달하여 웹 표준에 대한 전망은 다른 직업에 비하여 상대적으로 수요가 적다는 것을 도출하였다.

23) 웹 프로그래머 (▲)

<표 3-29> 웹 프로그래머 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	96	96	95	95	95	94	94	93	93	-0.45%
평균 변동률	119	119	119	119	120	120	120	120	121	0.20%
월별 공고수	122	117	112	108	104	100	96	92	88	-3.91%
누적월별 공고수	2809	3036	3281	3546	3832	4142	4476	4837	5228	8.07%

웹 프로그래머는 <표 3-29>과 같이 평균증감률의 CAGR이 -0.45%, 평균변동률 CAGR이 0.20%, 월별 공고 수의 CAGR이 -3.91%, 누적된 월별 공고 수의 CAGR이 8.07%로 도출되었다. 웹의 성장은 이제 성숙 단계에 접어 들었지만, 다른 채널 및 기기에 따른 웹과의 연동성에서 웹 프로그래머는 상대적으로 많은 수요를 갖고 있다. 평균변동률이 상승세를 보이고 있으며, 따라서 지속적인 수요가 존재 할 것이라고 예측된다. 모바일 기반의 기술들의 발전과 함께 웹에서도 같이 연동 될 수 있도록 설계 및 구축을 하면서 전반적인 웹을 활용할 수 있도록 지원 등의 역할을 하면서 지속적인 수요를 보일 것이라 예측된다.

24) UX/UI 디자이너 (-)

<표 3-30> UX/UI 디자이너 CAGR 분석결과

	2017	2018	2019	2020	2021	2022	2023	2024	2025	CAGR
평균 증감률	65	64	64	63	62	61	60	59	58	-1.42%
평균 변동률	89	88	88	87	87	86	86	86	85	-0.51%
월별 공고수	99	96	93	89	86	83	80	78	75	-3.47%
누적월별 공고수	2234	2421	2623	2841	3078	3335	3613	3915	4241	8.34%

UX/UI 디자이너는 <표 3-30>과 같이 평균증감률의 CAGR이 -1.42%, 평균변동률 CAGR이 -0.51%, 월별 공고 수의 CAGR이 -3.47%, 누적된 월별 공고 수의 CAGR이 8.34%로 도출되었다. 사용자의 경험을 통하여 인터페이스를 구현하고 보다 편리한 디자인을 제공함으로써 사용자의 이용 비율을 올리는 역할을 하는 직업으로써 상대적으로 평이한 수준의 성장률을 보이고 있으며, 평균변동률이 경우 상대적으로 낮은 감소율을 보이고 있으며, 추후 사용자 경험에 대한 연구와 지속적이 설계 및 구현을 통하여 사용자를 보다 편안하게 웹 및 어플리케이션을 운영 할 수 있도록 지원함으로써 지속적으로 수요가 유지될 것으로 예측된다.

<표 3-31>과 같이 CAGR을 4가지 평균증감률, 평균변동률, 월별 공고 수의 평균, 누적 월별 공고 수 의 평균으로 계산한 결과 상대적으로 수요 성장률이 상승하는 직업은 14개로 소셜 미디어 마케터, 모바일 어플리케이션 개발자, 정보 보안 전문가, 정보 시스템 운영자, 웹 운영자, 서버 개발자, AR/VR 콘텐츠 개발자, 웹 방송 엔지니어, 드론 소프트웨어 개발자, 자동화 설계자, 시스템 소프트웨어 개발자, 모바일 게임 개발자, 웹 호스팅 관리자, 웹 프로그래머이다. 기존과 유사하게 유지되어질 직업은 7개의 직업으로 사물인터넷 소프트웨어 개발자, 소셜커머스 관리자, 디지털 마케터, 로보틱스 시스템 개발자, 온라인 머천다이저, 임베디드 소프트웨어 개발자, UX/UI 디자이너이다. 상대적으로 하강하는 직업으로는 3개의 직업으로 멀티미디어 디자이너, 멀티미디어 콘텐츠

관리자, 웹 표준 전문가의 결론을 도출하였다.

<표 3-31> 24개 직업별 CAGR 분석결과 간 상대적 수요 증감 비교

직업	소셜미디어 마케터	모바일 어플리케이션 개발자	멀티미디어 디자이너	정보 보안 전문가	멀티미디어 콘텐츠 관리자	정보 시스템 운영자
전망	▲	▲	▼	▲	▼	▲
직업	웹 운영자	사물인터넷 소프트웨어 개발자	소셜커머스 관리자	서버 개발자	AR/VR 콘텐츠 개발자	웹 방송 엔지니어
전망	▲	-	-	▲	▲	▲
직업	드론 소프트웨어 개발자	자동화 설계자	디지털 마케터	로보틱스 시스템 개발자	온라인 머천다이저	시스템 소프트웨어 개발자
전망	▲	▲	-	-	-	▲
직업	임베디드 소프트웨어 개발자	모바일 게임 개발자	웹 호스팅 관리자	웹 표준 전문가	웹 프로그래머	UX/UI 디자이너
전망	-	▲	▲	▼	▲	-

3. CAGR의 결과에 따른 SW인력 수요 전망의 시사점

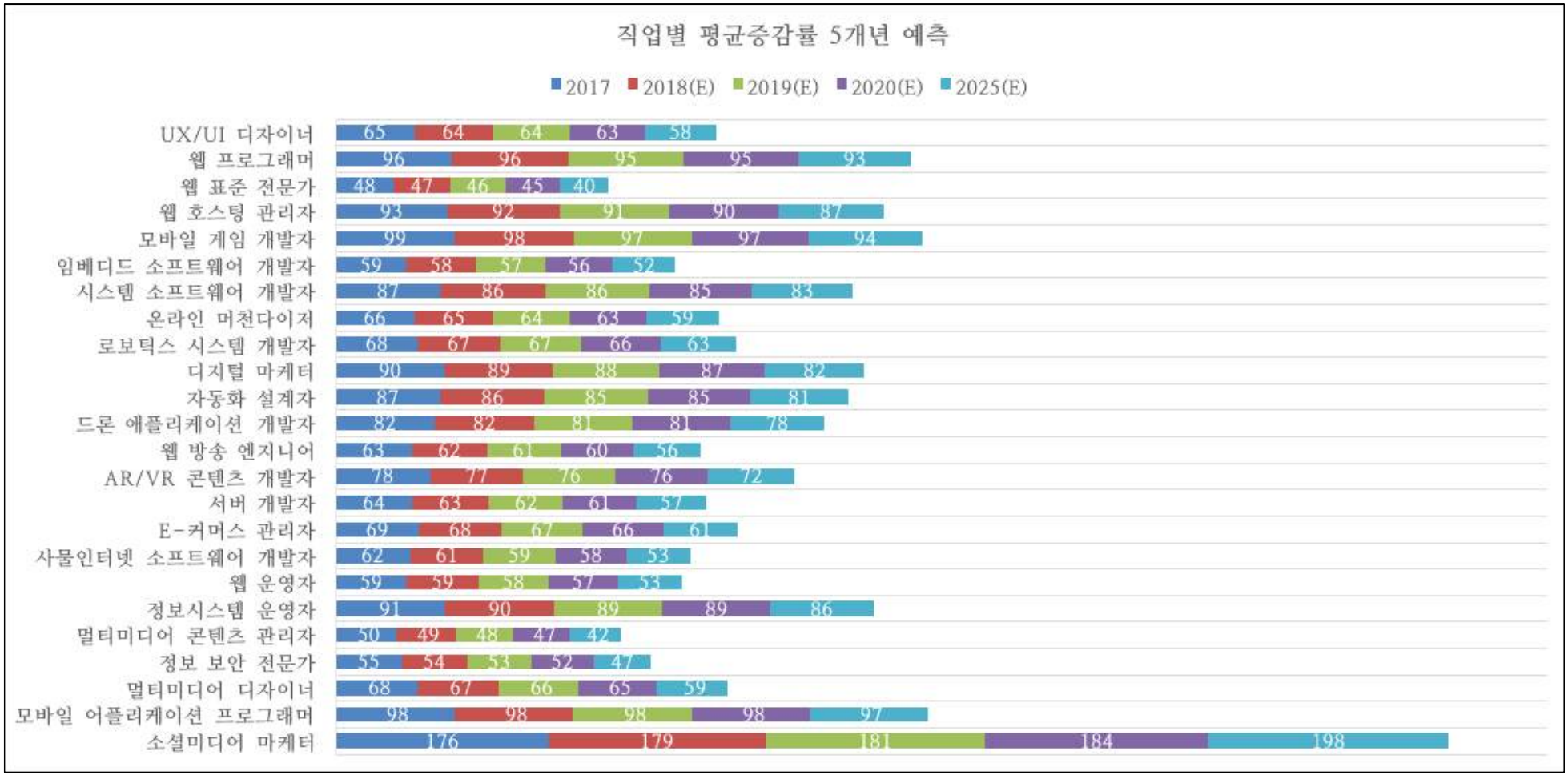
상기 분석을 토대로 2017, 2018, 2019, 2020, 2025년의 예측 값을 4가지 차원으로 비교해보면 [그림 3-45]와 같다. 소셜 미디어 마케터가 직업별 평균증감률에서 가장 빈도의 수가 높고 성장률도 높은 것으로 도출되었다. 또한 웹 프로그래머, 웹 호스팅 관리자, 모바일 게임 개발자, 정보시스템 운영자 등의 직종이 그 다음으로 빈도를 차지하고 있다. 상대적으로 웹 표준 전문가와 멀티미디어 콘텐츠 관리자 등의 직종은 낮은 빈도로 측정되고 연도별 증가량도 감소되고 있는 것을 확인 할 수 있다.

직업별 평균변동률 5개년 예측 비교는 [그림 3-46]과 같이 도출되었다. 직업별 평균 변동률에서도 소셜 미디어 마케터가 가장 많은 빈도수와 증가량을 보여줬으며, 웹 호스팅 관리자, 모바일 게임 개발자, 웹 프로그래머, 정보시스템 운영자, 자동화 설계자, 드론 애플리케이션 개발자 등이 그 다음의 빈도를 보였다. 수요가 작은 직업으로는 웹 표준 전문가, 멀티미디어 콘텐츠 관리자가 도출되었으며, 직업별 평균증감률과 비슷한 양상을 보였다.

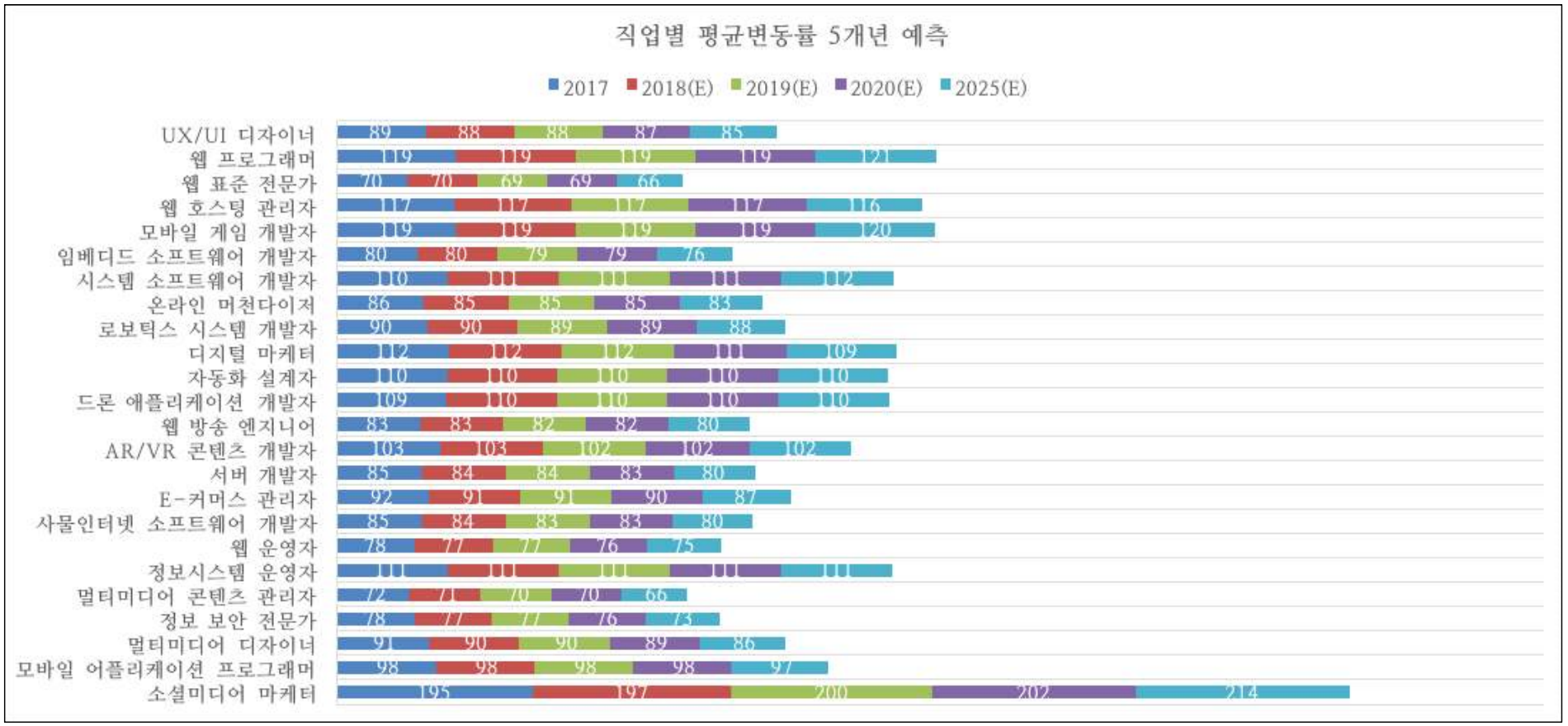
직업별 평균 월별 공고 수 5개년 예측은 [그림 3-47]과 같이 도출되었다. 직업별 평균 월별 공고 수에서도 소셜 미디어 마케터의 공고 수는 매우 높게 측정되었다. 또한 두드러지게 나타나는 직업으로는 사물인터넷 소프트웨어 개발자와 웹 호스팅 관리자가 그 뒤의 빈도를 나타냈다. 사물인터넷의 경우 4차 산업 혁명이 도래하고 난 이후에 반도체 기술의 발달과 정보 시스템의 발달로 인하여 산업 수요가 급증하고 있다는 것을 증명하고 있다.

반면에 웹 운영자가 가장 낮은 빈도를 보여줬으며, 웹 방송 엔지니어, 임베디드 소프트웨어 개발자 등의 순으로 낮은 빈도를 나타냈다. 웹 자체를 운영하는데 필요한 인력 보다는 산업 특성에 맞는 직업이 더 각광 받고 있다는 것을 시사해주고 있으며, 웹 방송 엔지니어의 경우도 1인 미디어로 변화하면서 점차 개인이 방송을 조율하고 엔지니어링 하는 기술력을 확보하고 있다는 것을 증명하고 있다.

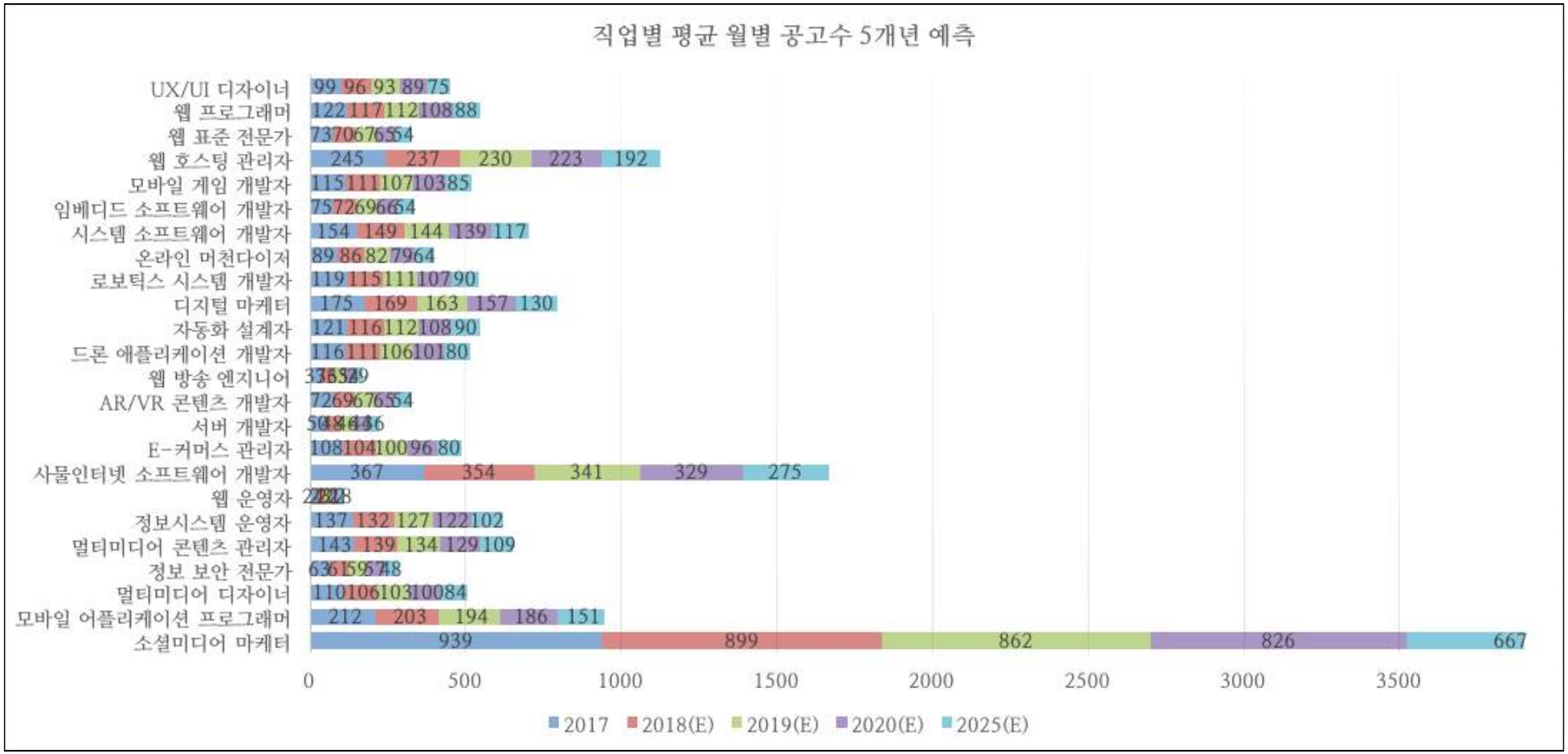
[그림 3-45] 직업별 평균증감률 5개년(2017~2020, 2025) 예측 비교



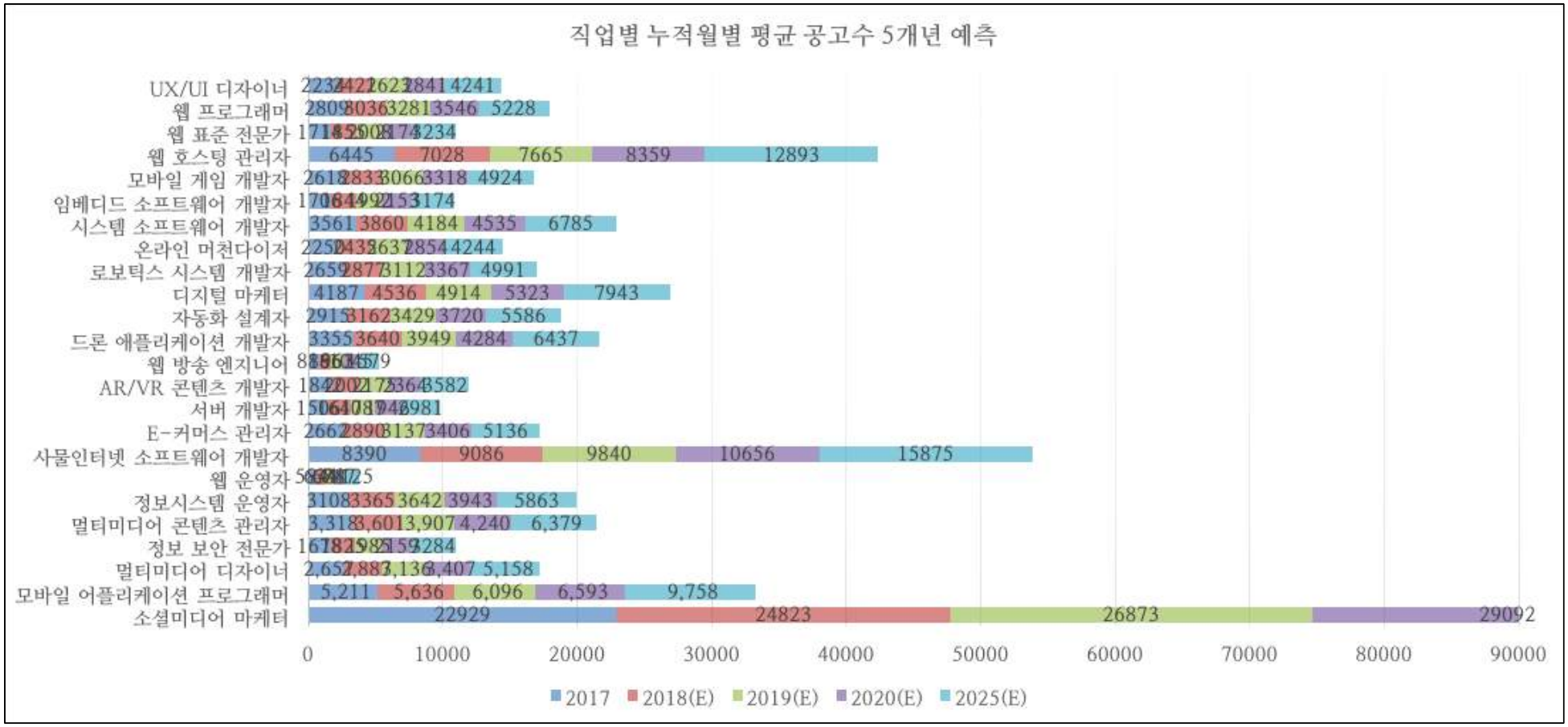
[그림 3-46] 직업별 평균변동률 5개년(2017~2020, 2025) 예측 비교



[그림 3-47] 직업별 평균 월별 공고 수 5개년(2017~2020, 2025) 예측 비교



[그림 3-48] 직업별 누적 월별 평균 공고 수 5개년(2017~2020, 2025) 예측 비교



직업별 누적 월별 공고 수 5개년 예측은 [그림 3-48]과 같이 도출되었다. 직업별 누적 월별 공고에서도 소셜 미디어 마케터가 가장 높은 빈도를 나타냈다. 즉, 소셜 미디어 마케터는 유망 직종으로써 지속적인 수요가 많은 직업으로 분석되었다. 그 다음으로 사물 인터넷 소프트웨어 개발자 등이 도출되었다. 수요가 적게 나타난 직업으로는 웹 운영자, 웹 방송 엔지니어, 서버 개발자 등이 도출되어 월별 평균과 상동 한다.

앞서 분석된 결과 들을 토대로 얻어낸 시사점으로는 첫째, 새로운 기술의 발전으로 소셜 미디어, AR/VR, 드론, 공정 자동화, 사물인터넷 등의 관련 직업이 각광 받고 있다는 점을 토대로 기존 분류체계와 기존의 시장의 현황과는 다르다는 것을 증명하였다. 더불어 전통적으로 존재하고 있던 기본적인 직업에 대해서도 꾸준하게 수요는 지속 될 것으로 전망된다.

제4장 시사점 및 제언

제1절 시사점 및 향후 연구 방향

본 연구과제는 SW 분야의 SW인력 수요에 대한 공공데이터와 시장데이터를 분석하고, 변화하는 산업·직업에 대응 가능한 수급전망 체계 및 인재정책을 제언함에 그 목적을 두었다. 그에 따라 인력 수요 정보(취업 포털 및 워크넷 등) 수집 데이터 크롤러(Crawler)를 개발하고 채용공고 자료를 전량 수집하였다. 이 과정에서 SW 직종 수요정보에 사용된 명사 단어 추출 및 데이터 분석을 위한 전처리(관련 용어 사전 만들기)과정을 거쳐 데이터 사전을 구성하였다. 이후 직종 도출을 위하여 DTM 토픽모델링을 통해 직종별 주요 키워드의 연관성을 추출하고 직종별 군집화를 수행하였다. 그에 따라 그룹별 키워드 기반 직종 및 직종 재정의 및 역량 정의를 통해 새롭게 도출된 직종을 확인하였다.

본 연구는 다음의 3가지 시사점을 갖는다.

1. 빅데이터 분석을 통한 SW 분야의 일자리 변화 확인

본 연구과제는 워크넷 데이터를 기반으로 SW산업 현장의 일자리 변화 형태를 분석하고 직종 분류를 재정립함으로써 이론과 실무 관점의 직종 전문성에 대한 이질감을 축소하였다. 또한 워크넷 채용공고 데이터에 대한 텍스트 마이닝을 통해 신규 직종의 도출 및 기존 직종의 특성을 확인하였다.

그에 따라 총 24종의 직종을 분류하였다. 특히 변화된 일자리로 4개 직종, 9개 직업을 구분하여 모바일 개발자, 서버 개발자, 웹 전문가, 임베디드 소프트웨어 전문가를 확인하였다. 또한 신규로 발견된 일자리로 미디어 콘텐츠 전문가, 디지털마케팅 전문가, 드론 애플리케이션 개발자, AR/VR 콘텐츠 개발자를 도출하였다. 이렇게 도출된 직종은 기존 분류체계인 ISCO, KSCO, KECO, NCS, ITSQF, O*Net, SFIA와의 매칭을 바탕으로 직종 및 직업의 특성을 재분류하였으며, SW 산업 분야 전문가들의 의견을 바탕으로 재정의한 직종 분류를

도출하였다. 또한 한국표준직종분류의 SW 분야 직종을 빅데이터 분석 기반 직종으로 재분류함으로서 분화, 변화, 발견 직종을 구분하여 향후 국가 경쟁력 재고를 위한 일자리 전략 도출이 가능한 기틀을 마련하였다.

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

본 연구과제는 기존 연구들이 활용한 설문 분석 및 실태 조사 또는 전문가 협의체에 의한 정성적 분류가 아닌, 실시간 텍스트 데이터를 기반으로 시간 흐름을 고려한 SW 분야의 신규 직종을 구분하였다. 그에 따라 워크넷 내의 채용공고 정보를 처리한 텍스트 마이닝을 통한 단어 구분 방식을 적용하였다.

추가적으로 기존 빅데이터 분석에서 활용되는 분류 기법인 LDA 알고리즘의 결과에 대한 한계점을 개선할 수 있는 DTM을 도입함으로서 SW산업 분야의 직종 분류체계 개발을 수행하였다.

특히 DTM을 바탕으로 도출된 직종·직업체계를 NCS와 워크넷의 분류와 비교 분석함으로서 기존 NCS와 워크넷 직종 기술에 관한 적합도가 비교적으로 떨어진다는 사실을 도출 할 수 있었고, 이를 바탕으로 데이터 분석 기반의 직종·직업 체계 방법론을 정립하여 향후 산업 수요 기반의 SW분야 일자리 육성방향 수립을 위한 기반 마련에 기여하였다. 이와 더불어 DTM으로 도출한 24개 직종 및 직업의 군집화 결과는 새로운 채용공고를 군집에 할당함으로써 실시간성을 띄는 직종 및 직업의 변화 추이를 관찰 할 수 있어 연속성을 가지는 대안통계로서 활용성을 보장하였다. 또한 각 직업별 CAGR을 바탕으로 향후 변화 가능성에 대한 예측을 도출해냈다는 점에서 본 연구는 향후 직업의 변화 가능성을 관찰 할 수 있는 체계를 마련하였다.

3. 일자리 변화 대처 전략 도출방안 제시

본 연구과제의 도출 결과는 기존 직업 및 직종의 체계를 산업계 수요에 맞는 체계로 재편하고, 그 변화상을 관찰 할 수 있다는 점에서 의의가 있다. 또한 연구과정에서 추후 SW 산업계의 육성 분야 및 일자리 창출 전략을 고려한 점으로 인해 관련 정책의 참고자료로서 활용하기에 용이 할 것으로 사료된다. 특히 직업 및 직종의 신규 발견 및 기존 직업의 분화 및 융합 사례 추출 결과는

그간 일자리 지형 파악에 어려움을 겪던 정책 입안자의 갈증을 해소하는 측면에서 연구 과제의 중요 시사점이라 할 수 있다.

제2절 제언

1. SW 분야 일자리 변화에 능동적 대응이 가능하고 일자리 문제 해결을 통해 국가경쟁력을 제고 할 수 있는 효과적인 데이터 기반 전략 도출 필요

현재까지 수행 되어 온 일자리 분야 연구의 직업 분석 결과는 신규 도출되는 직업 및 직종 도출의 근거로 활용하기에 어려운 측면이 존재하였다. 결과적으로 가까운 미래에 예상되는 직업 변화 대응에 어려움이 있었다. 본 연구과제가 개발한 빅데이터 분석 기반의 직업 도출 과정 및 일자리 변화 예측 방안을 활용은 향후 직종 및 직업에 대한 능동적 대응 정책을 도출하고 4차 산업혁명 시대의 국가 경쟁력을 강화함에 있어 도움을 줄 수 있다.

2. 全산업 분야의 직업 및 직종 변화예측을 통한 4차산업혁명 대응 체계 구축이 필요

본 연구의 SW분야 일자리 수요에 대한 분석 결과를 통해 실시간성, 직업의 변화상 확인의 중요성을 확인할 수 있었다. SW 분야 직업이 분화, 융합, 변화, 발견되는 양상으로 각기 다르게 변화하는 점을 볼 때, SW산업을 넘어 관련 산업의 변화상에 끼치는 영향을 함께 확인할 필요가 있다.

따라서 4차 산업사회를 대비하기 위하여 SW 산업뿐만 아니라 유관 산업에서 요구되는 직업 특성과 신규 직업의 도출, 기존 직업의 변화상을 지속적으로 관찰하는 것은 매우 중요하다. 특히 해당 정보는 일자리 변화에 대한 현재를 파악하고 향후 관련 산업 분야 전반의 확산을 도모하는 방안을 마련할 수 있다는 점에서, 달성되어야 할 과제이다.

참 고 문 헌

국내문헌

강만모, 김상락, 박상무 (2012), ‘특별기과-빅 데이터의 분석과 활용’, *정보과학학회지,* pp25-32

권우현, 공정승, 박광순, 심상완 (2014), ‘일반기계산업 인력수요전망’, *한국고용정보원*, pp.1-161

김난도, 전미영, 이향은, 이준영, 김서영, 최지혜 (2013), ‘트렌드 코리아 2014’, *서울:미래의 창*

김동규, 김중진, 장재호 (2015), ‘고용정보 활용성 제고를 위한 KECO-NCS 연계 및 개편 방안’, *한국고용정보원,* pp.1-182

김동연 (2017), ‘LDA 기반의 특허 트렌드 분석 도구의 설계 및 구현’, *숭실대학교 융합소프트웨어학과*, 석사학위 논문, pp. 1-47

김대건 (2014), ‘O2O동향과 시사점’, *정보통신정책연구원 정보통신 방송정책*, pp.22-26

김상겸 (2016). ‘토픽모델링을 이용한 국내 산업공학 연구동향 분석’, *서울과학기술대학교 일반대학원 데어터사이언스학과*, 석사학위 논문, pp. 1-43

김윤성, 이호창, 이석시, 이도길, 김한국, 김유일 (2017), ‘텍스트마이닝을 활용한 산업분석 방법론에 관한 연구: 문장 분류를 이용한 PEST와 긍정부정 분석’ *Journal of Knowledge Information Techonology and Systems*, 제 12권 1호, pp. 109-122

김태형 (2016), ‘토픽모델링을 이용한 핀테크 기술동향 분석’, *고려대학교대학원 산업경영공학과*, 석사학위 논문, pp. 1-45

미디어잇 기사 (2014.4.22.), ‘[SW가 힘이다] ① 경쟁의 룰 바꾸는 ‘소프트웨어’’.

미래창조과학부 (2017), ‘미래부, 17년 360억원 규모 디지털 콘텐츠 펀드 조성’, *미래창조과학부*

미래창조과학부 (2016), ‘미래부, 가상현실(VR) 전문펀드 조성 방안’, *미래창조과학부*

미래창조과학부 (2016), ‘소프트웨어 품질성능평가시험 의무화 제도 원년 결산 및 간담회 개최’,*미래창조과학부*

미래창조과학부 (2015), ‘공공 SW사업 첫 시범 분할발주로 SW산업 생태계 정상화 첫걸음’, *미래창조과학부*

미래창조과학부 보도자료 (2013.10.9.), ‘소프트웨어(SW), 창조경제 실현도구로 키운다’ *미래창조과학부.*

박대한 (2017), ‘토픽 모델링을 이용한 음악 추천 시스템’, *홍익대학교 일반대학원 산업공학과*, 석사학위 논문, pp. 1-52

박세정 (2015),‘워크넷 구인구직DB를 이용한 임금 동향’, *한국고용정보원,* pp. 1-22

박영숙, J. Glenn, T. Gorden, E. Florescu (2013), ‘유엔미래보고서 2040’, 교보문고

박자현, 송민, (2013), ‘토픽모델링을 활용한 국내 문헌정보학 연구동향 분석’, *정보관리학회지* , 제 30권 1호, pp. 7-34

배기형 (2016), ‘커뮤니케이션 이해총서’, *커뮤니케이션북스(주)*

산업통상자원부 (2017), ‘대한민국 로봇산업 기술로드맵 Ⅱ’, *산업통상자원부*

선대인 (2017), ‘일의 미래: 무엇이 바뀌고, 무엇이 오는가’, *인플루엔셜*

소프트웨어정책연구소(2016), ‘소프트웨어산업 인력수요 전망’, *소프트웨어정책연구소*, pp.1-64

안정국, 김소담, 김희웅 (2016), ‘텍스트 마이닝 기법을 이용한 정보시스템 분야 연구 동향’, *Information System Review*, 제 18권 3호, pp. 73-97

안주언 (2017), ‘카셰어링 오피니언 마이닝과 동적 토픽 모델링’, *서울과학기술대학교 SW분석 · 설계학과*, 석사학위 논문, pp.1-50

양승준, 이보현, 김희웅 (2016), ‘토픽모델링 기반 행복과 불행 이슈 분석 및 행복 증진 방안 연구, *지식경영연구*, 제 17권 2호, pp. 165-186

오동환, 채승완, 김수인 (2014), ‘창조경제 시대의 일자리 창출을 위한 인터넷 경제의 일자리 창출효과 분석 및 시사점’, *한국인터넷진흥원*, pp. 1-28

유종열(2014), ‘중소 SW기업의 직무만족, 조직몰입, 이직의도에 영향을 미치는 요인에 관한 연구’, *고려대학교 노동대학원*, 석사학위 논문, pp. 1-65

이동소 (2014), ‘세계 외식산업의 현황과 전망’, *한국농촌경제연구원*, pp.1-19

이호, 공영일 (2017), ‘빅데이터 분석 기반 SW산업 직종 · 직무 분류체계 개발 및 동향 연구’, *소프트웨어정책연구소*, pp. 1-272

임아영 (2014), ‘기초화장품 판매를 위한 온라인 커스터마이징 서비스 비즈니스 모델 연구’, *성균관대학교 대학원 석사논문*, pp25-26

임하균 (2017), ‘1인 미디어 캐릭터 디자인 : 개인 인터넷 방송을 중심으로’, *홍익대학교 국제디자인 전문대학원 디자인경영전공* 석사학위논문, p12

정보통신산업진흥원 (2014), ‘SW산업 주요통계’.

정보통신산업진흥원 (2013), ‘수요지향형 소프트웨어 산업 분류체계 개선 연구’.

조요한, 김동우, 문인철, 오혜연 (2009), ‘LDA를 이용한 대용량 블로그 문서 처리’, *한국과학기술원*, pp. 1-4

주기환 (2014), ‘디지털콘텐츠산업 활성화를 위한 투자자본의 역할과 과제’, *정보통신산업진흥원*, pp.1-167

최재원, 이호, 김정민, 송주호 (2017) ‘토픽모델링을 활용한 소프트웨어 분야 대학 교과과정 분석’, *한국전자거래학회*, 제 22권, 4호, pp.193-214

통계청 (2017), ‘한국표준직업분류’, *통계청,* pp.1-24

한국고용정보원 (2007), ‘한국고용직업분류’, *한국고용정보원,* pp. 1-407

한국언론진흥재단 (2017), ‘4차 산업혁명에 대한 국민들의 인식’, *미디어이슈*, 제 3권 제4호.

한국산업기술진흥원 (2016), ‘미래 유망 신산업의 시장 및 인력수요 전망’, *한국산업기술진흥원,* pp. 1-34

한국산업기술평가관리원 (2013), ‘2013 산업기술수준조사 보고서’.

한국소프트웨어산업협회 (2016), ‘IT 분야 역량인정체계(안) (ITSQF)’, *한국소프트웨어산업협회,* pp.1-23

한국소프트웨어산업협회 (2017) ‘국가직무능력표준(NCS) 안내 - 정보기술(SW)분야 26개 세분류’, *한국소프트웨어산업협회,* pp. 1-36

한국정보통신기술협회 (2010), ‘IT 서비스 및 자산관리 SW BMT 평가모델 개발’, *한국정보통신기술협회,* pp.1-264

황지현 (2015), ‘O2O 커머스를 넘어 On-Demand Economy로’, *KT경제연구소*

허성완 (2015), ‘한국어웹 문서 범주화를 위한 텍스트 마이닝 기법 연구’, *아주대학교 소프트웨어특성화학과,*

석사학위 논문, pp. 1-43

허재준(2017), ‘4차 산업혁명이 일자리에 미치는 변화와 대응’, *한국노동연구원,* 월간노동리뷰 3월호, pp. 62-72

홍성민, 이윤준, 김진하, 손경현, 정미나 (2016), ‘미래산업·신산업 분야 인재기반 조성을 위한 인적자원 양성 및 취·창업 지원방안 연구’, *과학기술정책연구원.*

홍성연, 최재원, (2017), ‘토픽 모델링 분석 기법을 활용한 대학의 학생 지원 연구 동향 분석’, *학습자중심교과교육연구,* 제17권 8호, pp. 21-48

국외문헌

Blei, M,, (2012) ‘Introduction to Probabilistic Topic Models’, *Princeton University,* pp.1-16

Blei, M., Andrew, Y., Michael I. (2003), ‘Latent Dirichlet Allocation’, *Journal of Machine Learning Research 3,* pp. 993-1022

Blei, M., Lafferty, J. (2006), ‘Dynamic Topic Models’, Carnegie Mellon University Research Showcase, pp. 1-9

Chen, Y., Zuo, W., Chen, K. and Lin, Y. (2015), ‘Time Label Topic Model’, *International Journal of Database Theory and Application*, Vol. 8, No. 1, pp. 213-226

Greene, D., Cross, J., (2016) ‘Exploring the Political Agenda of the European Parlianment Using a Dynamic Topic Modeling Approach’, *Cornell University Library*, pp. 1-47

International Labour Organization(ILO) (2007), ‘Updating the International Standard Classification of Occupation(ISCO) Summary of major changes between ISCO-88 and ISCO-08’, *ILO*, pp. 1-21

Iwata, T., Yamada, T., Yasushi, S. and Ueda, N. (2010), ‘Online Multiscale Dynamic Topic Models’, *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp.1-10

SFIA Foundation, (2011), ‘SFIA 5 Framework reference’, *Skills Framework for the Information Age*, pp.1-48

World Economic Forum (2017), ‘Future of Jobs’, World Economic Forum.

웹 문헌

O*net Online 홈페이지 : https://www.onetonline.org/
SFIA SW역량 소개 홈페이지 : sfia-online.org

주 의

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

93000
9 788961 084093
ISBN 978-89-6108-409-3

OPEN 공공누리 공공저작물 자유이용허락 출처표시 상업용금지 변경금지

[SW정책연구소]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.)