

박사학위논문

봇 기반 특허회피 설계안
자동 생성 방법에 관한 연구

2020년

한 성 대 학 교 대 학 원

스마트융합컨설팅학과

스마트융합제품전공

서 성 현

박 사 학 위 논 문
지도교수 노광현

봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구

A Study on a Bot-based Automatic Generation
Method of Patent Avoidance Design

2019년 12월

한 성 대 학 교 대 학 원

스마트융합컨설팅학과

스마트융합제품전공

서 성 현

박 사 학 위 논 문
지도교수 노광현

봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구

A Study on a Bot-based Automatic Generation
Method of Patent Avoidance Design

위 論文을 工學 博士學位論文으로 제출함

2019년 12월

한 성 대 학 교 대 학 원

스마트융합건설팅학과

스마트융합제품전공

서 성 현

서성현의 공학 박사학위논문을 인준함

2019년 12월 일

심사위원장 _____(인)

심 사 위 원 _____(인)

심 사 위 원 _____(인)

심 사 위 원 _____(인)

심 사 위 원 _____(인)

국 문 초 록

봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구

한 성 대 학 교 대 학 원
스 마 트 융 합 컨 설 링 학 과
스 마 트 융 합 제 품 전 공
서 성 현

특허회피 전략은 시제품을 출시하기 전 또는 연구 개발에 들어가기에 앞서 시장분석 및 선행조사에 대한 필요성을 느끼는 기업이나 개인 고객들에게 미리 공백기술 및 틈새시장을 겨냥한 양질의 특허를 확보하기 위한 방안을 모색하는데 주로 활용된다. 현재, 상대방이 이미 선점한 장벽특허 및 불완전 경쟁시장에서 일어나는 여러 가지 폐해로부터 조금이나마 벗어나기 위한 노력이 계속적으로 일어나고 있으나, 특허회피 전략 제시시 특허 전문가에 따라 각자 서로 다른 의견이 나올 우려가 있고, 각자 다른 견해를 갖는 특허 전문가의 의견들이 과연 회사나 개인 고객들에게 도움이 될 수 있게 설계된 것일지에 대하여 불확실한 문제점을 가지고 있다.

특허회피 전략은 향후 연구개발에 투입되는 인적, 물적 낭비를 최소화하고 상대방의 장벽특허를 회피할 수 있는 전략을 제시해 줌으로써, 공백특허 마련 및 독과점 사장이 아직 형성되어 있지 않은 미개척 시장 또는 새로운 틈새시장을 발굴해 상대방의 지식재산권에 대한 분쟁을 사전에 예방

할 수 있도록 도와주고, 고객이 관심 있어 하는 제품이나 관련 기술을 더욱 고도화시켜 기술 사업화 및 기술거래의 활성화를 촉진시켜 주는 역할을 한다.

본 논문은 특허회피 설계안을 필요로 하는 고객을 대상으로 봇 기반 특허회피 설계안 자동 생성 방법을 제안하는 연구이다. 특허회피 설계안 자동 생성시 적용되는 최종 주제어를 선정하기 위해 기업이 출시하고자 하는 시제품 및 그 해당 기술 분야를 대표하는 핵심 키워드를 적용하여 특허 건들을 검색하고, 검출된 선행 특허들에 기재된 특허청구범위를 분석해 특허 키워드를 색출하는 과정을 거친다. 이렇게 색출된 특허 키워드로 포털 사이트에 접근해 특허회피 설계안을 제시하는데 반드시 필요한 정보글을 색출하고 정보글로부터 후보 주제어를 추출한다. 추출된 후보 주제어를 대상으로 1차 필터링(유사성 판단), 2차 필터링(연관성 판단), 3차 필터링(적합성 판단)을 실시하여 특허회피 설계안으로 활용될 최종 주제어를 제시하게 된다.

봇 기반 자동 생성 프로그램 구현에 의해 선정된 최종 주제어를 대상으로 전문가에게 설문을 실시함을 통해 특허회피 설계안을 자동 생성하는데 적용되는 최종 주제어의 유용 가치성에 대한 검증을 실시하였다. 그 결과, 최종 주제어가 특허회피 설계안에 적극 반영되도록 봇을 통해 자동 생성시킴에 따라 특허회피 설계 마련 및 특허회피 전략으로 활용하는데 있어 최종 주제어가 갖는 정보가치 대한 중요성을 확인하였다.

봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구는 기업이나 개인 고객들로부터 나온 아이디어나 우수한 발명이 상대방의 지식재산권에 의해 인적으로나 물적으로 피해를 입지 않도록 함은 물론 공백특허의 확보 및 새로운 시장을 개척하는데 도움이 될 만한 다양한 특허회피 설계안을 봇을 통해 제공해 줄 수 있을 것이다.

【주요어】 봇, 특허회피 설계안, 핵심 키워드, 특허 키워드, 정보글, 후보 주제어, 최종 주제어

목 차

I. 서 론	1
II. 연구배경 및 필요성	6
2.1 연구배경	6
2.2 특허회피 설계 소개	12
2.2.1 특허회피 설계의 정의	12
2.2.2 특허회피 설계의 기준조건 및 사례	19
2.3 연구의 필요성	21
2.4 기존의 특허회피 설계방식의 한계	23
2.5 봇 소개 및 현황	24
2.6 특허회피 설계안에 관한 설문조사	32
2.6.1 빈도분석	32
2.6.2 교차분석	36
2.6.3 상관분석	62
III. 봇 기반 특허회피 설계안 자동 생성 방법	66
3.1 자동 생성 방법 소개	66
3.2 봇의 가상 모델	70
3.3 자동 생성 방법에 관한 알고리즘 설계	73
3.3.1 전체 알고리즘에 관한 설명	73
3.3.2 정보글 추출	75
3.3.3 유사성 판단	84
3.3.4 연관성 판단	91
3.3.5 적합성 판단	93
3.4 모의실험	95
3.4.1 사례1: 핵심 키워드가 ‘증강현실’, ‘가상현실’인 경우	95
3.4.2 사례2: 핵심 키워드가 ‘사물인터넷’, ‘IoT’인 경우	120

IV. 실험 및 검증	144
4.1 특허회피 설계안 자동생성 봇 구현	144
4.2 봇의 최종 주제어 선정 실험	149
4.3 실험 결과 검증	152
V. 결 론	159
참 고 문 헌	161
부 록	171
ABSTRACT	178

표 목 차

[표 2-1] 기존의 특허회피 설계안을 관점에서 본 SWOT 분석	22
[표 2-2] 챗봇의 발전단계	30
[표 2-3] 특허회피 설계의 필요성에 관한 빈도분석	33
[표 2-4] AI 플랫폼의 설계요건에 관한 빈도분석	34
[표 2-5] AI 플랫폼의 조건설정에 관한 빈도분석	35
[표 2-6] Q2-1와 Q2-7 간의 교차분석 결과표	37
[표 2-7] Q2-7와 Q2-8 간의 교차분석 결과표	38
[표 2-8] Q2-2과 Q2-1 간의 교차분석 결과표	40
[표 2-9] Q2-8와 Q2-11 간의 교차분석 결과표	41
[표 2-10] Q2-7과 Q2-2 간의 교차분석 결과표	42
[표 2-11] Q2-11과 Q3-1 간의 교차분석 결과표	44
[표 2-12] Q3-1과 Q3-18 간의 교차분석 결과표	45
[표 2-13] Q3-3과 Q3-4 간의 교차분석 결과표	46
[표 2-14] Q3-4와 Q3-5 간의 교차분석 결과표	48
[표 2-15] Q3-5과 Q3-6 간의 교차분석 결과표	49
[표 2-16] Q3-6과 Q3-7 간의 교차분석 결과표	51
[표 2-17] Q3-7과 Q3-18 간의 교차분석 결과표	52
[표 2-18] Q3-9과 Q3-10 간의 교차분석 결과표	53
[표 2-19] Q3-10와 Q3-12 간의 교차분석 결과표	54
[표 2-20] Q3-12와 Q3-4 간의 교차분석 결과표	56
[표 2-21] Q3-12과 Q3-14 간의 교차분석 결과표	57
[표 2-22] Q3-13과 Q3-12 간의 교차분석 결과표	58
[표 2-23] Q3-14와 Q3-16 간의 교차분석 결과표	59
[표 2-24] Q3-14와 Q3-7 간의 교차분석 결과표	61
[표 2-25] Q3-16와 Q3-4 간의 교차분석 결과표	62
[표 2-26] Q3-10 기준의 상관관계 분석결과	64
[표 2-27] 유의한 관점에서 본 상관관계 분석결과	65

[표 3-1] 유사성 판단으로 분류된 후보 주제어	98
[표 3-2] 유사성 판단에 따른 이상적 표본 안	101
[표 3-3] Ex 1-1의 유사성 통과 주제어	109
[표 3-4] Ex 1-2의 유사성 통과 주제어	113
[표 3-5] 사례1의 적합성 판단 실시	118
[표 3-6] Ex 2-1의 유사성 통과 주제어	133
[표 3-7] Ex 2-2의 유사성 통과 주제어	137
[표 3-8] Ex 2-3의 유사성 통과 주제어	140
[표 3-9] 사례2의 적합성 판단 실시	142
[표 4-1] 최종 주제어 제시	151
[표 4-2] 특허회피 설계안의 모범 예	154
[표 4-3] 5번 질의-응답결과	157

그 립 목 차

[그림 2-1] 자연어처리 일반	8
[그림 2-2] 붓 기반 특허회피 설계안 자동 생성 절차 개요	11
[그림 2-3] 기존의 특허회피 설계 절차와의 비교	12
[그림 2-4] 특허회피 설계론 소개	15
[그림 2-5] 특허회피 설계론 사례	19
[그림 3-1] 방법 알고리즘의 개략도	67
[그림 3-2] 붓의 제 1 가상 모델	71
[그림 3-3] 붓의 제 2 가상 모델	72
[그림 3-4] 붓의 가상 모델링 결과	73
[그림 3-5] 방법 알고리즘의 전체 블럭도	75
[그림 3-6] 정보글 추출	76
[그림 3-7] 붓 구동에 따른 정보글 추출	78
[그림 3-8] 조건식 변경에 따른 정당성 여부 타진	79
[그림 3-9] 변경된 조건식에 따른 붓의 정보글 추출	81
[그림 3-10] 정보글 추출에 관한 시나리오1	83
[그림 3-11] 정보글 추출에 관한 시나리오2	84
[그림 3-12] 유사성 판단	86
[그림 3-13] 유사성 판단의 예1	88
[그림 3-14] 유사성 판단의 예2(U그룹)	90
[그림 3-15] 유사성 판단의 예2(C그룹)	91
[그림 3-16] 연관성 판단	93
[그림 3-17] 적합성 판단	94
[그림 3-18] 사례1의 후보 주제어 색출	96
[그림 3-19] Ex 1-1의 피어슨 상관계수 및 정규분포	104
[그림 3-20] Ex 1-2~3의 피어슨 상관계수 및 정규분포	105
[그림 3-21] 정규분포 비교	105
[그림 3-22] Ex 1-1의 제 1 최적점수(OS ₁)	108

[그림 3-23] Ex 1-1의 제 2 최적점수(OS ₂)	108
[그림 3-24] Ex 1-1의 r, OS ₁ , OS ₂	109
[그림 3-25] Ex 1-2의 제 1 최적점수(OS ₁)	111
[그림 3-26] Ex 1-2의 제 2 최적점수(OS ₂)	112
[그림 3-27] Ex 1-2의 r, OS ₁ , OS ₂	112
[그림 3-28] 사례1의 최종 주제어 선정	120
[그림 3-29] 사례2의 후보 주제어 색출 및 유사성 판단	123
[그림 3-30] 삼각형 확률 분포도	125
[그림 3-31] 가중치 점수 부여에 관한 타당성 검증	126
[그림 3-32] 사례2에 적용된 가중치 점수	128
[그림 3-33] Ex 2-1의 제 1 최적점수(OS ₁)	132
[그림 3-34] Ex 2-1의 제 2 최적점수(OS ₂)	132
[그림 3-35] Ex 2-1의 r, OS ₁ , OS ₂	133
[그림 3-36] Ex 2-2의 제 1 최적점수(OS ₁)	135
[그림 3-37] Ex 2-2의 제 2 최적점수(OS ₂)	136
[그림 3-38] Ex 2-2의 r, OS ₁ , OS ₂	136
[그림 3-39] Ex 2-3의 제 1 최적점수(OS ₁)	139
[그림 3-40] Ex 2-3의 제 2 최적점수(OS ₂)	139
[그림 3-41] Ex 2-3의 r, OS ₁ , OS ₂	140
[그림 3-42] 사례2의 최종 주제어 설정	143
[그림 4-1] 특허회피 설계안 자동생성 봇의 핵심 키워드 입력	145
[그림 4-2] 특허회피 설계안 자동생성 봇의 특허 건 검출	146
[그림 4-3] 특허회피 설계안 자동생성 봇의 특허 키워드 색출	147
[그림 4-4] 특허회피 설계안 자동생성 봇의 정보글 색출	148
[그림 4-5] 특허회피 설계안 자동생성 봇의 최종 주제어 선정	149
[그림 4-6] 확장된 최종 주제어 추출	150
[그림 4-7] 1번 질의-응답결과	153
[그림 4-8] 3번 질의-응답결과	156

I. 서 론

세계 경제는 지속 가능한 발전의 개념을 도입하면서 지식기반 산업이 성장의 핵심으로 작용하게 되었고 특허권이 국제통상 마찰의 이슈로 등장하게 되었다. 선진국들이 개발도상국의 추격을 경계하기 위해 특허권을 국제 산업통상의 무기로 활용하게 됨에 따라, 지식재산권의 대명사격인 특허는 지식 기반 산업의 초석이 되었다. 특허는 국제주의를 취하고 있어 세계적으로 새로운 기술의 발명 여부를 바로 확인할 수 있으며, 특허공개 제도를 통해 국제적으로 공개되기 때문에 동일하거나 유사한 기술에 대하여 보호를 받을 수 있다.

특허의 목적은 산업상 이용 가능하고, 신규성과 진보성이 있는 발명에 대해 심사를 거친 후 일정기간 동안 출원인(권리자)에게 독점적이고, 배타적인 권리인 특허권을 부여하여 발명 창작의식을 고취시키고, 이를 통해 산업발전을 촉진하여 궁극적으로는 국가의 이익을 도모하기 위한 것이다. 객관적이고 표준적인 기술 정보인 특허는 기술수준과 기술혁신 흐름뿐만 아니라 기술혁신 동향 등을 조망하는데 유용하게 활용되며, 기술적, 사업적 정보제공의 측면에서 보았을 때도 풍부한 자료의 원천이라 할 수 있다. 연구개발의 주요한 산출물이라 할 수 있는 특허는 혁신과 기술적 변화를 위한 유익하며 새로운 기술을 창출할 수 있는 빅데이터 중 하나이다. 더욱이, 특허는 온라인 접근이 용이하며 특허의 작성 절차 및 양식이 명확하고 구체적으로 정립되어 있다.

특허 양식에는 특허의 수, 등록 연도, 등록 국가, 출원인 등과 같은 서지 정보와 통계정보가 포함되며, 이러한 정보는 특허의 기술적 가치와, 영향, 확산의 분석을 가능하게 하고, 특허 데이터는 기업의 혁신활동 조사, 새로운 기술 기회와 미래 기술 트렌드, 기업의 기술 경쟁력을 분석하는데 용이하다. 이에 따라, 여러 기업에서 특허 데이터를 활용하여 자사의 연구개발 과제의 탐색, 기술 동향 파악, 사업 전략 수립 등을 구체적으로 수립하며, 자사의 재무를 보완하고 연구개발의 증명을 통해 성과 및 가치를 창출한다.

특허는 새로운 아이디어 창출을 위한 투자를 촉진시키고, 삶의 질과 생산성 향상을 위한 새로운 기술 개발을 활성화시키기 때문에 경제 발전의 중요한 원동력 중 하나로 간주된다. 특허의 장점은 특허 자체가 가지고 있는 기술적 특징뿐만 아니라 이전에 출원되었던 특허에 대한 인용정보, 특허청구범위 등의 정성분석을 통해 기술변화 추세와 새로운 기술 출현의 방향을 제시할 수 있으며 특정 기술 영역에서의 기업이나 연구자의 기술 수준도 추출해 낼 수 있다.

이러한 지식재산권의 특허활용은 시장에서의 안정적인 사업을 보장하여 기업이윤 추구의 밑거름이 되고 있고, 지식재산권의 특허활용은 기업의 영속성을 보장하는 강력한 수단으로 자리 매김하고 있다. 이와 같은 지식재산권을 적극 활용하는 것을 다른 말로 일컬어 지식재산권 경영 또는 특허 전략이라고 표현한다. 특허전략 수립의 일환으로, 기업이 출원한 특허가 지식재산권으로 보호받게 되면, 이 특허는 라이선싱, 교차 라이선싱, 기술 무역 등 다양한 경로를 통해 다른 지역이나 국가를 넘나들면서 기업의 경제적 가치가 기업 수준에 그치는 것이 아니라 도시, 지역, 국가 차원까지 확산되어 영향력을 미치게 한다.

만약, 권리자의 보호가 미흡한 경우에는 특허제도 이용의 의욕을 상실시켜 기술을 비밀리에 보유하려는 경향을 조장하게 될 것이고, 역으로 과대 보호를 한다면 공개된 기술을 기초로 하여 새로운 기술을 개발하거나 그 기술을 이용해 사업하고자 하는 제 3 자에게 특허침해로 인한 예측할 수 없는 손해를 입게 함으로 어느 것이나 특허제도의 목적에 맞지 않게 된다.

즉, 특허권의 침해에 관한 판단여부는 특허권자의 이익과 제 3 자의 이익의 조화라는 근본적 이념을 고려하여 판단되어야 한다.

기업규모에 따라 대기업과 중소기업에서 지식재산권을 활용하는 형태에는 차이가 존재하며, 이러한 차이가 발생하는 근본적인 이유는 중소기업은 대기업이나 중견기업에 비해 상대적으로 인적자원 및 물질 자원이 제한적이기 때문이다. 대기업이나 중견기업의 경우는 연구개발을 통해 출시된 제품의 보호를 위해 전담 특허조직을 활용하여 자체 권리화, 기술이전 및 소송 등과 같은 형태로 지식재산 경영을 펼칠 수 있는 제반적 환경을 갖추

고 있다.

다시 말해, 대기업이나 중견기업은 사내에 충분한 자질을 갖춘 인력이 준비되어 있을 뿐만 아니라 개발 인프라도 체계적으로 갖추고 있어, 새로운 아이디어 발굴이나 지식재산권 출원-등록 그리고 사후 유지관리 및 보수가 원만하게 진행할 수 있다.

반면, 중소기업을 둘러싼 경영환경은 기술의 진보, 자유 무역 협정의 확산, 글로벌 아웃소싱의 확대 등으로 글로벌 경쟁이 심화되면서 더욱 빠르게 변화하고 있다. 국가 간 글로벌 경쟁이나 기업 간의 시장 경쟁에서 크게 우려되는 계층은 바로 중소기업이다. 중소기업이 생존하고 지속적으로 발전하기 위해서 기술개발은 필수 불가결하게 되었다. 실제 중소지역으로 갈수록 중소기업에서 소기업으로 갈수록 기술개발 전문 인력의 부족이 심각한 상황이며, 특히 고급 인력의 부족현상은 더욱 심한 것으로 나타나고 있다.

대다수의 중소기업의 경우에는 지식재산 경영을 위한 전담인원을 보유하지 못하거나, 보유하고 있더라도 소수의 1~2명 인원에 의존하는 경우가 대부분이기 때문에 대기업에서 실시하는 형태와 동일한 방법의 지식재산 경영을 실시한다는 것은 불가능한 일이다.

중소기업이 투자한 신규 개발 사업에서 실패할 경우, 중소기업은 투자한 자본이 실패비용이 되어 경제적인 치명타를 입게 되고, 자칫 기업영위가 어려워지는 상황에 처할 수 있다. 그렇다고 중소기업에서 준비한 신규 개발사업이 성공하였다고 해서 기업이 유지될 수 있는 것도 아니다. 중소기업은 우리 경제와 산업의 근간을 형성하는 중요한 자본조직이지만 경제적 약자의 속성을 지니고 있어, 무한경쟁에 방치될 경우 부정적 효과가 산업 전반으로 확산될 우려가 크다 할 것이다.

만약, 신규 기술개발을 뒷받침할 수 있는 지식재산권을 확보하지 못했을 경우 투자한 자본 이외에 경쟁사 등에게 막대한 실시료를 지급해야 하는 재정적 어려움을 겪는 상황에 처할 수 있으므로, 중소기업의 지식재산권 경영은 생존을 위해 반드시 성공적으로 이루어져야 할 중요 과제이다. 그럼에도 불구하고, 지식재산권 경영에 어려움을 겪고 있는 대다수의 중소기

업은 질적으로나 양적으로 전문 인력을 갖추고 있지 못해 대표자 자신이 직접 지식재산권 관련 모든 행정업무를 도맡아 처리해야 하는 현실적인 어려움을 겪고 있다.

특히, 지식재산권 분야는 관련 기술뿐만 아니라, 법적인 요소도 많이 알아야 하는 특화된 업무영역이라 지식재산권 확보 및 이에 관한 지식을 습득하는데 있어 많은 어려움을 겪을 수밖에 없다.

중소기업이나 개인 고객들은 지식재산권(특허, 실용신안, 상표, 디자인, 저작권) 관련해서 지속적으로 업무를 해오던 분야도 아니다 보니 여전히 생소함을 느끼고 있는 것 역시도 현실이다. 중소기업이나 개인 고객들은 기업영위에 도움이 될 만한 좋은 아이디어나 새롭게 출시하고픈 시제품이 아무런 법적보호 장치도 갖추지 않은 채 시장에 그냥 출시하는 경우도 비일비재하게 일어나고 있으며, 여전히 지식재산권 출원등록 상담 또는 지식재산권 과제 컨설팅은 주로 오프라인 상에서 회사 간의 업무 협약이나 일반 개인 고객 위주로 이루어지고 있는 추세이다.

그로 인해, 중소기업이나 개인 고객들은 지식재산권의 중요성에 대한 인식부족으로 인해 대기업이나 중견기업이 미리 선점해 둔 지식재산권의 독점권 행사에 영향을 받아 많은 시간, 자본, 노력이 들어간 시제품을 출시도 못한 채 그대로 폐기처분해야 하는 문제점이 빈번히 일어나고 있는 상황이며, 막대한 인적-물적 손해를 감내해야 하는 문제점을 여전히 안고 있다.¹⁾²⁾³⁾⁴⁾⁵⁾⁶⁾

따라서, 본 논문은 중소기업이나 개인 고객들이 많은 시간, 자본, 노력이

- 1) 서성현, 김승천, 노광현. (2018). 특허회피 설계를 위한 챗봇 및 알고리즘 연구. 『대한전자공학회 학술대회』, 1085-1088.
- 2) 김용석. (2008). “산학협력을 통한 중소기업에로기술 해결에 관한 연구”. 울산대학교 박사학위논문.
- 3) 박우중. (2013). “중소기업의 실효적인 지식재산 경영방안에 대한 연구”. 홍익대학교 석사학위논문.
- 4) 김재국. (2002). 특허권 침해의 유형. 『한국기업법학회』, 11, 249-281.
- 5) 윤병수. (2017). “기계학습을 이용한 핵심 특허문헌 추출시스템에 관한 연구”. 한국해양대학교 박사학위논문.
- 6) 이기은. (2019). “중소기업의 혁신을 위한 적합기술 탐색 방법론의 개발과 응용 연구”. 동국대학교 박사학위논문.

들어가야 하는 시제품을 양산하기에 앞서 시제품과 유사한 기존 제품의 출시여부를 확인하거나 기존 제품과 관련된 지식재산권의 보유여부를 미리 확인할 수 있으며, 대기업이나 중견기업이 미리 확보한 장벽특허의 독점권 행사, 이로부터 형성된 독과점 시장 즉, 불완전 경쟁시장에서 일어나는 각종 폐해로 인해 초래된 막대한 인적-물적 손해를 최대한 줄이는데 도움을 줄 뿐만 아니라, 대기업이나 중견기업이 이미 획득한 지식재산권에 대응할 수 있도록 다양하게 제시된 특허회피 설계안을 토대로 공백특허를 마련하고 이를 발판으로 형성한 미개척 시장을 육성시켜 가는데 도움이 되는 붓 기반 특허회피 설계안 자동 생성 방법을 제안하고자 한다.

Ⅱ. 연구배경 및 필요성

2.1 연구배경

본 논문은 새롭게 개척해야 하는 응용과학 분야 중 하나인 인공지능, 지식재산권 관련 빅데이터 및 이를 받쳐주는 클라우드 서비스 그리고 클라우드 서비스가 지원 가능한 5G 상에서의 통신 시스템 및 통신공학을 이론적 기초로 하며, 인공지능 붓 기반 하에 클라우드망을 통해 수집된 다량의 빅데이터를 맞춤 분석하고 자연어 처리하는데 동원되는 다양한 학문에 속한 과학기술이 서로 융복합해 새로운 학술적 가치를 창출하는 것을 모티브로 삼고 있다.

본 논문에 들어가기에 앞서, 그 이론적 배경이 되는 해당 연구 분야들의 기본 개념들을 좀 더 이해를 높고자 하는 취지에서 간단히 소개하고자 한다.

먼저, 4차 산업을 선도하는데 큰 주축인 5G는 초고속, 초저지연, 초연결이라는 세가지 특징으로 가지며, 주변의 모든 것들을 통신망으로 연결하여 즐겨 찾는 카페, 돈을 지불한 영수증, 자주 연락하는 친구와의 문자나 SNS, 즐겨듣는 음악이나 영화, 날씨에 따라 달라지는 생활 패턴 등 수많은 개인 정보들이 네트워크 신경망을 통해 전달되고 인공지능이 이를 학습하는 과정을 축적하게 될 시, 비로소 맞춤형 개인비서가 탄생하게 된다.

5G 기반 4차 산업과 관련하여 매우 중요하게 대두되고 있는 빅데이터 처리는 맞춤형 개인비서와 더불어 AR, VR, IoT, 자율주행, 스마트팩토리, 스마트시티 기술영역까지 그 범주를 넓혀가고 있으며, 단순히 개인 정보뿐만 아니라 이 세상에 존재하는 모든 데이터가 수집되고 활용되어야만 실현 가능하다 할 것이다. 날씨, 온도, 신호등, 교통량, 차량 움직임, 사람 움직임, 각종 IoT 기기 등 이전과 차원이 다른 종류와 양의 데이터가 필연적으로 발생되고 그에 따른 활용도 및 정보가치는 대용량 네트워크에 올라타는 디바이스들이 서로 연결되어 주체적인 의사 결정하는 형태로 발전

할 것으로 예상된다.⁷⁾

이러한 빅데이터를 적극적으로 활용한 기술인 자연어 처리는 보통 타이핑된 문장의 이해와 생성으로 한정될 순 있겠으나, 이에 국한되는 것은 아니며 [그림 2-1]과 같이, 형태소 분석-구문 분석-의미 분석-대화 분석-문장 생성으로 이루어져 있는 처리방식을 취함에 따라 인간이 사용하는 언어 중 음성, 문자로 표현된 언어를 컴퓨터로 분석하여 일관된 규칙으로 규정하기가 힘든 인간의 말의 독특한 구조와 형태, 의미를 공학적으로 이해하는데 도움을 준다.⁸⁾⁹⁾

또한, 인공지능 기술은 소득수준 향상, 고령화 사회 도래 등의 영향으로 인간의 편의와 안전을 중시하는 인간중시 가치산업으로 부상하고 있다. 특히, 저출산, 고령화 등에 따른 생산인구 감소에 대한 사회적 비용을 감소시킬 수 있는 대안으로 제시되고 있으며, 지능형 로봇, 무인항공기 등의 발전을 통해 인간의 접근이 어려운 위험 지역에서 활용 가능성이 확대되고 있다.

인공지능 기술은 미래 지식정보사회를 이끌어 갈 부가가치 창출의 새로운 원천으로 주목받고 있다. 데이터 관리 및 분석, 비즈니스 의사 결정 등에 활용되어 효율성이 증대되고 있으며 제조업 분야에서는 인간과 분리된 공간에서 주어진 프로그램에 따라 특정 위험/정밀 작업만을 수행하던 로봇이 인공지능 발전으로 인간과 함께 같은 공간에서 협업하는 형태로 발전되고 있다.

인공지능 기술은 금융, 교육, 유통업 등의 서비스 영역에서 일종의 질의 응답·컨설팅 에이전트가 되어 상황에 따라 맞춤형 정보 및 서비스를 제공하며 서비스 지능화를 촉진시키고 있다.

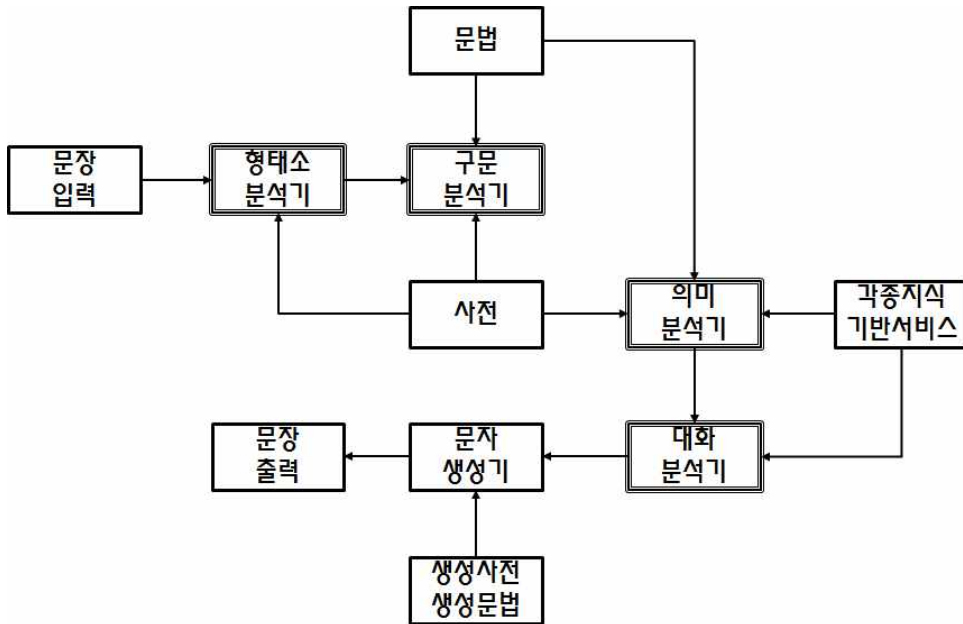
인공지능의 발달로 인해 문제 해결의 범위와 다양성이 확대되면서 인간 지능의 확장 효과로 컴퓨터 과학 등의 발전에도 큰 영향을 끼칠 것으로 예상되며, 인공지능 프로그램을 클라우드와 연결시킴으로 빅데이터와 인터

7) 커넥팅랩. (2019). 『지금 우리에게 5G란 무엇인가』. 서울 : 미래의 창.

8) 강아름. (2012). “온라인 게임 봇 탐지를 위한 사용자 행위 분석”. 고려대학교 석사학위논문.

9) 정소라. (2017). “특허 검색을 통한 인공지능 기반 음성인식 유망 기술 분석”. 고려대학교 석사학위논문.

넷의 효용성을 획기적으로 증대시킬 수도 있다.¹⁰⁾



[그림 2-1] 자연어처리 일반

이와 같이, 수학과 물리학을 근간으로 객관성 및 논리성이 반드시 수반되어야 하는 공학적 발상 및 인문학의 융복합적 사고가 어우러진 응용과학 분야의 다양한 연구는 앞으로도 계속 이루어져야 하는 바, 본 논문은 붓 기반 특허회피 설계안 자동 생성 방법에 관한 연구를 그 주제로 한다.

본 논문 주제와 대비되는 기존 연구에서는 특허청 정보검색 사이트인 키프리스에서 제공하는 특허 정보만을 이용해 신규 아이디어를 도출했기 때문에 시장 트렌드를 고려하지 못한 문제점이 있었다.¹¹⁾

또한, 기존 연구에서는 IPC 분류표를 이용해 유망기술을 도출하거나 F-term을 이용해 융합대상 기술 분야를 선정 및 아이디어를 도출한 것에 그쳐 이 역시 시장 트렌드를 고려하지 못한 문제점이 있었다.¹²⁾

10) 김윤정. (2016). 인공지능 기술 발전이 가져올 미래 사회 변화. 『한국과학기술기획평가원』, 12.

11) 장한수, 황성하, 김유빈. (2015). 특허분석을 통한 신기술의 잠재산업 발굴 - 핵융합 실증로 기획을 중심으로. 『한국기술혁신학회』, 206-211.

특허회피 설계목적은 권리침해 분쟁을 미리 예방하기 위한 일환 중 하나인 것으로, 우선 특허회피 설계를 하기 위해서는 특허청구범위의 정확한 파악이 선행되어야 한다. 대부분의 발명들은 기존의 기술 또는 발명을 개량한 개량발명 또는 상위개념인 기존발명을 구체화한 선택발명이 대부분이라 할 수 있다.

즉, 특허회피 설계(patent avoidance design)는 제 3 자가 기존발명의 구성요소의 일부 또는 전체를 대체, 삭제, 원용, 추가, 도입 등에 의하여 원 특허권자의 권리범위를 회피하면서 새로운 발명을 창작하는 것을 일컫는다.

특허침해를 회피하기 위한 효율적인 방법은 경쟁사와 유사한 제품을 생산하기 이전에 해당 기술에 대한 장벽특허가 있는지 여부를 찾아내는 특허선행조사가 먼저 이루어져야 할 필요가 있다.¹³⁾

특허선행조사는 특허검색 서비스를 활용한 특허 수집 과정을 필수 단계로 요구되며 이때 유효 특허 손실을 줄이기 위해, 변리사와 같은 대부분의 현업 전문가들은 개략적인 형태의 특허 검색식을 사용하여 초기 특허 집합을 구성하고, 주어진 기술과 관련성이 낮은 노이즈 특허를 제거하는 방식을 취하고 있다.

하지만, 이러한 특허 수집 과정은 전문가에 대한 의존도가 높으며, 업무의 대부분이 단순 반복 수작업으로 이루어짐으로, 고급전문 인력이 고도의 전문 작업에 집중적으로 시간을 투자할 수 없게 한다. 실제로 노이즈 특허 제거 과정은 특허 분석 소요시간의 60~80%를 차지하여, 결과적으로 전문 인력에 대한 많은 시간 및 비용 소요를 야기하는 문제점이 있다.¹⁴⁾

계속해서, 특허침해를 분석하는 방법과 특허회피 설계는 매우 유사하다. 차이점이 있다면, 특허침해 분석은 이미 특허회피 설계된 발명이 장벽특허를 침해하고 있는지에 대해 분석하는 것이라 하겠으나, 특허회피 설계는

12) 김방룡, 황성현, (2009). 특허 정보를 활용한 IT 유망기술 도출에 관한 연구. 『한국통신학회』, 34(10), 1021-1030.

13) 김광범. (2002). “화학발명에 관한 회피설계 사례 연구”. 연세대학교 석사학위논문.

14) 김무진. (2016). “특허 기반의 기술 트렌드 모니터링을 위한 의미론적 텍스트 마이닝 접근 방법”. 건국대학교 석사학위논문.

시제품 출시 전 즉, 아직 실시되지 않아 침해가 형성되어 있지 않은 발명에 대해 미리 장벽특허가 가진 권리를 벗어날 수 있는 방안을 찾아낸다는 점을 들 수 있다.

특허회피 전략수립은 특허 검색 결과를 대상으로 기업 활동 특성을 파악하거나 시장 트렌드에 맞게 잠재 산업(공백기술, 틈새시장)을 발굴하는 과정에서 필수적으로 이행되어야 하는 것 중 하나이다.

기존의 특허회피 설계안을 마련하는 방법에서는 유사 기술 분야에 속한 선행특허의 아이디어를 사업 분류에 따라 차용한 다음, 관련 기술에 대한 유효특허를 도출하는 작업이 반드시 선행되어야 했기 때문에 많은 시간과 노력이 필요한 문제점이 있었다.

더욱이, 연구자 간의 브레인스토밍(Brainstorming), 트리즈(TRIZ : The -ory of solving inventive problem)을 활용해 관련 기술에 대한 유효특허를 도출하고자 할 경우 솔루션 도출에 대한 성공 확률이 비교적 낮았다.

일부이긴 하나, 특허 전문가와 트리즈 전문가는 해당기술 분야에 대한 전문정보다는 경험이나 직관에 의해 결과물을 산출하는 경우가 많아 좋은 결과물을 도출해 내는데 까지 비효율적인 시행착오를 많이 겪어야 했다.¹⁵⁾

다시 말해, 브레인스토밍에서 나온 아이디어는 곧바로 좋은 아이디어로 평가될 수 있는 것이 아님으로 분류, 수정, 보강의 작업을 거쳐 보다 나은 아이디어로 유용화, 현실화되는 개략평가 단계를 밟게 되며, 냉정한 판단으로 평가받고 실행계획으로 결정되기까지 많은 시간과 노력이 요구된다.¹⁶⁾

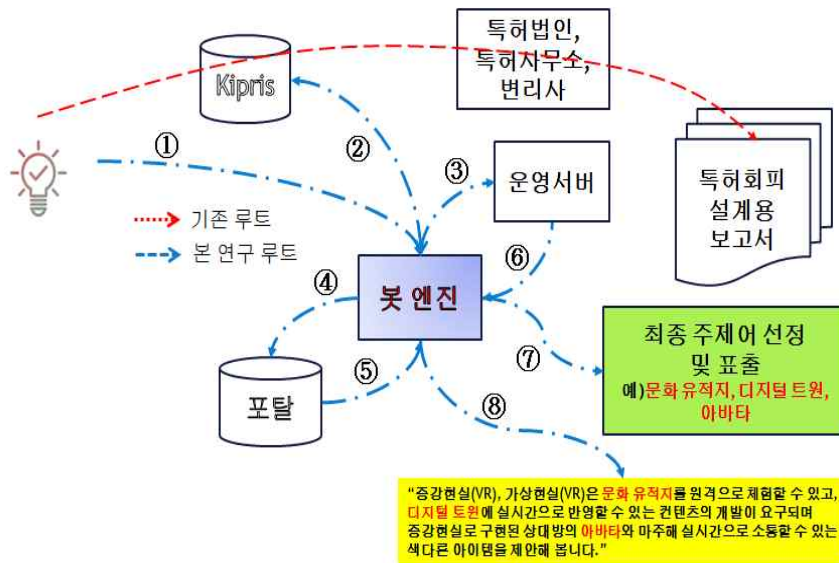
이를 해결하기 위한 목적으로, 본 논문의 붓 기반 특허회피 설계안 자동 생성 방법은 중소기업이나 개인 고객들이 제안한 뛰어난 아이디어, 신제품 또는 회사 내 업무와도 연관된 발명이 상대방의 지식재산권에 의해 인적, 물적 피해를 보지 않도록 특허회피 설계안을 붓을 통해 제시해 주는 방법에 관한 연구이다.

또한, 본 논문은 포털 사이트에서 제공하는 다양한 정보를 중 시장 트렌

15) 최옥희, 유명식. (2017). 특허 포토폴리오 구축을 위한 응용 제품에 대한 신규 아이디어 도출 알고리즘 연구. 『한국통신학회 학술대회』, 455-456.

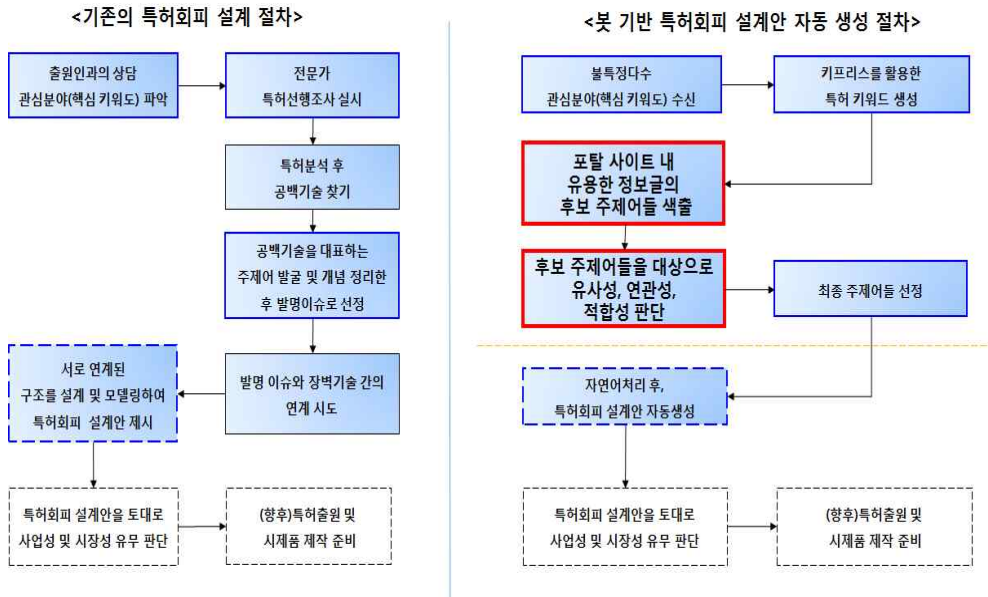
16) 박윤희. (2004). 『돈 버는 아이디어 엿보기』. 서울 : 한국생산성본부.

드에 부합함과 동시에 공백 기술로도 채택할 만하는 정보글만을 색출해 이를 특허회피 설계안을 자동 생성하는 자원(ex : 빅데이터)으로 활용한다는 내용을 담고 있으며, 봇을 통해 자동 생성된 특허회피 설계안을 기업이나 개인 고객들이 충분히 이해할 수 있는 자국어, 영어로 자연어 처리해 외부 표출하는 것을 최종 목표로 삼는다.



[그림 2-2] 봇 기반 특허회피 설계안 자동 생성 절차 개요

봇 기반 특허회피 설계안 자동 생성 절차는 [그림 2-2~3]로부터 알 수 있듯이, ①키프리스에 핵심 키워드 수신 ②다수의 특허건 조회 ③특허의 특허청구범위 크롤링 ④특허 키워드 생성 및 포털 사이트 내 정보글 검색 ⑤유용한 정보글 추출 및 그로부터 후보 주제어 색출 ⑥후보 주제어를 대상으로 1-2-3 필터링(유사성, 연관성, 적합성 판단) 및 ⑦최종 주제어 선정 ⑧자연어처리(NLP) 후 특허회피 설계안 자동 생성하는 절차적 프로세스를 실시한다.



[그림 2-3] 기존의 특허회피 설계 절차와의 비교

2.2 특허회피 설계 소개

2.2.1 특허회피 설계의 정의

특허회피 설계는 상대방의 특허권의 권리범위를 검토한 후 그 권리범위가 갖는 독점권을 벗어날 수 있도록 나의 권리범위를 개량 및 고찰해 향후 양산하고자 하는 시제품이 상대방의 특허권에 위배됨 없이 원활하게 시장에 출시될 수 있도록 나의 권리범위를 새롭게 설계하는 것을 말한다.

특허는 기술적 사상(아이디어 또는 제품)이 발명의 대상이지만, 실제 특허권으로 보호받는 것은 기술적 사상 자체가 아니라 “기술적 사상을 구현한 구체적인 구조”이다. 따라서, 상대특허와 기술적 사상은 동일하지만 그것을 구현하기 위한 구체적인 구조가 다를 경우 상대특허를 피할 수 있다.

특허침해에 대하여 일반적인 원칙을 적용하는 법원은 아래와 같은 조건을 만족시키면 침해를 인정하고 있지 않다.

첫째, 특허청구범위에 기재된 어느 구성요소 또는 한정범위가 장벽특허로부터 완전히 제거될 수 있어야 한다.(예를 들면, 한정범위가 장벽특허에서 찾아볼 수 없거나 각 구성요소 및 그 구성요소와 연관된 기능이 완전히 제거되거나 제거될 수 있는 것)

둘째, 특허회피 설계된 발명이 장벽특허 상의 특허청구범위의 한정 또는 각 구성요소와 그와 연관된 기능이 완전히 제거될 수 없다면, 장벽특허 내 권리범위로 설정된 구성요소를 벗어나 실질적으로 다른 방법(substantially different way or a physically different element)에 의하여 실행되는 다른 구성요소로 치환되거나 치환될 수 있는 것이어야 함을 판시하고 있다.¹⁷⁾

따라서, 특허회피 설계를 하기 위해서는 우선 회피하고자 하는 장벽특허의 특허청구범위를 살펴보고 그 권리범위와 의미를 분석한다. 그리고 특허청구범위를 벗어나거나 무효화할 수 있는 선행기술이 있는지 조사한 뒤 구성요소 완비의 원칙과 균등론에 기초해 비침해성을 확보할 수 있는지를 판단하는 것이 중요하다.

즉, 장벽특허를 대상으로 특허회피 설계하기 위한 기본적인 실무 프로세스는 다음과 같다.

첫째, 장벽특허 즉, 특허회피 설계하고자 하는 특허청구범위의 권리범위와 특허청구범위의 법리적 해석을 먼저 이해 분석한다.

둘째, 장벽특허가 속한 기술 분야를 대상으로 선행기술을 조사한다. 선행기술 조사 시, 시장경쟁의 경중 및 특허의 기술적 고도성 및 그 기술의 기술발전 추이(推移) 등을 따져 요구되는 만큼 넓게 시행한다.

특히, 특허맵(Patent Map)을 작성해 장벽특허가 속한 기술 분야의 위치를 파악할 필요가 있다. 이때, 특허권자가 특허청구범위에 포함되는 공지기술에까지 권리범위를 확장할 수 없도록 하기 위한 침해분석과 선행기술로 인하여 장벽특허를 무효시킬 수 있는지에 대한 유효성 검토도 이루어져야 할 것이다.

셋째, 장벽특허 내 기재된 광의(廣義)의 독립항의 구성요소들과 그 법적 실효성을 내포하고 있는 한정범위를 어떻게 회피할 수 있는지를 결정한다.

17) 장인규. (2016). “허가특허연계제도 시행 이후 제네릭 의약품 특허회피전략 분석”. 성균관대학교 석사학위논문.

즉, “장벽특허를 어떻게 하면 회피할 수 있을까?”에 대한 안(案)을 제시하는 특허회피 설계에 관한 방법론은 장벽특허 내 독립항의 구성요소들을 완전히 배제시키거나 또는 특허회피 설계안을 토대로 새롭게 설정한 구성요소들을 재구성, 재배치 형태로 포진시키는 것을 의미한다.

넷째, 법리적 회피와 기존 각 국가의 특허 침해 관련 중요 판례의 태도를 고려해 특허회피 설계안을 완성하고, 이 특허회피 설계안을 평가해 특허성이 있을시 발명의 출원전략에 따른 실질적인 출원과 등록 단계를 실시한다.¹⁸⁾

특허회피 설계의 궁극적인 목적은 타인의 특허권을 회피하는 것을 넘어 특허회피 설계안으로 특허권을 확보하고, 기업의 향후 외부의 특허 침해 소송 및 공격에 대비해 회피전략 내지 그 방지책을 마련함으로써 현실적 실익을 고려하기 위한 것이다.

좀 더 구체적으로, 특허회피 설계론에 대한 좀 더 자세한 사항은 이하와 같다.

[그림 2-4]로부터 알 수 있는 바와 같이, 발명의 기술적 사상을 보호받기 위해 발명에 관한 장치적 구성, 방법적 구성을 특허청구범위에 기재함에 따라 발명핵 A가 생성되며, 발명핵 A를 보호하는 가상의 울타리라 할 수 있는 발명의 동일성 범위가 특허법 중 균등론에 입각한 논리적 근거에 합당하도록 발명핵 A를 중심으로 그 주변에 설정된다.

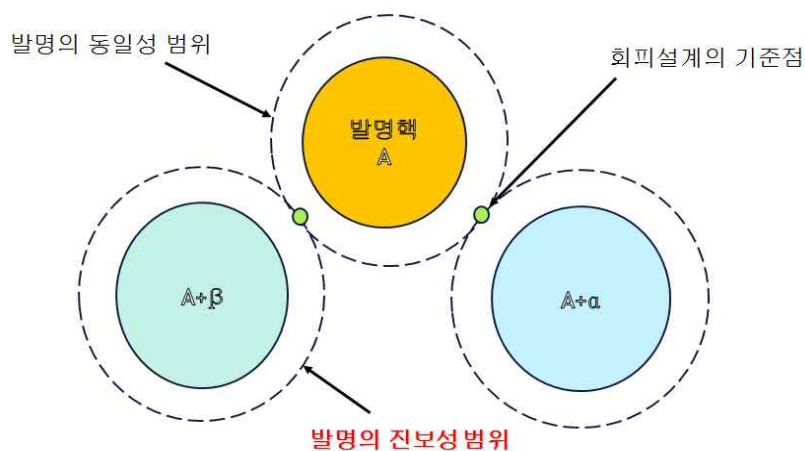
특허회피 설계전략은 이렇게 설정된 발명의 동일성 범위를 벗어나도록 α 라는 아이디어 또는 β 라는 아이디어를 발명핵 A와 융합하여 $A+\alpha$ 또는 $A+\beta$ 의 독자적인 영역이자 새로운 개념이 도입된 개량 발명을 생성케 함과 아울러 기존 발명핵 A가 갖는 발명의 동일성 범위를 벗어나 발명의 진보성 범위가 $A+\alpha$ 또는 $A+\beta$ 형태로 새롭게 조성될 수 있도록 특허 전략을 세우는 것을 말한다.

이때, 발명의 동일성 범위와 발명의 진보성 범위가 만나게 되는 교차점이 바로 회피설계의 기준점이라 할 수 있으며, 만약, α 아이디어와 β 아이디어가 발명의 동일성 범위에 속하게 되면 $A+\alpha$ 또는 $A+\beta$ 는 진보성이

18) 이지현. (2014). “특허회피설계 방법론과 그에 따른 발명의 특허성 확보 방안에 관한 연구”. 연세대학교 석사학위논문.

상실된 것으로 간주되나, α 아이디어와 β 아이디어가 발명의 진보성 범위에 속하게 되면 $A+\alpha$ 또는 $A+\beta$ 는 진보성이 인정된 것으로 간주되어 올바른 특허회피 설계전략이 세워졌다고 충분히 이해할 수 있다.

균등론에 입각해 특허청구범위를 평가하는 것은 기술의 가치를 현실적으로 평가하고 틀린 점을 회피하기 위해 유익하다. 특허청구범위를 대상으로 한 회피설계는 다음과 같은 체크 리스트에 따라 검토되는 것이 전형적이다. 첫째, 청구항의 구성요소 내 기술적 혹은 경제적인 관점에서 볼 때 필수요소가 아닌 것은 어떤 요소이며 청구된 구성요소 중 대신할만한 대체재는 어떠한 것인지를 결정한다. 둘째, 청구된 구성요소의 대체재로 가능한 것이 균등론적 범위 내 속한다고 주장할지를 정하기 위한 포대 내 자료를 면밀히 조사한다. 셋째, 선행기술에 의해 특허권자가 균등론을 주장하는 것이 불가능하게 되는 범위를 정한다. 넷째, 침해를 회피하거나 보다 중요한 기본 특허의 가치를 감소시키기 위한 개량이 가능한 범위를 검토한다. 특허회피 설계는 특허출원 전 또는 시제품 제작 전에 행해짐으로 장벽특허를 벗어나기 위한 나만의 권리범위를 필요이상 한정시키려는 우려로부터 벗어나 향후 대체재로도 채택 가능할 정도로 구성요소를 추가시킨 특허명세서 작성 및 개량된 시제품을 시장에 출시하는데 충분히 만족할 수 있도록 한다.¹⁹⁾



[그림 2-4] 특허회피 설계론 소개

19) 윤여선. (2007). “미국특허회피 설계 연구”. 연세대학교 석사학위논문.

원천특허는 기존기술이 없는 특허이거나 기존기술과 크게 차이가 있어 비교될 수 없는 특허를 말한다. 표준특허는 어떠한 표준을 사용하고자 할 때 피해갈 수 없는 기준특허를 말한다.²⁰⁾

개량특허는 원천특허, 표준특허를 적극 소화한 개량기술 또는 주변기술을 특허로 연결시켜 놓은 발명을 말한다. 개량특허는 특허전략상 선진기업들과 특허 분쟁이 발생할 경우에 크로스라이센스를 통한 로열티 부분을 절감할 수 있고, 특히 첨단기술 분야에 있어 심각한 기피현상을 타개할 수 있는 유용한 자원으로 사용할 수 있다.

특허회피 설계는 원천특허나 표준특허가 보유한 침해요소를 피해 향후 시제품에 반영될 수 있도록 구체적으로 다른 개량적 구조를 고안해 내는 것이다. 특허회피 설계를 이해하기 위한 전제조건으로는 발명의 핵을 알아야 하고 특허요건 중 동일한 판단 및 진보성 판단의 사고방법, 그리고 이들의 특허청구범위가 어떻게 해석되어 질 수 있는가를 이해하지 않으면 확신을 가질 수 없다 할 것이다.²¹⁾

앞에서도 언급한 바와 같이, 기본발명을 벗어날 수 있는 영역, 즉 특허회피 설계안을 제시할 만한 영역은 발명의 동일한 범위를 벗어나 최소한 발명의 진보성 범위에 있어야만 가능하다. 발명의 동일한 범위와 발명의 진보성 범위가 만나는 이 경계점이 바로 특허회피 설계안을 자동 생성하는 기준점이 된다.

일례로, IT 기반의 공유 플랫폼 서비스와 동일 유사한 서비스가 생긴 경우, 先 출원인(발명자)은 자신이 등록받은 특허권을 가지고 특허 침해를 주장할 것이다. 이때, 後 출원인(발명자)이 先 등록된 특허를 회피하는 방법은 이 등록특허를 구성하는 요소를 일부 생략하거나 이 등록특허를 구성하는 요소를 다른 구성요소로 대체하는 경우가 일반적이다.

대기업이나 중견기업은 특허회피 설계를 진행하는 것을 원칙으로 인지하고 있다. 일례로, 팬츠는 기술 아이템을 보유하고 있어 이를 사업화하는 방향으로 이어가고자 하나, 선행조사 결과 장벽특허가 존재해 특허침해와

20) 조경선, 임재용. (2014). 『지식재산금융과 가치평가실무』. 서울 : 한국금융연수원.

21) 주종호. (1998). 『미국 특허출원서 쓰는 방법』. 대전 : 한지재권 연구회.

같은 장애요소가 발생한 경우, 대기업이나 중견기업은 장벽특허를 무효화시킬 수 있는 잠재성을 지닌 선행 특허들을 찾아내기 위한 무효화 전략, 특허회피 설계, 비침해 논리개발과 같은 다양한 대응전략을 수립하고자 노력한다.²²⁾²³⁾²⁴⁾²⁵⁾²⁶⁾

또한, 대기업이나 중견기업은 시장에서 인기 있는 특허상품이 있다면 특허침해가 되지 않는 범위에서 인기 상품과 유사한 상품을 만들기 위해 노력의 일환으로, 무효화 전략, 특허회피 설계, 비침해 논리개발과 같은 다양한 대응전략을 수립한다. 대기업이나 중견기업은 타인의 특허회피 설계를 사전에 막고자 “발명 아이디어를 구현하는 다른 몇 가지 구조”를 특허로 출원하여 타인의 진입을 애초부터 막기 위한 특허 그물망을 계속적으로 치고 있다.

실제 생산하고자 하는 상품에 대한 특허는 1건이지만 이와 유사한 특허를 여러 건 등록받아 타인의 진입을 사전에 막을 수 있다.

특허권의 보호범위를 해석할 때, 구성요소완비의 원칙(All elements rule)과 균등론이 적용된다. 구성요소완비의 원칙(All elements rule)은 타인이 특허청구범위에 기재된 구성요소 전부를 실시하는 경우에 특허침해가 인정되며, 특허청구범위의 구성요소 전부를 포함하고 다른 구성요소를 더 부가한 경우에도 침해가 인정되지만(‘구성요소 부가’), 특허청구범위의 구성요소의 일부를 생략한 경우에는 침해가 인정되지 않는다(‘구성요소 생략’). <대법원 2000. 11. 14. 선고 98후2351 판결, 2001. 6. 1. 선고 98후 2856 판결 참조>

균등론은 특허청구범위를 해석해 발명의 보호범위를 정함에 있어 단순히 특허청구범위의 문언기재에 의한 문리해석의 범위를 넘어 문언과 균등 내

22) 정병기, 고남욱, 경진영, 최덕용, 윤장혁. (2017). 장벽특허의 무효화 분석을 위한 선행특허 기술 발굴시스템 개발. 『엔트루정보기술저널』, 16(1), 97-108.

23) 김태영. (2009). 『인터넷 판매자가 꼭 알아야 할 상표권과 상표권침해』. 경기 : 도서출판 비비컴.

24) 박현우, 성태웅, 변정은, 유연우. (2018). 『기술전략과 R&D 기획』. 서울 : 한성대학교출판부.

25) 남호현. (2019). 『태양 아래 모든 것이 특허대상이다』. 서울 : 도서출판 예가.

26) 김동현. (2015). 『팔리는 상품 끌리는 브랜드』. 서울 : 매일경제신문사.

지 등가의 발명도 권리보호범위에 속한다고 하는 침해소송의 경험칙으로부터 발전하여 온 이론이다.

다시 말해, 균등론은 원고의 특허발명과 피고의 실시형태를 비교하여 피고의 실시형태가 원고 특허발명의 실시예로서 명세서에 직접 기재되어 있지 아니하고 특허청구범위의 문언기재와도 반드시 일치하는 것은 아니지만, 그것이 원고의 특허발명에 개시된 내용으로부터 평균적 전문가가 쉽게 치환하여 동일한 효과를 얻을 수 있는 것일 때에는 이는 원고 특허발명의 기술사상과 균등한 것으로서 원고 특허발명의 기술적 범위에 포함된다고 보는 공평의 원리에 입각한 이론이다²⁷⁾²⁸⁾.

즉, 균등론은 “침해대상 제품 등에서 특허발명의 특허청구범위에 기재된 구성 중 치환 내지 변경된 부분이 있는 경우에도, 특허발명과 과제의 해결 원리가 동일하고, 그러한 치환에 의하더라도 특허발명에서와 같은 목적을 달성할 수 있고 실질적으로 동일한 작용효과를 나타내며, 그와 같이 치환하는 것이 발명이 속하는 기술 분야에서 통상의 지식을 가진 자(당업자)라면 누구나 용이하게 생각해 낼 수 있는 정도로 자명하다면 특별한 사정이 없는 한, 침해대상 제품 등은 전체적으로 특허발명의 특허청구범위에 기재된 구성과 균등한 것으로서 여전히 특허발명의 특허권을 침해한다고 보아야 한다(대법원 2011.09.29. 선고2010다65818 판결)”고 판시한 판례로부터 이해할 수 있다.

부연 설명으로, 상대방이 침해할 가능성이 높은 특허는 곧 기술가치가 높은 특허라 할 수 있으므로, 특허의 침해 가능성은 그 보호대상인 발명의 기술적 유용성과 그 보호 범위를 기재한 특허 명세서의 법적 강도에 의해 결정할 수 있다. 발명의 기술적 유용성은 기술 과제의 시장 적합성과 기술 구성의 산업적 핵심성에 의해 평가될 수 있으며, 특허 명세서의 법적 강도는 상세한 설명의 충실성과 특허청구범위의 견고성에 의해 평가될 수 있다.²⁹⁾

27) 임병웅. (2007). 『특허법 제6판』. 서울 : 한빛지적소유권센터.

28) 이시윤. (2014). 『신민사소송법 제8판』. 경기 : 도서출판 박영사.

29) 오병석. (2009). 『특허가치전략』. 서울 : 도서출판 페이퍼하우스.

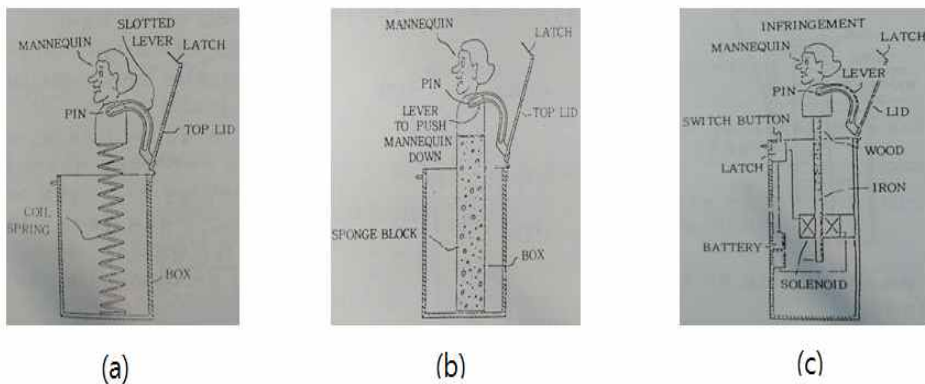
2.2.2 특허회피 설계의 기준조건 및 사례

다음은 특허회피 설계론을 좀 더 이해하는데 도움을 주는 사례이다.

[그림 2-5]에서 보여주는 예시는 어린이를 위한 놀이기구(놀람상자)에 대한 서로 다른 발명들에 관한 것으로, 특히 침해와 특허 비침해의 경우를 함께 파악할 수 있는 좋은 사례이다.

좌측에 보이는 도면(a)의 특허청구범위는 “어린이를 위한 놀이기구에 있어서, 래치를 가지며 힌지작동 가능한 상부뚜껑을 가진 상자, 코일 스프링, 상자의 하부에 설치된 코일스프링의 상부에 놓이는 마네킹, 마네킹에 의해 움직이는 핀을 수용하는 슬롯을 구비하며, 그의 일단이 상자뚜껑의 내측에 고정되는 레버로 이루어져서, 뚜껑이 닫힌 때는 레버가 마네킹을 상자로 밀고 뚜껑을 닫는데 필요한 에너지를 코일 스프링에 저장하고 있다가 뚜껑을 열 때 마네킹을 밀어 올리는 놀이기구”로 권리범위를 형성하고 있다.

발명의 명칭: 어린이를 위한 놀이기구(놀람상자)



[그림 2-5] 특허회피 설계론 사례

도면(a)와 도면(b)의 차이점은 도면(a)의 특허청구범위 중 마네킹 하부와 ‘코일 스프링’의 상부를 연결시켜 놔다는 점 대비 도면(b)의 특허청구범위는 코일 스프링 대신 ‘스폰지 블록’을 채택했다는 점에서 서로 상반된다고 주장하는 경우이다.

하나, 현행 특허법에서는 실질적으로 동일 기능인지, 실질적으로 동일 방법인지, 실질적으로 동일 결과를 내는지를 기준으로 특허침해 여부를 객관적으로 판단한다.

‘코일 스프링’ 대비 ‘스폰지 블록’이 현행 특허법에 비춰 특허침해 여부를 판단하는 기준에 부합하는지를 판단해 본 결과, ‘스폰지 블록’은 ‘코일 스프링’ 대신 기능적으로 대체가 가능하며 동일 방법으로 동작되고, 실질적으로 마네킹을 푸쉬업해 주는 기능까지 동일하다고 판단됨에 따라 특허침해인 것으로 결론내릴 수 있다고 보았다.

또한, 도면(b) 대비 도면(c)는 기능적 측면, 동작적 측면에서 봤을 때 로드-스위치-솔레노이드 부품들도 마네킹을 상자의 하부에서 상자의 상부위치로 밀어올림은 동일한 기능을 수행하고, 뚜껑을 열 때, 마네킹이 튀어오름 동작 역시도 동일한 동작을 수행함으로 특허침해인 것으로 보았다.

그러나, 도면(c)의 특허청구범위는 로드-스위치 부품은 스프링에 대항하며, 상자를 닫는데 필요한 힘을 저장에너지로 변환시킬 수 없다는 점을 강조하고 있다. 기존 기술인 도면(b)와 기능적 측면, 동작적 측면, 그리고 그 결과까지 동일한 기능, 동일한 동작, 동일한 결과를 가져온다고 볼 수 없다고 한 것이다. 뚜껑래치를 열 때, 스위치 버튼은 미끄러지면 상측으로 이동한다. 그 후 스위치는 솔레노이드를 통해 전기회로를 닫고, 철-나무로드의 철부위를 자성적으로 상부로 끌어당겨 마네킹을 밀어 올리는 특징을 가진다고 주장한 것이다.

또한, 상자를 닫으면, 흡인의 자성력에 대항하며 뚜껑을 밀면 스위치 버튼을 하강시켜 전기회로를 끊으면 솔레노이드가 작동하지 않게 됨으로써, 상자를 닫을 때에는 뚜껑을 미치는 힘이 없게 된다는 특징을 가진다고 주장한 것이다.

따라서, 도면(c)의 발명은 현행 특허법에 비춰 볼 때, 도면(b)으로부터 인지되는 특허와 대비해 ‘비침해’ 하다고 주장한 바를 수용한 것으로 두 발명은 서로 각기 다른 발명이라 정의내릴 수 있어, 이는 특허회피 설계를 성공적으로 보여준 좋은 사례라 할 것이다.³⁰⁾

2.3 연구의 필요성

일반적으로, SWOT 분석은 외부환경의 기회요인(Opportunity)과 위협요인(Threat)을 파악하고 내부 환경의 강점(Strength)과 약점(Weakness)을 분석해 전략적 옵션을 찾는 도구로서, 기업의 강점요인을 활용하여 새로운 사업기회를 포착하거나 약점을 최소화하여 위협요인에 대처하는 전략을 모색하는데 활용된다.³¹⁾

봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구에 대한 필요성은 [표 2-1]의 SWOT 분석을 통해 고찰된다.

기존의 특허회피 설계안을 제안했거나 제안받았던 방식은 새로운 틈새시장을 개척하기 위한 목적으로 공백특허를 확보하고자 하는 고객이 특허청, 법무법인, 특허법인, 특허사무소, 대학교의 산학협력단에서 종사하는 전문가들(변호사, 변리사, 교수, 박사, 기술사 등)에게 찾아가 1:1 오프라인으로 상담하는 방식을 취해 왔다.

SWOT 분석의 WT(Weakness and Threat)를 살펴보면, 고객이 고학력자로 구성된 특허 전문가들에게 의뢰하다 보니 아무래도 적잖은 비용적인 부담이 일어날 수밖에 없었고, 특허 전문가가 고객이 원하는 특허회피 설계안을 제시하기 위해 준비해야 하는 다양한 사전 조사를 하기 위해 물리적으로나 물질적으로 소요될 수밖에 없는 시간적인 투자와 노력이 반드시 요구되었다.

그러나, 특허 전문가라 할지라도 미래에 대한 확실성이 결여된 상태에서 특허회피 설계안을 제시하는 경우가 종종 있어 왔기 때문에, 이에 따른 기술성 분석, 시장성 분석 및 사업성 분석에 대한 그릇된 판단이나 잘못된 오류가 기업이 막대한 자본을 투자 했음에도 불구하고 잘못 양산된 시제품의 매출 저하, 그로 인한 경영난 악화 및 계속 이어지는 여파로 향후 기업에게 가져올 경제성 부담에 대한 우려를 지적한다.

30) 박대진. (1996). 『특허청구범위의 작성과 회피설계』. 서울 : 한빛지적소유권센터.

31) 나도성. (2017). 『한국 컨설팅시장의 이론과 실제3』. 서울 : 한성대학교출판부.

[표 2-1] 기존의 특허회피 설계안 관점에서 본 SWOT 분석

	강점(Strength)	약점(Weakness)
기회(Opportunity)	-틈새시장 개척으로 -시장선도 -공백특허 획득 -권리보호	공백특허 획득 외에 특허회피 설계할 필요가 없는 원천기술을 확보하는데 힘씀
위험(Threat)	공백특허를 많이 확보하여 타 업체에 기술이전 및 양도하여 매출증대를 향상시킴	-비용부담 -장기적인 시간 투자 必 -미래에 대한 불확실성 결여 -시장성, 사업성 분석 오류로 양산된 제품의 매출 저하로 인한 경영난 악화우려

본 논문의 봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구는 위에서 언급한 WT의 문제점을 해결할 수 있는 대처 방안을 다음과 같이 제시할 수 있다.

첫째, 특허회피 설계안을 제시하는 기존의 오프라인 방식에서 벗어나 온라인 방식으로 전환해 빠른 결과를 도출해 낼 수 있다. 둘째, 고비용(대리인 수수료)에서 저비용(봇 사용료)으로 특허회피 전략 서비스를 제공하는 주체를 달리 적용시켜 인건비로 인해 부담되었던 비용 문제를 해결해 줄 수 있다. 셋째, 봇을 통한 빠른 일처리로, 오프라인 상에서 시간이 많이 소요될 수밖에 없었던 물리적인 한계를 해결해 줄 수 있다.

넷째, 봇이 하나의 관심분야에 대해 각기 다른 특허회피 전략을 여러 번에 걸쳐 제시해 줄 수 있어, 고객이 다양하게 선택할 수 있는 기회를 제공해 줌으로 통해 미래에 대한 확실성을 갖도록 도와줄 수 있다. 다섯째, 봇의 지속적인 딥러닝(강화학습, 비강화학습)으로 잘못된 오류가 발생되지 않도록 업그레이드해 줌으로써, 봇이 제시하는 특허회피 설계안에 대한 신뢰성 구축 및 예측 가능한 시장 모멘텀에 따른 수익성 향상을 기대해 볼 수 있다.

2.4 기존의 특허회피 설계방식의 한계

기존의 특허회피 설계방식은 개별적인 선행특허들의 특허청구범위를 기준으로 구성요소를 부가, 치환, 변경, 생략하는 방식으로 특허회피 설계를 제안해 왔다. 하지만, 이는 다음과 같은 한계점을 안고 있다.

첫째, 기술 분야가 이미 정해진 어느 특정한 특허권에 대하여 특허회피 설계를 제안하는 정도에 그친다. 일례로, 전기자동차용 무선충전기술 분야에서 전기적 검출 방식의 특허기술이나 광학적 검출 방식의 특허기술 하나를 대상으로 특허회피 설계한다면, 전기적 검출 방식 혹은 광학적 검출 방식이라는 틀을 벗어나기가 좀처럼 쉽지 않다. 결국, 개별 특허기술의 권리범위를 벗어나기 위한 제한된 특허회피 설계하는 것에 머무를 수밖에 없다.

둘째, 해당 기술 분야의 기술특성 간 상호관계 혹은 결합관계를 통해 새롭고 진보된 결합기술을 도출하지 못한다는 한계가 있다.

즉, 금속 물체의 검출에 유리한 전기적 검출 방식과, 생물체 검출에 유리한 광학적 검출 방식을 결합하여 두 검출 방식의 상호 보완을 통해 금속 물체와 생물체를 동시에 검출할 수 있는 방식을 개발할 필요가 있는데 이에 대한 융복합적 연구가 미흡하다는 한계가 있다.

셋째, 해당 기술 분야에 속한 전체 특허기술을 모두 고려하지 못하므로 새로운 기술을 찾아내는데 한계가 있다. 이에 따라 해당 기술 분야에 제안된 특허들을 조직적이고도 체계적으로 연구 개발함에 신뢰도가 낮다.

넷째, 전체 특허기술을 대상으로 회피 설계할 순 없기 때문에 특허회피 설계하고자 하는 대상기술 간의 경제적 타당성을 고려해 볼 수가 없다.

다섯째, 물리적인 여건 조성의 미약으로, 연구 개발자들이 해당 기술 분야에 공개되거나 등록된 전체 특허들을 분석하여 다양한 기술의 결합을 통해 새로운 아이디어를 도출해 낼 수 있는 기회를 제공하지 못한다.

여섯째, 장벽특허들이 다수 개인 경우, 특허회피 설계의 개발 과정을 반복해야 한다. 따라서, 하나의 특허회피 대상특허의 권리범위를 벗어났다 하더라도, 다른 특허회피 대상특허의 권리범위에 속하게 될 가능성이 높

고, 이러한 경우, 특허회피 설계를 처음부터 다시 반복해야 하는 문제가 있다.

한편, 특허지도를 이용하여 특허를 포괄적이면서 전략적 접근을 시도하는 방법들도 있다. 예컨대, 특허청구범위나 특허지도를 작성하여 분석함으로써 과제 선정 또는 제품개발에 이용할 수 있다. 그러나, 이 역시 해당 기술분야와 관련된 특허청구범위를 나열하여 검토하는 수준에 머무를 뿐, 기술적으로 상호 보완성이 강한 기술들을 입체적으로 결합시켜 분석하지는 못한다.

이에 특허회피 설계를 제시하는 경우에 좀 더 체계적이고 조직적인 기술개발 전략을 접목시킬 필요가 있다.³²⁾

2.5 봇 소개 및 현황

봇(Bot)은 데이터를 찾아주는 소프트웨어 도구. 인터넷 웹 사이트를 방문하고 요청한 정보를 검색, 저장, 관리하는 에이전트의 역할을 하며 다른 프로그램 또는 사람의 행동을 흉내 내는 대리자로 동작하는 프로그램을 의미하기도 한다. 인터넷상에서 가장 보편적으로 존재하는 봇들은 스파이더, 크롤러라고도 불리는 프로그램들로서, 웹 사이트들에 주기적으로 방문하여 검색 엔진의 색인을 위한 콘텐츠를 모아오는 일을 한다.

현재 봇 종류에 관한 명확한 구분은 없으나 명령방식에 따라 채팅, 음성 인식, 개인비서로 구분되는 다중 서비스와 데이터를 적시에 찾아주는 검색 서비스를 기업과 개인 고객들에게 동시에 제공해 주는 역할을 한다. 봇의 대표적 사례 중 하나는 보이지 않는 곳에서 우리가 일상적으로 경험하고 있는 바로 검색 엔진용 봇이다.³³⁾

검색 엔진은 빠른 검색 결과를 제공해 주기 위해 사전에 크롤링(Crawling) 및 색인 생성(Indexing)이라는 두 단계를 거친다. 크롤링은 검

32) 정지원. (2016). “QFD 기법을 활용한 전기자동차 무선충전기의 이물질 검출 기술에 관한 회피설계 전략”. 성균관대학교 박사학위논문.

33) 장준희. (2016). 『인공지능 기반의 ‘챗봇(ChatBot)’ 서비스 등장과 발전 동향』. 대구 : NIA 한국정보화진흥원.

색 엔진의 문서 수집기(Crawler)가 공개 웹페이지에 접속하는 순간을 말하며, 색인 생성은 검색결과에 사용할 수 있도록 페이지에 대한 정보를 수집하는 것을 말한다. 이 두 기능은 모두 봇에 의해 자동으로 수행된다. 이렇듯 문서를 수집하고 색인하면 이용자가 입력한 검색어와 관련된 문서들을 분석하는 매칭(Matching) 작업과 이 매칭된 문서들을 분석해 순위화(Ranking)하는 작업이 이어진다.

구글, 네이버, 다음카카오, Bing(Bing) 등 다양한 검색 서비스들은 각자 보유한 검색 엔진을 통해 검색 결과를 제공하고 있지만, 문서수집, 색인, 매칭, 순위화 등 검색 엔진의 기본적인 작동 원리는 비슷하다고 할 수 있다. 이들 회사들은 각자 추측해 놓은 자신만의 고유한 검색엔진 알고리즘을 적용하기 때문에 같은 검색어에 대한 결과가 각각의 검색 서비스마다 다를 수 있다.

검색 엔진용 봇 뿐만 아니라 이용자 맞춤형으로 이루어지고 있는 뉴스 추천, 자동 기사 생성, 자동 뉴스 배열 등 미디어 영역에는 이미 많은 봇들이 작동하고 있다. 눈에 보이지는 않지만 소프트웨어를 통해 자동으로 수행되고 있는 많은 미디어 기능들은 미디어용 봇에서 수행된다. 미디어용 봇은 이미 검색 결과 상단에 위치하기 위해, 눈에 띄는 뉴스 배열의 위치를 차지하기 위해서 눈에 보이지 않을 뿐 많은 작동을 하고 있으며, 우리의 행위에 많은 영향을 미치고 있다. 어떤 의미에서 우리는 이미 봇과 경쟁을 하고 있기도 하다.³⁴⁾³⁵⁾

또한, 봇을 대표하는 사례 중 또 하나는 챗봇(Chatbot)이 있다. 챗봇은 사용자가 메신저를 이용해 친구와 대화하듯 자연스럽게 질문을 입력하면, 인공지능 기술 기반의 채팅 플랫폼이 입력된 대화를 분석하여 마치 사람과 대화하는 것 같은 가상 개인비서 서비스를 제공한다.³⁶⁾³⁷⁾³⁸⁾

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

35) 오세욱. (2016). 미디어로서의 봇(bot) : 뉴스 챗봇에 대한 시론적 논의. 『한국언론정보학회』, 79(10), 70-103.

36) 이대근, 나승유. (2018). 중소상인을 위한 AI 챗봇 플랫폼의 설계 및 구현. 『한국전자통신학회』, 13(3), 561-570.

37) 김보경, 김현아, 심채인, 정진우. (2017). 소셜 챗봇을 활용한 클라우드 소싱 기반 질의응답

가상비서 서비스의 개념은 스마트폰, 웨어러블 디바이스 등 각종 IT 기기가 사용자의 평소 습관과 행동 방식을 패턴화하여 학습함으로써 개인별 필요에 따라 마치 실제 인간 비서처럼 사용자의 언어를 이해하고 사용자가 원하는 지시사항을 수행하는 맞춤형 서비스를 제공하는 소프트웨어 애플리케이션으로, 주로 음성인식 센서를 통해 사용자가 원하는 정보를 인식한 뒤, 주변 편의시설의 위치나 날씨 등 해당 정보를 검색하여 제공해 준다.³⁹⁾⁴⁰⁾

챗봇은 사람과의 대화에서 시스템이 가능한 인간다운 대화를 하는데 그 목적을 둔다. 스마트폰 또는 스마트 스피커를 이용한 가상 개인비서 서비스에서는 Apple의 Siri, Amazon의 Alexa, SoundHound의 Hound 등과 같이 목적지향 대화 시스템을 챗봇과 결합하여 특정된 목적을 달성하면서 일반대화도 수행 가능하도록 설계하였다. 이런 가상 개인비서 서비스는 향후 몇 년 사이 가장 주목할 인공지능 기술 중의 하나로 간주되고 있다.^{41) 42)}

또한, 챗봇은 유통 금융서비스 등과 같이 고객 상담이 많은 분야에서 유용하게 쓰일 수 있고 간단한 상담을 챗봇으로 대체하면 인건비를 절감할 수 있는 장점도 있어, 그 파급효과는 매우 크다고 할 것이다.

국외에는 업무지원을 위한 지능형 비서 챗봇인 ‘탈라’, 주문부터 결제까지 처리하는 챗봇인 ‘마스터카드 봇’, 영국 최초의 로봇 변호사 챗봇인 ‘Do not pay’, 우울증 환자에게 친구가 되어주는 챗봇인 ‘Woebot’이 대표적이라 할 것이다. 국내에는 대구시 민원상담 챗봇인 ‘뚜봇’, 온라인 쇼핑

시스템의 설계 및 구현. 『한국정보기술학회』, 15(11), 125-138.

38) 챗봇 관련 홈페이지(<http://www.Chatbots.org>), (2019년 11월 11일자 검색)

39) 윤수진. (2016). “뉴 실버세대의 감성 커뮤니케이션과 맞춤형 가상비서(Virtual Assistant) 콘텐츠 연구”. 경희대학교 박사학위논문.

40) KB금융지주연구소, (2016). 인공지능 기술의 발달과 가상 개인비서 서비스의 진화. 『KB 지식비타민』, 16(53).

41) 황금하, 권오욱, 이경순, 김영길. (2017). 목적지향 대화 시스템을 위한 챗봇 연구. 『정보처리학회』, 6(11), 499-506.

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

몰용 마케팅 챗봇인 ‘넛지’, 패션 정보를 알려주는 챗봇인 ‘패션을 부탁해’, 지능형 자동응대 챗봇인 ‘거버넌스’가 등장하고 있으나, 5G를 기반으로 한 VR, AR, IoT 등의 응용 기술을 융합해 국민이 체감할 수 있는 감성 챗봇 서비스의 발전이 요구되고 있는 상황이다.⁴³⁾

한국, 미국, 중국의 상품 주문결제 챗봇 서비스를 비교해 보면, 미국은 ‘킵 봇’이라는 플랫폼 상에서 개발된 ‘Sephora’, ‘H&M’과 페이스북 메신저 플랫폼 상에서 개발된 ‘Epitome Stylist’를 소개할 수 있겠고, 중국은 ‘위챗’ 플랫폼 상에서 개발된 ‘Chumenwenwen’, 한국의 경우엔 ‘카카오톡’ 플랫폼 상에서 개발된 ‘GS shop 톡’을 소개할 수 있겠다.

이에 대한 품질 테스트에 관한 설문조사를 실시한 결과, 전체적으로 만족도가 높은 플랫폼은 미국의 ‘킵 봇’이 우수평가를 받았다. ‘킵 봇’ 플랫폼 상에서 운영되는 ‘Sephora’는 사용자 취향 분석의 정확도가 가장 높다는 평가를 받았고, ‘H&M’의 경우는 버튼 입력방식과 제시된 상품 정보 관련항목에서 만족도가 가장 높은 것으로 조사되었다.

‘페이스북 메신저’ 플랫폼 상에서 운영되는 ‘Epitome Stylist’는 검색결과를 수평으로 넘겨보는 방식과 버튼식 대화갱신 인터페이스 방식이 타 챗봇에 비해 우수하다는 평가를 받았다. ‘위챗’ 플랫폼 상에서 운영되는 ‘Chumenwenwen’은 취향 탐색 등의 개인화 서비스를 제공하지 않아 관련 항목에서 최저점을 받았으며 선별되지 않은 다수의 검색 결과들을 제시하기 때문에 검색결과 수의 적절성 면에선 최저점수를 받았다.

한국의 ‘카카오톡’ 플랫폼 상에서 운영되는 ‘GS SHOP 톡’은 1회 발화당 글자 수가 지나치게 많고 상품 정보에 이미지가 배제되어 관련 항목들에서 최저점을 받았으며, 더욱이 챗봇의 대화체가 실제 사람과 가장 거리가 먼 챗봇으로 나타나 이에 대한 분발이 반드시 필요하다고 지적되었다.

44)

다음은 구글, 페이스북 메신저 플랫폼 상에서 운영되는 모바일 챗봇 서비스에 대한 사용자 경험평가를 비교한 경우를 소개한다.

43) 박정호. (2019). 『챗봇서비스 구현을 위한 기술로드맵』. 대전 : 중소기업기술정보진흥원.

44) 최수민, 최용순. (2017). 모바일 메신저기반 인공지능 챗봇의 상품주문 결제 인터페이스 연구. 『한국 HCI학회 학술대회』, 237-240.

사용자 경험평가 대상항목은 접근성, 가시성, 신뢰성, 검색성, 유용성, 사용성 이렇게 총 6가지 측면에서 사용자 평가를 시행하였다. 그 결과, 구글의 ‘Allo’는 실시간으로 구체적인 정보를 제공하고 유대감을 느낄 수 있었으나, 검색 이외 다른 서비스가 불가능하고 AI만의 자기생각이 없었다는 평가를 받았다.

페이스북의 ‘M’은 유저간의 메신저 기능이 좋고 채팅 중 일정잡기, 콜택시 예약서비스와 같은 부가 서비스를 제공한다는 점에선 좋았으나 버튼을 클릭해야만 검색정보를 제공한다는 점과 사용자와의 유대감을 못 느끼겠다는 점을 지적하였다. 또한, 두 챗봇(‘Allo’, ‘M’) 모두는 단순질문은 가능하나 파생질문은 불가능하다는 점도 지적되었다. 이는 동일한 질문에 대해 똑같은 응답을 하기 때문에 발화의 다양성이 부족하다는 결론이다.⁴⁵⁾⁴⁶⁾

향후 이를 개선하기 위한 다양한 연구가 이루어져야 할 필요성이 있다고 여겨지는 대목이다.

계속해서, 약(藥)에 대한 정보를 전달하는 챗봇은 채팅형 응답방식인 카카오톡 메신저 플랫폼 상에서 운영되는 봇으로, 식품의약품 안전처에서 제공하는 공공데이터를 활용해 다양한 의약품에 대한 이미지를 인식한다. 이 챗봇은 의약품에 대한 올바른 사용법과 적절한 의학 정보를 제공해 줌으로써, 안전한 의약품 복용 및 약에 대한 유용하고 정확한 헬스케어 서비스를 지원한다. 의약품 이미지를 이미지 인식 모듈에 전달하거나, 텍스트 정보(ex : URL 주소)를 데이터베이스에 저장하며, 이미지 인식 모듈인 Opencv를 활용해 의약품 이미지에 있는 색상과 모양, 분할선 등 약에 대한 정보를 식품의약품 안전처에 구축된 리눅스 서버에 전달함에 따라 알고자 하는 의약품에 대한 약 정보를 받아와 데이터베이스에 저장한다.

다시 말해, 이 챗봇은 이미지 분석된 결과를 가지고 식품의약품 안전처에서 제공하는 공공데이터를 활용해 약에 대한 정보를 가져와 이에 관한 약 정보를 챗봇을 통해서 사용자에게 제공하는 방식으로 동작된다.⁴⁷⁾

45) 강희주, 김승인. (2017). 메신저 기반의 모바일 챗봇 서비스 사용자 경험 평가 -구글(Allo)과 페이스북(M messenger)을 중심으로-. 『한국융합학회』, 8(9), 271-276.

46) 전원표, 김학수. (2011). 음성 기반 챗봇을 위한 중복 발화 회피 기법. 『한국정보과학회 학술대회』, 225-227.

자가혈당 관리지원용 챗봇은 환자의 생체정보를 기반으로 한 당뇨병증 사망 예측지표에 초점을 맞춰 가상 개인비서 서비스를 제공한다. 자가혈당 관리지원용 챗봇은 당뇨병증으로 인한 사망 예측지표를 제공하기 위해 먼저 관련 정보(성별, 연령, 수축기혈압, 이완기혈압, 공복혈당, 총콜레스테롤, 흡연여부, 음주여부, 체중, 신장, 운동여부, 주간운동 횟수)를 고객으로부터 전달받는다. 고객으로부터 입력이 완료되면, 자가혈당 관리지원용 챗봇은 예측지표 알고리즘을 통해 당뇨병증으로 인한 사망 예측지표를 그래프와 아이콘으로 표시해 주면서 생활습관이나 운동습관에 대한 안내도 해준다.⁴⁸⁾

챗봇은 헬스케어 분야에서도 지속적인 건강관리 서비스를 제공하기 위해 많이 도입하고 있는 상태다.⁴⁹⁾⁵⁰⁾

특히, 한의학계에서는 많은 한의학 관련 연구자들의 관심과 사업으로 한의학 관련 사이트들이 생겨나고 있으며, 그 중에서 한국 한의학 연구원에서 연구된 오아시스(OASIS, Oriental Medicine Advanced Searching Integrated System)는 논문검색 시스템에서 각 검색어의 패턴을 분석하고 연관성 있는 검색어를 추출하는 알고리즘을 설계해 전통의학 관련 논문을 보다 쉽게 찾을 수 있도록 서비스를 제공하고 있다.⁵¹⁾

계속해서, 현재 국내 중소기업들의 기술 수준은 [표 2-2]에서도 알 수 있듯이, 1-2 단계 수준에 머물러 있지만, 향후 인공지능 기술이 성숙함에 따라 향후 챗봇 서비스도 3단계의 지능형 감성 비서로 발전하여 더욱 전문적이고 개인에게 맞춰진 타입으로 출시되리라 예상된다. 1단계의 챗봇 서비스는 서비스와 관련된 데이터를 축적하여 사용자가 질문하면 미리 설

47) 김건우, 김태민, 이창환, 김지은, 홍정희. (2017). 이미지 인식과 대화를 통한 의약품 정보 제공 챗봇. 『한국시뮬레이션학회 학술대회』, 21-23.

48) 김민영. (2018). 당뇨병환자의 자가관리를 위한 챗봇 서비스. 『대한전자공학회 학술대회』, 1711-1714.

49) 김유정, 한상규, 윤종목, 허은영, 김정훈, 이중식. (2017). 챗봇 에이전트 정체성에 따른 사용자의 인식 및 행동 차이에 대한 연구. 『한국HCI학회』, 12(4), 45-55.

50) 박종진. (2018). 자연어 처리를 이용한 감정 스트레스 인지 및 관리 챗봇 개발. 『대한전기학회』, (67)7, 954-961.

51) 김익태. (2015). “전통의학 정보포털 논문검색을 위한 연관검색어 추출 알고리즘에 대한 연구”. 한밭대학교 석사학위논문.

계된 시나리오 기반으로 답변하나, 그 외의 질문은 답변이 어려운 기술수준에 있는 단계이다.

현재 ETRI를 비롯한 민간부문, 공공부문에서 1단계 구현을 위한 기초 연구를 수행 중에 있으나, 2022년도 전후로부터는 2단계인 지능형 비서(자연어처리 기술로 단순한 형태의 소통이 가능하며, 스스로 검색을 통해 질문에 답변하고 신고, 신청, 예약 등 간단한 업무일 경우 직접 수행이 가능함)와, 3단계인 감성비서(공공·민간의 각종 빅데이터, 감성인지 기술 등이 더해져 사용자의 마음까지 읽는 서비스를 제공하고, 서비스 전달 방식 또한 텍스트와 음성에 국한되지 않고 다양화될 것으로 예상됨)까지 발전 가능성이 있다고 예측된다.

또한, AI 활용 방향에도 시사점을 도출할 수 있다. 어떤 기술을 개발하는 것과 이를 구체화하는 것은 다르다. 기술을 어떻게 적용할지가 중요하다.

[표 2-2] 챗봇의 발전단계

구분	1단계 챗봇서비스 (Chatbot)	2단계 지능형 비서 (Intelligent Assistant)	3단계 감성 비서 (conscious Assistant)
제공 방식	텍스트, 음성	텍스트, 음성, 시각자료	텍스트, 음성, 시각자료, 행동인식
입력 방식	폐쇄형, 개방형 일부	폐쇄형 일부, 개방형	개방형-폐쇄형 복합사용
주요 기술	패턴매칭, 키워드 및 연관어 추출 등	딥러닝, 머신러닝, 자연어처리, 기타 신기술 융합 등	감성인지기술, 빅데이터 정형화 기술
내용	-학습된 내용에 대한 질의응답 -사용자와 단순한 형태의 소통 -검색을 통한 결과 제공	-사용자의 패턴 상황을 고려한 개인 맞춤형 서비스제공 -간단한 업무처리	-감성 교류를 통한 서비스 및 각종 서비스에 대한 선제적 대응

첫째, 4차 산업혁명 시대를 맞아 민첩하고 정확하게 움직일 수 있는 컨트를 타워가 필요하다. 일례로, 4차 산업혁명의 핵심기술 중에 하나인 3D 프린팅 기술이 일반제조업 뿐만 아니라 일상생활에도 보편화되어 다양하게 확산되어가고 있다.⁵²⁾

둘째, 4차 혁명이 클라우드, 빅데이터, 모바일 등 ICT 기술을 바탕으로 인터넷 플랫폼 상에서 벌어지는 영토전쟁이라는 점에서 어떻게 디지털 주권을 확보할 것인가가 중요하며, 클라우드 산업이나 빅데이터 산업의 활성화를 위해 엄격한 개인정보나 위치정보 규제 완화가 요구된다.

셋째, 빅데이터 활용을 위한 개인정보 보호법 개정에 여러 진전이 있고 논의가 있지만, 선진국 기업들이 하고 있는 수준의 법제정비와 규제완화, 그리고 적극적인 홍보를 통해 전반적인 문화개선도 필요하다.

넷째, 보안 산업이 규제 때문에 성장했으나 이제는 규제 때문에 가로막혀 더 이상 커지지를 못하고 있다. 또한, 4차 혁명에 맞춰 정부조직 체계부터 바꾸는 작업이 필요하고 이렇게 되어야 규제도 네거티브 규제 방식으로 바뀔 것이다.

다섯째, 이러한 네거티브 방식의 규제개선 시스템 도입을 위한 규제 프리존 특별법 정비와 새로운 ICT 융합서비스가 법 미비로 시장에 출시되지 못하는 것을 방지하기 위한 신속처리 임시허가 제도의 도입이 요구된다.⁵³⁾

52) 이용규. (2019). “Deep Learning을 활용한 3D 모델링 디자인권의 지식재산권 침해 예측 평가 연구”. 호서대학교 박사학위논문.

53) 김영인. (2017). 『국방 인공지능(AI) 활용방안 연구』. 서울 : 한국융합안보연구원.

2.6 특허회피 설계안에 관한 설문조사

본 논문은 설문조사를 통해 붓 기반 특허회피 설계안 자동 생성 방법에 관한 연구의 필요성 및 이로부터 자동 생성되는 특허회피 설계안을 고찰해 보고자 하였다.

설문조사의 취지는 “지식재산권 분야에서 국내외적으로 필수 불가결하게 일어나고 있는 특허침해 문제를 해결하기 위해 특허회피 전략이 요구된다. 이러한 특허회피 전략을 모색해 주는 특허회피 설계안은 당사자들 간의 법적 분쟁을 완충시켜 주는데 도움을 주며, 기업의 미래성장 동력을 발굴 육성토록 하는데 있어서도 나침반 역할을 담당하고 있다 이를 좀 더 명확하게 연구 및 고찰하고자 다음과 같은 설문을 준비하게 되었습니다.”라는 내용의 취지를 담고 있다.

본 논문의 붓 기반 특허회피 설계안 자동 생성 방법에 관한 연구의 필요성을 강조하고자 특허 전문가(변호사, 변리사, 교수, 박사, 기술사, 협업 종사자 등)분들의 다양한 의견을 여쭙보는 “특허회피 설계안 고찰을 위한 설문조사”를 다음과 같이 실시하였다.〈총 52명, 남성(44명), 여성(8명)〉

객관식 26문항과 주관식 4문항을 포함하여 총 30개 문항에 대한 응답결과를 받아 보았다.〈부록 참조〉

2.6.1 빈도분석

[표 2-3], [표 2-4], [표 2-5]는 30개 질문 중 선별된 20개 문항에 대한 응답결과를 토대로 빈도분석을 한 결과를 정리한 것이다.

단서 조항1로, 2-4, 2-5, 2-9, 2-10, 3-15, 3-17은 일반질의, 중복질의임으로 본 빈도분석에선 제외되었다.

[표 2-3] 특허회피 설계의 필요성에 관한 빈도분석

문항 번호	질문(Q)	답변(A)	빈도율
2-1	지식재산권의 필요성1에 관한 질의	특허분쟁 및 권리보호를 받기 위해	50%
2-2	지식재산권의 영향력에 관한 질의	정부 R&D 사업 과제 추진 및 수행으로 기업가치와 성장성에 도움이 됨	40.4%
2-3	지식재산권의 필요성2에 관한 질의	특허분쟁의 예방 및 권리 보호를 받을 수 있음으로	48.1%
2-7	특허회피 설계의 필요성에 관한 질의	[1]특허 등록 [2]자사기술 확보 및 보호장벽 마련(복수선택)	[1]55.8 % [2]53.8 %
2-8	특허회피 설계를 하기 위한 사전 필수요건에 관한 질의	키프리스를 활용한 선행기술조사	55.8%
2-11	기존 특허회피 설계의 문제점 및 대처방안에 관한 질의	[1]출원 후에 등록받은 특허 의 효용성 및 가치성 부재 [2]등록된 특허의 제품화 추진 후, 시장의 반응미비(복수선택)	[1]38.5 % [2]28.8 %

Q2-6 : 당신은 특허회피 설계안 또는 특허회피 전략 보고서를 받아 보았거나 향후 받아 본다면, 도움이 되었거나 도움이 될 거라 보는지요? 라는 주관식 질의1

A2-6 : 주관식 질의1에 대한 답변 중 중요시 여길만한 의견은 “특허를 분석하고 그에 따른 회피를 하는 것은 특허분쟁의 예방 및 권리 보호를 받을 수 있는 특허등록을 위한 최상의 전략이라 생각함”, “특허회피 설계안을 토대로 특허출원하였고, 향후 특허를 준비하는데 많은 도움이 됨으로 특허출원 및 특허등록에 효과적임”, “특허회피 설계를 통한 기술개발 방향 설정 및 개발제품의 권리보호, 경쟁력 강화”라는 코멘트가 도움이 되었다.

[표 2-4] AI 플랫폼의 설계요건에 관한 빈도분석

문항 번호	질문(Q)	답변(A)	빈도율
3-1	기존 방식 대비 AI 플랫폼[특허회피 설계안 자동생성 봇]의 가치성에 관한 질의	[1]다양한 특허회피 설계안을 제공 [2]언제, 어디서든 접근 및 사용가능(복수선택)	[1]46.2% [2]38.5%
3-2	AI 플랫폼이 갖춰야 할 필수요건에 관한 질의	[1]다양한 특허회피 설계안을 제시해 줄 수 있어야 함 [2]유저와 AI 플랫폼 간의 원활한 소통(복수선택)	[1]59.6% [2]44.2%
3-3	AI 플랫폼의 맞춤형 설계요건1에 관한 질의	정보의 대표성을 높이기 위해 특허청구범위 중 독립항을 중점으로 스크린할 필요가 있다고 봄	42.3%
3-4	AI 플랫폼의 맞춤형 설계요건2에 관한 질의	추출된 정보글들의 공통분모를 찾아 카테고리별로 분류	46.2%
3-5	AI 플랫폼의 맞춤형 설계요건3에 관한 질의	IPC의 특허분류코드를 부여받은 유저의 산업군과 주제어들 간의 비교방식	46.2%
3-6	AI 플랫폼의 맞춤형 설계요건4에 관한 질의	가중치 값 부여에 따른 객관적인 기준 설정 모색	57.7%
3-7	AI 플랫폼의 맞춤형 설계요건5에 관한 질의	주제어들에 대한 최종 선정 기준을 산술 평균치 이외 다른 기준치로 채택	34.6%

Q3-8 : 특허회피 전략을 제시해 주는 AI 플랫폼의 출시를 알고 필요성을 느껴 사용하게 됐을 시, 기능적으로 혹은 사업적으로 어떤 점들을 더 보완하거나 강화하면 좋을지 조언 부탁드립니다. 라는 주관식 질의2

A3-8 : 주관식 질의2에 대한 답변 중 중요시 여길만한 의견은 “모든 특허를 대상으로 함으로 회피할 수 없는 특허에 대한 대책 수립”, “일반적 사실 추출 보다는 특정 기업/기술의 맞춤형 전략 제시 모듈화”, “사용자와 AI 플랫폼 간의 인터페이스의 고급화”라는 코멘트가 도움이 되었다.

[표2-5] AI 플랫폼의 조건설정에 관한 빈도분석

문항 번호	질문(Q)	답변(A)	빈도율
3-9	사전준비1에 관한 질의	특허건수 10건 내지 30건	36.5%
3-10	사전준비2에 관한 질의	특허건수 3건 내지 10건	28.8%
3-12	만족 조건설정1에 관한 질의	정보글 3회 내지 5회	32.7%
3-13	산업군별 오차범위 조절에 관한 질의	ICT, IoT, AI, 가전, 로봇, 정보처리 분야	48.1%
3-14	만족 조건설정2에 관한 질의	주제어 10개 내지 30개 미만	40.4%
3-16	만족 조건설정3에 관한 질의	과거부터 현재까지 이슈로 상위 검색순위에 여러 차례 올라오는 정보글이어야 함(10년 기준)	38.5%
3-18	미래의 AI 플랫폼의 요구사항에 관한 질의	[1]데이터 마이닝(많은 데이터 가운데 숨겨져 있는 유용한 상관관계를 발견하여, 미래에 실행 가능한 정보를 추출해 내고 의사 결정에 이용하는 기법) [2]시멘틱 분석(컴퓨터가 정보자원의 뜻을 이해하고, 논리적 추론까지 할 수 있는 기법)(복수선택)	[1]44.2 % [2]34.6 %

Q3-11 : 특허정보 검색사이트(www.kipris.or.kr)를 통해 검색된 특허
건수 중 몇 건 정도가 특허회피 전략을 제시해 주는 AI 플랫폼에 적용되
어야 내가 원하는 특허회피 설계안을 충분히 제시해 줄 수 있을 것이
라는 3-10의 질의를 정한 이유를 묻는 주관식 질의3

A3-11 : 주관식 질의3에 대한 답변 중 중요시 여길만한 의견은 “특허
회피 전략을 제시하기 위한 검색된 특허가 너무 많을 경우 필수 구성요소
가 원하지 않은 방향으로 도출되어 고객의 의도와 다른 특허가 나올 수
있고, 너무 적을 경우 다른 선행특허로 인해 신규성이나 진보성이 없어 특

회피 설계가 이루어지지 않을 수 있을 듯함”, “보통 수만 건을 필터링하여 수백 건을 보고, S 급으로 분류된 정보는 세부분석을 해야 한다”, “한두건은 회피전략에 참고삼기에는 너무 적은 것 같고, 그렇다고 너무 많으면 데이터 축적 등에 따른 코스트를 고려해야하기 때문임. 중요한 점이 부각되어있고, 연관성이 깊은 회피전략이 제시된다면 4~5건도 무난할 것으로 보임”라는 코멘트가 도움이 되었다.

Q3-19 : 설문 후기~<자유로운 평가글 부탁드립니다.>주관식 질의4

A3-19 : 주관식 질의4에 대한 답변 중 중요시 여길만한 의견은 “특히 회피 설계안은 늘 데이터의 신선도와 관계가 많았습니다! 공개되지 않은 특허도 위험 가능성이 높은 사항은 지속적으로 고객에게 경고를 주어야합니다! 감사합니다! 좋은 논문 쓰세요!”, “객관성, 다양성 확보 차원의 deep learning이 필요하겠습니다.”, “특허회피 플랫폼이라는 새로운 영역을 개척해 나가심에 박수를 보내드립니다.”라는 코멘트가 도움이 되었다.

2.6.2 교차분석

교차분석(카이제곱 분석 : χ^2)은 검정하고자 하는 변수들 간의 독립성과 관련성을 분석하는데 이용하며, 교차분석에서는 변수들 간의 빈도를 [표 2-3] 내지 [표 2-22]와 같이 교차표를 작성할 수 있으며, 카이제곱 (Chi-Square)의 유의확률로 변수들 간의 독립성과 관련성을 검정했다면 카이제곱 검정이라 할 것이다. 카이제곱 분석에서 기억해야 할 유의사항은 분석결과의 해석이며, 변수들이 명목척도 및 서열척도임으로 변수들 간의 상관관계를 충분히 반영하지 않는다. 단지, 변수들 간의 분포의 차이에 중점을 두어 변수 간의 독립성을 검증한다.⁵⁴⁾

단서 조항2로, 3-8, 3-11, 3-15, 3-17, 3-19은 일반상식에 관한 질의이거나 주관식 질의임으로 본 교차분석에선 제외되었다.

(1) Q2-1과 Q2-7 간의 χ^2 검정결과

54) 송지준. (2015). 『SPSS/AMOS 통계분석방법』. 경기 : 도서출판 : 21세기사.

(2) Q2-7과 Q2-8 간의 χ^2 검정결과

Q2-7의 특허회피 설계의 필요성과 Q2-8의 특허회피 설계 사전 필요요건과 간의 χ^2 검정결과, χ^2 통계량은 12.022(p=.150)로 통계적 유의수준 하에서 두 변수 간에는 분포의 차이는 없었다.

특허회피 설계 사전 필수요건 중 가장 높은 분포를 보이는 문항은 키프리스를 활용한 선행기술 조사(n=35, 68.6%)이고, 특허회피 설계의 필요성 중 가장 높은 빈도를 보이는 문항은 특허등록(n=29, 56.9%)으로 나타났다. 또한 두 문항 간의 관련성은 다른 문항들과 비교해서 가장 높은 빈도수를 보였다(n=24, 47.1%).

[표 2-7]에 반영되는 Q2-7의 선택 지문은 다음과 같다.

- ① 특허등록 ② 기술사업화
③ 틈새시장 발굴을 통한 시장 선점 ④ 자사기술 확보/보호 장벽마련

[표 2-7]에 반영되는 Q2-8의 선택 지문은 다음과 같다.

- ① 발명자의 직무 발명 신고서
- ② 키프리스를 활용한 선행기술조사
- ③ 포털 사이트에 올라온 검색정보 활용
- ④ 외부 전문가의 코멘트 및 방향 제시

[표 2-7] Q2-7과 Q2-8 간의 교차분석 결과표

			특허회피 설계의 필요성(Q2-7)						x^2/p
			①	②	③	④	기타	전체	
특허회피 설계 사전 필수 요건 (Q2-8)	①	빈도	3	3	2	1	0	9	12.022 / .150
		전체	5.9%	5.9%	3.9%	2.0%	.0%	17.6%	
	②	빈도	24	3	4	3	1	35	
		전체	47.1%	5.9%	7.8%	5.9%	2.0%	68.6%	
	④	빈도	2	1	1	3	0	7	
		전체	3.9%	2.0%	2.0%	5.9%	.0%	13.7%	
	전체	빈도	29	7	7	7	1	51	
		전체	56.9%	13.7%	13.7%	13.7%	2.0%	1000%	
* $p < .05$, ** $p < .01$									

여기서, 특허회피 설계 사전 필수 요건의 선택지문③은 설문 응답자 중 어느 분도 선택하지 않아 [표 2-7]에는 표기되지 않음.

(3) Q2-2와 Q2-1 간의 χ^2 검정결과

Q2-2의 지식재산권의 영향력과 Q2-1의 지식재산권의 필요성 간의 χ^2 검정결과, χ^2 통계량은 7.599($p=.575$)로 통계적으로 두 변수 간에는 분포의 차이는 보이지 않았다.

지식재산권의 영향력 중 가장 높은 분포를 보이는 문항은 정부 R&D사업 과제 추진 및 수행으로 기업 가치와 성장성에 도움이 됨이 $n=21$, 46.7%로 가장 높은 분포를 보였고, 지식재산권의 필요성 중에서는 특허분쟁 및 권리보호를 받기 위해가 $n=26$, 57.8%의 수치를 보였다. 또한 두 문항간의 관련성은 $n=12$, 26.7%로 가장 높게 나타났다.

[표 2-8]에 반영되는 Q2-2의 선택 지문은 다음과 같다.

- ① 정부 R&D 사업 과제 추진 및 수행으로 기업가치와 성장성에 도움이 됨
- ② 정부출연자금 또는 담보대출과 같은 자본금 확충에 도움이 됨
- ③ 우수제품 개발 및 조달등록으로 시장 진입에 도움이 됨
- ④ 제품 마케팅의 성공으로 기업 수익성 창출에 도움이 됨

[표 2-8]에 반영되는 Q2-1의 선택 지문은 다음과 같다.

- ① 시장에서 독점권지위를 확보하기 위해
- ② 특허분쟁 및 권리보호를 받기 위해
- ③ R&D 투자회수 및 향후 추가기술 개발에 대한 기틀 마련을 위해
- ④ 정부의 각종 정책자금 및 다양한 세제 혜택을 제공받기 위해

[표 2-8] Q2-2와 Q2-1 간의 교차분석 결과표

			지식재산권의 필요성(Q2-1)					χ^2/p
			①	②	③	④	전체	
지식 재산권의 영향력 (Q2-2)	①	빈도	5	12	3	1	21	7.599/ .575
		전체	11.1%	26.7%	6.7%	2.2%	46.7%	
	②	빈도	0	3	0	0	3	
		전체	.0%	6.7%	.0%	.0%	6.7%	
	③	빈도	1	4	2	2	9	
		전체	2.2%	8.9%	4.4%	4.4%	20.0%	
	④	빈도	3	7	2	0	12	
		전체	6.7%	15.6%	4.4%	.0%	26.7%	
	전체	빈도	9	26	7	3	45	
		전체	20.0%	57.8%	15.6%	6.7%	100.0%	
* $p<.05$, ** $p<.01$								

(4) Q2-8과 Q2-11 간의 χ^2 검정결과

Q2-8의 특허회피 설계 사전 필수요건과 Q2-11의 기존 특허회피 설계의 문제점과 대처방안 간의 χ^2 검정결과, χ^2 통계량은 2.408($p=.966$)로 통계적으로 두 변수 간에는 관련성이 나타나지 않았다.

특허회피 설계 사전 필수요건 중 가장 높은 분포를 보이는 문항은 키프리스를 활용한 선행기술 조사($n=35$, 68.6%)이고, 기존 특허회피 설계의 문제점과 대처방안에서는 출원 후에 등록받은 특허의 효용성 및 가치성 부재($n=17$, 33.3%)로 나타났다. 또한 두 문항간의 관련성은 $n=11$, 21.6%로 가장 높은 수치를 보였다.

[표 2-9]에 반영되는 Q2-8의 선택 지문은 다음과 같다.

- ① 발명자의 직무 발명 신고서
- ② 키프리스를 활용한 선행기술조사
- ③ 포털 사이트에 올라온 검색정보 활용
- ④ 외부 전문가의 코멘트 및 방향 제시

[표 2-9]에 반영되는 Q2-11의 선택 지문은 다음과 같다.

- ① 주관적 관점의 안 제시로, 일부 편협된 사업모델로 인한 미래 방향성

상실

- ② 제시된 보고서에 따라 출원되었던 특허의 등록거절
- ③ 출원 후에 등록받은 특허의 효용성 및 가치성 부재
- ④ 등록된 특허의 제품화 추진 후, 시장의 반응 미비

여기서, 특허회피 설계 사전 필수 요건의 선택지문③은 설문 응답자 중 어느 분도 선택하지 않아 [표 2-9]에는 표기되지 않음.

[표 2-9] Q2-8과 Q2-11 간의 교차분석 결과표

			기존 특허회피 설계의 문제점과 대처방안(Q2-11)						χ^2/p
			①	②	③	④	기타	전체	
특허 회피 설계 사전 필수 요건 (Q2-8)	①	빈도	2	2	3	1	1	9	2.408/ .966
		전체	3.9%	3.9%	5.9%	2.0%	2.0%	17.6%	
	②	빈도	8	4	11	5	7	35	
		전체	15.7%	7.8%	21.6%	9.8%	13.7%	68.6%	
	④	빈도	2	0	3	1	1	7	
		전체	3.9%	.0%	5.9%	2.0%	2.0%	13.7%	
	전체	빈도	12	6	17	7	9	51	
		전체	23.5%	11.8%	33.3%	13.7%	17.6%	100.0%	
* $p < .05$, ** $p < .01$									

(5) Q2-7과 Q2-2 간의 χ^2 검정결과

Q2-7의 특허회피 설계의 필요성과 Q2-2의 지식재산권의 영향성 간의 χ^2 검정결과, χ^2 통계량은 12.822($p=.171$)로 통계적으로 두 변수 간에는 분포의 차이가 없었다.

특허회피 설계의 필요성 중 가장 높은 분포를 보이는 문항은 특허등록 ($n=25$, 55.6%)이고, 지식재산권의 영향성에서는 정부 R&D사업 과제 추

진 및 수행으로 기업 가치와 성장성에 도움이 됨(n=21, 46.7%)으로 나타났다. 또한 두 문항간의 관련성은 n=15, 33.3%로 가장 높은 수치를 보였다.

[표 2-10]에 반영되는 Q2-7의 선택 지문은 다음과 같다.

- ① 특허등록
- ② 기술사업화
- ③ 틈새시장 발굴을 통한 시장 선점
- ④ 자사기술 확보 및 보호 장벽 마련

[표 2-10]에 반영되는 Q2-2의 선택 지문은 다음과 같다.

- ① 정부 R&D 사업 과제 추진 및 수행으로 기업가치와 성장성에 도움이 됨
- ② 정부출연자금 또는 담보대출과 같은 자본금 확충에 도움이 됨
- ③ 우수제품 개발 및 조달등록으로 시장 진입에 도움이 됨
- ④ 제품 마케팅의 성공으로 기업 수익성 창출에 도움이 됨

[표 2-10] Q2-7과 Q2-2 간의 교차분석 결과표

			지식재산권의 영향성(Q2-2)					x^2/p
			①	②	③	④	전체	
특허 회피 설계의 필요성 (Q2-7)	①	빈도	15	1	4	5	25	12.822 / .171
		전체	33.3%	2.2%	8.9%	11.1%	55.6%	
	②	빈도	3	0	1	1	5	
		전체	6.7%	.0%	2.2%	2.2%	11.1%	
	③	빈도	3	0	2	3	8	
		전체	6.7%	.0%	4.4%	6.7%	17.8%	
	④	빈도	0	2	2	3	7	
		전체	.0%	4.4%	4.4%	6.7%	15.6%	
	기타	빈도	0	0	0	0	0	
		전체	.0%	.0%	.0%	.0%	.0%	
	전체	빈도	21	3	9	12	45	
		전체	46.7%	6.7%	20.0%	26.7%	100.0%	
* $p<.05$, ** $p<.01$								

(6) Q2-11과 Q3-1 간의 χ^2 검정결과

Q2-11의 기존 특허회피 설계의 문제점 및 대처방안과 Q3-1의 기존방식 대비 AI 플랫폼의 가치성 간의 χ^2 검정결과, χ^2 통계량은 16.643(p=.409)로 통계적으로 두 변수 간에는 관련성이 나타나지 않았다.

기존 특허회피 설계의 문제점과 대처방안 중 가장 높은 분포를 보이는 문항은 출원 후에 등록받은 특허의 효용성 및 가치성 부재(n=17, 32.7%)이고, 기존 방식대비 AI 플랫폼의 가치성에서는 변리사의 컨설팅비용 vs 저렴한 수수료(n=19, 36.5%)로 나타났다. 또한 두 문항간의 관련성은 n=6, 11.5%의 수치를 보였으며, 주관적 관점의 안 제시로, 일부 편협된 사업 모델로 인한 미래 방향성 상실과 변리사의 컨설팅비용 vs 저렴한 수수료에서도 n=6, 11.5%의 수치로 나타났다.

[표 2-11]에 반영되는 Q2-11의 선택 지문은 다음과 같다.

- ① 주관적 관점의 안 제시로, 일부 편협된 사업모델로 인한 미래 방향성 상실
- ② 제시된 보고서에 따라 출원되었던 특허의 등록거절
- ③ 출원 후에 등록받은 특허의 효용성 및 가치성 부재
- ④ 등록된 특허의 제품화 추진 후, 시장의 반응 미비

[표 2-11]에 반영되는 Q3-1의 선택 지문은 다음과 같다.

- ① 변리사의 컨설팅 비용 vs 저렴한 수수료
- ② 다양한 특허회피 설계안을 제공
- ③ 언제, 어디서든 접근 및 사용가능
- ④ 지속적인 AI의 지능화로 결과치 신뢰

[표 2-11] Q2-11과 Q3-1 간의 교차분석 결과표

			기존방식대비 AI플랫폼의 가치성(Q3-1)						x^2/p
			①	②	③	④	기타	전체	
기존 특허 회피 설계의 문제점 및 대처 방안 (Q2-11)	①	빈도	6	5	2	0	0	13	16.64 3/ .409
		전체	11.5%	9.6%	3.8%	.0%	.0%	25.0%	
	②	빈도	4	2	0	0	0	6	
		전체	7.7%	3.8%	.0%	.0%	.0%	11.5%	
	③	빈도	6	4	5	1	1	17	
		전체	11.5%	7.7%	9.6%	1.9%	1.9%	32.7%	
	④	빈도	2	4	1	0	0	7	
		전체	3.8%	7.7%	1.9%	.0%	.0%	13.5%	
	기타	빈도	1	2	3	1	2	9	
		전체	1.9%	3.8%	5.8%	1.9%	3.8%	17.3%	
	전체	빈도	19	17	11	2	3	52	
		전체	36.5%	32.7%	21.2%	3.8%	5.8%	100.0%	
* $p<.05$, ** $p<.01$									

(7) Q3-1과 Q3-18 간의 χ^2 검정결과

Q3-1의 기존방식 대비 AI 플랫폼의 가치성과 Q3-18의 미래의 AI플랫폼의 요구사항 간의 χ^2 검정결과, χ^2 통계량은 50.265($p=.001$)로 통계적 유의수준 하에서 두 변수 간에는 분포의 차이는 있는 것으로 나타났다.

기존방식 대비 AI 플랫폼의 가치성 중 가장 높은 분포를 보이는 문항은 변리사의 컨설팅 비용 vs 저렴한 수수료($n=19$, 36.5%)이고, 미래의 AI 플랫폼의 요구사항 중 가장 높은 분포를 보이는 문항은 시멘틱 분석($n=17$, 32.7%)으로 나타났다. 또한 기존방식 대비 AI 플랫폼의 가치성과 미래의 AI 플랫폼의 요구사항 간 관련성은 $n=9$, 17.3%로 가장 높게 나타났다.

[표 2-12]에 반영되는 Q3-1의 선택 지문은 다음과 같다.

- ① 변리사의 컨설팅 비용 vs 저렴한 수수료
- ② 다양한 특허회피 설계안을 제공

- ③ 언제, 어디서든 접근 및 사용가능
- ④ 계속적인 AI의 지능화로 결과치 신뢰

[표 2-12]에 반영되는 Q3-18의 선택 지문은 다음과 같다.

- ① 오피니언 마이닝 ② 시맨틱 분석
- ③ 텍스트 마이닝 ④ 데이터 마이닝
- ⑤ 온톨로지 ⑥ 메타분석

[표 2-12] Q3-1과 Q3-18 간의 교차분석 결과표

			미래의 AI플랫폼의 요구사항(Q3-18)								χ^2/p
			①	②	③	④	⑤	⑥	기타	전체	
기존 방식 대비 AI 플랫폼 의 가치성 (Q3-1)	①	빈도	2	9	4	4	0	0	0	19	50.26 5/ .001*
		전체	3.8%	17.3%	7.7%	7.7%	.0%	.0%	.0%	36.5%	
	②	빈도	1	7	3	2	1	3	0	17	
		전체	1.9%	13.5%	5.8%	3.8%	1.9%	5.8%	.0%	32.7%	
	③	빈도	3	0	0	7	0	0	1	11	
		전체	5.8%	.0%	.0%	13.5%	.0%	.0%	1.9%	21.2%	
	④	빈도	0	0	1	0	1	0	0	2	
		전체	.0%	.0%	1.9%	.0%	1.9%	.0%	.0%	3.8%	
	기타	빈도	0	1	0	0	1	0	1	3	
		전체	.0%	1.9%	.0%	.0%	1.9%	.0%	1.9%	5.8%	
	전체	빈도	6	17	8	13	3	3	2	52	
		전체	11.5%	32.7%	15.4%	25.0%	5.8%	5.8%	3.8%	100.0%	

* $p < .05$, ** $p < .01$

(8) Q3-3과 Q3-4 간의 χ^2 검정결과

Q3-3의 AI플랫폼의 맞춤형 설계요건1과 Q3-4의 AI플랫폼의 맞춤형 설계요건2 간의 χ^2 검정결과, χ^2 통계량은 9.076($p=.430$)으로 통계적으로 두 변수 간에는 분포의 차이가 나타나지 않았다.

AI플랫폼의 맞춤형 설계요건1 중 가장 높은 분포를 보이는 문항은 정보의 대표성을 높이기 위해 특허청구범위 중 독립항을 중점으로 스크린 할 필요가 있다고 봄($n=22$, 42.3%)이고, AI플랫폼의 맞춤형 설계요건2에서는

추출된 정보글들의 공통분모를 찾아 카테고리별로 분류(n=25, 48.1%)로 나타났다. 또한 두 문항간의 관련성은 n=13, 25.0%로 가장 높은 수치를 보였다.

[표 2-13]에 반영되는 Q3-3의 선택 지문은 다음과 같다.

- ① 정보의 다양화를 높이기 위해 특허청구범위 외 다른 구분영역도 스크린할 필요가 있다고 봄
- ② 정보의 대표성을 높이기 위해 특허청구범위 중 독립항을 중점으로 스크린 할 필요가 있다고 봄
- ③ 정보의 표준화를 높이기 위해 특허청구범위의 전체를 스크린 할 필요가 있다고 봄
- ④ 정보의 간단 명료성을 높이기 위해 부호의 설명, 요약서로만 스크린 해도 충분하리라 봄

[표 2-13]에 반영되는 Q3-4의 선택 지문은 다음과 같다.

- ① 유저의 산업군과 추출된 정보글 간의 연관성 판단
- ② 추출된 정보글들의 공통분모를 찾아 카테고리별로 분류
- ③ 유저의 산업군에 속한 현재 관심도가 높은 정보글 보여주기
- ④ 과거 이슈화 되었던 검색어가 기재된 정보글 보여주기

[표 2-13] Q3-3과 Q3-4 간의 교차분석 결과표

			AI플랫폼의 맞춤형 설계요건2((Q3-4)					x^2/p
			①	②	③	④	전체	
AI 플랫폼 의 맞춤형 설계 요건1 (Q3-3)	①	빈도	5	7	1	0	13	9.076/ .430
		전체	9.6%	13.5%	1.9%	.0%	25.0%	
	②	빈도	4	13	3	2	22	
		전체	7.7%	25.0%	5.8%	3.8%	42.3%	
	③	빈도	3	4	4	1	12	
		전체	5.8%	7.7%	7.7%	1.9%	23.1%	
	④	빈도	1	1	2	1	5	
		전체	1.9%	1.9%	3.8%	1.9%	9.6%	
	전체	빈도	13	25	10	4	52	
		전체	25.0%	48.1%	19.2%	7.7%	100.0%	
* $p<.05$, ** $p<.01$								

(9) Q3-4와 Q3-5 간의 χ^2 검정결과

Q3-4의 AI 플랫폼의 맞춤형 설계요건2와 Q3-5의 AI플랫폼의 맞춤형 설계요건3 간의 χ^2 검정결과, χ^2 통계량은 6.420($p=.697$)으로 통계적으로 두 변수 간에는 분포의 차이가 나타나지 않았다.

AI 플랫폼의 맞춤형 설계요건2 중 가장 높은 분포를 보이는 문항은 추출된 정보글들의 공통분모를 찾아 카테고리별로 분류($n=24$, 48.0%)이고, AI 플랫폼의 맞춤형 설계요건3 중 가장 높은 분포를 보이는 문항은 IPC의 특허분류코드를 부여받은 유저의 산업군과 주제어들 간의 비교방식($n=25$, 50.0%)으로 밝혀졌다. 또한 두 문항간의 관련성은 $n=12$, 24.0%로 다른 문항들보다 높은 수치로 나타났다.

[표 2-14]에 반영되는 Q3-4의 선택 지문은 다음과 같다.

- ① 유저의 산업군과 추출된 정보글 간의 연관성 판단
- ② 추출된 정보글들의 공통분모를 찾아 카테고리별로 분류
- ③ 유저의 산업군에 속한 현재 관심도가 높은 정보글 보여주기
- ④ 과거 이슈화 되었던 검색어가 기재된 정보글 보여주기

[표 2-14]에 반영되는 Q3-5의 선택 지문은 다음과 같다.

- ① IPC의 부정확성에 대한 고찰로 다른 분류방식 채택 권유
- ② IPC의 특허분류코드를 부여받은 유저의 산업군과 주제어들 간의 비교방식
- ③ 4개의 카테고리로 분류된 주제어들의 활용도 모색
- ④ 코드간 비유사로 저가중치 집단으로 구분된 주제어들의 분류상태에 대한 재검증

[표 2-14] Q3-4와 Q3-5 간의 교차분석 결과표

			AI플랫폼의 맞춤형 설계요건3(Q3-5)					x^2/p
			①	②	③	④	전체	
AI 플랫폼 의 맞춤형 설계 요건2 (Q3-4)	①	빈도	4	7	1	1	13	6.420/ .697
		전체	8.0%	14.0%	2.0%	2.0%	26.0%	
	②	빈도	3	12	8	1	24	
		전체	6.0%	24.0%	16.0%	2.0%	48.0%	
	③	빈도	2	4	2	1	9	
		전체	4.0%	8.0%	4.0%	2.0%	18.0%	
	④	빈도	0	2	2	0	4	
		전체	.0%	4.0%	4.0%	.0%	8.0%	
	전체	빈도	9	25	13	3	50	
		전체	18.0%	50.0%	26.0%	6.0%	100.0%	
* $p<.05$, ** $p<.01$								

(10) Q3-5와 Q3-6 간의 χ^2 검정결과

Q3-5의 AI플랫폼의 맞춤형 설계요건3과 Q3-6의 AI 플랫폼의 맞춤형 설계요건4 간의 χ^2 검정결과, χ^2 통계량은 14.203($p=.115$)으로 통계적 유의 수준 하에서 유의미한 분포의 차이가 나타나지 않았다.

AI 플랫폼의 맞춤형 설계요건3 중 가장 높은 분포를 보이는 문항은 IPC의 특허분류코드를 부여받은 유저의 산업군과 주제어 들 간의 비교방식($n=24$, 49.0%)이고, AI 플랫폼의 맞춤형 설계요건4 중 가장 높은 빈도를 보이는 문항은 가중치 값 부여에 따른 객관적인 기준 설정 모색($n=29$, 59.2%)으로 나타났다. 또한 두 문항간의 관련성은 $n=15$, 30.6%의 수치를 보였다.

[표 2-15]에 반영되는 Q3-5의 선택 지문은 다음과 같다.

- ① IPC의 부정확성에 대한 고찰로 다른 분류방식 채택 권유
- ② IPC의 특허분류코드를 부여받은 유저의 산업군과 주제어 들 간의 비교방식

- ③ 4개의 카테고리로 분류된 주제어들의 활용도 모색
- ④ 코드간 비유사로 저가중치 집단으로 구분된 주제어들의 분류상태에 대한 재검증
- [표 2-15]에 반영되는 Q3-6의 선택 지문은 다음과 같다.
- ① 0.5를 기준으로 한 피어슨상관계수의 일반적 접근에 대한 용이성 판단
- ② 가중치 값 부여에 따른 객관적인 기준 설정 모색
- ③ 전문가가 선정한 S급 특허의 채택 기준에 대한 고시 의무화
- ④ 상관식으로 구한 피어슨 상관계수 값으로 알아낸 최적점수로 제거한 주제어들 이의제기

[표 2-15] Q3-5와 Q3-6 간의 교차분석 결과표

			AI플랫폼의 맞춤형 설계요건4(Q3-6)					x^2/p
			①	②	③	④	전체	
AI 플랫폼 의 맞춤형 설계 요건3 (Q3-5)	①	빈도	0	6	3	0	9	14.203/ .115
		전체	.0%	12.2%	6.1%	.0%	18.4%	
	②	빈도	1	15	6	2	24	
		전체	2.0%	30.6%	12.2%	4.1%	49.0%	
	③	빈도	4	6	3	0	13	
		전체	8.2%	12.2%	6.1%	.0%	26.5%	
	④	빈도	0	2	0	1	3	
		전체	.0%	4.1%	.0%	2.0%	6.1%	
	전체	빈도	5	29	12	3	49	
		전체	10.2%	59.2%	24.5%	6.1%	100.0%	
* $p<.05$, ** $p<.01$								

(11) Q3-6과 Q3-7 간의 χ^2 검정결과

Q3-6의 AI 플랫폼의 맞춤형 설계요건4와 Q3-7의 AI 플랫폼의 맞춤형

설계요건5 간의 χ^2 검정결과, χ^2 통계량은 12.329($p=.195$)로 통계적으로 의미 있는 분포의 차이가 없었다.

AI플랫폼의 맞춤형 설계요건4 중 가장 높은 분포를 보이는 문항은 가중치 값 부여에 따른 객관적인 기준 설정 모색($n=30$, 60.0%)이고, AI플랫폼의 맞춤형 설계요건5 중 가장 많은 빈도를 보인 문항은 주제어들에 대한 최종 선정 기준을 산술 평균치 이외 다른 기준치로 채택($n=12$, 24.0%)으로 밝혀졌다. 또한 두 문항간의 관련성은 $n=12$, 24.0%의 수치로 나타났다.

[표 2-16]에 반영되는 Q3-6의 선택 지문은 다음과 같다.

- ① 0.5를 기준으로 한 피어슨상관계수의 일반적 접근에 대한 용이성 판단
- ② 가중치 값 부여에 따른 객관적인 기준 설정 모색
- ③ 전문가가 선정한 S급 특허의 채택 기준에 대한 고시 의무화
- ④ 상관식으로 구한 피어슨 상관계수 값으로 알아낸 최적점수로 제거한 주제어들 이의제기

[표 2-16]에 반영되는 Q3-7의 선택 지문은 다음과 같다.

- ① 추출된 정보글에 표현된 긍정, 부정, 중립 여부에 대한 판단기준 모호
- ② 주제어들에 대한 최종 선정 기준을 산술 평균치 이외 다른 기준치로 채택
- ③ 기존 오피니언 마이닝 기법에 형용사를 더 추가시킨 이유에 대한 적합한 설명
- ④ 최종 선택된 주제어들의 자연어처리 및 자국어, 영어로 표현될 특허 회피 전략의 언어 구사력

[표 2-16] Q3-6과 Q3-7 간의 교차분석 결과표

			AI플랫폼의 맞춤형 설계요건5(Q3-7)					x^2/p
			①	②	③	④	전체	
AI 플랫폼의 맞춤형 설계 요건4 (Q3-6)	①	빈도	2	2	1	0	5	12.329 / .195
		전체	4.0%	4.0%	2.0%	.0%	10.0%	
	②	빈도	8	12	7	3	30	
		전체	16.0%	24.0%	14.0%	6.0%	60.0%	
	③	빈도	1	4	5	2	12	
		전체	2.0%	8.0%	10.0%	4.0%	24.0%	
	④	빈도	1	0	0	2	3	
		전체	2.0%	.0%	.0%	4.0%	6.0%	
	전체	빈도	12	18	13	7	50	
		전체	24.0%	36.0%	26.0%	14.0%	100.0%	
* $p<.05$, ** $p<.01$								

(12) Q3-7과 Q3-18 간의 χ^2 검정결과

Q3-7의 AI플랫폼의 맞춤형 설계요건5와 Q3-18의 미래의 AI플랫폼의 요구사항 간의 χ^2 검정결과, χ^2 통계량은 21.103($p=.274$)으로 통계적으로 유의미한 관련성이 나타나지 않았다.

AI플랫폼의 맞춤형 설계요건5 중 가장 높은 분포를 보이는 문항은 주제어들에 대한 최종 선정기준을 산술 평균치 이외 다른 기준치로 채택($n=18$, 36.0%)이고, 미래의 AI플랫폼의 요구사항 중 가장 많은 빈도수를 보인 문항은 시맨틱 분석($n=17$, 34.0%)으로 나타났다. 또한 두 문항간의 관련성은 $n=8$, 16.0%의 수치로 나타났다.

[표 2-17]에 반영되는 Q3-7의 선택 지문은 다음과 같다.

- ① 추출된 정보글에 표현된 긍정, 부정, 중립 여부에 대한 판단기준 모호
- ② 주제어들에 대한 최종 선정 기준을 산술 평균치 이외 다른 기준치로 채택
- ③ 기존 오피니언 마이닝 기법에 형용사를 더 추가시킨 이유에 대한 적

합한 설명

④ 최종 선택된 주제어들의 자연어처리 및 자국어로 표현될 특허회피 전략의 언어 구사력

[표 2-17]에 반영되는 Q3-18의 선택 지문은 다음과 같다.

- ① 오피니언 마이닝 ② 시멘틱 분석
- ③ 텍스트 마이닝 ④ 데이터 마이닝
- ⑤ 온톨로지 ⑥ 메타분석

[표 2-17] Q3-7과 Q3-18 간의 교차분석 결과표

			미래의 AI플랫폼의 요구사항(Q3-18)								x^2/p
			①	②	③	④	⑤	⑥	기타	전체	
AI 플랫폼 의 맞춤형 설계 요건5 (Q3-7)	①	빈도	3	2	2	2	1	1	1	12	21.10 3/ .274
		전체	6.0%	4.0%	4.0%	4.0%	2.0%	2.0%	2.0%	24.0%	
	②	빈도	1	8	3	6	0	0	0	18	
		전체	2.0%	16.0%	6.0%	12.0%	.0%	.0%	.0%	36.0%	
	③	빈도	2	5	3	2	0	1	0	13	
		전체	4.0%	10.0%	6.0%	4.0%	.0%	2.0%	.0%	26.0%	
	④	빈도	0	2	0	2	2	1	0	7	
		전체	.0%	4.0%	.0%	4.0%	4.0%	2.0%	.0%	14.0%	
	전체	빈도	6	17	8	12	3	3	1	50	
		전체	12.0%	34.0%	16.0%	24.0%	6.0%	6.0%	2.0%	100.0%	
* $p<.05$, ** $p<.01$											

(13) Q3-9와 Q3-10 간의 χ^2 검정결과

Q3-9의 사전준비1과 Q3-10의 사전준비2 간의 χ^2 검정결과, χ^2 통계량은 41.289($p=.001$)로 통계적 유의수준 하에서 유의미한 관련성이 있는 것으로 나타났다.

계량은 57.559($p=.000$)으로 통계적으로 의미 있는 관련성이 있는 것으로 밝혀졌다.

사전준비2 중 가장 높은 분포를 보이는 문항은 3건 내지 5건($n=16$, 30.8%)이고, 만족조건설정1 중 가장 높은 빈도수를 보인 문항은 3회 내지 5회($n=17$, 32.7%)로 나타났다. 또한 두 문항간의 관련성은 $n=8$, 15.4%의 수치로 가장 높게 나타났다.

[표 2-19]에 반영되는 Q3-10의 선택 지문은 다음과 같다.

- ① 1건 내지 2건 ② 3건 내지 5건
- ③ 6건 내지 10건 ④ 10건 이상

[표 2-19]에 반영되는 Q3-12의 선택 지문은 다음과 같다.

- ① 2회 이하 ② 3회 내지 5회
- ③ 6회 내지 10회 ④ 10회 이상

[표 2-19] Q3-10과 Q3-12 간의 교차분석 결과표

			만족조건설정1(Q3-12)						x^2/p
			①	②	③	④	기타	전체	
사전 준비2 (Q3-10)	①	빈도	0	0	0	1	0	1	57.559/ .000**
		전체	.0%	.0%	.0%	1.9%	.0%	1.9%	
	②	빈도	3	8	1	4	0	16	
		전체	5.8%	15.4%	1.9%	7.7%	.0%	30.8%	
	③	빈도	2	5	8	0	0	15	
		전체	3.8%	9.6%	15.4%	.0%	.0%	28.8%	
	④	빈도	1	4	2	4	0	11	
		전체	1.9%	7.7%	3.8%	7.7%	.0%	21.2%	
	기타	빈도	0	0	1	1	7	9	
		전체	.0%	.0%	1.9%	1.9%	13.5%	17.3%	
전체	빈도	6	17	12	10	7	52		
	전체	11.5%	32.7%	23.1%	19.2%	13.5%	100.0%		
* $p<.05$, ** $p<.01$									

(15) Q3-12와 Q3-4 간의 χ^2 검정결과

Q3-12의 만족조건설정1과 Q3-4의 AI플랫폼의 맞춤형 설계요건2 간의 χ^2 검정결과, χ^2 통계량은 14.241($p=.286$)로 통계적으로 의미 있는 관련성이 없었다. 만족조건설정1 중 가장 높은 분포를 보이는 문항은 3회 내지 5회($n=17$, 32.7%)이고, AI플랫폼의 맞춤형 설계요건2 중 가장 높은 분포를 보인 문항은 추출된 정보 글들의 공통분모를 찾아 카테고리별로 분류($n=25$, 48.1%)로 밝혀졌다.

또한, 두 문항간의 관련성은 $n=4$, 7.7%의 수치로 나타났으며, 이 보다 더 높은 빈도수를 보인 문항들은 6회 내지 10회와 추출된 정보 글들의 공통분모를 찾아 카테고리별로 분류($n=7$, 13.5%)와 2회 이하/10회 이상과 추출된 정보 글들의 공통분모를 찾아 카테고리별로 분류($n=5$, 9.6%)로 밝혀졌다. <표 15>에서 가장 높은 관련성을 보인 문항은 3회 내지 5회와 유저의 산업군과 추출된 정보글 간의 연관성 판단($n=8$, 15.4%)으로 나타났다.

[표 2-20]에 반영되는 Q3-12의 선택 지문은 다음과 같다.

- ① 2회 이하
- ② 3회 내지 5회
- ③ 6회 내지 10회
- ④ 10회 이상

[표 2-20]에 반영되는 Q3-4의 선택 지문은 다음과 같다.

- ① 유저의 산업군과 추출된 정보글 간의 연관성 판단
- ② 추출된 정보글들의 공통분모를 찾아 카테고리별로 분류
- ③ 유저의 산업군에 속한 현재 관심도가 높은 정보글 보여주기
- ④ 과거 이슈화 되었던 검색어가 기재된 정보글 보여주기

			AI플랫폼의 맞춤형 설계요건2(Q3-4)					x^2/p
			①	②	③	④	전체	
만족 조건 설정1 (Q3-12)	①	빈도	0	5	0	1	6	14.241/ .286
		전체	.0%	9.6%	.0%	1.9%	11.5%	
	②	빈도	8	4	4	1	17	
		전체	15.4%	7.7%	7.7%	1.9%	32.7%	
	③	빈도	3	7	2	0	12	
		전체	5.8%	13.5%	3.8%	.0%	23.1%	
	④	빈도	1	5	3	1	10	
		전체	1.9%	9.6%	5.8%	1.9%	19.2%	
	기타	빈도	1	4	1	1	7	
		전체	1.9%	7.7%	1.9%	1.9%	13.5%	
전체	빈도	13	25	10	4	52		
	전체	25.0%	48.1%	19.2%	7.7%	100.0%		
* $p < .05$, ** $p < .01$								

Q3-12의 만족조건설정1과 Q3-14의 만족조건설정2 간의 χ^2 검정결과, χ^2 통계량은 17.736(p=.038)으로 통계적 유의수준 하에서 유의미한 관련성이 있는 것으로 밝혀졌다.

만족조건설정1 중 가장 높은 빈도수를 보이는 문항은 3회 내지 5회 (n=16, 36.4%)이고, 만족조건설정2 중 가장 높은 분포를 보인 문항은 10 개 내지 30개 미만(n=21, 47.7%)로 나타났다. 또한 두 문항간의 관련성은 n=9, 20.5%의 수치로 가장 높게 나타났다.

① 2회 이하 ② 3회 내지 5회
③ 6회 내지 10회 ④ 10회 이상

① 10개 미만 ② 10개 내지 30개 미만
③ 30개 내지 50개 미만 ④ 50개 이상

			만족조건설정2(Q3-14)					x^2/p
			①	②	③	④	전체	
만족 조건 설정1 (Q3-12)	①	빈도	4	1	1	0	6	17.736/ .038*
		전체	9.1%	2.3%	2.3%	.0%	13.6%	
	②	빈도	7	9	0	0	16	
		전체	15.9%	20.5%	.0%	.0%	36.4%	
	③	빈도	0	9	2	1	12	
		전체	.0%	20.5%	4.5%	2.3%	27.3%	
	④	빈도	4	2	3	1	10	
		전체	9.1%	4.5%	6.8%	2.3%	22.7%	
	기타	빈도	0	0	0	0	0	
		전체	.0%	.0%	.0%	.0%	.0%	
전체	빈도	15	21	6	2	44		
	전체	34.1%	47.7%	13.6%	4.5%	100.0%		

Q3-13의 산업군별 오차범위 조절과 Q3-12의 만족조건설정1 간의 χ^2 검정결과, χ^2 통계량은 44.824 ($p=.001$)로 통계적으로 매우 유의미한 관련성이 있는 것으로 나타났다.

[표 2-22]에 반영되는 Q3-12의 선택 지문은 다음과 같다.

- [표 2-22]에 반영되는 Q3-13의 선택 지문은 다음과 같다.

- 57 -

- ② ICT, IoT, AI, 가전, 로봇, 정보처리 분야
- ③ 기계, 메카트로닉스, 자동차, SoC 분야
- ④ SW, 법률, 플랫폼 서비스, 문화예술 분야
- ⑤ 농수산, 유통물류, 조선해양, 항공우주 분야

[표 2-22] Q3-13과 Q3-12 간의 교차분석 결과표

			만족조건설정1(Q3-12)						x^2/p
			①	②	③	④	기타	전체	
산업 군별 오차 범위 조절 (Q3-13)	①	빈도	0	4	2	3	0	9	44.824 / .001**
		전체	.0%	7.8%	3.9%	5.9%	.0%	17.6%	
	②	빈도	3	11	7	4	0	25	
		전체	5.9%	21.6%	13.7%	7.8%	.0%	49.0%	
	③	빈도	1	2	1	1	0	5	
		전체	2.0%	3.9%	2.0%	2.0%	.0%	9.8%	
	④	빈도	0	0	2	1	2	5	
		전체	.0%	.0%	3.9%	2.0%	3.9%	9.8%	
	⑤	빈도	1	0	0	0	5	6	
		전체	2.0%	.0%	.0%	.0%	9.8%	11.8%	
	기타	빈도	0	0	0	1	0	1	
		전체	.0%	.0%	.0%	2.0%	.0%	2.0%	
	전체	빈도	5	17	12	10	7	51	
		전체	9.8%	33.3%	23.5%	19.6%	13.7%	100.0%	
* $p < .05$, ** $p < .01$									

(18) Q3-14와 Q3-16 간의 χ^2 검정결과

Q3-14의 만족조건 설정2와 Q3-16의 만족조건 설정3 간의 χ^2 검정결과, χ^2 통계량은 9.872($p=.361$)로 통계적으로 의미 있는 관련성이 나타나지 않았다.

만족조건 설정2 중 가장 높은 빈도수를 보이는 문항은 10개 내지 30개 미만($n=21$, 47.7%)이고, 만족조건설정3 중 가장 많은 분포를 보인 문항은 과거부터 현재까지 이슈로 상위검색 순위에 여러 차례 올라오는 정보글이

Q3-14의 만족조건 설정2와 Q3-7의 AI 플랫폼의 맞춤형 설계요건5 간의 χ^2 검정결과, χ^2 통계량은 5.945 ($p=.745$)로 통계적으로 유의미한 관련성이 없었다.

만족조건 설정2 중 가장 높은 분포를 보인 문항은 10개 내지 30개 미만 ($n=21$, 48.8%)이고, AI 플랫폼의 맞춤형 설계요건5에서 가장 높은 빈도수를 보인 문항은 주제어들에 대한 최종 선정 기준을 산술 평균치 이외 다른 기준치로 채택($n=16$, 37.2%)으로 나타났다. 또한 두 문항간의 관련성은 $n=7$, 16.3%로 가장 높게 나타났다.

[표 2-24]에 반영되는 Q3-14의 선택 지문은 다음과 같다.

- ① 10개 미만
- ② 10개 내지 30개 미만
- ③ 30개 내지 50개 미만
- ④ 50개 이상

[표 2-23]에 반영되는 Q3-7의 선택 지문은 다음과 같다.

- ① 추출된 정보글에 표현된 긍정, 부정, 중립 여부에 대한 판단기준 모호
- ② 주제어들에 대한 최종 선정 기준을 산술 평균치 이외 다른 기준치로 채택
- ③ 기존 오피니언 마이닝 기법에 형용사를 더 추가시킨 이유에 대한 적합한 설명
- ④ 최종 선택된 주제어들의 자연어처리 및 자국어로 표현될 특허회피 전략의 언어 구사력

[표 2-24] Q3-14와 Q3-7 간의 교차분석 결과표

			AI플랫폼의맞춤형설계요건5(Q3-7)					x^2/p
			①	②	③	④	전체	
만족 조건 설정2 (3-14)	①	빈도	3	4	5	2	14	5.945/ .745
		전체	7.0%	9.3%	11.6%	4.7%	32.6%	
	②	빈도	4	7	6	4	21	
		전체	9.3%	16.3%	14.0%	9.3%	48.8%	
	③	빈도	1	4	1	0	6	
		전체	2.3%	9.3%	2.3%	.0%	14.0%	
	④	빈도	0	1	0	1	2	
		전체	.0%	2.3%	.0%	2.3%	4.7%	
	전체	빈도	8	16	12	7	43	
		전체	18.6%	37.2%	27.9%	16.3%	100.0%	
* $p<.05$, ** $p<.01$								

(20) Q3-16과 Q3-4 간의 χ^2 검정결과

Q3-16의 만족조건 설정3과 Q3-4의 AI 플랫폼의 맞춤형 설계요건2 간의 χ^2 검정결과, χ^2 통계량은 13.963($p=.124$)로 통계적 유의수준 하에서 의미 있는 관련성을 보이지 않았다. 만족조건설정3 중 가장 높은 분포를 보인 문항은 과거부터 현재까지 이슈로 상위검색 순위에 여러 차례 올라오는 정보글이어야 함($n=20$, 44.4%)이고, AI 플랫폼의 맞춤형 설계요건2에서 가장 높은 빈도수를 보인 문항은 추출된 정보글들의 공통분모를 찾아 카테고리별로 분류($n=21$, 46.7%)로 밝혀졌다. 또한 두 문항간의 관련성은 $n=10$, 22.2%로 모든 문항들과 비교해서 가장 높은 수치를 보였다.

[표 2-25]에 반영되는 Q3-4의 선택 지문은 다음과 같다.

- ① 유저의 산업군과 추출된 정보글 간의 연관성 판단
- ② 추출된 정보글들의 공통분모를 찾아 카테고리별로 분류
- ③ 유저의 산업군에 속한 현재 관심도가 높은 정보글 보여주기
- ④ 과거 이슈화 되었던 검색어가 기재된 정보글 보여주기

[표 2-25]에 반영되는 Q3-16의 선택 지문은 다음과 같다.

- ① 과거부터 현재까지 이슈로 상위검색순위에 여러 차례 올라오는 정보글이어야 함(10년 기준)
- ② 현재 이슈로, 2-3회 정도 상위검색순위에 올라온 정보글이어야 함(10년 기준)
- ③ 과거 이슈로, 상위 검색순위에 여러 차례 올라온 정보글이어야 함(10년 기준)
- ④ 과거 이슈로, 2-3회 정도에 한해 상위 검색순위에 올라왔던 정보글이어야 함(10년 기준)

[표 2-25] Q3-16과 Q3-4 간의 교차분석 결과표

			AI플랫폼의맞춤형설계요건2(Q3-4)					x^2/p
			①	②	③	④	전체	
만족 조건 설정3 (Q3-16)	①	빈도	6	10	1	3	20	13.963 / .124
		전체	13.3%	22.2%	2.2%	6.7%	44.4%	
	②	빈도	2	6	6	0	14	
		전체	4.4%	13.3%	13.3%	.0%	31.1%	
	③	빈도	4	3	1	0	8	
		전체	8.9%	6.7%	2.2%	.0%	17.8%	
	④	빈도	0	2	1	0	3	
		전체	.0%	4.4%	2.2%	.0%	6.7%	
	전체	빈도	12	21	9	3	45	
		전체	26.7%	46.7%	20.0%	6.7%	100.0%	
* $p<.05$, ** $p<.01$								

2.6.3 상관분석

상관관계란 변수들 간의 관계를 말하는 것으로, 두 개 이상의 변수에 있어서 한 변수가 변화함에 따라 다른 변수가 어떻게 변화하는지와 같은 변화의 강도와 방향을 상관관계라고 한다.

상관관계의 정도는 0에서 ± 1 사이로 나타나며, ± 1 에 가까울수록 상관

관계는 높아지고 0에 가까울수록 상관관계는 낮아진다. 즉, 변화의 강도는 절대값에 1에 가까울수록 높고, 변화의 방향은 +는 정의 방향, -는 음의 방향이라고 한다.

상관관계 분석은 변수들 간의 관련성을 나타내는 것임으로, 인과관계로 파악해서는 안되며, 변수들 간의 인과관계는 독립변수가 종속변수에 어떠한 영향을 미치는지를 파악하는 것이어서, 상관관계분석은 독립변수와 종속변수와의 영향관계가 아니라 두 변수 간의 관련성의 정도와 방향을 보여주는 것이다.⁵⁵⁾

[표 2-23], [표 2-24]은 본 논문에서 사용한 모든 변수 간 상관관계의 수준 및 방향성을 측정한 피어슨 상관관계 분석 결과이다.

상관관계 분석은 변수들 간의 관계의 강도에 따른 관련성에 대한 대체적인 윤곽을 제시해 준다. 또한, 상관계수의 정도는 절대치를 기준으로 0.7이상~0.9미만 높은 상관관계, 0.4이상~0.7미만 다소 높은 상관관계, 0.2이상~0.4미만 낮은 상관관계, 마지막으로 0.2미만은 상관관계 거의 없다고 한다.

통계적으로 의미 있는 관련성이 있는 변수들 중 가장 높은 상관계수 값을 보인 변수는 사전준비2와 산업군별 오차범위 조절 간 상관계수인 0.580이고, 가장 낮은 상관계수 값은 사전준비2와 미래의 AI 플랫폼의 요구사항간 0.274로 나타났다.

55) 이일현. (2016). 『EasyFlow 회귀분석』. 서울 : 한나래 아카데미.

[표 2-26] Q3-10 기준의 상관관계 분석결과

구분	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1.00																		
2	.166	1.00																	
3	.012	.229	1.00																
4	.150	.179	.184	1.00															
5	.092	-.197	.180	.025	1.00														
6	.169	.100	.136	-.072	.394**	1.00													
7	-.021	.064	.161	-.099	.340*	.363**	1.00												
8	.168	.018	.065	-.153	.155	.116	.071	1.00											
9	-.045	-.150	-.030	-.260	.140	.151	-.045	.297*	1.00										
10	-.027	.111	.008	.142	.069	-.040	-.293*	.065	.179	1.00									
11	-.055	-.181	.486**	.114	.144	.197	.104	-.184	.031	-.093	1.00								
12	-.091	-.148	.285*	.128	.242	.212	.220	-.053	.112	.029	.315*	1.00							
13	.121	.112	-.082	.123	.118	.298*	.316*	-.203	-.165	-.237	.005	-.020	1.00						
14	.094	.420**	-.089	-.066	.255	.340*	.218	.172	.019	.055	.004	-.197	.523**	1.00					
15	.132	.383**	.048	-.068	.207	.260	.152	.169	.118	-.016	.025	-.128	.173	.520**	1.00				
16	.041	.293*	.021	-.254	.170	.303*	.236	-.090	.004	.131	.042	-.066	.175	.580**	.462**	1.00			
17	.012	.076	-.181	-.073	.177	.114	-.037	-.180	-.158	-.367*	.098	-.010	.510**	.450**	.343*	.213	1.00		
18	.138	.030	.012	.190	-.072	.095	-.013	.094	-.045	.072	-.083	.007	-.052	-.267	.304*	-.033	-.050	1.00	
19	.437**	.291*	.142	.124	.403**	.361**	.241	.188	.178	-.010	-.029	.088	.156	.274*	.392**	.166	.081	.066	1.00
1. 지식재산권의 필요성 2. 지식재산권의 영향력 3. 특허회피설계의 필요성 4. 특허회피설계 사전필수요건 5. 기존 특허회피설계의 문제점과 대안 6. 총래방식 대비 AI 플랫폼의 가치성 7. AI 플랫폼이 갖춰야 할 필수요건 8. AI 플랫폼의 맞춤형 설계요건1 9. AI 플랫폼의 맞춤형 설계요건2 10. AI 플랫폼의 맞춤형 설계요건3 11. AI 플랫폼의 맞춤형 설계요건4 12. AI 플랫폼의 맞춤형 설계요건5 13. 사전 준비1 14. 사전 준비2 15. 만족 조건 설정1 16. 산업군별 오차범위 조절 17. 만족 조건 설정2 18. 만족 조건 설정3 19. 미래의 AI 플랫폼의 요구사항																			

[표 2-27] 유의한 관점에서 본 상관관계 분석결과

구분	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1.00																		
2	.166	1.00																	
3	.012	.229	1.00																
4	.150	.179	.184	1.00															
5	.092	-.197	.180	.025	1.00														
6	.169	.100	.136	-.072	.394 ^{**}	1.00													
7	-.021	.064	.161	-.099	.340 ^{**}	.363 ^{**}	1.00												
8	.168	.018	.065	-.153	.155	.116	.071	1.00											
9	-.045	-.150	-.030	-.260	.140	.151	-.045	.297 [*]	1.00										
10	-.027	.111	.008	.142	.069	-.040	.293 [*]	.065	.179	1.00									
11	-.055	-.181	.486 ^{**}	.114	.144	.197	.104	-.184	.031	-.093	1.00								
12	-.091	-.148	.285 [*]	.128	.242	.212	.220	-.053	.112	.029	.316 [*]	1.00							
13	.121	.112	-.082	.123	.118	.298 [*]	.316 [*]	-.203	-.165	-.237	.005	-.020	1.00						
14	.094	.420 ^{**}	-.089	-.066	.255	.340 [*]	.218	.172	.019	.055	.004	-.197	.623 ^{**}	1.00					
15	.132	.383 ^{**}	.048	-.068	.207	.260	.152	.169	.118	-.016	.025	-.128	.173	.520 ^{**}	1.00				
16	.041	.293 [*]	.021	-.254	.170	.303 [*]	.236	-.090	.004	.131	.042	-.066	.175	.580 ^{**}	.462 ^{**}	1.00			
17	.012	.076	-.181	-.073	.177	.114	-.037	-.180	-.158	-.367 [*]	.098	-.010	.510 ^{**}	.450 ^{**}	.343 [*]	.213	1.00		
18	.138	.030	.012	.190	-.072	.095	-.013	.094	-.045	.072	-.083	.007	-.052	-.267	.304 [*]	-.033	-.050	1.00	
19	.437 ^{**}	.291 [*]	.142	.124	.403 ^{**}	.361 ^{**}	.241	.188	.178	-.010	-.029	.088	.156	.274 [*]	.392 ^{**}	.166	.081	.066	1.00

14. 사전준비2는 2, 6, 13, 15, 16, 17, 19와 통계적으로 유의한 상관관계가 있으며, 그 중에 16과의 상관관계가 가장 높으며(.580^{**}) 19와의 상관관계가 가장 낮게(.274^{*}) 나타남

파란색 표시도 모두 통계적으로 유의한 상관관계를 나타냄
추가로 상관관계분석은 변수들 간의 방향성과 밀접정도를 보여주는 것
인과관계를 나타내는 것이 아님

1.지식재산권의 필요성
6.종래방식 대비 시틀래품의 가치성
11.시틀래품의 맞춤형 설계요건4
16.산업군별 최적 임무조절

2.지식재산권의 영향력
7.시틀래품이 갖춰야 할 필수요건
12.시틀래품의 맞춤형 설계요건5
17.만족조건별정2

3.특허회피 설계의 필요성
8.시틀래품의 맞춤형 설계요건1
13.사전준비1
18.만족조건별정3

4.특허회피 설계 사전 필수요건
9.시틀래품의 맞춤형 설계요건2
14.사전준비2
19.미래의 시틀래품의 요구사항

5.기존 특허회피 설계의 문제점과 대안
10.시틀래품의 맞춤형 설계요건3
15.만족조건별정1

** 상관이 0.01 수준에서 유의
* 상관이 0.05 수준에서 유의

Ⅲ. 봇 기반 특허회피 설계안

자동 생성 방법

3.1 자동 생성 방법 소개

본 논문은 기업이나 개인 고객들이 찾고자 하는 신동력 사업을 적극 추진하는데 동기 부여하고, 이에 따른 발명 아이디어를 발굴해 내는데 있어서도 도움이 될 특허회피 설계안을 봇을 통해 자동 생성하는 방법에 관한 연구이다.

본 논문은 아이디어 발굴 또는 기업에서 요구하는 신사업 추진에 있어 도움이 될 만한 특허회피 설계안을 봇을 통해 고객에게 제시해 줌으로써, 기업발전, 산업발전 및 사회발전에 크게 이바지하기 바라며, 향후 중소기업이나 개인 고객들이 우량 기업으로 성장하는데 교두보 역할을 해 줄 공백기술 발굴 및 그에 따른 지식재산권 확보, 기술노하우 법적보호 및 틈새시장의 선점에 따른 기업이윤 창출에 대한 기대감을 고무시켜 주는데 있어 도움을 되기 바란다.

본 논문은 특허회피 설계안을 필요로 하는 고객을 대상으로 봇 기반 특허회피 설계안 자동 생성 방법을 제안하는 연구에 관한 것으로, 봇을 구동시켜 키프리스와 포탈에 접근해 우선순위가 낮은 정보글(또는 우선순위가 가장 낮은 정보글) 중에 후보 주제어를 대상으로 하여 1차 필터링(유사성 판단), 2차 필터링(연관성 판단), 3차 필터링(적합성 판단)을 실시해 유사성 통과 주제어 및 연관성 통과 주제어를 색출한다.

그런 다음, 본 논문은 1, 2, 3차 필터링을 거쳐 선정된 적합성 통과 주제어인 최종 주제어를 대상으로 대화형 구문형식에 맞도록 단어 조합하는 자연어 처리함으로써, (비)강화 학습된 특허회피 설계안을 제시하는 봇 기반 특허회피 설계안 자동 생성 방법을 제안하고자 한다.〈그림 3-1 참조〉



[그림 3-1] 방법 알고리즘의 개략도

본 논문에서 소개되는 봇은 다음과 같은 프로세싱을 통해 좀 더 구체적인 동작을 실시한다.

먼저, 본 논문의 봇은 2가지 경우에 입각해 각기 다른 동작을 실시한다. Case 1은 핵심 키워드를 직접 포털 사이트에 입력해 핵심 키워드와 연관된 정보를 색출하는 방안이고, Case 2는 핵심 키워드가 아닌 N개의 선행특허 내 기재된 특허청구범위를 크롤링한 후 이로부터 색출된 특허 키워드로 포털 사이트에 접근해 포털 사이트에서 제공하는 유용한 정보를 색출하는 방안이다.

본 논문의 봇 기반 특허회피 설계안 자동 생성 방법은 특허 키워드로 정보를 색출해 이 정보로부터 수집된 최종 주제어로 특허회피 설계안을 제시하는 Case 2의 경우를 대표연구로 한다.

봇은 Case 1과 Case 2에 따라 각각 달리 구분 동작하며, 하기 내용은 Case 1을 설명한 것으로 본 논문에서 소개하는 대표연구는 아니나, 본 논문의 대표연구와 관련성이 있어 간단히 소개한다.

Case 1 :

- ① 핵심 키워드가 봇에 입력됨
- ② 봇이 특허정보 검색서비스(www.kipris.or.kr) 접속

- ③ 봇은 공개 및 등록된 다수의 특허 건 조회 및 자동 검색
- ④ 봇은 다수의 특허 중 N개의 선행특허 추출
- ⑤ 봇은 핵심 키워드, N개의 선행특허 기록 및 파일 저장
- ⑥ 봇은 핵심 키워드를 ‘네이버’, ‘구글’, ‘다음카카오’, ‘네이트’, ‘빙’ 과 같은 포털 사이트에 입력 후 자동 검색
- ⑦ 봇은 포털 사이트에 올라온 다수의 정보들 중 핵심 키워드가 많은 기재된 M개의 정보글 추출
- ⑧ 봇은 N개의 선행특허를 하나씩 오픈해 특허청구범위 크롤링 실시
- ⑨ 봇은 반복 체크된 용어를 특허 키워드로 선정
- ⑩ 봇은 ⑦에서 취합된 M개의 정보글로부터 특허 키워드의 기재 상태를 체크
- ⑪ 봇은 특허 키워드가 가장 적게 표기된 1개의 정보글 혹은 비교적 적게 기재된 P개의 정보글만 추려냄
- ⑫ 봇은 N개의 선행특허 수만큼 ⑧⑨⑩⑪을 반복 실행
- ⑬ 봇은 ⑫를 통해 얻어진 1개의 정보글 또는 P개의 정보글로부터 최종 주제어를 뽑아냄
- ⑭ 봇은 최종 주제어가 모두 들어갈 수 있도록 자연어 처리해 자국어 혹은 영어로 표출되는 특허회피 설계안을 제시함

다음은 본 논문의 특허회피 설계안 자동 생성 방법에 관한 연구를 대표하는 Case 2에 대한 구체적인 설명이다.

먼저, 특허회피 설계안을 필요 요청하는 고객(ex : 출원인, 발명자, 고안자, 창작자 등)은 우선적으로 본업을 영위하는데 중요한 현 업무분야 혹은 타 기업들과 체결한 여러 계약 프로젝트와 상관되는 적어도 하나 이상의 검색어를 찾아 골라낸 후 이를 핵심 키워드로 선정한다.

봇은 고객으로부터 핵심 키워드를 전달받아 이미 연계해 놓은 특허청 정보검색 사이트인 키프리스(www.kipris.or.kr)에 접근해 핵심 키워드를 검색 창에 입력하고 특허 검색을 실시한다.

다시 말해, 특허회피 설계전략을 필요로 하는 고객은 우선적으로 자신의

관심분야 혹은, 업무상 현재 진행하고 있는 중요한 계약 프로젝트와 관련된 핵심 키워드를 선정해 봇에 입력한다.

봇은 이미 연계해 놓은 키프리스에 접속해 핵심 키워드를 검색 창에 입력하고 특히 검색을 실시하여 등록되거나 공개된 다수의 특허 건들을 조회함을 통해 N개의 선행특허를 추출하며, 키프리스로부터 추출된 다수의 특허 건들 중 불필요한 특허들을 제외하기 위해 핵심 키워드를 토대로 관심분야에 속해 있거나 현 업무상 연관성이 높은 N개의 선행특허만을 선별한 다음, 고객으로부터 전달된 핵심 키워드와 키프리스를 통해 획득한 N개의 선행특허를 기록 저장한다.

봇은 N개의 선행특허를 하나씩 열어 문헌 내 기재되어 있는 발명의 명칭, 요약, 발명의 효과(목적), 발명의 상세한 설명, 특허청구범위를 확인한 후, 이 중 발명에 대한 법적 권리를 나타내는 특허청구범위를 크롤링함에 따라 특허청구범위에서 반복적으로 표현된 용어들을 색출해 특허 키워드로 선정한다.

봇은 특허 키워드를 포털 사이트(대표적으로는 네이버, 구글, 다음카카오, 네이트, 빙 등)에 입력해 M개의 정보글을 추출한 후, 포털 사이트로부터 확보된 M개의 정보글 중 특허 키워드의 기재 상태를 1:1 매칭 선별동작을 통해 체크한 다음, M개의 정보글(1, 2, ..., M순위) 중에서 미리 가정해 놓은 조건식에 부합하는 특허 키워드가 가장 적게 기재된 1개의 정보글 혹은 비교적 적게 기재된 P개의 정보글만을 추려낸다.

봇은 N개의 선행특허 수만큼 상기 동작을 동일하게 반복 실행한다.

본 논문의 봇 기반 특허회피 설계안 자동 생성 방법은 특허 키워드가 가장 적게 기재된 1개의 정보글 혹은 비교적 적게 기재된 P개의 정보글 중 후보 주제어를 색출하고, 후보 주제어를 대상으로 1, 2, 3차 필터링을 실시함으로써 “특허회피 설계안을 제시하기 위한 목적성을 갖는 명사”라 호칭할 만한 최종 주제어를 선정한다.

본 논문의 봇 기반 특허회피 설계안 자동 생성 방법은 최종 주제어가 모두 들어갈 수 있도록 어휘력(전제 : 이미 언어 학습능력을 기본적으로

갖춘 인공지능형 봇을 기반으로 함 : 자연어 처리기술 탑재 완비)을 구사해 자국어 혹은 영어로 표출되는 특허회피 설계안을 제시한다. 이때, 봇은 인간이 인지할 수 있는 적합한 언어 구사력을 동원해 영상(아바타 출현), 음성으로도 디스플레이함도 지원 가능할 것이다.

이하 기재된 Case 2는 본 논문의 대표연구인 봇 기반 특허회피 설계안 자동 생성 방법을 좀 더 쉽게 이해할 수 있도록 스텝별로 구분해 정리한 내용이다.

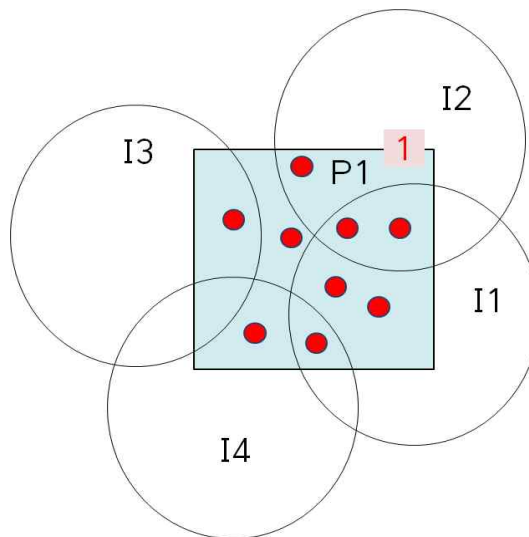
Case 2 :

- ① 핵심 키워드가 봇에 입력됨
- ② 봇이 특허정보 검색서비스(www.kipris.or.kr) 접속
- ③ 봇은 공개 및 등록된 다수의 특허 건 조회 및 자동 검색
- ④ 봇은 다수의 특허 중 N개의 선행특허 추출
- ⑤ 봇은 핵심 키워드, N개의 선행특허 기록 및 파일 저장
- ⑥ 봇은 N개의 선행특허를 하나씩 오픈해 특허청구범위 크롤링 실시
- ⑦ 봇은 반복 체크된 용어를 특허 키워드로 선정
- ⑧ 봇은 특허 키워드를 포털 사이트의 검색창에 입력해 다수의 정보글을 검색하고 다수의 정보글 중에서 M개의 정보글을 추출
- ⑨ 봇은 M개의 정보글 중 특허 키워드의 기재 상태를 1:1 매칭 선별동작을 통해 체크한 다음, 미리 정해놓은 가정 조건식에 부합하는 1개의 정보글 혹은 비교적 적게 기재된 P개의 정보글만을 추려냄
- ⑩ 봇은 N개의 선행특허 수만큼 ⑥⑦⑧⑨을 반복 실행
- ⑪ 봇은 ⑩를 통해 얻어진 1개의 정보글 또는 P개의 정보글로부터 색출된 후보 주제어를 대상으로 1, 2, 3차 필터링 후 최종 주제어를 선정함
- ⑫ 봇은 최종 주제어가 모두 들어갈 수 있도록 어휘력(단, 자연어처리 기법들로부터 이미 언어 학습능력을 갖춘 봇임을 전제함)을 구사해 모국어 혹은 영어로 표현되는 특허회피 설계안을 제시함

3.2. 봇의 가상 모델

본 논문은 특허회피 설계안 자동생성 봇에 대한 프로세싱을 좀 더 쉽게 접근할 수 있기 위해 가상의 구현 모델을 [그림 3-2], [그림 3-3], [그림 3-4]와 같이 제시하고, 각 단계별로 동작되는 세부 과정들을 이해해 봇 기반 특허회피 설계안 자동 생성 방법에 대한 개념을 파악한다.

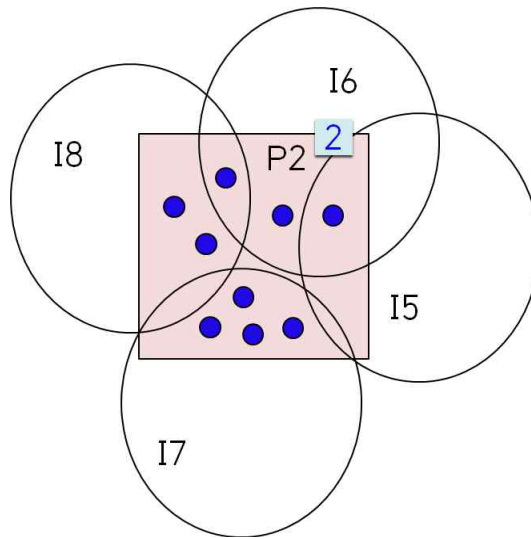
- Step 1 : 봇에 핵심 키워드 입력
- Step 2 : 봇은 키프리스에서 N개의 선행특허 추출
- Step 3 : 이 중 샘플특허(P1) 선정
- Step 4 : 샘플특허로(P1)부터 특허 키워드(●) 색출
- Step 5 : 포털 사이트의 검색창에 특허 키워드 입력
- Step 6 : 다수의 정보글(I1~I4) 제공
- Step 7 : 이 중 특허 키워드가 적게 기재된 정보글(I3)을 특허회피 설계안 제시 시 적용되는 빅데이터(I3)로 정함
- Step 8 : 빅데이터(I3)의 특허 키워드를 최종 주제어로 설정(가정)



[그림 3-2] 봇의 제 1 가상 모델

- Step 1' : 봇에 핵심 키워드 입력

- Step 2' : 봇은 키프리스에서 N개의 선행특허 추출
 Step 3' : 이 중 샘플특허(P2) 선정
 Step 4' : 샘플특허로(P1)부터 특허 키워드(●) 색출
 Step 5' : 포털 사이트의 검색창에 특허 키워드 입력
 Step 6' : 다수의 정보글(I5~I8) 제공
 Step 7' : 이 중 특허 키워드가 적게 기재된 정보글(I5)을 특허회피 설계안 제시 시 적용되는 빅데이터(I5)로 정함
 Step 8' : 빅데이터(I5)의 특허 키워드를 최종 주제어로 설정(가정)

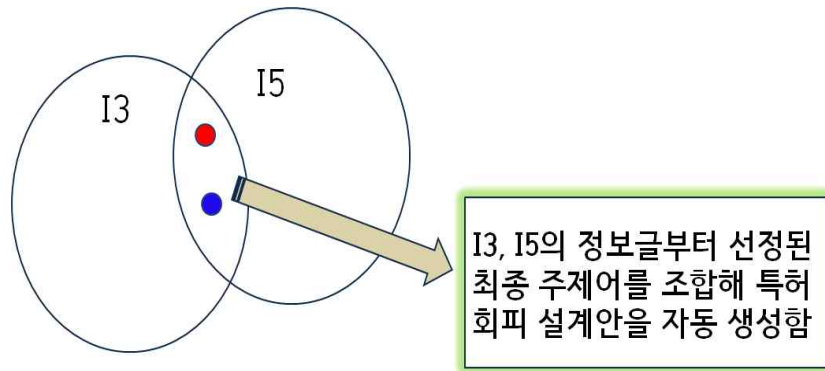


[그림 3-3] 봇의 제 2 가상 모델

Step 8과 Step 8' 진행 후, Step 9가 이어 진행된다. 즉, 최종 주제어는 정보글(I3)과 정보글(I5)로부터 추출된 것임으로, 봇은 완전한 문장으로 형성된 특허회피 설계안을 제시해 주기 위해 정보글(I3)과 정보글(I5) 내 기재된 형태소, 단어 또는 여러 품사(동사(문장 내 서술어), 접속사, 부사, 형용사)를 검출한다. 그 다음, 봇은 최종 주제어를 기준으로 최종 주제어와 가까운 위치에 있는 형태소, 단어 또는 여러 품사를 가져와 언어적 1:1 매칭 동작을 실시해 최종 주제어가 모두 들어갈 수 있도록 자연어 처리한다.

그에 따라, 봇은 [그림 3-4]와 같이, 완전한 문장으로 형성된 특허회피 설계안을 자동 생성시켜 제시한다. 단, 특허 키워드는 후보 주제어에 속하

면서 최종 주제어에 선정됨을 전제한다.



[그림 3-4] 붓의 가상 모델링 결과

3.3 자동 생성 방법에 관한 알고리즘 설계

3.3.1 전체 알고리즘에 관한 설명

일반적으로, 규칙기반 전문가 시스템은 기반지식, 데이터베이스, 추론 엔진, 해설 설비, 사용자 인터페이스로 이루어지며, 규칙기반 전문가 시스템은 문제를 해결하고자 할 때, 심벌 추론(Symbolic reasoning)을 채택한다. 사실, 개념, 규칙 같은 다른 종류의 지식을 표현하는데 심벌을 사용하며, 지식을 처리하거나 질적인 데이터를 쉽게 다룰 수 있다.⁵⁶⁾

본 논문의 붓 기반 특허회피 설계안 자동 생성 방법에 관한 연구는 규칙기반 전문가 시스템에 적용되는 알고리즘 연구 중 하나이다.

본 논문의 전체 알고리즘에 관한 플로우차트는 [그림 3-5]에서 보여지는 바와 같이 구성되며, 이를 좀 더 이해하는데 도움이 되는 구체적인 설명은 다음과 같다.

즉, 고객은 관심분야 또는 회사 프로젝트로부터 정확히 인지되는 핵심 키워드를 키프리스(www.kipris.or.kr)에 입력한 후, 그로부터 추출된 다수

56) 마이클 네그네베타스키. (2017). 『인공지능 개론』. 서울 : 한빛아카데미.

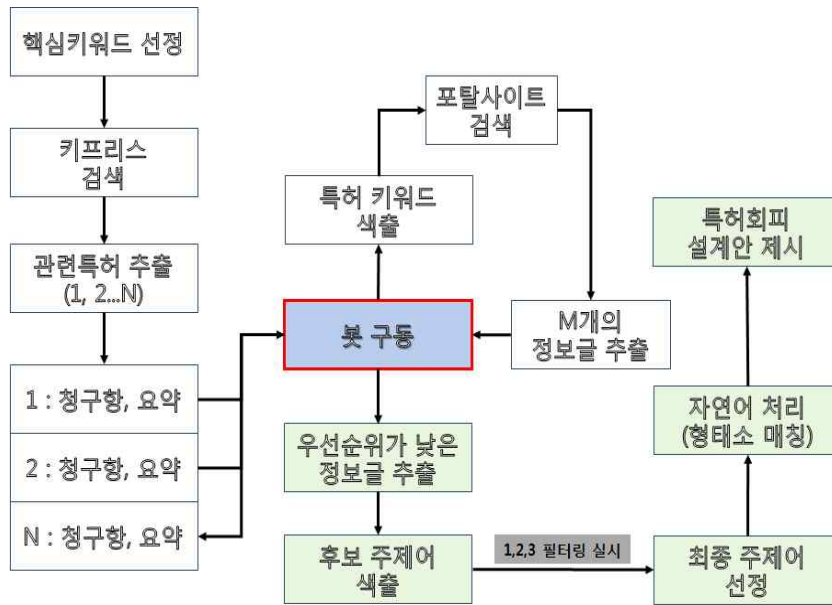
의 특허 건 중 업무 연계성이 높은 N개의 선행 특허를 선택한다.

본 논문의 알고리즘은 N개의 선행 특허를 하나씩 순차적으로 연 후, 그 선행 특허의 기술적 사상을 담고 있는 특허 명세서의 특허청구범위(청구항) 또는 그 선행 특허의 발명의 요지를 빨리 이해할 수 있는 기재된 요약을 자동으로 스캔하고, 스캔된 명사들 중 반복적으로 표현된 용어를 색출해 특허 키워드로 선정한다.

본 논문의 알고리즘은 특허 키워드를 포털 사이트(대표적으로, 구글, 네이버, 다음카카오, 네이트, 빙 등)의 ‘뉴스 영역’ 혹은 ‘지식 iN 영역’에 입력해 자동 검색을 실시함에 따라 포털 사이트에 게재된 많은 정보글 중 특허 키워드가 많이 포함된 M개의 정보글을 추출하며, 포털 사이트로부터 얻은 M개의 정보글 중 특허 키워드가 기재된 상태를 자동 체크하고, M개의 정보글 (1, 2, ..., M순위) 중 우선순위가 낮은 정보글만을 추려낸다.

본 논문의 알고리즘은 M개의 정보글 중 우선순위가 낮은 정보글 중에서 “특허회피 설계안을 제시하는데 용이하게 사용될 명사” 즉, 목적성을 갖는 단어를 후보 주제어(정제되지 않은 raw 데이터)로 색출해 낸 후, 후보 주제어 중 1, 2, 3 필터링에 의해 걸러진 최종 주제어를 선정한 후, 최종 주제어가 모두 들어가 있는 어휘력(전제 요건 : (비)강화 학습을 통해 이미 언어적 능력을 기본적으로 갖추고 있어 최종 주제어로부터 유도된 형태소, 단어, 여러 품사 간의 매칭이 가능한 인공지능형 봇임 : 자연어 처리기술 완비)을 구사해 특허회피 설계안을 제시하는 방법이라 할 것이다.

본 논문의 알고리즘은 특허회피 설계안 제시 시, 미리 구축해 둔 플랫폼을 이용해 자신이 이해할 수 있는 자국어, 영어로 쉽게 표현해 주는 것을 목표로 한다.



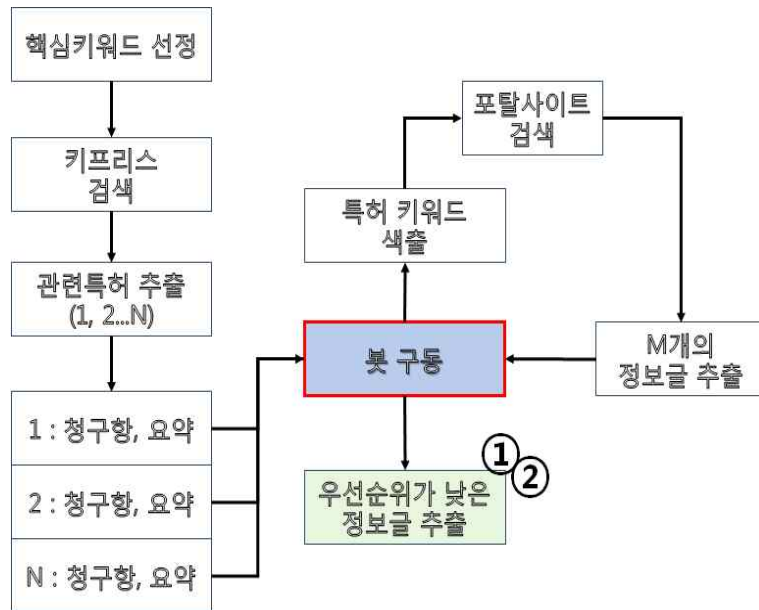
[그림 3-5] 방법 알고리즘의 전체 블록도

3.3.2 정보글 추출

계속해서, 본 논문에 대한 이해를 돕고자 [그림 3-5]의 전체 플로우차트를 좀 더 세부적으로 나눠 순서에 맞게 하나씩 소개한다.

먼저, 정보글 추출은 앞서 [그림 3-5]에서도 설명한 바와 같이, 핵심 키워드를 키프리스(www.kipris.or.kr)에 입력하고 그로부터 추출된 다수의 특허 건 중 업무 연계성이 높은 N개의 선행 특허를 추출한 후, 봇을 이용해 N개의 선행특허마다 기재된 특허청구범위(청구항) 또는 요약을 크롤링하여 특허 키워드를 색출한다.

본 논문의 봇은 [그림 3-6]에서도 알 수 있는 바와 같이, 특허 키워드를 포털 사이트의 검색창에 입력해 M개의 정보글을 추출하고, 시나리오①,② 중 선택된 하나의 방식을 이용해 우선순위가 낮은 정보글을 추출한다.



[그림 3-6] 정보글 추출

봇 구동에 따른 정보글 추출은 [그림 3-7]에서도 알 수 있듯이, 발명의 기술적 사상을 담고 있는 특허명세서의 특허청구범위(청구항) 중 특히 독립항을 중요 블록으로 설정한다는 가정 하에 특허 키워드를 색출하는 동작을 시작한다.

키워드 추출(Keyword Extraction)은 정보검색(Information Retrieval), 문서분류(Text Categorization), 주제탐색(Topic Detection), 문서요약(Document Summarization) 등을 포함하는 텍스트 마이닝 분야의 기본이 되는 기술로써, 키워드 추출기법에 의거해 색출된 특허 키워드는 자연어처리 기법 중 하나인 텍스트 마이닝 기술이 적용된 결과에 가장 중요한 영향을 미친다.

또한, 특허명세서는 일반문서와 다르게 <발명의 명칭>, <도면의 간단한 설명>, <발명의 상세한 설명>, <청구항>과 같은 여러 개의 섹션들로 구성되며, 해당 섹션은 그 섹션명에 맞는 내용으로 구성되어야 하고 각 섹션에 포함되는 용어들은 각 섹션에 맞는 주요 용어들로 배치되어야 한다.

특허명세서로부터 효율적으로 특허 키워드를 추출하기 위해서는 특허명세서의 특성을 잘 반영해야 한다.⁵⁷⁾

특히, 특허청구범위는 등록 전이면 권리요구서로서 출원인이 보호받고자 하는 기술에 대한 독점권 특허청에 요구하는 역할을 하고, 등록 후엔 출원인에게 부여된 독점권을 행사하거나 보호해 주는 막강한 권리서로서의 역할을 하기 때문에 이로부터 특허 키워드를 색출해 내는 것은 특허회피 설계안을 마련하기 위한 중요한 시작점이 된다.

즉, 봇 구동에 따른 정보글 추출은 독립항에서 특허 키워드를 색출하고자 할 시, 특허 키워드를 색출하는 조건으로 첫째, 독립항 내에서 반드시 출현되는 동일한 목적어[O]이어야 하거나/하고, 둘째, 독립항 내 출현된 목적어[O]로써 적어도 5회 이상 중복 기재됨을 만족하는 특허 키워드만을 색출한다.

또한, 봇 구동에 따른 정보글 추출은 특허 키워드와 특허 키워드를 이용해 포털 사이트로부터 검색되어진 M개의 정보글을 1:1 매칭시켜 M개의 정보글이 어떤 특허 키워드에 의해 조회된 정보글인지 여부를 빠르게 선별하는 동작을 실시한다.

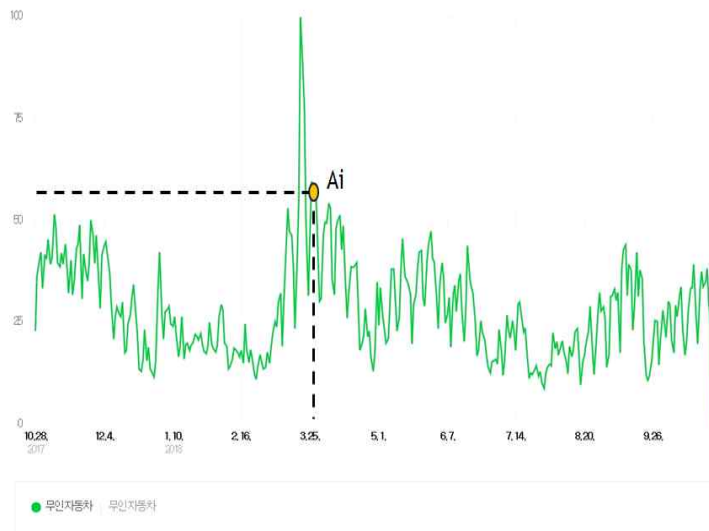
이때, 봇 구동에 따른 정보글 추출은 M개의 정보글 중 (1) : 동일한 목적어[O]가 2개 이하로 표현되어 있는 정보글이어야 하거나/하고, (2) : 그 중 목적어[O]의 중복 기재수가 2~3회로만 표기되어 있는 정보글만 선별한다는 조건 동작을 실시한다.

57) 이순근. (2016). “특허문서의 키워드 추출을 위한 텍스트 마이닝 기반의 관계성 그래프 모델”. 강릉원주대학교 박사학위논문.

나, 대중에게 잘 알려 있지 않거나 잘 모르는 제한된 정보, 대중에게 이슈화되어 널리 알려져 있었으나 이후 소외된 정보, 즉 저평가된 우량 정보를 찾아 이를 반영하는 것이 가장 중요하다 할 것이다.

일례로, 저평가된 우량 정보를 찾는데 있어 해당 업종에 따라 다를 순 있겠으나 [그림 3-8]에서 알 수 있듯이, 일정한 기간 내에 한번 이상은 최고점을 찍은 경험이 있는 저평가된 우량 정보를 특허회피 설계안으로 사용할 만한 빅데이터로 선정한다.

[그림 3-8]의 예는 ‘무인자동차’를 핵심키워드나 특허키워드로 삼아 포털 사이트에 넣고 검색한 경우 2017년도 3월 25일자의 해당하는 클릭수를 나타낸 것으로 전체 클릭수는 알 순 없겠으나, 클릭율이 55%임으로 무인자동차가 들어간 정보글 역시 저평가된 우량 정보로 활용할 만한 가치가 있다고 할 것이다.



[그림 3-8] 조건식 변경에 따른 정당성 여부 타진

따라서, 특허회피 설계안에 반영할 우량 정보를 찾기 위한 바람직한 조건식은 [그림 3-9]에서 보여지는 바와 같이, 고객의 관심분야이면서 (1)전체 클릭율을 1로 봤을 때 어느 한 시점의 해당 클릭율이 적어도 0.5 이상 일어난 정보, (2)3년을 기준으로 적어도 1회에서 10회 이내에서 급상승했

던 이력이 있는 정보여야 할 것이라 본다. 그래야지만, 이 저평가된 우량 정보를 파악해 이를 토대로 장벽특허를 회피해 고객이 원하는 공백기술을 안내해 주거나 독과점이 아직 존재하지 않은 미개척 시장 혹은 새로운 틈새시장을 발굴해 줄 수 있을 것이다.

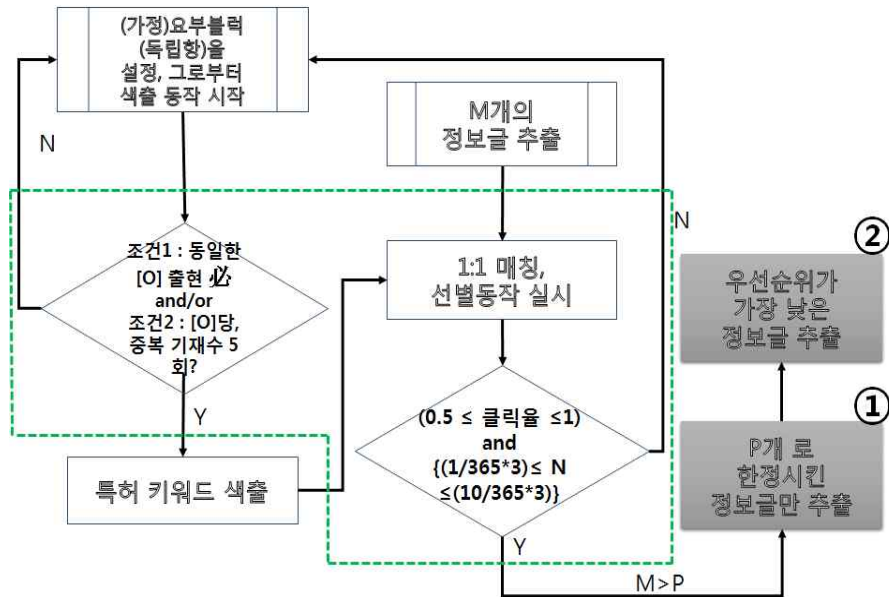
특허회피 설계안은 상대방이 먼저 선점해 놓은 특허권의 권리범위를 검토해 그 권리범위가 갖는 독점권을 벗어날 수 있도록 고객의 권리범위를 개량 및 고찰하여 향후 생산하고자 하는 제품이 상대방의 특허권에 위배됨 없이 원활히 시장에 출시될 수 있게 고객의 권리범위를 설계해 주는 것을 말한다.

그러기 위해, 고객의 관심분야에 속한 선 등록된 특허의 선행조사가 필요할 것이며, 선행조사를 통해 얻어진 여러 회피 안들이 올바른 방향으로 설정되어 상대방의 권리범위를 침해하지 않고 나의 기술적 노하우가 반영되면서 더욱 진보된 형태로 개량되어진 공백특허의 확보가 필수적이라 할 것이다.

또한, 특허회피 설계안은 기존의 선행조사 뿐만 아니라, VLDB(Very Large Database)에 많은 정보를 담고 있고, 5G망을 통해 대중에게 많은 정보를 제공하고 있는 유용한 포털 사이트의 활용이 무엇보다 중요하다.

즉, 본 논문에서 특허회피 설계안을 마련하는데 활용되는 정보글은 특허 키워드를 포털 사이트에 입력해 그로부터 검색된 빅데이터 중에 (1)전체 클릭율을 1로 봤을 때 어느 한 시점의 해당 클릭율이 적어도 0.5 이상 일어난 정보, (2)3년을 기준으로 적어도 1회에서 10회 이내에서 급상승했던 이력이 있는 정보를 그 대상으로 삼는다.

이외에도 특허회피 설계안을 마련하는데 활용되는 정보글을 골라내기 위한 다른 조건이 산업별, 분야별, 업종별, 최신 트렌드에 따라 기간, 클릭율, 클릭수, 출원빈도, 가중치 등을 고려해 설정될 수 있다.



[그림 3-9] 변경된 조건식에 따른
봇의 정보글 추출

정보글 추출에 관한 시나리오1은 [그림 3-10]으로도 알 수 있듯이, M개의 정보글 중 우선순위가 낮은 정보글을 추출하는 대표적인 사례 중 시나리오 ①일 경우에 이에 합당한 $P(M>P)$ 개의 한정된 정보글만 추출하는 세부 프로세싱을 진행한다.

정보글 추출에 관한 시나리오1은 특허회피 설계안을 자동 생성하는데 가장 중요한 요소인 최종 주제어를 다수개 선정하는 과정에서 일단 선제적으로 M개의 정보글 중 조건식($0.5 \leq \text{클릭율} \leq 1$) and $\{(1/365 \cdot 3) \leq \text{급상승 횟수} \leq (1/365 \cdot 3)\}$ 을 만족하는 P개의 정보글을 추출한 후, P개 정보글로부터 정제되지 않은 raw 데이터라 할 수 있는 후보 주제어를 적어도 하나 이상 색출한다.

이때, 후보 주제어는 P개의 정보글에 기재된 명사들 중 특허 키워드, 중복된 명사 및 핵심 키워드를 배제시킨 상태에서 색출된 것임을 전제로 한다.

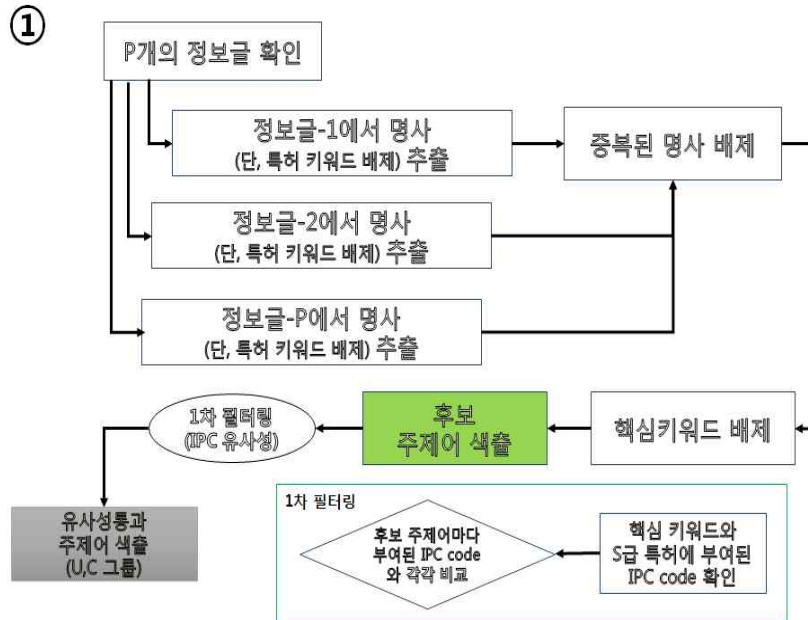
왜냐하면, 특허 키워드, 중복된 명사 및 핵심 키워드는 특허회피 설계안을 사용자 인터페이스 맞게 프로그램으로 로직 설계된 플랫폼을 통해 보

여주기 전에 이미 알려진 오픈 정보라 할 수 있으므로, 이미 알게 되었다는 것만으로도 특허회피 설계안을 제시하는데 적용시킬만한 활용가치 측면에서 보더라도 주제어로 채택하기엔 다소 의미적 역할을 상실하였다고 간주할 수 있기 때문이기도 하겠다.

이어서, 정보글 추출에 관한 시나리오1은 IPC 분류표를 활용해 이로부터 추출된 적어도 하나 이상의 후보 주제어를 대상으로 IPC 코드를 각각 알아내고, 이와 마찬가지로 핵심 키워드와 S급 특허에도 부여된 IPC 코드를 조회해 낸다.

그런 다음, 후보 주제어마다 부여된 IPC 코드와 핵심 키워드의 IPC 코드 간의 유사도를 비교해 유사 정도에 따라 4개의 모집단으로 레벨 분류하여 U그룹을 구성하고, 후보 주제어마다 부여된 IPC 코드와 S급 특허에 부여된 IPC 코드 간의 유사 여부를 판단하는 유사도를 측정하여 서로 간의 유사 정도에 따라 U그룹처럼 4개의 모집단으로 레벨 분류한 C그룹을 구성하는 1차 필터링을 실시하게 된다.

여기서, S급 특허란 기술적 개념을 최초 구현한 원천특허 또는 독자적 기술을 구체화, 고도화하여 제품양산으로도 가능하도록 전략화한 특허로, 국제출원 및 등록된 특허의 주인 격이면서, (피)인용지수도 높게 나타나는 특허를 말한다. S급 특허를 많이 보유한 국가일수록 국가 경쟁력은 당연히 높다고 할 것이다. 국가 경쟁력이 우수한 국가로는 미국, 핀란드, 일본, 프랑스, 한국 등이 있다.



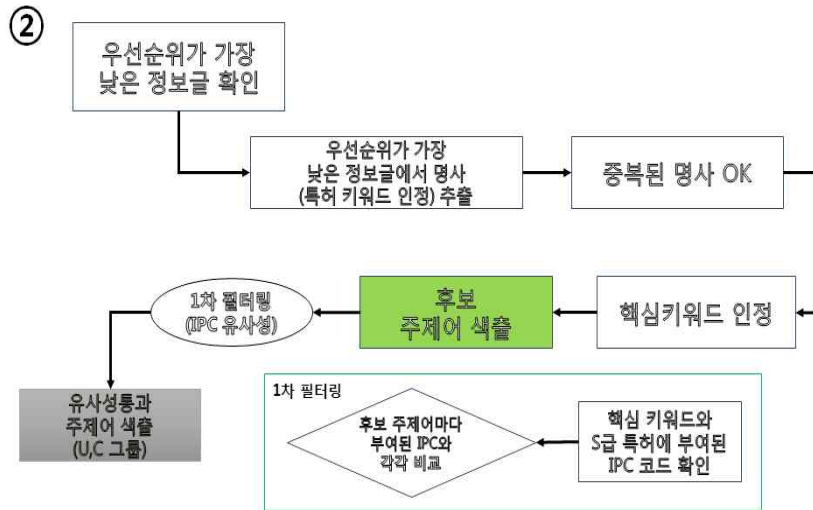
[그림 3-10] 정보글 추출에 관한 시나리오1

또한, 정보글 추출에 관한 시나리오2는 [그림 3-11]에서도 보여지는 바와 같이, M개의 정보글 중 우선순위가 낮은 정보글을 추출하는 사례 이외에 다른 시나리오 ②에 해당되는 경우로 우선순위가 가장 적은 정보글 (P=1)만 추출하는 세부 프로세싱을 진행한다.

즉, 정보글 추출에 관한 시나리오2는 특허회피 설계안을 제시하는데 가장 중요한 키인 최종 주제어를 다수개 선정하기 위한 일환으로, M개의 정보글 중 우선순위가 가장 정보글, 다시 말해, 클릭율 1에 가깝거나 급상승 횟수가 3년 기준 1회 내지 2회에 한해 만족하는 P가 1인 정보글을 추출한 후, 이 정보글 내 기재된 후보 주제어를 적어도 하나 이상 색출한다.

이때, 후보 주제어는 하나의 정보글에 기재된 명사들 중 특허 키워드와 중복된 명사 그리고 핵심 키워드 모두를 인정하는 것을 전제로 한다. 그 이유는 특허회피 설계안을 자동 생성시키기 위해 후보 주제어가 반드시 필요한 상황에서 P=1인 하나의 정보글이 갖는 희귀성 면에선 대단히 우수하다고 하겠으나, 핵심 키워드와 특허 키워드로부터 유도된 이 정보글이 특허회피 설계안을 제시해 주는데 있어 이를 명확히 하는 해법이구나라고

누구도 장담할 수 없음이 분명하기에 이를 해결하기 위한 방안으로써, 이미 알려진 오픈 정보 즉, 특허 키워드, 중복된 명사 및 핵심 키워드도 후보 주제어로 삼아 인정하는 것이 특허회피 설계안을 필요로 하는 고객의 관심사에서 동떨어지지 않으면서 P=1인 하나의 정보글이 내포하고 있는 숨은 내용을 바르게 적용해 볼 수 있기 때문이다.



[그림 3-11] 정보글 추출에 관한 시나리오2

계속해서, 정보글 추출에 관한 시나리오2는 ①과 동일하게 IPC 분류표를 활용해 이로부터 추출된 적어도 하나 이상의 후보 주제어를 대상으로 IPC 코드를 각각 알아내고, 이와 마찬가지로 핵심 키워드와 S급 특허에도 부여된 IPC 코드를 조회해 낸다.

그런 다음, 후보 주제어마다 부여된 IPC 코드와 핵심 키워드의 IPC 코드 간의 유사도를 비교해 유사 정도에 따라 4개의 모집단으로 레벨 분류하여 U그룹을 구성하고, 후보 주제어마다 부여된 IPC 코드와 S급 특허에 부여된 IPC 코드 간의 유사 여부를 판단하는 유사도를 측정하여 서로 간의 유사 정도에 따라 4개의 모집단으로 레벨 분류한 C그룹을 구성하는 1차 필터링을 실시한다.

3.3.3 유사성 판단

본 논문의 1차 필터링에서 실시되는 유사성 판단은 [그림 3-12]에서 보여지는 바와 같이 간단히 요약된다.

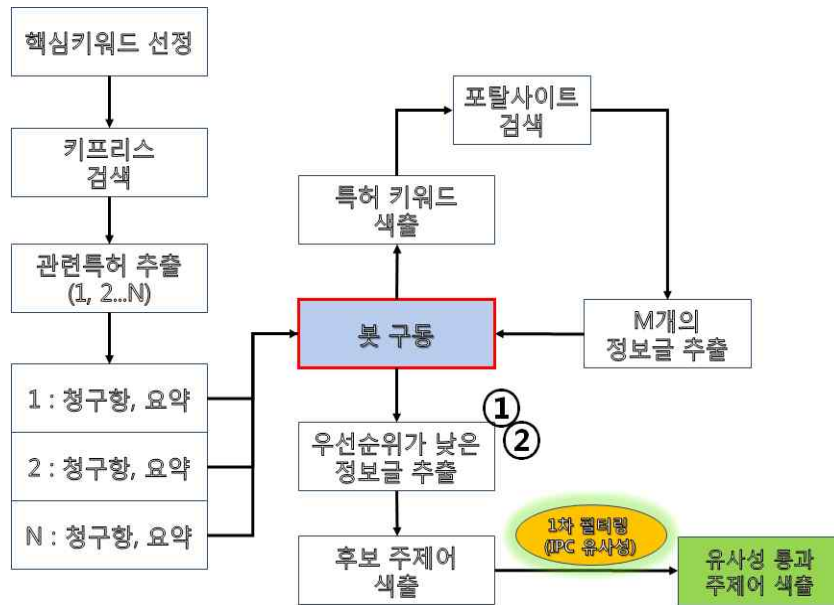
즉, 고객은 자신이 정한 핵심 키워드를 키프리스(www.kipris.or.kr)입력 후 검색해 N개의 선행 특허를 찾아낸다.

본 논문의 유사성 판단은 N개의 선행 특허를 하나씩 순차적으로 연 후, 그 선행 특허의 특허청구범위(청구항) 또는 요약을 자동으로 스캔하고, 스캔된 명사들 중 반복적으로 표현된 용어를 색출해 특허 키워드로 선정한다.

본 논문의 유사성 판단은 특허 키워드를 포털 사이트에 입력해 자동 검색을 실시함에 따라, 포털 사이트에 게재된 많은 정보글 중 특허 키워드가 많이 포함된 M개의 정보글을 추출한다.

본 논문의 유사성 판단은 포털 사이트로부터 얻은 M개의 정보글 중 특허 키워드가 기재된 상태를 자동 체크하고, M개의 정보글 (1, 2, ..., M순위) 중 우선순위가 낮은 정보글만을 추려낸다.

본 논문의 유사성 판단은 M개의 정보글 중 우선순위가 낮은 정보글 중에서 "특허회피 설계안을 제시하는데 용이하게 사용될 명사"를 후보 주제어(정제되지 않은 raw 데이터)로 색출한다.



[그림 3-12] 유사성 판단

이하, 본 논문의 1차 필터링에서 실시되는 유사성 판단은 후보 주제어를 색출하는 과정과 연결되는 부분임으로 본 논문에서는 이해를 돕기 위하여 좀 더 구체적으로 설명하기로 한다.

유사성 판단의 예1은 데이터베이스에 이미 저장되어 있는 핵심 키워드를 특허청에서 제공하는 IPC 분류표(IPC Chart : International Patent Classification Chart)에 대입해 이미 국제적으로 정리해 놓은 기술트리를 인지하고 핵심 키워드에 해당하는 관심분야에 관한 IPC 코드를 조회한다 .58)59)60)

유사성 판단의 예1은 특허회피 설계안을 자동 생성하는데 가장 중요한 요소인 최종 주제어를 선정하기 위해 진행되는 중간 단계로, [그림 3-13]에서 보여지는 바와 같이, 만일 고객이 ‘혼합현실’(증강현실, 가상현실)을

58) 김현우, 김종찬, 이준혁, 박상성, 장동식. (2015). 핵심 기술 및 특허 추출을 위한 IP 마이닝에 관한 연구. 『한국지능시스템학회』, 25(4), 392-397.

59) 전기억. (2001). “신기술 영역에서 특허보호와 관련된 한국의 기술경쟁력”. 서울대학교 대학원 박사학위논문.

60) 이민정, 송창현, 김연배. (2018). 국제특허분류(IPC) 공동출현 네트워크 분석을 통해 살펴본 기업의 지식융합 특성이 혁신성과에 미치는 영향. 『지식재산연구』, 13(1), 245-284.

핵심 키워드로 설정해 ‘혼합현실’이 적용되는 당해기술을 자신의 관심분야인 것으로 정한 경우, ‘혼합현실’을 IPC 분류표에 대입해 ‘혼합현실’이라는 핵심 키워드가 속하는 IPC 코드를 색출한다.

이로써, ‘혼합현실’에 해당하는 IPC 코드는 G06 T 19/00, G02 B 27/01, G02 B 27/14, G02 B 3/00, G06 Q 50/16, G06 Q 50/10, G06 Q 30/06으로 조회된다.

유사성 판단의 예1은 ‘혼합현실’이라는 핵심 키워드를 활용한 검색으로 키프리스를 거쳐 포털 사이트로부터 추출된 후보 주제어들을 모아놓은 후보 주제어 그룹(트래킹, welfare, 인증, 대중문화, 랜더링, 애니메이션, HMD, 몰입감, 게임, 시각화, 아바타, 양안시차, LBS, 입체영상, 시뮬레이션, 애너그리프, 이미지맵, 영상접합)과 ‘혼합현실’이라는 핵심 키워드 간의 유사성 여부를 판단하기 위하여 서로 간의 IPC 코드를 각각 비교한다.

후보 주제어 그룹 중 ‘HMD’는 IPC 분류표를 통해 확인한 바 G06 Q 50/10라는 IPC 코드를 가지며, ‘양안시차’는 G02 B 27/02라는 IPC 코드를 가짐을 확인할 수 있다.

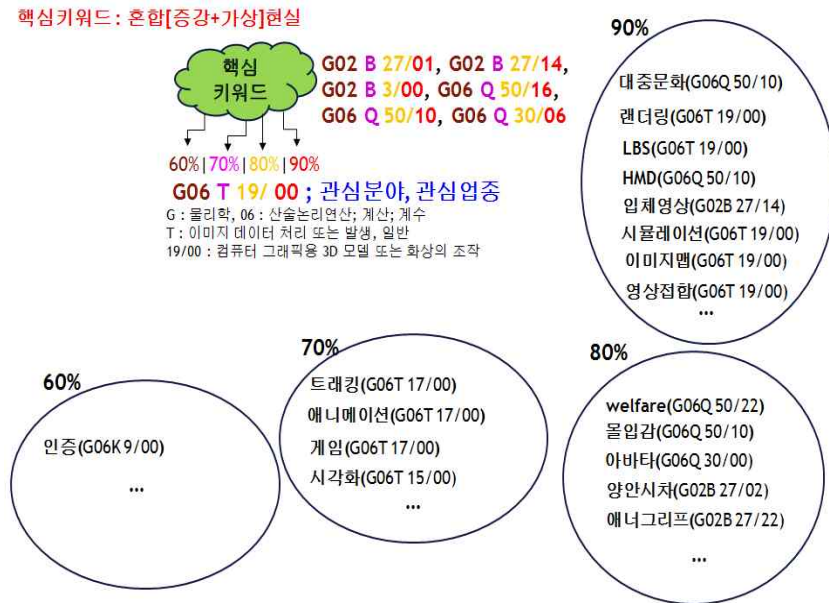
유사성 판단의 예1은 ‘혼합현실’이 갖는 IPC 코드와 ‘HMD’, ‘양안시차’가 갖는 IPC 코드를 비교해 레벨 분류를 실시한다. 레벨 분류는 IPC 분류표에 의거해 확률적으로 90% 유사도를 갖는 제 1 모집단에서부터 60% 유사도를 갖는 제 4 모집단까지 분류되어진다.

따라서, 유사성 판단의 예1은 IPC 코드를 활용한 서로 간의 비교 동작을 통해 ‘HMD’를 제 1 모집단에 ‘양안시차’를 제 2 모집단으로 분류해 내며, 그밖에 후보 주제어들(트래킹, welfare, 인증, 대중문화, 랜더링, 애니메이션, 몰입감, 게임, 시각화, 아바타, LBS, 입체영상, 시뮬레이션, 애너그리프, 이미지맵, 영상접합)도 제 1~4 모집단 중 하나에 속할 수 있도록 분류해 낸다.

유사성 판단의 예1은 하나의 트랙 그리고, 유사성 판단만으로 최종 주제어들을 선정해 그에 따른 특허회피 설계안을 모의 제시하는 것으로 제한한다.

왜냐하면, 본 논문에 관한 유사성 판단의 예1과 유사성 판단의 예2와의

큰 차이점을 보여주기 위해서이다.



[그림 3-13] 유사성 판단의 예1

즉, 본 논문의 유사성 판단의 예2는 핵심 키워드(ex : IoT, 사물인터넷)와 후보 주제어 그룹 간의 유사성 여부를 판단하기 위해 IPC 분류표를 활용해 'IoT'의 IPC 코드와 후보 주제어 그룹에 속한 후보 주제어들마다 부여된 IPC 코드를 비교한다.

유사성 판단의 예2는 IPC 분류표를 토대로 삼아 [그림 3-14]에서 보여주는 바와 같이, 'IoT'의 IPC 코드와 후보 주제어 그룹에 속한 모든 후보 주제어들(아두이노, 클라우드, 헬스케어, 비콘, 라즈베리파이, 유비쿼터스, 웨어러블 디바이스, 빅데이터, 융합, 인공지능, 3D 프린터, 블루투스, RFID, 클러스터, 3D 모델링, USN, 지그비, 상황인지)마다 각기 부여된 IPC 코드를 상호 비교함을 통해 4개의 모집단 중 하나에 속하도록 레벨 분류하는 유사성 판단을 실시한다.

유사성 판단의 예2는 특허회피 설계안을 제시하는데 유용하게 쓰이게 되는 최종 주제어들을 선정하기 위해 IPC 코드를 활용해 유사도를 측정한다.

‘IoT’의 IPC 코드(G06 F 13/14, G06 F 13/38, G06 F 9/44, H04 L 12/66, H04 L 29/08, H04 M 3/42, H04 M 11/04)와 ‘아두이노’의 IPC 코드 H04 L 29/08와 비교해 유사도를 측정한 결과, ‘IoT’에서 부여된 여러 IPC 코드 중 동일한 IPC 코드가 ‘아두이노’라는 후보 주제어들에게 부여되어 있음을 확인할 수 있어 ‘아두이노’를 유사도 90%에 속한다고 본 제 1 모집단에 편입시키도록 한다.

동일한 방법으로 ‘IoT’의 IPC 코드와 ‘클라우드’라든지, ‘헬스케어’라든지 상기에서 이미 언급된 후보 주제어 그룹에 속한 후보 주제어들에게 부여된 IPC 코드를 서로 비교함으로써, 4개의 모집단 중 하나에 속하도록 레벨 분류를 실시한다.

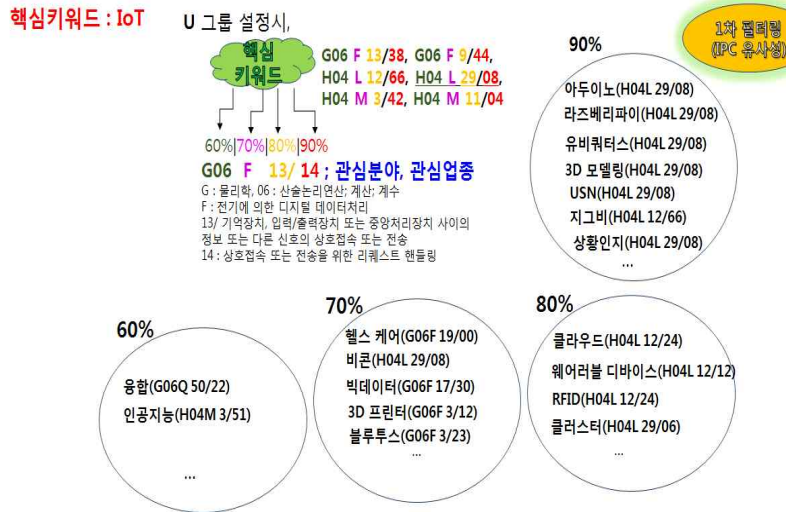
‘클라우드’의 IPC 코드는 H04 L 12/04인 것으로 ‘IoT’의 여러 IPC 코드와 비교한 결과, H04 L 12가 같아 유사도 80%에 속한다고 본 제 2 모집단에 편입시키도록 한다. ‘헬스케어’의 IPC 코드는 G06 F 19/00인 것으로 ‘IoT’의 여러 IPC 코드와 비교한 결과, G06 F가 같아 유사도 70%에 속한다고 본 제 3 모집단에 편입시키도록 한다.

유사성 판단의 예2는 이와 같은 방법을 반복적으로 적용하여 “클라우드, 헬스케어, 비콘, 라즈베리파이, 유비쿼터스, 웨어러블 디바이스, 빅데이터, 융합, 인공지능, 3D 프린터, 블루투스, RFID, 클러스터, 3D 모델링, USN, 지그비, 상황인지” 까지 포함된 후보 주제어 그룹을 모두 레벨 분류한다.

여기서, 90% 유사도에 속한 제 1 모집단으로 분류된 후보 주제어들, 80% 유사도에 속한 제 2 모집단으로 분류된 후보 주제어들, 70% 유사도에 속한 제 3 모집단으로 분류된 후보 주제어들을 채택하는 것이 더욱 바람직하다.

왜냐하면, 1차 필터링에 해당하는 유사성 판단은 고객이 선택한 관심분야를 대표하는 핵심 키워드와 후보 주제어 그룹에 속한 후보 주제어들 간의 유사성 여부를 판단하는 것으로, 유사도가 높다는 의미는 서로를 연계시키는 시너지가 큼을 뜻하거나 상관관계 측면에서도 연관성이 분명 존재함을 내포하고 있어, 향후 유사도가 높은 모집단에 속한 후보 주제어들을 선호하는 것이 양질의 특허회피 설계안을 제시할 수 있도록 하는데 더욱

유리하기 때문이라 하겠다.



[그림 3-14] 유사성 판단의 예2(U그룹)

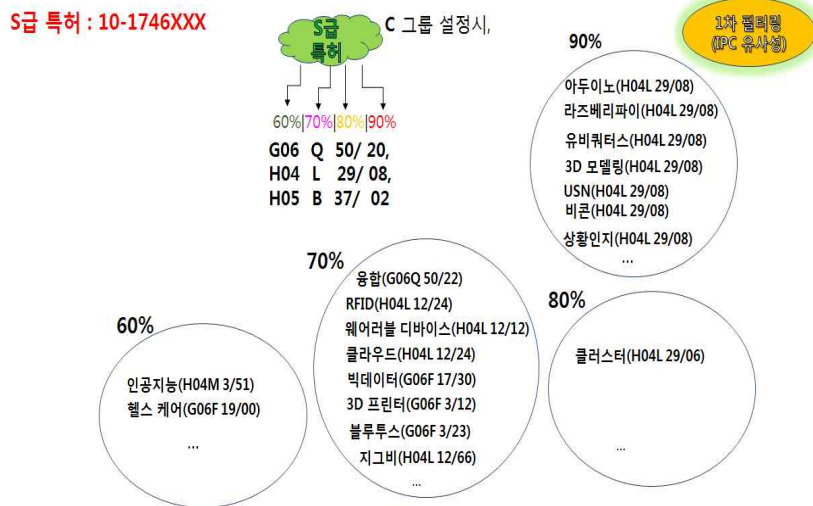
유사성 판단의 예2는 이에 국한되지 않고 IPC 분류표를 활용해 이로부터 추출된 후보 주제어들을 대상으로 IPC 코드를 각각 알아내고, 전문가에 의해 이미 선정된 S급 특허에 부여된 IPC 코드를 조회한다.

그런 다음, 후보 주제어들마다 부여된 IPC 코드와 S급 특허에 부여된 IPC 코드 간의 유사 여부를 판단하는 유사도를 측정하여 서로 간의 유사 정도에 따라 4개의 모집단으로 레벨 분류한 C그룹을 조성한다.

C그룹은 앞서 IPC 코드를 이용해 상호 비교하는 채택방식에 의해 조성됨에 있어서 U그룹이 조성되는 방식과 동일하나, S급 특허에 부여된 IPC 코드를 후보 주제어들마다 부여된 IPC 코드와 비교해 레벨 분류한다는 점에서 U그룹과 차이가 있다고 할 것이다.

따라서, 유사성 판단 사례2는 후보 주제어 그룹을 4개의 모집단으로 1차 필터링을 실시하고, 향후 양질의 특허회피 설계안을 제시하는데 적용되는 형태소 매칭이 주로 이행되는 자연어 처리를 하게 된다. 이때, 반드시 간과해서는 안되는 최종 주제어 선정에 관한 명확한 근거를 모색하는 방법 중 하나는 IPC 코드를 이용해 유사성 판단을 실시하는 것이겠고, 그에 따라, [그림 3-15]처럼 유사도 면에서 비교적 높은 모집단에 존재하는 후보

주제어들을 1차 필터링을 통해 유사성 통과 주제어들로 색출해 낸다.



[그림 3-15] 유사성 판단의 예2(C그룹)

3.3.4 연관성 판단

두 변수 간에 연관성을 조사하기 위해선 피어슨 상관계수 식이 필수적으로 적용된다. 피어슨 상관계수 식으로부터 구해지는 피어슨 상관계수(r)는 두 변수 간의 연관성에 관한 강도를 측정하는 지표라 할 수 있다.

피어슨 상관계수(r)는 1과 -1 사이 값을 갖으며, 1은 두 변수가 완전히 연관되었음을 의미하고, 0은 연관이 전혀 없음을 의미하며, -1은 완전 반대로 연관되었음을 뜻한다.

본 논문의 연관성 판단은 [그림 3-16]에서 보여지는바와 같이, 1차 필터링을 통해 U그룹으로 조성된 유사성 통과 주제어들과 C그룹으로도 조성된 유사성 통과 주제어들을 대상으로 가중치 점수를 부여해 유사성 판단으로 제 1~4 모집단($G_1, \dots, G_4$)으로 레벨 분류된 유사성 통과 주제어들을 수치로 정형화하고 이를 이용해 U그룹과 C그룹간의 연관성 여부를 파악하는데 유용한 피어슨 상관계수($R^U_{G_1, \dots, G_4}$, $R^C_{G_1, \dots, G_4}$)를 구한다.

이때, 수치로 정형화된 유사성 통과 주제어들은 피어슨 상관계수 식에

대입되는 파라미터 값을 생성시키는 백데이터로 활용된다.

또한, 본 논문의 연관성 판단은 피어슨 상관계수(r)가 0.5보다 크면 U그룹과 C그룹은 서로 연관성이 있다고 할 것이고, 앞서 구한 피어슨 상관계수(r)와 대비해 1:1 대응되는 최적점수를 찾은 후 이 최적점수보다 큰 가중치 점수를 갖는 유사성 통과 주제어들을 골라 연관성 통과 주제어들로 선정한다. 이에 따른, 2차 필터링에 적용되는 피어슨 상관계수(r)의 공식은 아래와 같다.

$$r = \frac{cov(U, C)}{\sqrt{Var(U)} \sqrt{Var(C)}} = \frac{E[(U - ux)(C - cx)]}{\sqrt{E[(U - ux)^2]} \sqrt{E[(C - cx)^2]}} \quad (1)$$

즉,

$$r = \frac{1/n \left(\sum_{k=1}^n (U_k - ux)(C_k - cx) \right)}{\sqrt{1/n \left(\sum_{k=1}^n [(U_k - ux)^2] \right)} \sqrt{1/n \left(\sum_{k=1}^n [(C_k - cx)^2] \right)}}$$

여기서,

$$cov(U, C) = E[(U - ux)(C - cx)] = 1/n \left(\sum_{k=1}^n (U_k - ux)(C_k - cx) \right)$$

cov(U,C)은 U와 C가 열 벡터값을 가지는 확률 변수일 때의 공분산을 나타낸 것이다.

$$var(U) = E[(U - ux)^2] = 1/n \left(\sum_{k=1}^n [(U_k - ux)^2] \right)$$

$$var(C) = E[(C - cx)^2] = 1/n \left(\sum_{k=1}^n [(C_k - cx)^2] \right)$$

var(U) : U의 분산, var(C) : C의 분산

$$ux = \frac{1}{n} \sum_{k=1}^n U_k \quad cx = \frac{1}{n} \sum_{k=1}^n C_k$$

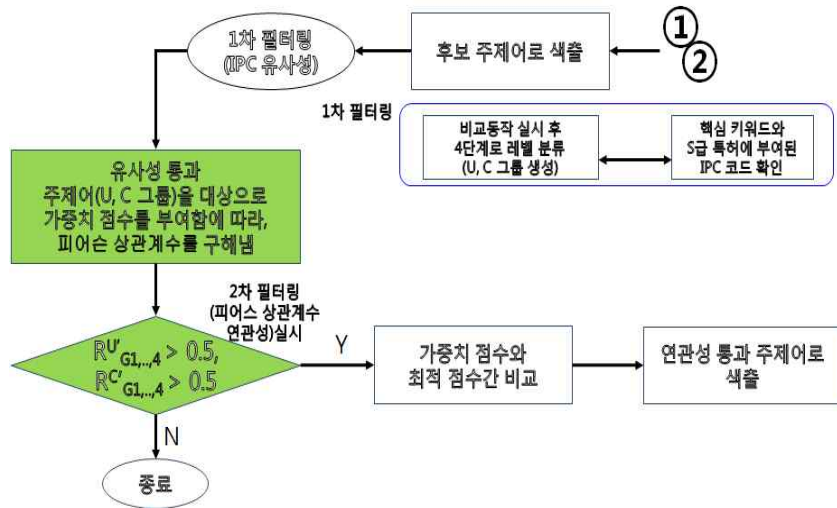
u_x : U의 평균값, c_x : C의 평균값

$$U_n = \{u_1, u_2, \dots, u_n\}$$

U 그룹은 핵심 키워드의 IPC에 의거해 가중치 점수를 부여받은 유사성 통과 주제어들의 집합이다.

$$C_n = \{c_1, c_2, \dots, c_n\}$$

C 그룹은 핵심 키워드로 검색된 다수의 특허 건들 중 전문가가 찾아 낸 S급 특허(등록번호 : 10-1234567, 20-2345678)의 IPC에 의거해 가중치 점수를 부여받은 유사성 통과 주제어들의 다른 집합이다.



[그림 3-16] 연관성 판단

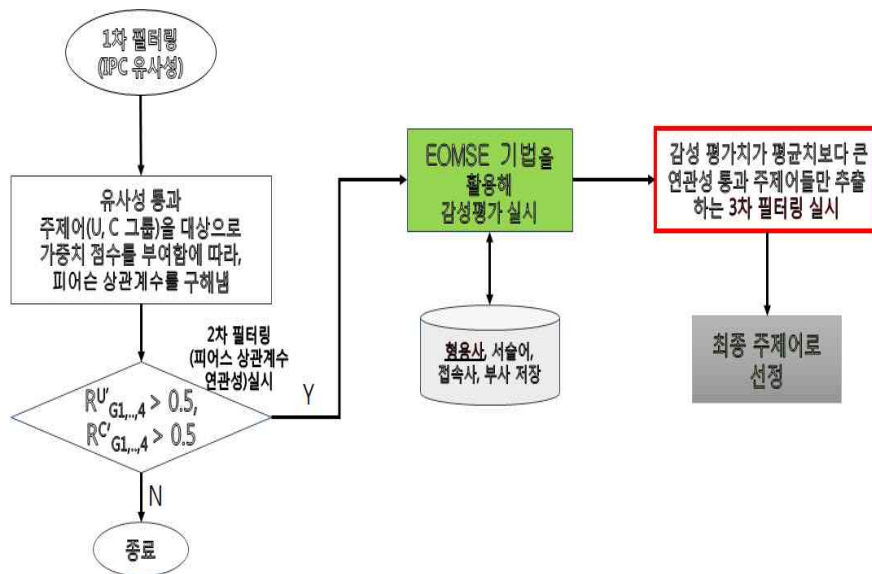
3.3.5 적합성 판단

본 논문의 적합성 판단은 피어슨 상관계수 식에 의해 나온 피어슨 상관계수를 반영해 2차 필터링 완료된 연관성 통과 주제어들을 대상으로 본 논문에서 제안하는 EOMSE 기법을 적용하여 3차 필터링을 실시한다.

즉, 본 논문의 적합성 판단은 [그림 3-17]을 통해 알 수 있듯이, EOMSE 기법(Expension Opinion Mining Sensitivity Evaluation Scheme)을 활용해 감성평가를 실시하고 이로부터 구해진 감성 평가치와 이등의 평균을 구한 평균치를 비교해 평균치보다 큰 감성 평가치를 갖는 연관성 통과 주제어들만을 색출하는 3차 필터링을 실시함에 따라 최종 주제어들을 선정한다.

그런 후, 적합성 판단은 특허회피 설계안을 마련하기 위한 방법으로 P개의 정보글 혹은 하나의 정보글 중 최종 주제어를 중심으로 여러 형태소들이 기재되어 있는 근접 위치를 파악하고, 최종 주제어와 최종 주제어 주변에 근접해 있는 형태소 간의 매칭을 실시함으로써, 최종 주제어가 모두 들어 있으면서 형태소 매칭에 충분히 이루어져 일반 대중이 알아들 수 있는 어휘력이 발휘된 문장으로 형성된 특허회피 설계안을 자국어 혹은 영어로 표현될 수 있도록 구현하는 것을 최종 목표로 한다.

여기서, 최종 주제어가 표기된 P개의 정보글 혹은 하나의 정보글 그리고, 최종 주제어를 기준으로 그 주변에 존재하는 여러 품사들(형용사, 서술어, 접속사, 부사)은 이미 데이터베이스에 저장되어 있음을 전제로 한다.



[그림 3-17] 적합성 판단

적합성 판단에서 제기되는 EOMSE 기법은 서술어, 접속사, 부사가 가지고 있는 극성값을 계산해 고객이 관심 있어 하는 어느 한 주제어에 대한 감성 평가치를 알아내는 기존의 오피니언 마이닝 기법을 그대로 적용하는 방식을 채택하면서 품사 부분에서 중요한 역할을 함에도 불구하고 기존 기법에는 전혀 반영되지 않았던 ‘형용사’를 추가시켜 연관성 통과 주제어에 대한 감성 평가치를 더욱 정밀하게 구해낼 수 있다.

이에 따라, EOMSE 기법은 유사성 통과 주제어 중에서 어떤 주제어가 최종 주제어로 선정될 만한 것인지를 판명해 주는 적합성에 대한 검증력을 기존의 오피니언 마이닝 기법보다 더 확장 및 강화시켜 놓은 기법이라 할 것이다.

연관성 통과 주제어의 감성 평가치(o.value)는 연관성 통과 주제어가 기재된 정보글로부터 추출한 서술어, 접속사, 부사 그리고 형용사를 포함하는 4개의 품사에 대한 감성 평가치를 각각 구하고 이들을 모두 합한 결과치인 것으로, o.value를 구하는 공식은 다음과 같다.

$$o.value = v.value + c.value + ad.value + a.value \quad (2)$$

[(> 0, 긍정), (= 0, 중립), (< 0, 부정)]

3.4 모의실험

3.4.1 사례1: 핵심 키워드가 ‘증강현실’, ‘가상현실’인 경우

3.4.1.1 유사성 판단만으로 결과도출

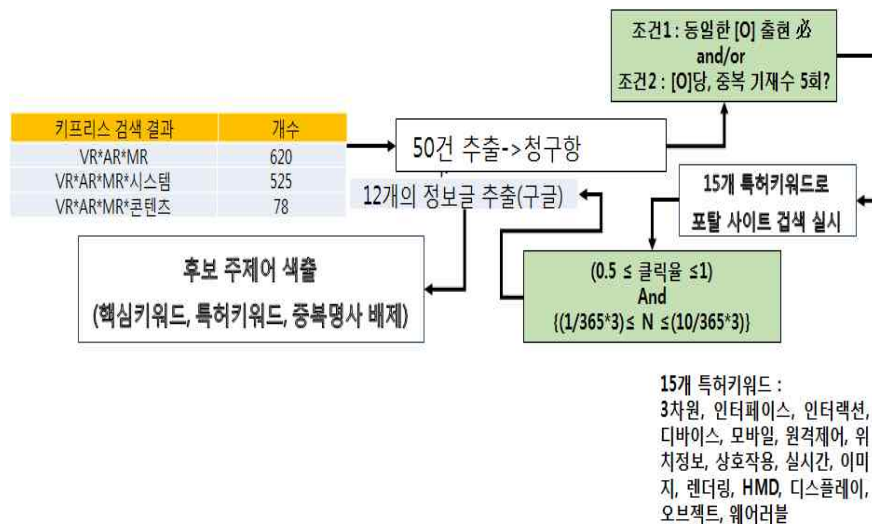
먼저, 모의실험의 사례1을 실시하는 논문 연구자는 핵심 키워드를 ‘증강현실’, ‘가상현실’로 설정한 후 특허청 정보검색 사이트인 키프리스 <www.kipris.or.kr>에 접속해 검색창에 ‘VR*AR*MR’로 된 1차 검색식을 입력해 특허검색 건수 620를 조회하거나 ‘VR*AR*MR*시스템’으로 2차 검색식을 입력해 특허검색 건수 525를 조회하거나, ‘VR*AR*MR*콘텐츠’

로 된 3차 검색식을 입력해 특허검색 건수 78를 조회한다.

그 다음, 주요특허로 여길만한 50건을 추려내고 50건의 주요특허에 기재된 특허청구범위 중 독립항을 탐독해 독립항에 기재된 권리범위 내에서 특허 키워드를 15개 추출해 낸다.(단, 특허 키워드는 <조건 1>동일한 목적어가 반드시 출현된 것이어야 하며, <조건 2>동일한 목적어가 5회 이상 중복 기재된 것임을 전제로 함)

사례1을 통해 추출되는 15개의 특허 키워드는 “3차원, 인터페이스, 인터랙션, 디바이스, 모바일, 원격제어, 위치정보, 상호작용 실시간, 이미지, 렌더링, HMD, 디스플레이, 오브젝트, 웨어러블”이 그 대상이 된다.

계속해서, 사례1은 15개의 특허 키워드를 포털 사이트 ‘구글’ 검색창에 삽입해 특허회피 설계안을 제시하는데 유용되는 정보글을 12개 추출해 낸다. 이때, 단순 조건식은 (1)‘구글’에서 제공되는 다수의 정보글 중에서 동일한 목적어가 2개 이하로 표현된 정보글일 것, (2)‘구글’에서 제공되는 다수의 정보글 중에서 목적어당 중복 기재수가 2 내지 3회 이하인 정보글일 것으로 임의 설정한다.<그림 3-18 참조>



[그림 3-18] 사례1의 후보 주제어 색출

하나, 바람직한 조건식은 앞서 3.3.2장에서 설명한 바와 같이 고객의 관

심분야이면서 (1)전체 클릭율을 1로 봤을 때 어느 한 시점의 해당 클릭율이 적어도 0.5 이상 일어난 정보, (2)3년을 기준으로 적어도 1회에서 10회 이내에 급상승했던 이력이 있는 정보이어야 할 것이다.

사례1은 ‘구글’로부터 추출해 낸 12개의 정보글을 바탕으로 핵심 키워드, 특허 키워드, 중복 명사를 배제시킨 후보 주제어들을 집합한 후보 주제어 그룹을 생성시킨다.

이로부터 생성된 후보 주제어 그룹은 정비지원을 위한 시스템, 모바일기기, 문화 유적지, 원격체험, AR 모바일 애플리케이션, 수술 시뮬레이션, 뷰티 미래, 증강현실 레이싱, AR놀이 플랫폼, 디지털 트윈, 아바타, 농작물이다.

사례1은 핵심 키워드인 ‘증강현실’, ‘가상현실’과 후보 주제어 그룹 내 후보 주제어들 간의 관련성 여부를 파악하기 위한 유사성 판단을 실시함으로써, 두 트랙으로 나눈 후 각 트랙마다 4개의 모집단으로 레벨 분류한다.

즉, 사례1의 유사성 판단은 [표 3-1]에서 보여지는 바와 같이, 후보 주제어 그룹 내 후보 주제어들마다 부여받은 IPC 코드와 핵심 키워드의 IPC 코드 간을 비교하는 핵심 키워드 중심 채택방식(U)과 후보 주제어들마다 부여받은 IPC 코드와 S급 특허의 IPC 코드 간을 비교하는 S급 특허 중심 채택방식(C)의 두 트랙으로 구분하고, 각 트랙마다 4개의 모집단으로 레벨 분류할 시, D급 이하의 후보 주제어들에 대해선 제거시킴을 의미한다.

사례1의 유사성 판단은 IPC 분류표에서 확인되는 H04 N 7/15, G06 T 19/00, G03 H 1/00로 된 ‘증강현실’, ‘가상현실’과, “정비지원을 위한 시스템(G06 Q 50/10), 모바일기기(H04 W 88/02), 문화 유적지(H04 N 7/14), 원격체험(G06 Q 30/06), AR 모바일 애플리케이션(H04 N 13/30), 수술 시뮬레이션(G09 B 23/28), 뷰티 미래(H04 N 21/443), 증강현실 레이싱(H04 M 1/725), AR놀이 플랫폼(H04 N 13/30), 디지털 트윈(G06 F 3/01), 아바타(G06 T 19/00), 농작물(A01 G 7/06)”로 된 후보 주제어 그룹의 IPC 코드를 각기 비교하여 정규 분포 및 레벨 분류에 따른 등급산정을 객관적인 지표로 명확히 하고자 4개의 모집단을 80점,

60점, 50점, 40점으로 각기 구분해 점수화시킨다.(단, 60%이하의 유사도를 갖는 번외 모집단은 10점의 점수로 할당한다)

핵심 키워드 중심 채택방식(U)은 H04 N 7/15, G06 T 19/00, G03 H 1/00로 된 ‘증강현실’, ‘가상현실’ 핵심 키워드의 IPC 코드와 후보 주제어 그룹에 속한 후보 주제어들을 모두 비교해 4개 모집단으로 분류해 내는 채택방식이라 할 것이다.

즉, 80점의 가중치 점수로 환산되는 90% 유사도를 갖는 A급 모집단은 아바타(G06 T 19/00)이 되며, 60점의 가중치 점수로 환산되는 80% 유사도를 갖는 B급 모집단은 문화유적지(H04 N 7/14)이 되고, 50점의 가중치 점수로 환산되는 70% 유사도를 갖는 C급 모집단은 AR 모바일 애플리케이션(H04 N 13/30), 뷰티미러(H04 N 21/443), AR놀이 플랫폼(H04 N 13/30)이 된다.

또한, 40점의 가중치 점수로 환산되는 60% 유사도를 갖는 D급 모집단은 모바일기기(H04 W 88/02), 원격체험(G06 Q 30/06), 정비지원을 위한 시스템(G06 Q 50/10), 증강현실 레이싱(H04 M 1/725), 디지털 트윈(G06 F 3/01)이 되고, 10점의 가중치 점수로 환산되는 번외 모집단은 수술 시뮬레이션(G09 B 23/28), 농작물(A01 G 7/06)이 된다 할 것이다.

[표 3-1] 유사성 판단으로 분류된 후보 주제어

후보 주제어 그룹	
정비지원을 위한 시스템, 모바일기기, 문화 유적지, 원격체험, AR 모바일 애플리케이션, 수술 시뮬레이션, 뷰티 미러, 증강현실 레이싱, AR놀이플랫폼, 디지털 트윈, 아바타, 농작물	
U(유사도)	[1] IPC 코드(유사성 판단) : 핵심 키워드 중심 채택방식
A급, 90%(80점)	아바타(G06T 19/00)
B급, 80%(60점)	문화유적지(H04N 7/14)
C급, 70%(50점)	AR 모바일 애플리케이션(H04N 13/30), 뷰티미러(H04N 21/443), AR놀이 플랫폼(H04N 13/30)
D급, 60%(40점)	모바일기기(H04W 88/02), 원격체험(G06Q 30/06), 정비지원을 위한 시스템(G06Q 50/10), 증강현실 레이싱(H04M 1/725), 디지털 트윈(G06F 3/01)

60%미만 (10점)	수술 시뮬레이션(G09B 23/28), 농작물(A01G 7/06)
핵심키워드 의 IPC	H04N 7/15, G06T 19/00, G03H 1/00
C(유사도)	[1] IPC 코드(유사성 판단) : S급 특허 중심 채택방식
A급, 90%(80점)	아바타(G06T 19/00), 정비지원을 위한 시스템(G06Q 50/10)
B급, 80%(60점)	모바일기기(H04W 88/02)
C급, 70%(50점)	원격체험(G06Q 30/06)
D급, 60%(40점)	문화유적지(H04N 7/14), AR 모바일 애플리케이션(H04N 13/30), 뷰티미러(H04N 21/443), AR놀이 플랫폼(H04N 13/30), 증강현실 레이싱(H04M 1/725), 디지털 트윈(G06F 3/01)
60%이하 (10점)	수술 시뮬레이션(G09B 23/28), 농작물(A01G 7/06)
S특허의 IPC	G06Q 50/10, H04W 88/06, G06T 19/00, G06Q 50/00

S급 특허 중심 채택방식(C)은 키프리스를 통해 검색된 다수의 주요특허 중에서 특허 전문가가 엄선한 핵심특허(S급 특허 : 20-0486347)로부터 확인되는 IPC 코드[G06 Q 50/10, H04 W 88/06, G06 T 19/00, G06 Q 50/00]와 후보 주제어 그룹 속한 후보 주제어들을 모두 비교해 4개의 모집단으로 분류해 내는 채택방식이라 할 것이다.

즉, 80점의 가중치 점수로 환산되는 90% 유사도를 갖는 A급 모집단은 아바타(G06 T 19/00), 정비지원을 위한 시스템(G06 Q 50/10)이 되며, 60점의 가중치 점수로 환산되는 80% 유사도를 갖는 B급 모집단은 모바일 기기(H04 W 88/02)이 되고, 50점의 가중치 점수로 환산되는 70% 유사도를 갖는 C급 모집단은 원격체험(G06 Q 30/06)이 된다.

또한, 40점의 가중치 점수로 환산되는 60% 유사도를 갖는 D급 모집단은 문화유적지(H04 N 7/14), AR 모바일 애플리케이션(H04 N 13/30), 뷰티미러(H04 N 21/443), AR놀이 플랫폼(H04 N 13/30), 증강현실 레이싱(H04 M 1/725), 디지털 트윈(G06 F 3/01)이 되고, 10점의 가중치

점수로 환산되는 변외 모집단은 수술 시뮬레이션(G09 B 23/28), 농작물(A01 G 7/06)이 된다 할 것이다.

정리하자면, 사례1의 유사성 판단은 “증강현실, 가상현실, 혼합현실”이라는 핵심 키워드를 키프리스로 검색해 검색식에 따라 달리 검색된 620건, 525건, 78건으로부터 50개의 관련 특허들을 추출한다.

또한, 사례1의 유사성 판단은 50개의 관련 특허들마다 기재되어 있는 특허청구범위를 탐독함으로부터 조건1, 조건2에 부합된 15개의 특허키워드들을 포털 사이트(ex : 구글)에 넣고 검색을 실시한다.

사례1의 유사성 판단은 검색 실시로부터 다른 조건1, 다른 조건2에 부합된 12개의 정보글로부터 뽑아낸 12개의 주제어들을 대상으로 국제특허 분류표(IPC Chart)를 기준으로 추정 가능한 정규분포를 객관적으로 레벨 분류하겠다는 검증 취지에 맞게 이를 점수로 치환한 핵심 키워드 중심 채택방식(U), S급 특허 중심 채택방식(C)에 적용한다.

사례1의 유사성 판단은 12개의 주제어들 중 D급 미만에 속하는 주제어들을 필터링하고 나머지 IPC 코드를 이용한 유사성 판단에서 높은 점수를 받아 상위레벨로 구분된 유사성 통과 주제어들(아바타, 문화유적지, 정비지원 시스템, 모바일 기기, 원격체험, 디지털 트윈 등)을 조합해 자연어 처리한다.

그에 따라, 사례1의 유사성 판단은 유사성 통과 주제어들을 대상으로 자연어 처리함을 통해 대화형 구문형식에 가장 걸맞는 자국어로 표현해 주는 특허회피 설계안을 봇을 통해 제시함을 알 수 있다.<표 3-2 참조>

즉, 사례1의 유사성 판단은 유사도가 비교적 낮은 부류(60% 이하)는 필터링하고 유사도가 상대적으로 높은 부류(60% 이상)에 속한 유사성 통과 주제어들(아바타, 문화유적지, 정비지원 시스템, 모바일 기기, 원격체험, 디지털 트윈 등)을 조합하고 자연어 처리(데이터 마이닝, 텍스트 마이닝, 온톨로지, 시멘틱, 메타 분석, 오피니언 마이닝 등)함을 통해 일반 대중이 이해할 수 있는 자국어로 표현되는 특허회피 설계안을 봇을 통해 자동 생성시킨다.

[표 3-2] 유사성 판단에 따른 이상적 표본 안

"증강현실(VR), 가상현실(VR)은 문화 유적지를 원격으로 체험할 수 있고, 디지털 트윈에 실시간으로 반영할 수 있는 콘텐츠의 개발이 요구되며 증강현실로 구현된 상대방의 아바타와 마주해 실시간으로 소통할 수 있는 색다른 아이템을 제안해 봅니다. 또한, 휴대가 용이한 모바일 기기를 기반으로 하여 정비사가 증강현실을 활용한 선박과 같은 장비를 정비할 수 있는 기술적인 지원이 가능한 정비지원 시스템도 제안해 봅니다."

3.4.1.2 사례1의 연관성 판단

앞서, 3.4.1장의 사례1에서 실시한 유사성 판단은 후보 주제어 그룹에 속한 후보 주제어들이 제각기 부여받은 IPC 코드와 핵심 키워드 즉, ‘증강현실’, ‘가상현실’로부터 인지되는 IPC 코드를 비교해 동일 유사여부에 따라 A급(90% 유사, 가중치 점수 : 80점)부터 B급(80% 유사, 가중치 점수 : 60점), C급(70% 유사, 가중치 점수 : 50점), D급(60% 유사, 가중치 점수 : 40점)까지 4개의 모집단으로 분류하고, D급 이하(유사도 60% 이하, 가중치 점수 : 10점)을 걸러내는 핵심 키워드 중심 채택방식(U)의 1차 필터링을 진행하였다.

또한, 3.4.1장의 사례1에서 실시한 유사성 판단은 후보 주제어 그룹에 속한 후보 주제어들마다 각기 부여받은 IPC 코드와 핵심 키워드로 키프리스를 통해 검색된 50건의 특허들 중 S급 특허(20-0486347)로부터 인지되는 IPC 코드와 비교해 동일 유사여부에 따라 A급(90% 유사, 가중치 점수 : 80점)부터 D급(60% 유사, 가중치 점수 : 40점)까지 후보 주제어들을 IPC 분류 코드를 활용해 4개의 모집단으로 분류하고, D급 이하(유사도 60%이하, 가중치 점수 : 10점)을 걸러내는 S급 특허 중심 채택방식(C)의 1차 필터링도 진행하였다.

본 논문의 2차 필터링은 핵심 키워드 ‘증강현실’, ‘가상현실’과 1차 필터링을 통해 구분 정렬된 후보 주제어 그룹에 속한 후보 주제어들을 대상으

로 가중치 점수를 부여한 후, 피어슨 상관계수 식을 이용해 피어슨 상관계수 값(r)을 구하고 피어슨 상관계수 값(r)으로부터 유추된 연관 최적점수와 가중치 점수 간의 비교를 통해 연관성 통과 주제어들을 색출한다.

2차 필터링인 연관성 판단은 피어슨 상관계수 값(r)이 기준치인 0.5보다 크면 유사성 통과 주제어들을 핵심 키워드인 ‘증강현실’, ‘가상현실’과 상관관계가 있다고 보는 모집단으로 판정하고, 기준치인 0.5보다 작으면 유사성 통과 주제어들을 ‘증강현실’, ‘가상현실’과 상관관계가 없다고 보는 모집단으로 판정한다.

[그림 3-19]는 1차 필터링을 통해 핵심 키워드 중심 채택방식(U)에 따라 4개의 카테고리 분류된 U 그룹과, S급 특허 중심 채택방식(C)에 따라 4개의 다른 카테고리 분류된 C 그룹으로부터 피어슨 상관계수(r)를 구해 나타낸 것이다.

피어슨 상관계수(r)를 구하기 위해, 피어슨 상관계수 식에 적용되는 변수들 즉, U 그룹에 속한 후보 주제어들마다 부여된 가중치 점수들을 평균한 평균값(u_x), C 그룹에 속한 후보 주제어들마다 부여된 가중치 점수들을 평균한 평균값(c_x), U 그룹의 분산값($\text{var}(U)$)과 C 그룹의 분산값($\text{var}(C)$), 그리고 U 그룹과 C 그룹 간의 공분산값($\text{con}(U, C)$)도 알아내야 할 것이다.

사례1의 연관성 판단은 유사성 통과 주제어들마다 부여된 자기 몫(가중치 점수)에 따라 피어슨 상관계수(r)가 어떻게 달라지는지 즉, 상관관계가 높게 나타나게 할 상승적인 요인으로 작용하는지 아니면 그 반대로 작용하는지 대해서 알아보기로 한다.

또한, 모의실험의 사례1은 두 주체<핵심 키워드와 S급 특허>간의 연관성 여부를 판단하는 것에 대한 척도인 피어슨 상관계수(r)가 0.5 보다 큰지를 피어슨 상관계수 식을 통해 미리 계산해 봄으로써, 두 주체 간의 연관성 있음을 증명해 보는 실험이기도 하다.

모의실험의 사례1은 [그림 3-19], [그림 3-20]에서도 알 수 있듯이, D급 이하의 후보 주제어들에게 부여된 가중치 점수를 0점으로 부여하는 대신 D급 이하에 있는 개체 수를 총 개체 수에서 배제시키지 않고 포함시키

는 것이 피어슨 상관계수(r)를 1에 보다 가깝게 접근함을 확인하였다.

즉, 피어슨 상관계수 식은 두 주체 간의 연관성 여부를 알아보기 위해 통계학에서 적용되는 수식인 바, [Ex 1-1]에서 구해진 피어슨 상관 계수(r)는 $0.7096(r > 0.5)$ 으로 산출되었고 [Ex 1-2]에서 구해진 피어슨 상관 계수는 $0.78(r > 0.5)$ 으로 산출되었음을 확인할 수 있었다.

사례1의 결과로, 1차 필터링에 의해 배제된 D급 이하의 후보 주제어들의 개체 수를 총 개체 수의 합에 포함시켜 피어슨 상관계수 식에 적용하되, D급 이하의 후보 주제어들에게 부여된 가중치 점수는 리셋한 값(0점)으로 적용하는 것이 두 주체간의 연관성을 높이는 방법 중 하나임을 알아냈다.

이어서, 사례1은 3가지의 각기 다른 경우 즉, 1차 필터링에 의해 제거된 후보 주제어들에게 부여된 몫(가중치 점수)에 따라 정규분포(가우시안 확률밀도 함수 : Gaussian Probability Density Function)가 과연 어떻게 나타나는지에 대해서 알아보았다.

다시 말해, 1차 필터링에 의해 제거된 후보 주제어들이 가지고 있는 몫을 어떻게 적용하는 것이 확률적인 입장에서의 시스템 안정화를 기대할 수 있을지에 대해서도 알아보았다.⁶¹⁾

[Ex 1-1]은 1차 필터링에 의해 제거된 후보 주제어들이 가지고 있는 몫으로 평균값(μ)과 표준편차(σ)를 각각 구한 뒤 정규분포를 구한 것이다. <그림 3-19 참조>

61) 강훈, 유지상, 이재진. (2001). 『확률, 랜덤변수 및 랜덤신호 원리』. 경기 : 한국맥그로힐 출판사.

$$r = \frac{272.9}{384.6} = 0.7096 \quad (r > 0.5); \text{식 (1) 적용}$$

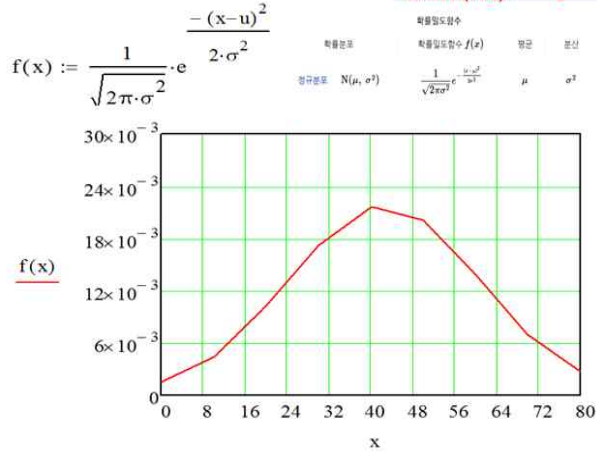
OK

$$\sigma := 18.3$$

$$u := 42.5$$

+

f(x) =	x =
1.47·10 ⁻³	0
4.504·10 ⁻³	10
0.01	20
0.017	30
0.022	40
0.02	50
0.014	60
7.048·10 ⁻³	70
2.671·10 ⁻³	80



[그림 3-19] Ex 1-1의 피어슨 상관계수 및 정규분포

[Ex 1-2]은 1차 필터링에 의해 제거된 후보 주제어들이 가지고 있는 몫을 0점을 부여하고, 그 대신에 60% 이하에 있는 개체 수를 총 개체 수에 포함시키는 상태에서 평균값(u)과 표준편차(σ)를 각각 구하고 정규분포를 구한 것이다.〈그림 3-20 참조〉

[Ex 1-3]의 U 그룹은 1차 필터링에 의해 제거된 후보 주제어들이 가지고 있는 몫을 0점을 주고, 60% 이하에 있는 개체 수를 총 개체 수에서 제외시킨 상태에서 평균값(u)과 표준편차(σ)를 각각 구한 다음, 정규분포를 구한 것이다.〈그림 3-20 참조〉

[Ex 1-2]

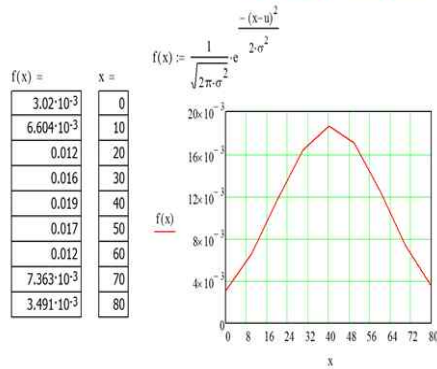
OK

$$r1' = \frac{397.9}{510.2} = 0.78$$

($r > 0.5$)

U(유사도)	후보 주제어수 (12개)
A급, 90%(80점)	1
B급, 80%(60점)	1
C급, 70%(50점)	3
D급, 60%(40점)	5
60%미만(0점)	2

$$\sigma := 21.4 \quad u := 40.83$$



[Ex 1-3]

NO

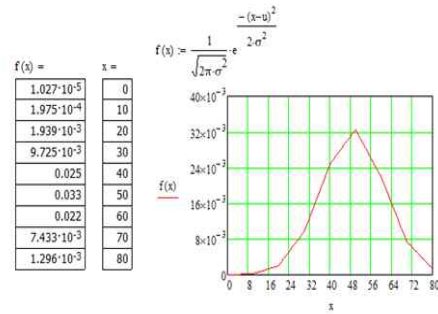
$$r1'' = \frac{61}{160.5} = 0.3167$$

($r < 0.5$)

U(유사도)	후보 주제어수 (10개)
A급, 90%(80점)	1
B급, 80%(60점)	1
C급, 70%(50점)	3
D급, 60%(40점)	5
60%미만(0점)	2

$$x := 0, 10, 80$$

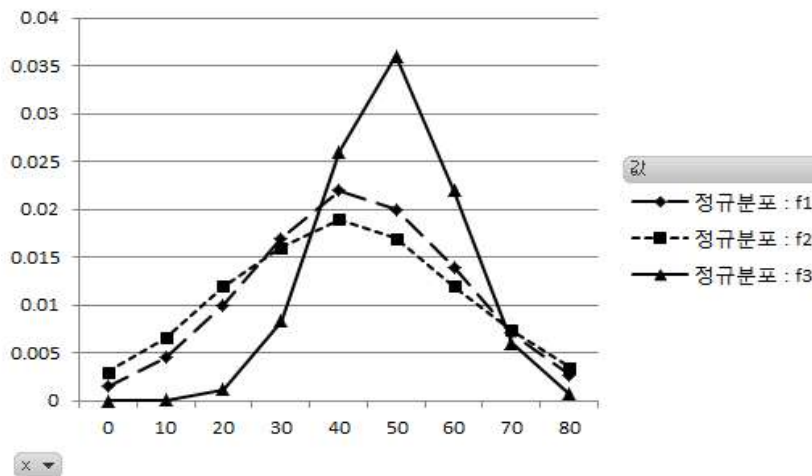
$$\sigma := 12.2 \quad u := 49$$



[그림 3-20] Ex 1-2~3의 피어슨 상관계수 및 정규분포

그 결과, [Ex 1-1], [Ex 1-2] 및 [Ex 1-3] 중 확률적인 중간 극한 정리(Central-limit theorem) 입장에서의 시스템 안정화를 기대할 수 있는 경우는 [그림 3-21]을 통해 알 수 있듯이, [Ex 1-1]이 이에 해당된다.

정규분포 : f1 정규분포 : f2 정규분포 : f3



[그림 3-21] 정규분포 비교

즉, D급 이하에 속한 후보 주제어들에게 부여된 점수 10점을 그대로 적용시키면서 D급 이하에 있는 개체 수를 총 개체 수에서 포함시키는 것이 시스템 안전성에 확률적으로 도움이 됨을 미시적으로나마 확인할 수 있었다.

정리하자면, 특허회피 설계안 자동 생성 방법 및 봇 기반에서 구현되는 시스템의 안정화를 위한 방안은 D급 이하의 주제어들에게 부여된 점수를 반영하는 것이 적합하리라 여겨지며, 두 주제(핵심 키워드와 S급 특허)간의 연관성을 향상시키기 위한 방안으로서는 D급 이하의 주제어들에게 부여된 가중치 점수를 배제시키는 것이 적합하다 할 것이다.

봇 기반의 시스템 안전성과 두 주제간의 연관성 즉, 이 둘 간의 공통점은 D급 이하의 주제어들의 개체 수를 그대로 반영시켰다는 점이라 하겠다.

다음은 [Ex 1-1]을 대상으로 피어슨 상관계수(r)를 구하고, 유사성 통과 주제어들을 연관성 통과 주제어들로 골라내는 2차 필터링에 관한 내용을 정리한 사항이다.

[그림 3-22~24]는 2차 필터링을 위해 적용되는 피어슨 상관계수 식으로 피어슨 상관계수(r)를 구하고, 이로부터 확인되는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂)를 알아내는 과정을 나타내고 있음을 바로 인지할 수 있다.

U의 평균값(u_x)은 고객의 관심분야에 속한 핵심 키워드의 IPC 코드를 기준으로 4단계에 걸쳐 레벨 분류하고 유사성 통과 주제어들마다 부여된 가중치 점수들($u_1, u_2, \dots, u_n$)를 모아 평균을 낸 값이다.

C의 평균값(c_x)은 고객의 관심분야 속하면서 특허 전문가가 주관적으로 찾아낸 S급 특허(예 : 한국 실용신안 등록번호 : 20-0486347)의 IPC 코드를 기준으로 4단계에 걸쳐 레벨 분류하고 유사성 통과 주제어들마다 부여받은 가중치 점수들($c_1, c_2, \dots, c_n$)를 모아 평균을 낸 값이다. 그로부터 정리되는 수학적식과 해당 결과값은 다음과 같다.

[Ex 1-1]에서는 U의 평균값(u_x) : 42.5, C의 평균값(c_x) : 44.2, U의 분산(σ_1) : 335.4, C의 분산(σ_2) : 441.0, U의 표준편차($\sqrt{\sigma_1}$) : 18.3,

C의 표준편차($\sqrt{\sigma^2}$) : 21.0이 되면, [Ex 1-1]의 피어슨 상관계수(r)는 식 (1)를 근거로 아래에 구한 바와 같이 0.7096로 나타냄을 확인할 수 있다.

$$r = \frac{272.9}{384.6} = 0.7096$$

이를 통해, [Ex 1-1]에서 구한 피어슨 상관계수(r1)가 0.5 이상임으로 두 주제(핵심 키워드와 S급 특허)는 서로 연관성이 있다고 간주된다.

유사성 통과 주제어들에게 부여한 가중치 점수의 설정 범위가 80점에서 10점까지 주어짐으로, [Ex 1-1]에서 구한 피어슨 상관계수(r)인 0.7096과 부합하는 제 1 최적점수(OS₁)를 1:1 매칭을 통해 찾아낼 수 있다. 이에 부합하는 제 1 최적점수(OS₁)는 약 73점이 된다.

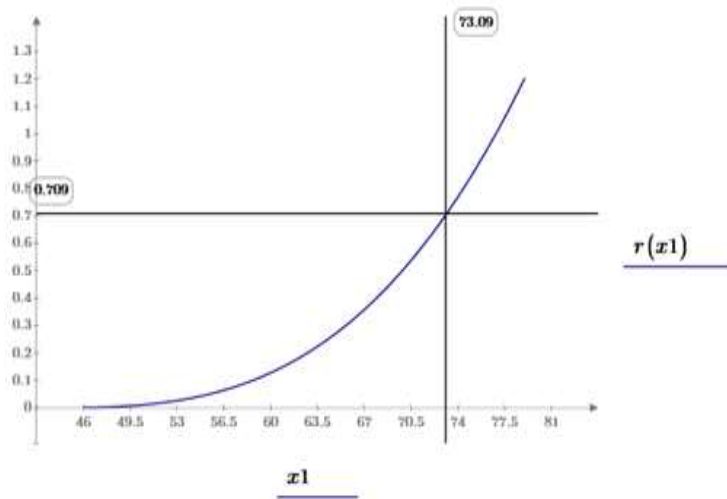
제 1 최적점수(OS₁)가 고정되면, 피어슨 상관계수(r)는 0과 1의 범위를 유지하기 위해 가중치 점수의 상대성 허용 범위를 46점 내지 73점을 갖게 된다.

그에 따라, 제 1 최적점수(OS₁)가 약 73점 일 경우, 이와 대비되는 제 2 최적점수(OS₂)는 약 68점으로 나타냄을 알 수 있다.

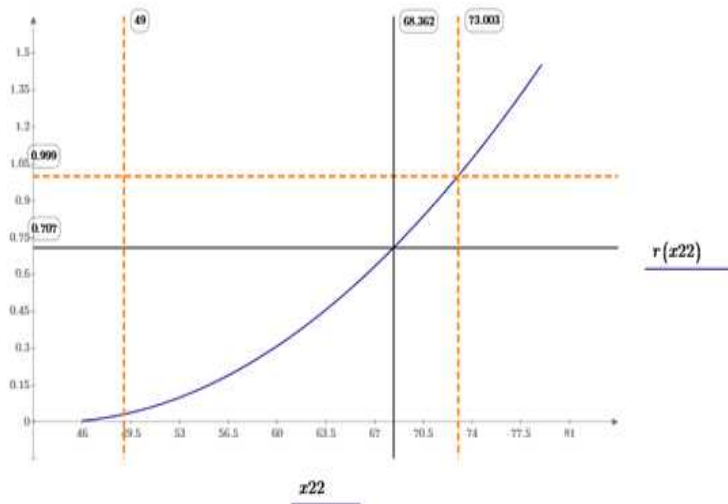
2차 필터링은 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 어느 하나를 골라 연관 최적점수로 정한 뒤, 연관 최적점수와 유사성 통과 주제어들마다 부여받은 가중치 점수를 비교해 연관 최적점수보다 높은 가중치 점수를 갖는 연관성 통과 주제어들을 색출한다. 즉, 2차 필터링의 연관성 판단은 피어슨 상관계수 식을 통해 유사성 통과 주제어들 중에서 연관성 통과 주제어들을 선별해 내는 과정이라 할 것이다.

제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 낮은 점수를 채택한 이유는 유사성 통과 주제어들의 채택 가능성을 높여 양질의 특허회피 설계안을 대량으로 자동 생성될 시 고객이 수용할 수 있을 만한 여러 案이 다양한 각도에서 나올 수 있도록 하기 위해서라 하겠다.

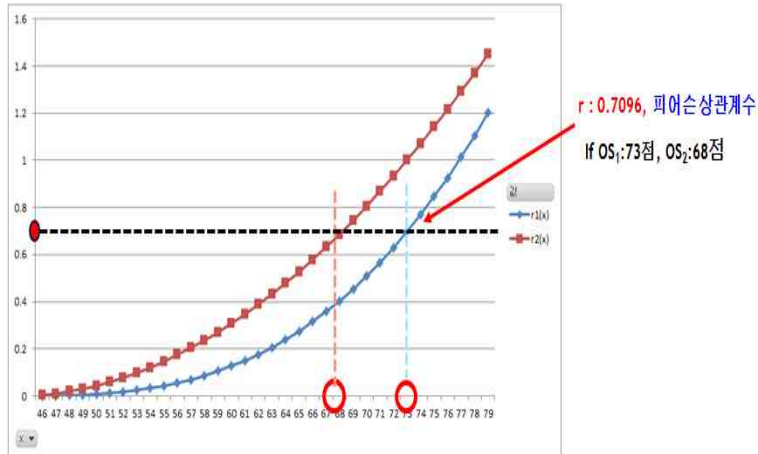
[Ex 1-1]에서는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 낮은 점수로 판명된 약 68점의 연관 최적점수를 기준으로 유사성 통과 주제어들을 판별해 낸다.



[그림 3-22] Ex 1-1의 제 1 최적점수(OS₁)



[그림 3-23] Ex 1-1의 제 2 최적점수(OS₂)



[그림 3-24] Ex 1-1의 r , OS_1 , OS_2

[Ex 1-1]의 2차 필터링은 피어슨 상관계수 식을 이용해 연관성 판단을 실시한 결과, ‘아바타’, ‘정비지원을 위한 시스템’ 주제어들은 가중치 점수가 모두 68점 이상인 것으로 확인되어 특허회피 설계안 자동생성 봇은 좀 더 신뢰할 수 있는 연관성 통과 주제어들을 색출한다.

즉, 봇은 [표 3-3]에서 알 수 있는 바와 같이, 혼합현실(가상현실, 증강현실)분야에 관심 있는 고객에게 ‘아바타’, ‘정비지원을 위한 시스템’이 삽입된 특허회피 설계안을 제시할 수 있을 것이다.

[표 3-3] Ex 1-1의 유사성 통과 주제어

U(유사도)	[1] IPC 코드(유사성 판단) : 핵심 키워드 중심 채택방식
A급, 90%(80점)	아바타(G06T 19/00)
B급, 80%(60점)	문화유적지(H04N 7/14)
C급, 70%(50점)	AR 모바일 애플리케이션(H04N 13/30), 뷰티미러(H04N 21/443), AR놀이 플랫폼(H04N 13/30)
D급, 60%(40점)	모바일기기(H04W 88/02), 원격체험(G06Q 30/06), 정비지원을 위한 시스템(G06Q 50/10), 증강현실 레이싱(H04M 1/725), 디지털 트윈(G06F 3/01)
60%미만 (10점)	수술 시뮬레이션(G09B 23/28), 농작물(A01G 7/06)
핵심키워드	H04N 7/15, G06T 19/00, G03H 1/00

의 IPC	
C(유사도)	[1] IPC 코드(유사성 판단) : S급 특허 중심 채택방식
A급, 90%(80점)	아바타(G06T 19/00), 정비지원을 위한 시스템(G06Q 50/10)
B급, 80%(60점)	모바일기기(H04W 88/02)
C급, 70%(50점)	원격체험(G06Q 30/06)
D급, 60%(40점)	문화유적지(H04N 7/14), AR 모바일 애플리케이션(H04N 13/30), 뷰티미러(H04N 21/443), AR놀이 플랫폼(H04N 13/30), 증강현실 레이싱(H04M 1/725), 디지털 트윈(G06F 3/01)
60%미만 (10점)	수술 시뮬레이션(G09B 23/28), 농작물(A01G 7/06)
S특허의 IPC	G06Q 50/10, H04W 88/06, G06T 19/00, G06Q 50/00

다음은 [Ex 1-2]를 대상으로 피어슨 상관계수(r)를 구하고, 유사성 통과 주제어들 중에서 연관성 통과 주제어들을 골라내는 2차 필터링에 관한 내용을 정리한 사항이다.

[그림 3-25~27]는 2차 필터링을 위해 적용되는 피어슨 상관계수 식을 이용해 피어슨 상관계수(r)를 구하고, 이로부터 확인되는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂)를 알아내는 과정을 나타내고 있음을 바로 인지할 수 있다.

[Ex 1-2]에서는 U의 평균값(ux) : 40.8, C의 평균값(cx) : 42.5, U의 분산(σ^2) : 457.6, C의 분산(σ^2) : 568.8, U의 표준편차 ($\sqrt{\sigma^2}$) : 21.4, C의 표준편차($\sqrt{\sigma^2}$) : 23.8 이 되면, [Ex 1-2]의 피어슨 상관계수(r)는 식(1)를 근거로 아래에 구한 바와 같이 0.78로 나타냄을 확인할 수 있다.

$$r = \frac{397.9}{510.2} = 0.78$$

다시 말해, [Ex 1-2]에서 구한 피어슨 상관계수(r)가 0.5 이상임으로 두 주제(핵심 키워드와 S급 특허)는 서로 연관성이 있다고 간주된다.

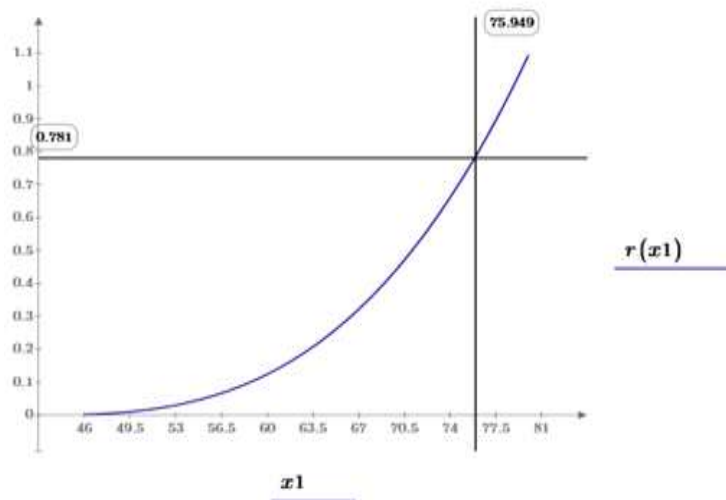
유사성 통과 주제어들에게 부여한 가중치 점수의 설정 범위가 80점에서

0점까지 주어짐으로, [Ex 1-2]에서 구한 피어슨 상관계수(r) 0.78과 부합하는 제 1 최적점수(OS_1)를 1:1 매칭을 통해 찾아 낼 수 있다. 이에 부합하는 제 1 최적점수(OS_1)는 약 76점이 된다.

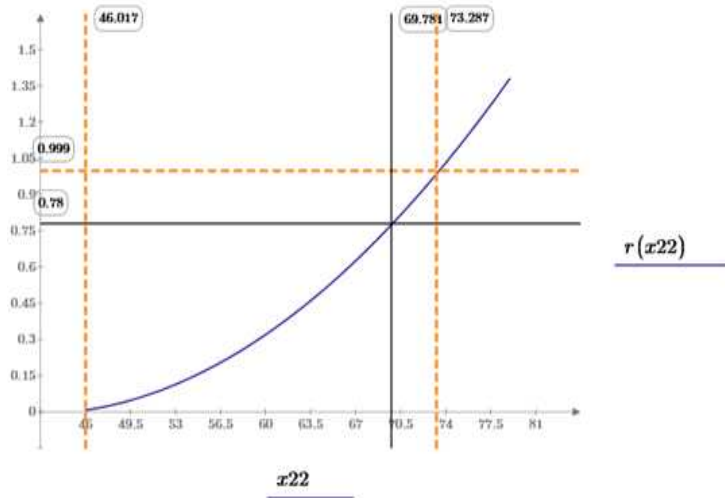
제 1 최적점수(OS_1)가 고정되면, 피어슨 상관계수(r)는 0과 1의 범위를 유지하기 위해 가중치 점수의 상대성 허용 범위를 46점 내지 74점을 갖게 된다. 그에 따라, 제 1 최적점수(OS_1) 약 76점일 경우, 이와 대비되는 제 2 최적점수(OS_2)는 약 70점으로 나타냄을 알 수 있다.

2차 필터링은 제 1 최적점수(OS_1)와 제 2 최적점수(OS_2) 중 어느 하나를 골라 연관 최적점수를 정한 뒤, 연관 최적점수와 유사성 통과 주제어들마다 부여받은 가중치 점수를 비교해 연관 최적점수보다 높은 가중치 점수를 갖는 연관성 통과 주제어들을 색출한다. 즉, 2차 필터링의 연관성 판단은 피어슨 상관계수 식을 통해 유사성 통과 주제어들 중에서 연관성 통과 주제어들을 선별해 내는 과정이라 할 것이다.

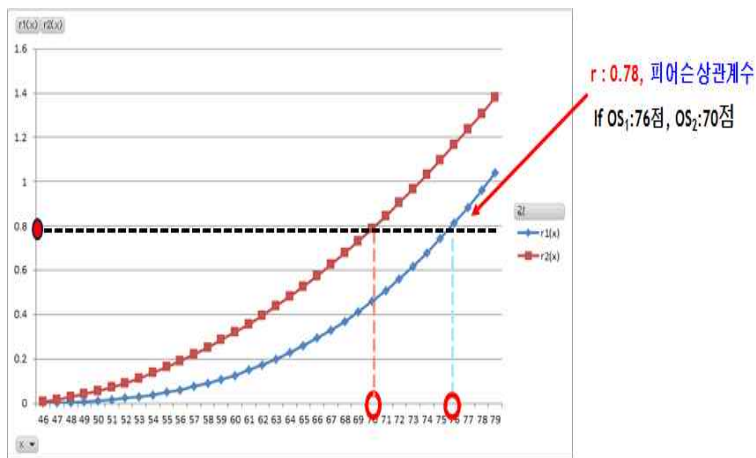
[Ex 1-2]에서는 제 1 최적점수(OS_1)와 제 2 최적점수(OS_2) 중 낮은 점수로 판명된 약 70점의 연관 최적점수를 기준으로 유사성 통과 주제어들만 판별해 낸다.



[그림 3-25] Ex 1-2의 제 1 최적점수(OS_1)



[그림 3-26] Ex 1-2의 제 2 최적점수(OS₂)



[그림 3-27] Ex 1-2의 r , OS₁, OS₂

[Ex 1-2]의 2차 필터링은 피어슨 상관계수 식을 이용한 연관성 판단을 실시한 결과, ‘아바타’, ‘정비지원을 위한 시스템’ 주제어들은 모두 70점 이상인 것으로 확인된 바, 특허회피 설계안 자동생성 봇은 좀 더 신뢰할 수 있는 유사성 통과 주제어들을 색출한다.

즉, 봇은 [표 3-4]에서 알 수 있듯이, 혼합현실(가상현실, 증강현실)분야

에 관심 있어 하는 고객에게 ‘아바타’, ‘정비지원을 위한 시스템’만을 연관성 통과 주제어들로 제시한다.

[표 3-4] Ex 1-2의 유사성 통과 주제어

U(유사도)	[1]IPC 코드(유사성 판단) : 핵심 키워드 중심 채택방식
A급, 90%(80점)	아바타(G06T 19/00)
B급, 80%(60점)	문화유적지(H04N 7/14)
C급, 70%(50점)	AR 모바일 애플리케이션(H04N 13/30), 뷰티미러(H04N 21/443), AR놀이 플랫폼(H04N 13/30)
D급, 60%(40점)	모바일기기(H04W 88/02), 원격체험(G06Q 30/06), 정비지원을 위한 시스템(G06Q 50/10), 증강현실 레이싱(H04M 1/725), 디지털 트윈(G06F 3/01)
60%미만 (0점)	수술 시뮬레이션(G09B 23/28), 농작물(A01G 7/06)
핵심키워드 의 IPC	H04N 7/15, G06T 19/00, G03H 1/00
C(유사도)	[1]IPC 코드(유사성 판단) : S급 특허 중심 채택방식
A급, 90%(80점)	아바타(G06T 19/00), 정비지원을 위한 시스템(G06Q 50/10)
B급, 80%(60점)	모바일기기(H04W 88/02)
C급, 70%(50점)	원격체험(G06Q 30/06)
D급, 60%(40점)	문화유적지(H04N 7/14), AR 모바일 애플리케이션(H04N 13/30), 뷰티미러(H04N 21/443), AR놀이 플랫폼(H04N 13/30), 증강현실 레이싱(H04M 1/725), 디지털 트윈(G06F 3/01)
60%미만 (0점)	수술 시뮬레이션(G09B 23/28), 농작물(A01G 7/06)
S특허의 IPC	G06Q 50/10, H04W 88/06, G06T 19/00, G06Q 50/00

여기서, [Ex 1-3]은 피어슨 상관계수(r)가 기준치 0.5보다 낮은 0.3167이 나와 연관성 통과 주제어들을 색출하는 2차 필터링의 실험 대상에서 제외시키기로 한다.

정리하자면, [Ex 1-1]과 [Ex 1-2]는 피어슨 상관계수(r)가 각각 다르게

나왔음에 주목해야 할 것이다. [Ex 1-1]의 경우에는 D급 이하의 유사성 통과 주제어들의 가중치 점수를 10점으로 부여하고 D급 이하의 유사성 통과 주제어들을 총 개체 수에 포함시켰으나, [Ex 1-2]의 경우에는 D급 이하의 유사성 통과 주제어들의 가중치 점수를 0점을 부여했다는 점에서 서로 구별되는 차이점이 있다.

[Ex 1-1]과 [Ex 1-2]는 수동방식으로 결과를 미리 추정해 본 것이어서, [Ex 1-1]의 피어슨 상관관계수(r)와 [Ex 1-2]의 피어슨 상관관계수(r)는 각기 다르게 나왔으나, [Ex 1-1]과 [Ex 1-2]에서의 연관성 통과 주제어들의 변경 사항은 없음으로 확인하였다.

만약, 다른 실험에서 M개의 정보글로부터 색출된 후보 주제어 그룹의 개체 수가 늘어나 4개의 모집단(A~D급)으로 각기 분류되는 후보 주제어들의 수도 많아진 경우가 발생되었다면, 4개의 모집단(A~D급)으로 각기 분류되는 후보 주제어들은 피어슨 상관관계수(r) 값의 변화에 영향을 끼칠 것이고, 그에 따른 핵심 키워드와 S급 특허 간의 상관관계에도 영향을 줄 수밖에 없음으로 개체 수의 변동은 붓이 연관성 통과 주제어들을 수용하는데 있어 중요 변수로 작용하지 않을까 예측된다.

따라서, 후보 주제어 그룹의 개체 수가 많으면 많을수록 피어슨 상관관계수(r)은 유사성 통과 주제어들을 대상으로 연관성 통과 주제어들로 받아들일 중요한 기준치가 될 수 있음으로 최종 주제어들을 선정하기 위한 중간 매개 역할을 한다고 본다.

3.4.1.3 사례1의 적합성 판단

사례1를 대상으로 적합성 판단을 위한 3차 필터링은 유사성 통과 주제어들을 확장된 오피니언 마이닝 감성평가 기법(EOMSE Scheme : Expansion Opinion Mining Sensitivity Evaluation Scheme)에 적용하여 최종 주제어들을 정한다.(긍정반응 선호)

일반적으로, 오피니언 마이닝이란 키워드 입력에 따라, 선정된 주제어를 대상으로 i개의 의견 데이터를 수집하여 의견 데이터에 내포된 긍정적 표

현, 중립적 표현 또는 부정적 표현을 극성값 +3, +1, 0, -1, -3로 감성 평가하는 자연어 처리기법 중 하나이다.

오피니언 마이닝은 선정된 주제어를 대상으로 감성평가를 실시하기 위해 i 개의 수집된 의견 데이터를 명사, 서술어, 접속사 또는 부사 중 어느 하나 이상의 품사로 구분해 각기 다른 토큰을 생성한다.

의견 데이터 중 서술어(혹은, 동사)의 경우는 서술어의 긍정적 표현, 중립적 표현 또는 부정적 표현 여부에 따라 극성값을 설정하게 된다. 즉, 서술어의 극성값($v.value$)은 [(+1), if 긍정], [(0), if 중립] 및 [(-1), if 부정] 중 선택된 어느 하나로 결정된다.

접속사의 극성값($c.value$)은 서술어 전후의 등위 접속 또는 반의 접속에 따라 접속사의 극성 가중치(+1, -1 중 어느 하나)를 정한 다음, 서술어의 극성값과 접속사의 극성 가중치를 곱해 구하게 된다. 다시 말해, 접속사의 극성값($c.value$)은 $v.value$ 와, [(+1), if 서술어이전 : 긍정, and 등위 접속사], [(-1), if 서술어이전 : 긍정, and 반의어] 및 [(+1), if 서술어이전 : 부정 and 반의어] 중 선택된 하나를 곱해 구한다.

부사의 극성값($ad.value$)은 서술어의 극성값에 따라 부사의 극성 가중치(+3, -3 중 어느 하나)를 정한 다음, 서술어의 극성값과 부사의 극성 가중치를 더해 구하게 된다. 즉, 부사의 극성값($ad.value$)은 $v.value$ 와, [(+3), if 긍정] 및 [(-3), if 부정] 중 선택된 하나를 더해 구하면 된다.⁶²⁾

본 논문에서의 확장된 오피니언 마이닝 감성평가 기법(EOMSE Scheme)은 형용사를 필수적으로 적용한다. 형용사의 극성값($a.value$)은 앞에서 이미 언급된 서술어의 극성값을 구하는데 적용되는 극성 가중치의 설정 포맷과 동일한 형태로 설정되는 것이 바람직하다.

왜냐하면, 형용사의 극성값($a.value$)는 명사를 꾸며주는 역할을 하는 품사로 절과 절을 이어주는 접속사나, 앞과 뒤에 표현된 명사가 의미하는 바에 부합되도록 부가 혹은 부연 설명하는 부사와는 전혀 다른 중요성을 내포하고 있다.

장문이나 단문을 읽거나 해석할 때, 접속사나 부사가 빠져 있더라도 다

62) 정은희, 이병관. (2015). 오피니언 마이닝 기반 SNS 감성 정보 분석 전략 설계, 『한국정보전자통신기술학회』, 8(6), 544-550.

소 불편하긴 하겠지만, 전체 문장을 이해하는데 있어 이 문장이 어떤 내용을 담고 있는지 파악하는데 전혀 불편함을 못 느낄 수 있겠지만, 만약 명사를 꾸며주는 형용사가 없는 문장(ex : 서술어와 명사만으로 구성된 문장)을 읽게 된 경우, 전체 문장을 이해하는데 큰 난관에 봉착할 우려가 있다.

쉬운 예로, “학생이 공부한다”에서 ‘학생’이라는 명사와 ‘공부한다’라는 서술어만으로 형성된 문장이 있을 시, ‘학생’을 꾸며주는 형용사가 어떤 용어가 삽입되느냐에 따라 “학생이 공부한다”라는 원문 전체가 갖는 본래의 의미와 전혀 달라질 수 있을 것이다.

즉, ‘학생’이라는 명사 앞에 현재 신분을 나타내는 용어, “재학 중인, 휴학한, 복학한, 졸업한, 입학한” 단어들 중 어떤 단어를 선택하느냐에 따라 공부하는 주체가 원래 의도했던 바가 아닌 다르게 왜곡되어진 문장으로 충분히 해석될 수 있다.

따라서, 형용사가 갖는 중요성은 서술어와 비교해보더라도 전혀 손색없는 언어학적인 가치성을 서술어만큼 충분히 갖추고 있다고 여겨짐으로, 본 논문에서는 형용사의 극성값의 극성 가중치를 서술어의 극성값이 갖는 극성 가중치와 동일한 범주에 두는 것을 전제로 한다.

EOMSE 기법에 적용되는 형용사는 형용사의 긍정적 표현, 중립적 표현 또는 부정적 표현 여부에 따라 극성값을 설정하게 된다.

긍정적 표현의 형용사이면 (+1), 중립적 표현의 형용사이면 (0), 부정적 표현의 형용사이면 (-1)로 설정한다. 다시 말해, 형용사의 극성값은 서술어와 마찬가지로 [(+1), if 긍정], [(0), if 중립] 및 [(-1), if 부정] 중 선택된 어느 하나로 결정된다.

a.value : [(+1), if 긍정], [(0), if 중립] 및 [(-1), if 부정] 중 선택된 어느 하나로 결정

EOMSE 기법에서 활용되는 각 주제어들마다 부여되는 극성값[o.value]은 서술어의 극성값, 접속사의 극성값, 부사의 극성값 그리고 형용사의 극성값의 합으로 구성되며, 4개의 품사의 합이 0보다 큰 극성값이면 그 주제어는 긍정적 의미를 갖는 명사라 할 수 있고, 4개의 품사의 합이 0이면

그 주제는 중립적 의미를 갖는 명사라 할 수 있겠고, 4개의 품사의 합이 0보다 작은 극성값이면 그 주제는 부정적 의미를 갖는 명사라 할 것이다.

본 논문의 EOMSE 기법은 유사성 통과 주제어들 또는 연관성 통과 주제어들을 식(2)에 적용시켜 감성평가를 실시하며, n개의 주제어들의 극성값을 모두 합하며, 모두 합한 값을 주제어들의 수 n으로 나눠 산술평균치를 구한다.

그런 다음 각 주제어들마다 개별 측정된 극성값(기대치)와 산술 평균치를 각각 비교하는 동작을 실시한다. 만약, k번째 주제가 갖는 기대치($o.value_k$)가 평균치($Opinion_{average}$)보다 크면, k번째 주제는 EOMSE 감정평가가 우수한 것으로 평가되어 붓이 향후 특허회피 설계안을 제시하는데 용이하게 쓸 최종 주제로 활용된다.

$$o.value_k > Opinion_{average} = \sum_{i=1}^n o.value_i \quad (3)$$

i : 발췌한 주제어 수

만약, k번째 주제가 갖는 기대치($o.value_k$)가 평균치보다 작거나 같으면, k번째 주제는 EOMSE 감정평가가 비우수한 것으로 구별되어 향후 특허회피 설계안을 제시하는데 쓰지 않을 목적으로 선별된다.

정리하자면, EOMSE 기법은 1차, 2차 필터링 과정을 거쳐 발췌된 유사성 통과 주제어들 또는 연관성 통과 주제어들의 수가 n개인 경우, 발췌된 주제어들의 극성값의 모두를 합한 값을 n으로 나눠 산술 평균치를 구한 다음 각 주제어들이 갖는 기대치와 산술 평균치를 비교해 특허회피 설계안으로 유용할 최종 주제어들만 채택한다는 것이다.

사례1의 EOMSE 기법은 유사성 통과 주제어들 또는 연관성 통과 주제어들로부터 발췌된 주제의 기대치가 평균치보다 크면, EOMSE 기법에 부합하는 주제어인 것으로 간주해 [표 3-5]처럼, 식(3)을 적용함에 따라 평균치 2.556 이상인 주제어들만 최종 선발한다.

즉, 문화 유적지의 극성값(기대치)은 평균치 2.556 보다 큰 5, 디지털 트

원의 극성값(기대치)도 5, 아바타의 극성값(기대치)은 최고점수 7로 책정됨에 따라, 최종 주제어들은 “문화 유적지, 디지털 트윈, 아바타”로 선정된다는 것이다.

[표 3-5] 사례1의 적합성 판단 실시

[3]EOMSE 기법 적용(적합성 판단)			
	o.value	v.value + c.value + ad.value + a.value	값
1	정비지원 을 위한 시스템	휴대가 용이한 모바일 기기를 기반으로 정비사에게 정비할 부분에 있어서 기술적인 지원이 가능하도록 해주는 증강현실을 활용한 새로운 선박에서의 정비 지원을 위한 시스템	
		0+0+0+1 = 1	1
2	모바일 기기	휴대가 용이한 모바일 기기를 기반으로 정비사에게 정비할 부분에 있어서 기술적인 지원이 가능하도록 해주는 증강현실을 활용한 새로운 선박에서의 정비 지원을 위한 시스템	
		0+0+0+1 = 1	1
3	문화 유적지	케이 걸쳐 타임머신 어플리케이션은 창덕궁을 대상으로 한 시범 서비스로 사용자는 VR기기에 스마트폰을 장착해 제공되는 360도 비디오로 문화 유적지를 원격으로 체험	
		1+0+4+0 = 5	5
4	원격 체험	케이 걸쳐 타임머신 어플리케이션은 창덕궁을 대상으로 한 시범 서비스로 사용자는 VR기기에 스마트폰을 장착해 제공되는 360도 비디오로 문화 유적지를 원격으로 체험	
		0+0+0+0 = 0 ; 오류	x
5	AR 모바일 애플리 케이션	고객은 자신의 휴대폰에 AR 모바일 애플리케이션(앱)을 다운로드 받아 이용할 수 있다. 휴대폰을 매장에 설치된 센서에 갖다 대면 컬렉션 의류를 착용하고 포즈를 취하는 모델 영상을 보여준다	
		1+0+0+0 = 1	1
6	뷰티 미러	'뷰티 미러' 어플은 스마트폰 카메라를 통해 가상으로 라네즈 제품을 적용한 메이크업을 체험	
		1+0+0+0 = 1	1
7	증강	증강현실 콘텐츠를 기반으로 한 최첨단 레이싱	

	현실 레이싱	$1+0+0+0 = 1$	1
8	AR 놀이플 랫폼	AR(증강현실) 놀이플랫폼 '생생체험학습' 부스에서는 아이가 그린 그림이 ... 속으로 나타나는 '물고기그리 기' 등의 AR 콘텐츠를 이용해 볼 수 있다	
		$1+0+0+0 = 1$	1
9	디지털 트윈	살아있는 인체가 환경 속 돌발 요인으로 손상을 입으 면 ... 신호가 태블릿 속 디지털 트윈에 실시간으로 반영될 수 있도록 하는 것이다	
		$1+0+4+0 = 5$	5
10	아바타	증강현실로 구현된 상대방의 아바타와 마주하고 실시 간으로 소통할 수 있으며	
		$1+1+4+1 = 7$, 최고점수	7
평 균		$23 / 9 = 2.556$	23

[그림 3-28]은 2차 필터링의 결과로 선정된 유사성 통과 주제어들인 ‘정비지원을 위한 시스템’, ‘아바타’와, 3차 필터링의 결과로 선정된 연관성 통과 주제어들인 ‘아바타’, ‘문화유적지’, ‘디지털 트윈’을 조합해 3가지의 경우로 특허회피 설계안을 제시해 줄 수 있음을 예시한다.

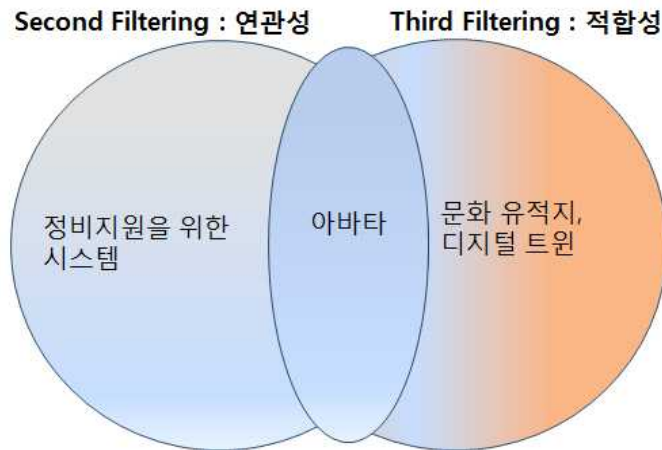
첫째, Second Filtering(S)의 유사성 통과 주제어들 and Third Filtering(T)의 연관성 통과 주제어들 간의 교집합으로 결정된 ‘아바타’가 최종 주제어로 선정되는 경우이다.

둘째는 S or T 간의 합집합으로 결정된 ‘정비지원을 위한 시스템’, ‘아바타’, ‘문화유적지’, ‘디지털 트윈’이 최종 주제어로 선정되는 경우이다.

셋째는 T만으로 결정된 ‘아바타’, ‘문화유적지’, ‘디지털 트윈’이 최종 주제어들로 선정되는 경우라 할 것이다.

이와 같이, 3가지 경우로 자동 생성되는 특허회피 설계안은 핵심 키워드 즉, ‘증강현실’, ‘가상현실’이 표기되어야 바람직하다.

또한, 2차 필터링(연관성 판단)과 3차 필터링(적합성 판단)을 통과한 최종 주제어의 앞뒤에 표기된 문구들이나 어구들은 M개의 정보글 중에서 S에 속한 유사성 통과 주제어들, T에 속한 연관성 통과 주제어들이 기재된 정보글로부터 가져와 적용하는 것이 바람직하다고 할 것이다.



[그림 3-28] 사례1의 최종 주제어 선정

왜냐하면, S에 속한 유사성 통과 주제어들, T에 속한 연관성 통과 주제어들은 특허회피 설계안을 제시해 주는데 아주 중요한 단어이나, 이 단어가 아무 근거 없이 얻어진 것이 아니라 포털 사이트로부터 확보된 정보글로부터 가져온 것이다.

이 정보글 내 기재된 문구들이나 어구들은 S에 속한 유사성 통과 주제어들, T에 속한 연관성 통과 주제어들을 언어적으로 좀 더 쉽게 이해하고 글의 맥락을 서로 이어주는 역할까지 한다고 볼 수 있으므로 특허회피 설계안의 언어적 품격을 조절하는 자원으로 활용 가치가 있을 뿐만 아니라, S에 속한 유사성 통과 주제어들, T에 속한 연관성 통과 주제어들의 조합을 실질적으로 분석하는데 도움을 주기 때문이라 하겠다.

3.4.2 사례2: 핵심 키워드가 ‘사물인터넷’, ‘IoT’인 경우

3.4.2.1 사례2의 유사성 판단

3.4.1장과 다른 모의실험의 사례2는 핵심 키워드를 ‘사물인터넷’, ‘IoT’로 설정한 후 특허청 정보 검색사이트인 키프리스에 접속해 검색창에 “사물인

터넷*IoT*IPC”로 된 1차 검색식을 입력해 특허검색 건수 784를 조회하거나, “사물인터넷*IoT*상황인지*IPC”로 2차 검색식을 입력해 특허검색 건수 238를 조회하거나, “사물인터넷*IoT*M2M*IPC”로 된 3차 검색식을 입력해 특허검색 건수 75를 조회한다.

그 다음, 주요특허로 여길만한 13건을 추려내고 13건의 주요특허에 기재된 특허청구범위 중 독립항을 탐독해 독립항에 기재된 권리범위 내에서 특허 키워드를 13개 추출해 낸다.(단, 특허 키워드는 <조건 1>동일한 목적어가 반드시 출현된 것이어야 할 것, <조건 2>동일한 목적어가 5회 이상 중복 기재된 것임을 전제로 함)

사례2를 통해 추출되는 13개의 특허 키워드는 “기기식별정보, 맥내, 암호화, 센싱 데이터, 통신채널, 리셋모드, 게이트웨이, 디바이스, 네트워크, 제어상태, 생체정보, 고유식별정보, 비대칭키”가 그 대상이 된다.

모의실험의 사례2는 13개의 특허 키워드를 포털 사이트(ex : 네이버) 검색창에 삽입해 특허회피 설계안을 제시하는데 유용되는 정보글을 15개 추출해 낸다.

이때, 단순 조건식은 (1)‘네이버’에서 제공되는 다수의 정보글 중에서 동일한 목적어가 2개 이하로 표현된 정보글일 것, (2)‘네이버’에서 제공되는 다수의 정보글 중에서 목적어당 중복 기재수가 2 내지 3회 이하인 정보글일 것으로 임의 설정한다.

그러나, 바람직한 조건식은 앞서 3.3.2장에서 설명한 바와 같이 고객의 관심분야이면서 (1)전체 클릭율을 1로 봤을 때 어느 한 시점의 해당 클릭율이 적어도 0.5 이상 일어난 정보, (2)3년을 기준으로 적어도 1회에서 10회 이내에서 급상승했던 이력이 있는 정보이어야 할 것이다.<그림 3-29 참조>

모의실험의 사례2는 ‘네이버’로부터 추출해 낸 15개의 정보글을 바탕으로 핵심 키워드, 특허 키워드, 중복 명사를 배제시킨 후보 주제어들을 집합한 후보 주제어 그룹을 생성시킨다.

이로부터 생성된 후보 주제어 그룹은 “게임기기, 의용생체, 행동분석, ICT, 게이트웨이, 데이터베이스, AES(Advanced Encryption Standard),

RF 차폐, 유전막, 임베디드, 센싱정보, DMB, 탄성체, RFID, MDM(Mobile Device Management)”이다.

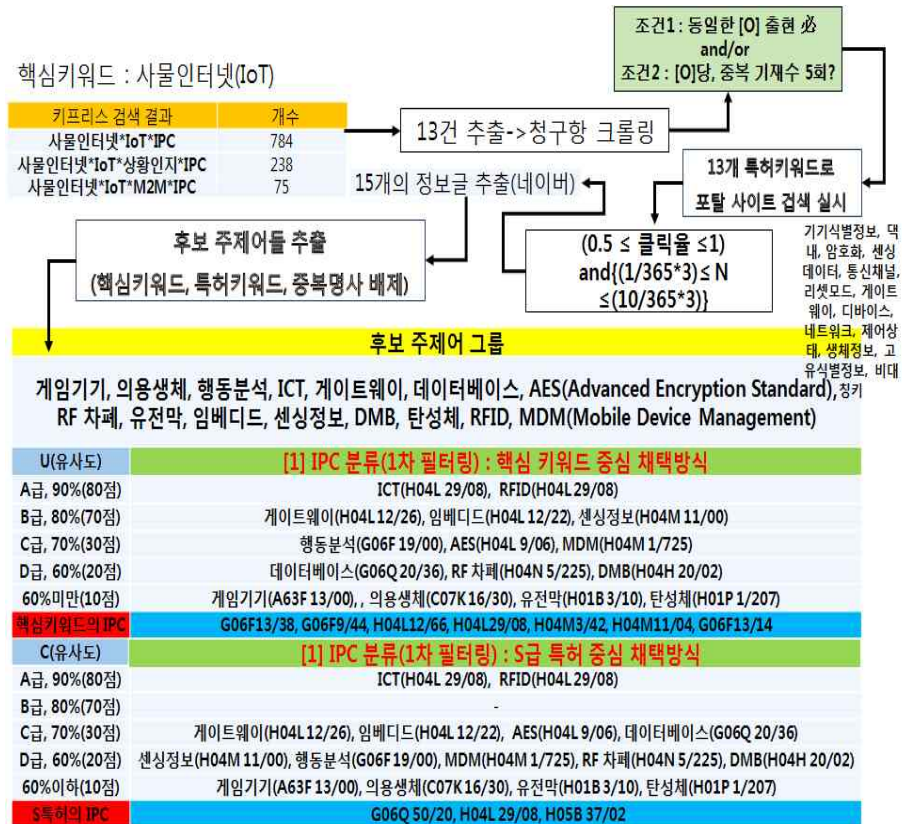
즉, 사례2에서 채택한 핵심 키워드인 ‘사물인터넷’, ‘IoT’와 후보 주제어 그룹 내 후보 주제어들 간의 관련성 여부를 파악하기 위한 유사성 판단을 실시함으로써 통해 4개의 모집단으로 레벨 분류되는 1차 필터링 동작을 실시한다. 사례2의 유사성 판단은 IPC 분류표를 활용해 핵심 키워드 중심 채택방식(U), S급 특히 중심 채택방식(C)으로 나눠 각기 실시된다.

사례2의 유사성 판단은 IPC 분류표에서 확인되는 G06 F 13/38, G06 F 9/44, H04 L 12/66, H04 L 29/08, H04 M 3/42, H 04 M 11/04, G06 F 13/14를 갖는 ‘사물인터넷’, ‘IoT’와, ICT(H04 L 29/08), RFID(H04 L 29/08), 게이트웨이(H04 L 12/26), 임베디드(H04 L 12/22), 센싱정보(H04 M 11/00), 행동분석(G06 F 19/00), AES(H04 L 9/06), MDM(H04 M 1/725), 데이터베이스(G06 Q 20/36), RF 차폐(H04 N 5/225), DMB(H04 H 20/02), 게임기기(A63 F 13/00), 의용생체(C07 K 16/30), 유전막(H01 B 3/10), 탄성체(H01 P 1/207)로 된 후보 주제어 그룹의 IPC 코드를 각기 비교하여 정규 분포 및 레벨 분류에 따른 등급산정을 객관적인 지표로 명확히 하고자 4개의 모집단을 80점, 70점, 30점, 20점으로 각기 구분해 점수화시킨다.(단, 60%이하의 유사도를 갖는 변외 모집단은 10점의 점수로 할당한다)

유사성 판단은 핵심 키워드 중심 채택방식(U), S급 특히 중심 채택방식(C)으로 각각 나눠 IPC 분류표에 의거해 4개의 모집단으로 각각 확률적으로 레벨 분류하게 된다. 4개의 모집단은 90% 유사도를 갖는 제 1 모집단, 80% 유사도를 갖는 제 2 모집단, 70% 유사도를 갖는 제 3 모집단 및 60% 유사도를 갖는 제 4 모집단으로 분류된다.

이 때, 유사성 판단은 확률적으로 레벨 분류된 4개의 모집단을 대상으로 가중치 점수를 부여하게 되는데, 가중치 점수 부여에 대한 타당성은 다음과 같은 실증적 검증을 통해 점수화시킬 수 있음을 충분히 이해할 수 있을 것이다.⁶³⁾

63) 송근원. (2003). 조사 결과의 평가점수 부여에 대한 새로운 접근. 『사회과학연구』, 19(1), 179-197.



[그림 3-29] 사례2의 후보 주제어 색출 및 유사성 판단

즉, 일반적인 평가기준에서 바라봤을 때, 5점 척도에 4점이 나오면 ‘우수’로 평가받는다. 이를 100점 기준으로 환산하게 되면 75점이 되게 되는데 이는 일반적인 평가기준에서 ‘보통’ 수준으로 취급받게 된다. 예컨대, XX신문에 oo구청 공무원들의 친절도 평가점수가 61점으로 대서 특필되면, 대부분의 일반 국민들은 이를 두고 “형편없는 점수를 받았구나”라고 인식하기 쉽다.

이러한 사고는 일부 편파성 보도나 방송, 어떤 기사에 실린 내용 그대로를 믿은 나머지 자신의 이성적 판단이나 통찰적 이해, 객관적 관점에서 바라보고자 하는 노력이라곤 전혀없이 어떤 선입견이나 편견에 빠져 집단적 오류에 빠질 가능성이 높다는 데에서 그 원인을 찾을 수 있을 것이다.

이러한 편파성 오류를 벗어나기 위한 노력의 일환으로는 평가점수가 왜 이렇게 나왔을까를 궁금해 하는 과학적 발상이자 비판적 사고를 가져야

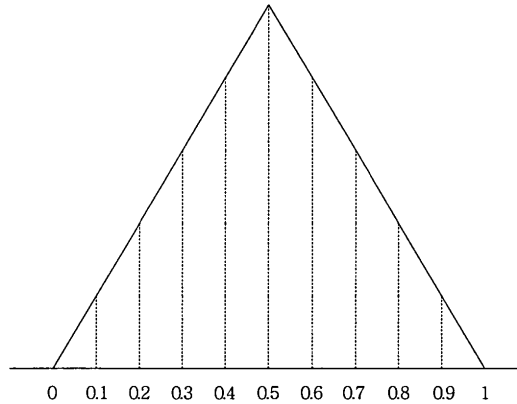
할 것이며, 해당 원인이 발생하게 된 계기나, 주변배경, 동기 그리고, 그 외 여러가지 상황, 외적요인 등을 고려함과 더불어 평가기준을 설정하는데 있어서도 단지 상대평가 혹은 절대평가로만 국한시키는 것이 아니라 보다 일반인의 정서에 맞는 평가 점수 방법의 개발이 요구되어 왔다.

따라서, [그림 3-30]의 삼각형 확률 분포도처럼 응답의 중앙화 경향을 고려하여 응답자의 비율이 중앙을 중심으로 비례적으로 감소한다고 가정하고 등급 분포표에 맞추어 평가 점수를 올바르게 환산해 주는 S 등급분포표 및 S 점수환산표 적용방법은 일반인의 정서에 맞는 평가 점수 방법 중 하나에 속하면서 상기 문제를 해결하는데 도움이 되는 좋은 사례이다.(삼각형에서 점선으로 표시된 높이는 응답자의 수를 상징한다)

삼각형 확률 분포도에서 보듯이, 밑변의 길이가 1이고, 높이가 y , 전체 넓이는 100일 때, 삼각형의 높이 구하는 공식에서의 y 는 200이 된다. x 가 0.5일 때 $y = 200$ 이라면, $x : y' = 0.5 : 200$ 라는 비례식이 성립이 된다고 할 것이다. 이를 다시 풀이하면, y' 는 $400x$ 이 된다.

따라서, 10% 면적에 해당하는 밑변의 길이가 x , 높이를 y' 라 하면 $10 = xy'/2$ 가 되고, 다시 풀면 $10 = x(400x)/2$ 가 되어 $x^2 = 0.05$ 즉, $x = 0.2236$ 이 나온다.

같은 방식으로 20% 면적일 때 x 는 31.42, 30% 면적일 때 x 는 38.73, 40% 면적 : 0.4472, 50% 면적 : 0.5, 60% 면적 : 0.5528, 70% 면적 : 0.6127, 80% 면적 : 0.6838, 90% 면적 : 0.7764, 100% 면적일 때 x 는 1이 나옴을 알 수 있다.



[그림 3-30] 삼각형 확률 분포도

S 등급분포표 및 S 점수환산표 적용방법은 본 논문의 유사성 판단시 레벨 분류된 4개의 모집단을 대상으로 가중치 점수를 확률적으로 부여하는 가중치 점수 부여에 대한 타당성을 실증적 입증하는데 도움을 준다.

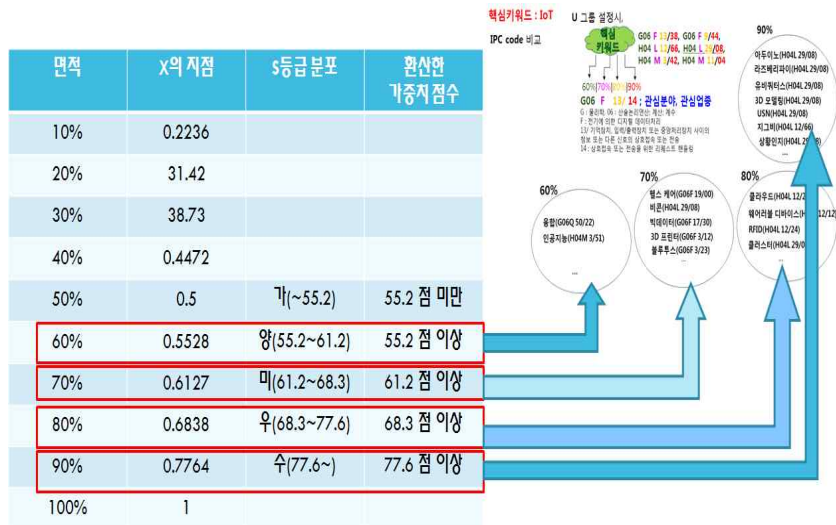
S 등급분포표 및 S 점수환산표 적용방법은 전체 100% 삼각형 면적을 점수화시켜 100점으로 환산함을 전제로 두고 [그림 3-31]에서 보듯이, 삼각형 확률 분포도의 면적이 60%이면, 60% 삼각형 면적으로부터 유추되는 해당 분포지점의 x 값은 0.5528이 됨을 확인할 수 있고 점수화로는 55.28 점으로 환산됨을 알 수 있다.

또한, 70% 삼각형 면적에 대한 해당 분포지점 x 값이 0.6127로 확인되면, 환산된 점수는 61.27이 됨을 파악할 수 있겠고, 80% 삼각형 면적에 대한 해당 분포지점 x 값이 0.6838로 확인되면, 환산된 점수는 68.38이 됨을 파악할 수 있으며, 삼각형 90% 면적에 대한 해당 분포지점 x 값이 환산된 점수는 0.7764로 됨을 파악할 수 있을 것이다.

S 등급분포표 및 S 점수환산표 적용방법에 준하여 삼각형 면적분포에 따라 차등 적용되는 x 지점을 찾아내고 x 지점을 점수화시켜 환산한 가중치 점수는 90%이면 77.6점 이상을 반드시 주어야 하고, 80%이면 77.6 미만 68.3점 이상을 주어야 하며, 70%이면 68.3 미만 61.2점 이상을 주어야 하고, 60%이면 61.2점 미만 55.2점 이상을 주어야 함을 의미하는 것으로 한정하는 것이 아니라, 좀 더 포괄적인 개념으로 봐야하는 것이 옳다고 할

것이다.

즉, S 등급분포표 및 S 점수환산표 적용방법이 제안된 원 취지를 포함하는 포괄적인 개념이라 함은 보다 객관적인 입장에서 봤을 때, 90%의 신뢰도 구간에 들어온 어떤 집단이나 어떤 부류에게 가중치 점수를 부여하기 위해 환산하고자 할시, 반드시 그 해당값을 90점으로 주지 않아도 됨을 입증해 낸 학설로, 어떤 집단이나 어떤 부류에게 90% 신뢰도 구간에 들어왔다 하여도 가중치 점수를 90점이 아닌 다른 가중치 점수를 부여 하더라도 90% 신뢰도 구간은 그대로 유지된다는 것이 올바른 해석이라 하겠다.



[그림 3-31] 가중치 점수 부여에 관한 타당성 검증

예를 들어, 60명이 한 반인 고등학생 학급에서 수학시험을 본다고 가정하고 수학시험 30문항, 한 문제당 3.3점 배점으로 이 수학시험 결과를 가지고 내신성적을 내기 위해 학생 분포표를 이 학급의 담임선생님이 만들어 보겠다고 한 경우, 이 학생 분포표는 어떻게 형성되어야만 바람직할지 고민해봐야 할 사항이라 할 것이다.

시험 결과, 수학시험을 본 60명 학생들의 점수는 다양하게 나올 것이라 당연히 예측된다. 100 점부터 0.1 기본 점수까지 나올 수 있는 경우의 수는 101개라고 할 수 있다. 예를 들어 100점 1명, 93.3점 1명, 86.8점 2명,

73.6점은 3명 70.3점 7명, 63.7점은 10명, 57.1점은 18명 47.2점은 15명 47.2점 이하는 3명이 나왔다고 가정해 본다.

90%, 80%, 70%, 60%에 각각 해당하는 학생 분포는 과연 어느 점수대까지를 봐야 할 것인지가 궁금해 질 것이다.

기존 통념상, 90% 신뢰도에 들어오는 학생 수는 2명(90점이상, 1등급), 80% 신뢰도에 들어온 학생 수는 2명(90점미만 80점이상, 2등급), 70% 신뢰도에 들어온 학생 수는 10명(80점미만 70점이상, 3등급) 등이라 할 것이다. 그러나 1등급부터 10등급까지 내신성적을 줘야하는 선생님 입장에선 이 방식대로 내신성적을 주는 것이 과연 합당한 방법일지 반문해 볼 필요가 있다고 할 것이다.

왜냐하면, 60명의 학생이 시험 본 수학 정답율은 수학 문제의 난이도, 출제선생님의 출제경향에 따라 달라질 수 있기 때문에 내신등급 산정에 있어 이를 반영하지 못하는 문제점이 있어 왔다는 것이다.

그럼으로, S 등급분포표 및 S 점수환산표 적용방법을 인용해 학생 분포도를 만들어 본다고 한다면, 1등급부터 10등급까지 매겨지는 내신등급 산정은 가우시안 확률분포를 더욱 발전시킨 삼각형 확률 분포도에 의거해 보다 객관적으로 산정하였다고 봐도 무관할 것이다.

S 등급분포표 및 S 점수환산표 적용방법의 계산식에 적용해 1등급부터 10등급 중 1등급을 줄 수 있는 가중치 점수는 77.6점 이상이라 할 수 있어, 1 등급을 받을 수 있는 학생 수는 2명에서 4명으로 늘어나 불합리한 기준에 의해 부적절한 등급을 판정받을 수 있었던 통상적 과오로부터 2명의 학생을 구제받을 수 있게 했음을 간파한 것이라 하겠다.

따라서, [그림 3-32]와 같이, 본 논문의 유사성 판단시 적용되는 가중치 점수를 핵심 키워드 중심 채택방식(U) 및 S급 특히 중심 채택방식의 구분 없이 IPC 분류에 따른 유사도가 90%를 가지면 80점을 부여하고, 80%이면 70점, 70% 유사도이면 30점, 60% 유사도면 20점, 60% 이하이면 10점을 부여한다고 해서 S 등급분포표 및 S 점수환산표 적용방법의 본 취지를 어긋나거나 잘못된 논거가 절대 아니라 할 것이다.

이러한 점에서, 본 논문에 S 등급분포표 및 S 점수환산표 적용방법의 포

은 데이터베이스(G06 Q 20/36), RF 차폐(H04 N 5/225), DMB(H04 H 20/02)이 되고, 10점의 가중치 점수로 환산되는 변외 모집단은 게임기기(A63 F 13/00), 의용생체(C07 K 16/30), 유전막(H01 B 3/10), 탄성체(H01 P 1/207)이 된다 할 것이다.

S급 특허 중심 채택방식(C)은 키프리스를 통해 검색된 다수의 주요특허 중에서 특허 전문가가 엄선한 핵심특허(S급 특허 : 10-1746215)로부터 확인되는 IPC 코드[G06 Q 50/20, H04 L 29/08, H05 B 37/02]와 후보 주제어 그룹 속한 후보 주제어들을 모두 비교해 4개 모집단으로 분류해 내는 채택방식이다.

즉, 80점의 가중치 점수로 환산되는 90% 유사도를 갖는 A급 모집단은 ICT(H04 L 29/08), RFID(H04 L 29/08)이 되며, 70점의 가중치 점수로 환산되는 80% 유사도를 갖는 B급 모집단은 그 대상이 없고, 30점의 가중치 점수로 환산되는 70% 유사도를 갖는 C급 모집단은 게이트웨이(H04 L 12/26), 임베디드(H04 L 12/22), AES(H04 L 9/06), 데이터베이스(G06 Q 20/36)이 된다.

또한, 20점의 가중치 점수로 환산되는 60% 유사도를 갖는 D급 모집단은 센싱정보(H04 M 11/00), 행동분석(G06 F 19/00), MDM(H04 M 1/725), RF 차폐(H04 N 5/225), DMB(H04 H 20/02)이 되고, 10점의 가중치 점수로 환산되는 변외 모집단은 게임기기(A63 F 13/00), 의용생체(C07 K 16/30), 유전막(H01 B 3/10), 탄성체(H01 P 1/207)이 된다 할 것이다.

사례2의 유사성 판단은 15개의 주제어들 중 D급 미만에 속하는 주제어들을 필터링하고 나머지 IPC 코드를 이용한 유사성 판단에서 높은 점수를 받아 상위레벨로 구분된 유사성 통과 주제어들(ICT, RFID, 게이트웨이, 임베디드, 센싱정보, 행동분석, AES, MDM, 데이터베이스, RF 차폐, DMB)을 조합해 자연어 처리한다.

그에 따라, 사례2의 유사성 판단은 1차 필터링된 유사성 통과 주제어들을 대상으로 자연어 처리함을 통해 대화형 구문형식에 가장 걸맞는 자국어로 표현해 주는 특허회피 설계안을 제시해 줄 수 있을 것이다.

여기서, 사례2의 유사성 판단은 특허회피 설계안 자동 생성 방법으로부터 예측되는 결과치를 수동 방식으로 미리 추정해 본다는 제한적 여건 하에서 진행되는 것이어서, 2차 필터링에 대한 연관성 검증 및 유사성 통과 주제어들의 개체 수가 적을 수밖에 없는 물리적인 한계를 극복하기 위해 서라도 IPC 코드에 의해 D급 미만으로 분류된 후보 주제어들을 배제시키지 않고 3.4.2.2장에 그대로 적용하기로 한다.

3.4.2.2 사례2의 연관성 판단

다음은 [Ex 2-1]을 대상으로 피어슨 상관계수(r)를 구하고, 유사성 통과 주제어들 중 연관성 통과 주제어들을 골라내는 2차 필터링에 관한 내용을 정리한 사항이다.

[그림 3-33~35]은 2차 필터링을 위해 적용되는 피어슨 상관계수 식을 이용해 피어슨 상관계수(r)를 구하고, 이로부터 확인되는 제 1 최적점수(Optimal Score 1 : OS₁)와 제 2 최적점수(Optimal Score 2 : OS₂)를 알아내는 과정을 나타내고 있음을 바로 인지할 수 있다. 그로부터 정리되는 수학적식과 해당 결과값은 다음과 같다.

[Ex 2-1]에서는 U의 평균값(ux) : 37.3, C의 평균값(cx) : 28.0, U의 분산(σ^2) : 726.2, C의 분산(σ^2) : 469.3, U의 표준편차($\sqrt{\sigma^2}$) : 26.9, C의 표준편차($\sqrt{\sigma^2}$) : 21.7이 되면, [Ex 2-1]의 피어슨 상관계수(r)는 식(1)을 근거로 아래에 구한 바와 같이 0.7559로 나타냄을 확인할 수 있다.

$$r = \frac{441.3}{583.8} = 0.7559$$

여기서,

$$U_n = \{u_1, u_2, \dots, u_n\}$$

U 그룹 : 핵심 키워드의 IPC 코드에 의거해 가중치 점수를 부여받은 유사성 통과 주제어들의 집합이고,

$$C_n = \{c1, c2, \dots, cn\}$$

C 그룹 : 핵심 키워드로 검색된 다수의 특허 건들 중 전문가가 찾아 낸 S급 특허(등록번호 : 10-1746215)의 IPC 코드에 의거해 가중치 점수를 부여받은 유사성 통과 주제어들의 다른 집합임을 전제로 한다.

이를 통해, [Ex 2-1]에서 구한 피어슨 상관계수(r)가 0.5 이상임으로 두 주제(핵심 키워드와 S급 특허)는 서로 연관성이 있다고 간주된다.

유사성 통과 주제어들에게 부여한 가중치 점수의 설정 범위가 90점에서 10점까지 주어짐으로, [Ex 2-1]에서 구한 피어슨 상관계수(r) 0.7559와 부합하는 제 1 최적점수(OS₁)를 1:1 매칭을 통해 찾아낼 수 있다. 이에 부합하는 제 1 최적점수(OS₁)는 약 69점이 된다.

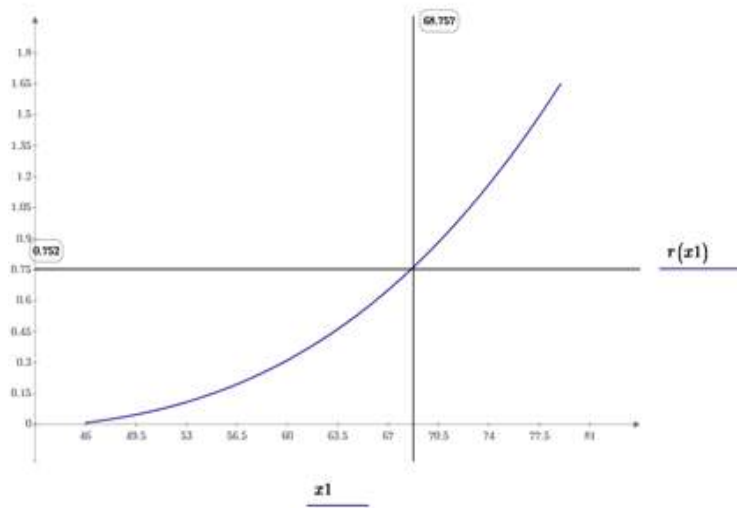
제 1 최적점수(OS₁)가 고정되면, 피어슨 상관계수(r)는 0과 1의 범위를 유지하기 위해 가중치 점수의 상대성 허용 범위를 46점 내지 67점을 갖게 된다.

그에 따라, 제 1 최적점수(OS₁)가 약 69점일 경우, 이와 대비되는 제 2 최적점수(OS₂)는 약 63점으로 나타냄을 알 수 있다.

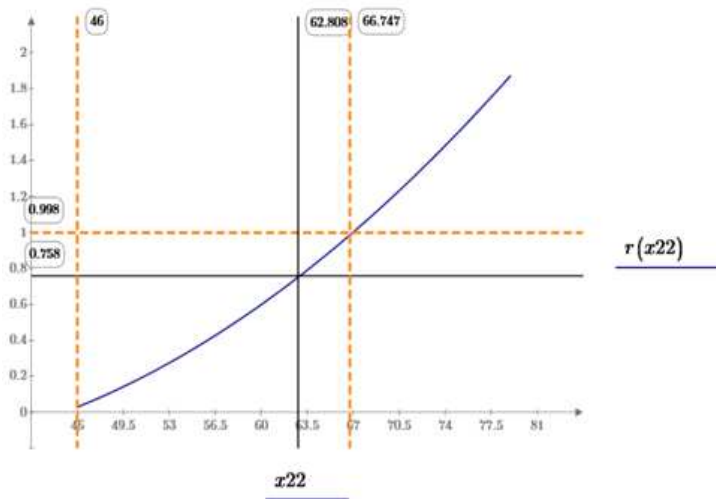
2차 필터링은 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 어느 하나를 골라 연관 최적점수로 정한 뒤, 연관 최적점수와 유사성 통과 주제어들마다 부여받은 가중치 점수를 비교해 연관 최적점수보다 높은 가중치 점수를 갖는 연관성 통과 주제어들을 색출한다.

즉, 2차 필터링의 연관성 판단은 피어슨 상관계수 식을 통해 유사성 통과 주제어들 중에서 연관성 통과 주제어들을 선별해 내는 과정이라 할 것이다.

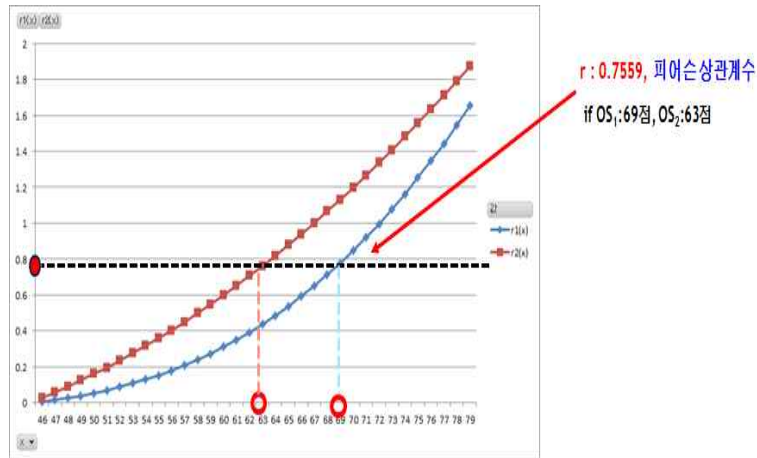
[Ex 2-1]에서는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 낮은 점수로 판명된 약 63점의 연관 최적점수를 기준으로 유사성 통과 주제어들만 판별해 낸다.



[그림 3-33] Ex 2-1의 제 1 최적점수(OS_1)



[그림 3-34] Ex 2-1의 제 2 최적점수(OS_2)



[그림 3-35] Ex 2-1의 r , OS_1 , OS_2

[Ex 2-1]의 2차 필터링은 피어슨 상관계수 식을 이용해 연관성 판단을 실시한 결과, ‘ICT’, ‘RFID’, ‘게이트웨이’, ‘임베디드’, ‘센싱정보’ 주제어들은 모두 63점 이상인 것으로 확인된 바, 특허회피 설계안 자동생성 봇은 좀 더 신뢰할 수 있는 연관성 통과 주제어들을 색출한다.

즉, 봇은 [표 3-6]에 나열된 Ex 2-1의 유사성 통과 주제어들 중에서 사물인터넷(IoT) 분야에 관심 있는 고객에게 ‘ICT’, ‘RFID’, ‘게이트웨이’, ‘임베디드’, ‘센싱정보’만을 연관성 통과 주제어들로 제시한다.

[표 3-6] Ex 2-1의 유사성 통과 주제어

U(유사도)	[1] IPC 코드(유사성 판단) : 핵심 키워드 중심 채택방식
A급, 90%(90점)	ICT(H04L 29/08), RFID(H04L 29/08)
B급, 80%(70점)	게이트웨이(H04L 12/26), 임베디드(H04L 12/22), 센싱정보(H04M 11/00)
C급, 70%(30점)	행동분석(G06F 19/00), AES(H04L 9/06), MDM(H04M 1/725)
D급, 60%(20점)	데이터베이스(G06Q 20/36), RF 차폐(H04N 5/225), DMB(H04H 20/02)
60%미만	게임기기(A63F 13/00), 의용생체(C07K 16/30),

(10점)	유전막(H01B 3/10), 탄성체(H01P 1/207)
핵심키워드 의 IPC	G06F 13/38, G06F 9/44, H04L 12/66, H04L 29/08, H04M 3/42, H04M 11/04, G06F 13/14
C(유사도)	[1] IPC 코드(유사성 판단) : S급 특허 중심 채택방식
A급, 90%(90점)	ICT(H04L 29/08), RFID(H04L 29/08)
B급, 80%(70점)	-
C급, 70%(30점)	게이트웨이(H04L 12/26), 임베디드(H04L 12/22), AES(H04L 9/06), 데이터베이스(G06Q 20/36)
D급, 60%(20점)	센싱정보(H04M 11/00), 행동분석(G06F 19/00), MDM(H04M 1/725), RF 차폐(H04N 5/225), DMB(H04H 20/02)
60%미만 (10점)	게임기기(A63F 13/00), 의용생체(C07K 16/30), 유전막(H01B 3/10), 탄성체(H01P 1/207)
S특허의 IPC	G06Q 50/20, H04L 29/08, H05B 37/02

다음은 [Ex 2-2]를 대상으로 피어슨 상관계수(r)를 구하고, 유사성 통과 주제어들 중에서 연관성 통과 주제어들을 골라내는 2차 필터링에 관한 내용을 정리한 사항이다.

[그림 3-36~38]은 2차 필터링을 위해 적용되는 피어슨 상관계수 식을 이용해 피어슨 상관계수(r)를 구하고, 이로부터 확인되는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂)를 알아내는 과정을 나타내고 있음을 바로 인지할 수 있다.

[Ex 2-2]에서는 U의 평균값(ux) : 34.7, C의 평균값(cx) : 25.3, U의 분산(σ^2) : 891.6, C의 분산(σ^2) : 584.9, U의 표준편차 ($\sqrt{\sigma^2}$) : 29.9, C의 표준편차($\sqrt{\sigma^2}$) : 24.2이 되면, [Ex 2-2]의 피어슨 상관계수(r)는 식 (1)를 근거로 아래에 구한 바와 같이 0.8056로 나타냄을 확인할 수 있다.

$$r = \frac{581.8}{722.1} = 0.8056$$

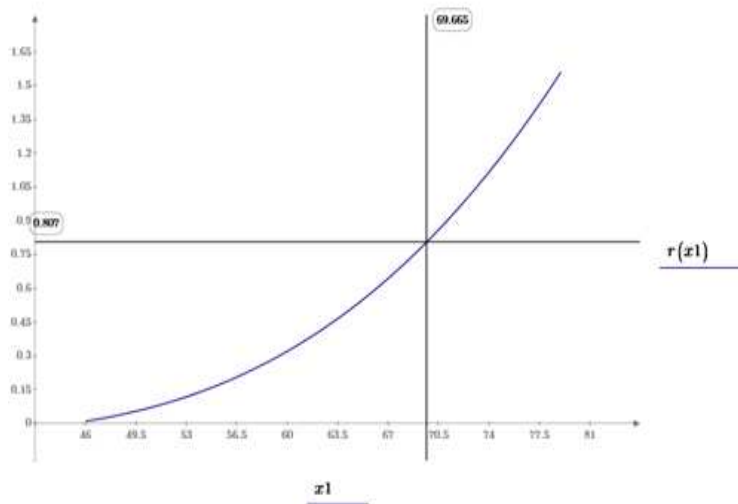
이를 통해, [Ex 2-2]에서 구한 피어슨 상관계수(r)가 0.5 이상임으로 두 주제(핵심 키워드와 S급 특허)는 서로 연관성이 있다고 간주된다.

유사성 통과 주제어들에게 부여한 가중치 점수의 설정 범위가 90점에서 0점까지 주어짐으로, [Ex 2-2]에서 구한 피어슨 상관계수(r) 0.8056과 부합하는 제 1 최적점수(OS_1)를 1:1 매칭을 통해 찾아낼 수 있다. 이에 부합하는 제 1 최적점수(OS_1)는 약 70점이 된다.

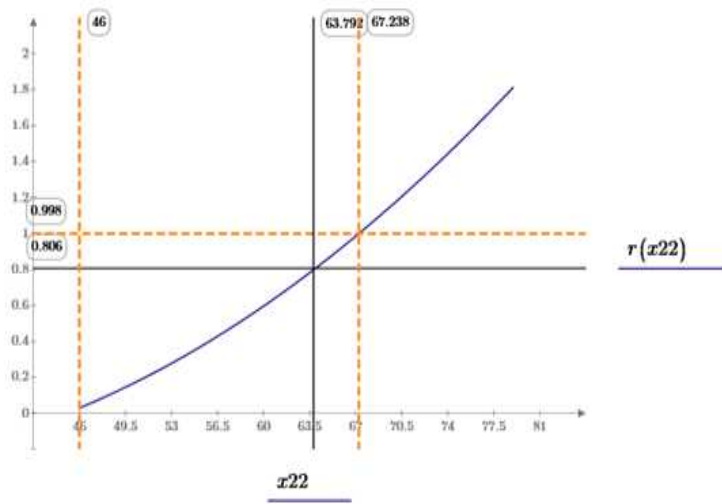
제 1 최적점수(OS_1)가 고정되면, 피어슨 상관계수(r)는 0과 1의 범위를 유지하기 위해 가중치 점수의 상대성 허용 범위를 46점 내지 68점을 갖게 된다. 그에 따라, 객관적 최적점수가 약 70점일 경우, 이와 대비되는 제 2 최적점수(OS_2)는 약 64점으로 나타냄을 알 수 있다.

2차 필터링은 제 1 최적점수(OS_1)와 제 2 최적점수(OS_2) 중 어느 하나를 골라 연관 최적점수로 정한 뒤, 연관 최적점수와 유사성 통과 주제어들마다 부여받은 가중치 점수를 비교해 연관 최적점수보다 높은 가중치 점수를 갖는 연관성 통과 주제어들을 색출한다.

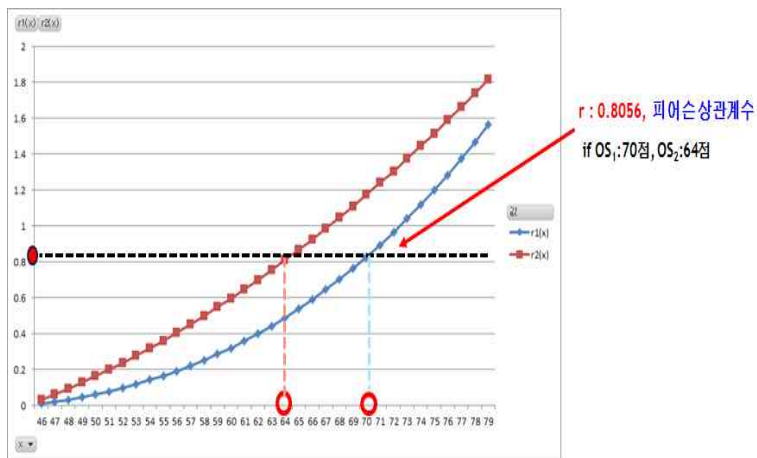
즉, 2차 필터링의 연관성 판단은 피어슨 상관계수 식을 통해 유사성 통과 주제어들 중에서 연관성 통과 주제어들을 선별해 내는 과정이라 할 것이다. [Ex 2-2]에서는 제 1 최적점수(OS_1)와 제 2 최적점수(OS_2) 중 낮은 점수로 판명된 약 64점의 연관 최적점수를 기준으로 유사성 통과 주제어들만 판별해 낸다.



[그림 3-36] Ex 2-2의 제 1 최적점수(OS_1)



[그림 3-37] Ex 2-2의 제 2 최적점수(OS₂)



[그림 3-38] Ex 2-2의 r , OS₁, OS₂

[Ex 2-2]의 2차 필터링은 피어슨 상관계수 식을 이용한 연관성 판단을 실시한 결과, ‘ICT’, ‘RFID’, ‘게이트웨이’, ‘임베디드’, ‘센싱정보’ 주제어들은 모두 64점 이상인 것으로 확인된 바, 특허회피 설계안 자동생성 봇은 좀 더 신뢰할 수 있는 유사성 통과 주제어들을 추천받는다.

즉, 봇은 [표 3-7]처럼, 사물인터넷(IoT) 분야에 관심 있는 고객에게 ‘ICT’, ‘RFID’, ‘게이트웨이’, ‘임베디드’, ‘센싱정보’가 삽입된 특허회피 설계안을 제시한다.

[표 3-7] Ex 2-2의 유사성 통과 주제어

U(유사도)	[1] IPC 코드(유사성 판단) : 핵심 키워드 중심 채택방식
A급, 90%(90점)	ICT(H04L 29/08), RFID(H04L 29/08)
B급, 80%(70점)	게이트웨이(H04L 12/26), 임베디드(H04L 12/22), 센싱정보(H04M11/00)
C급, 70%(30점)	행동분석(G06F 19/00), AES(H04L 9/06), MDM(H04M 1/725)
D급, 60%(20점)	데이터베이스(G06Q 20/36), RF 차폐(H04N 5/225), DMB(H04H 20/02)
60%미만 (0점)	게임기기(A63F 13/00), 의용생체(C07K 16/30), 유전막(H01B 3/10), 탄성체(H01P 1/207)
핵심키워드 의 IPC	G06F 13/38, G06F 9/44, H04L 12/66, H04L 29/08, H04M 3/42, H04M 11/04, G06F 13/14
C(유사도)	[1] IPC 코드(유사성 판단) : S급 특허 중심 채택방식
A급, 90%(90점)	ICT(H04L 29/08), RFID(H04L 29/08)
B급, 80%(70점)	-
C급, 70%(30점)	게이트웨이(H04L 12/26), 임베디드(H04L 12/22), AES(H04L 9/06), 데이터베이스(G06Q 20/36)
D급, 60%(20점)	센싱정보(H04M 11/00), 행동분석(G06F 19/00), MDM(H04M 1/725), RF 차폐(H04N 5/225), DMB(H04H 20/02)
60%미만 (0점)	게임기기(A63F 13/00), 의용생체(C07K 16/30), 유전막(H01B 3/10), 탄성체(H01P 1/207)
S특허의 IPC	G06Q 50/20, H04L 29/08, H05B 37/02

다음은 [Ex 2-3]을 대상으로 피어슨 상관계수(r)를 구하고, 유사성 통과 주제어들을 연관성 통과 주제어들로 골라내는 2차 필터링에 관한 내용을 정리한 사항이다.

[그림 3-39~41]은 2차 필터링을 위해 적용되는 피어슨 상관계수 식을 이용해 피어슨 상관계수(r)를 구하고, 이로부터 확인되는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂)를 알아내는 과정을 나타내고 있음을 바로 인지할 수 있다.

[Ex 2-3]에서는 U의 평균값(ux) : 47.3, C의 평균값(cx) : 34.5, U의 분산(σ^2) : 568.2, C의 분산(σ^2) : 439.4, U의 표준편차($\sqrt{\sigma^2}$) : 23.8, C의 표준편차($\sqrt{\sigma^2}$) : 21.0이 되면, [Ex 2-3]의 피어슨 상관계수(r)는 식 (1)을 근거로 아래에 구한 바와 같이 0.7162로 나타냄을 확인할 수 있다.

$$r = \frac{357.9}{499.7} = 0.7162$$

이를 통해, [Ex 2-3]에서 구한 피어슨 상관계수(r)가 0.5 이상임으로 두 주제(핵심 키워드와 S급 특허)는 서로 연관성이 있다고 간주된다.

유사성 통과 주제어들에게 부여한 가중치 점수의 설정 범위가 90점에서 20점까지 주어짐으로, [Ex 2-3]에서 구한 피어슨 상관계수(r) 0.7162와 부합하는 제 1 최적점수(OS₁)를 1:1 매칭을 통해 찾아 낼 수 있다. 이에 부합하는 제 1 최적점수(OS₁)는 약 75점이 된다.

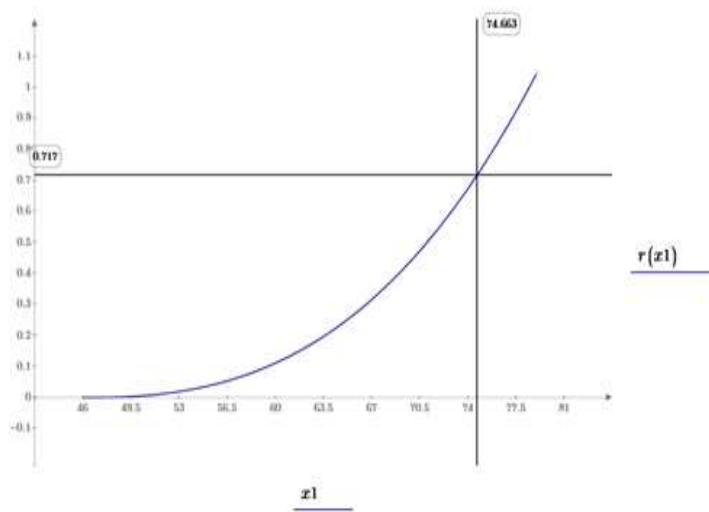
제 1 최적점수(OS₁)가 고정되면, 피어슨 상관계수(r)는 0과 1의 범위를 유지하기 위해 가중치 점수의 상대성 허용 범위를 46점 내지 71점을 갖게 된다. 그에 따라, 제 1 최적점수(OS₁)가 약 75점일 경우, 이와 대비되는 제 2 최적점수(OS₂)는 약 66점으로 나타냄을 알 수 있다.

2차 필터링은 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 어느 하나를 골라 연관 최적점수로 정한 뒤, 연관 최적점수와 유사성 통과 주제어들마다 부여받은 가중치 점수를 비교해 연관 최적점수보다 높은 가중치 점수를 갖는 연관성 통과 주제어들을 색출한다.

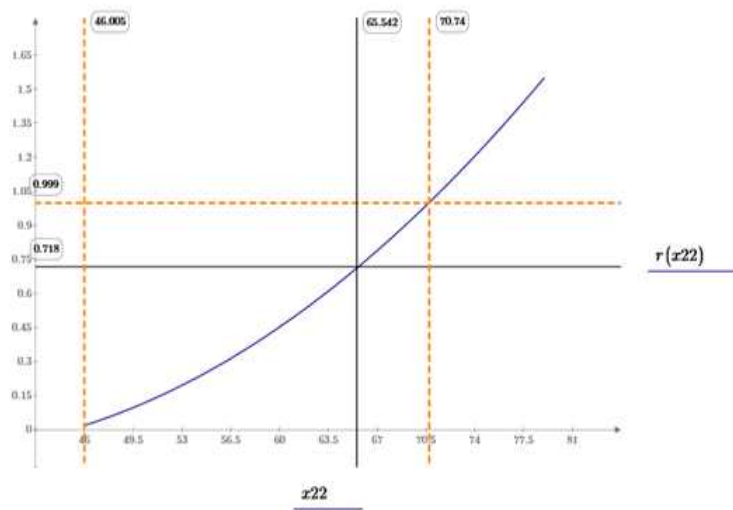
즉, 2차 필터링의 연관성 판단은 피어슨 상관계수 식을 통해 유사성 통

과 주제어들을 연관성 통과 주제어들로 선별해 내는 과정이라 할 것이다.

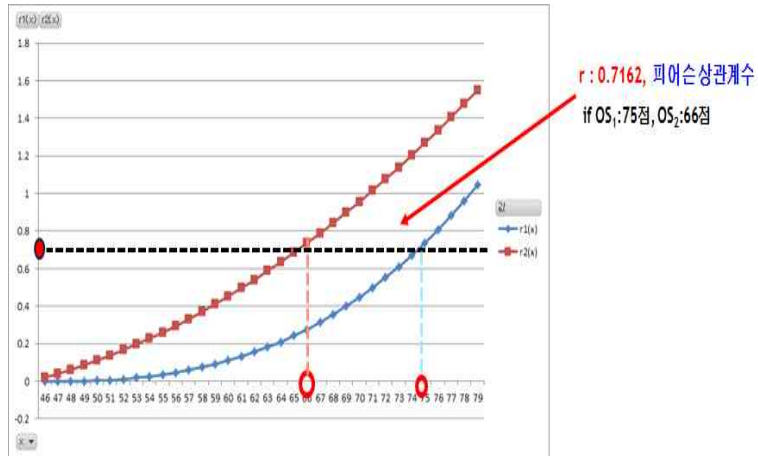
[Ex 2-3]에서는 제 1 최적점수(OS₁)와 제 2 최적점수(OS₂) 중 낮은 점수로 판명된 약 66점의 연관 최적점수를 기준으로 유사성 통과 주제어들만 판별해 낸다.



[그림 3-39] Ex 2-3의 제 1 최적점수(OS₁)



[그림 3-40] Ex 2-3의 제 2 최적점수(OS₂)



[그림 3-41] Ex 2-3의 r , OS_1 , OS_2

[Ex 2-3]의 2차 필터링은 피어슨 상관계수를 이용한 연관성 판단을 실시한 결과, ‘ICT’, ‘RFID’, ‘게이트웨이’, ‘임베디드’, ‘센싱정보’ 주제어들은 모두 66점 이상인 것으로 확인된 바, 붓에게 신뢰할 수 있는 연관성 통과 주제어들을 추천받는다.

즉, 붓은 [표 3-8]에서 알 수 있는 바와 같이, 사물인터넷(IoT) 분야에 관심 있는 고객에게 ‘ICT’, ‘RFID’, ‘게이트웨이’, ‘임베디드’, ‘센싱정보’가 삽입된 특허회피 설계안을 제시할 수 있을 것이다.

[표 3-8] Ex 2-3의 유사성 통과 주제어

U(유사도)	[1] IPC 코드(유사성 판단) : 핵심 키워드 중심 채택방식
A급, 90%(90점)	ICT(H04L 29/08), RFID(H04L 29/08)
B급, 80%(70점)	게이트웨이(H04L 12/26), 임베디드(H04L 12/22), 센싱정보(H04M 11/00)
C급, 70%(30점)	행동분석(G06F 19/00), AES(H04L 9/06), MDM(H04M 1/725)
D급, 60%(20점)	데이터베이스(G06Q 20/36), RF 차폐(H04N 5/225), DMB(H04H 20/02)
핵심키워드 의 IPC	G06F 13/38, G06F 9/44, H04L 12/66, H04L 29/08, H04M 3/42, H04M 11/04, G06F 13/14
C(유사도)	[1] IPC 코드(유사성 판단) : S급 특허 중심 채택방식
A급,	ICT(H04L 29/08), RFID(H04L 29/08)

90%(90점)	
B급, 80%(70점)	-
C급, 70%(30점)	게이트웨이(H04L 12/26), 임베디드(H04L 12/22), AES(H04L 9/06), 데이터베이스(G06Q 20/36)
D급, 60%(20점)	센싱정보(H04M 11/00), 행동분석(G06F 19/00), MDM(H04M 1/725), RF 차폐(H04N 5/225), DMB(H04H 20/02)
S특허의 IPC	G06Q 50/20, H04L 29/08, H05B 37/02

[Ex 2-1], [Ex 2-2], [Ex 2-3]은 피어슨 상관계수(r)가 각각 다르게 나왔다는 점에서 큰 차이점이 있으며, 이는 [Ex 2-1]의 경우 D급 이하의 유사성 통과 주제어들을 가중치 점수를 10점으로 부여하고 총 개체 수에 포함시켰으며, [Ex 2-2]는 D급 이하의 유사성 통과 주제어들을 가중치 점수를 0점을 부여하고 총 개체 수에는 포함시키는 실험 환경을 조성하였고, [Ex 2-3]은 D급 이하의 유사성 통과 주제어들을 총 개체 수에서 아예 배제시키는 실험 환경을 조성하였기 때문이라 하겠다.

3.4.2.3 사례2의 적합성 판단

사례2의 EOMSE 기법은 유사성 통과 주제어들 또는 연관성 통과 주제어들로부터 발췌된 주제어의 기대치가 평균치보다 크면, EOMSE 기법에 부합하는 주제어인 것으로 간주해 [표 3-9]에서 보여지는 바와 같이, 식 (2)를 적용해서 구한 평균치 1.727 이상인 주제어만 최종 선정한다.

즉, 붓은 평균치 1.727 큰 ‘ICT’, ‘RFID’, ‘센싱정보’, ‘AES’, ‘MDM’, ‘DMB’을 최종 주제어들로 선정한다.

[표 3-9] 사례2의 적합성 판단 실시

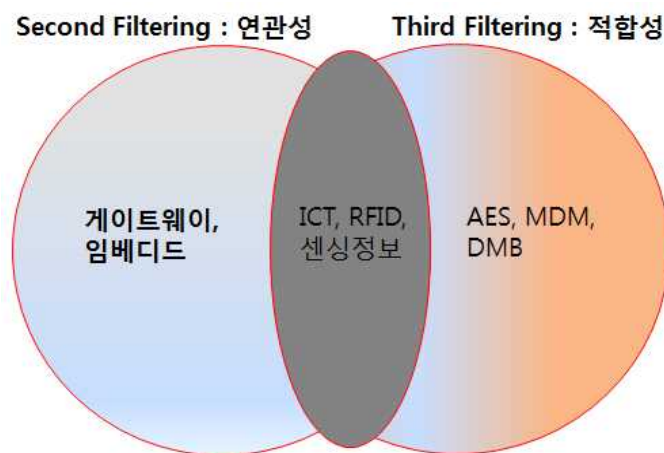
[3]EOMSE 기법 적용(적합성 판단)			
	o.value	v.value + c.value + ad.value + a.value	값
1	ICT	정보통신 인프라 구축, ICT 인력자원 개발, 법 및 제도 정비를 기반으로	
		$1+1+0+0 = 2$	2
2	RFID	RFID 시스템(무선식별시스템)과 같은 유비쿼터스 기술을 활용하여 관람객에게 개인화	
		$1+1+0+0 = 2$	2
3	게이트웨이	게이트웨이로부터 서버로 전달되는 데이터에 대한 보안기능을 지원하여야 한다	
		$1+0+0+0 = 1$	1
4	임베디드	모든 정보기기. 영상 인식. 생체 인식. 필기체 인식. <표 1-3> 임베디드 기본/공통 응용 ...	
		$1+0+0+0 = 1$	1
5	센싱 정보	수집 기기(디바이스) 온도/습도/압력 등 각종 센싱 정보를 포함하여 각종 데이터를 ...	
		$1+0+0+1 = 2$	2
6	행동 분석	빅데이터 분석 산업 핵심 국가 산업 정부 정책 행동 분석 산업 개인정보 오용 남용	
		$0+0+0+0 = 0$	0
7	AES	데이터 보안을 위해 AES 같은 암호화 알고리즘이 널리 쓰이고 있지만	
		$1-1+3+0 = 3$	3
8	MDM	회사 업무용 전산시스템 스마트폰 적용 MDM Mobile Device Management 앱 App	
		$1+0+0+1 = 2$	2
9	데이터베이스	데이터베이스 시스템 사용자의 관점에서 보고자 하는 정보의 집합	
		$1+0+0+0 = 1$	1
10	RF차폐	암호화,. 양방향인증,. 키관리 서버 개발. RF 차폐. 무선신호. 차폐기술. 무선신호 암호화	
		$0+0+0+0 = 0$	0
11	DMB	기존 T-DMB의 채널 용량을 2배로 증대시킴으로써 고전송효율, 비디오 서비스	
		$1+1+3-1 = 5$	5
평균		$19 / 11 = 1.727$	19

[그림 3-42]는 2차 필터링의 결과로 선정된 유사성 통과 주제어들인 ‘게이트웨이’, ‘임베디드’, ‘ICT’, ‘RFID’, ‘센싱정보’와 3차 필터링의 결과로 선정된 연관성 통과 주제어들인 ‘ICT’, ‘RFID’, ‘센싱정보’, ‘AES’, ‘MDM’, ‘DMB’을 조합해 3가지의 경우로 특허회피 설계안을 제시해 줄 수 있음을 예시한다.

첫째, Second Filtering의 유사성 통과 주제어들과 Third Filtering의 연관성 통과 주제어들 간의 교집합으로 결정된 ‘ICT’, ‘RFID’, ‘센싱정보’가 최종 주제어들로 선정되는 경우이다.

둘째는 S or T 간의 합집합으로 결정된 ‘게이트웨이’, ‘임베디드’, ‘ICT’, ‘RFID’, ‘센싱정보’, ‘AES’, ‘MDM’, ‘DMB’가 최종 주제어들로 선정되는 경우이다. 셋째는 T만으로 결정된 ‘AES’, ‘MDM’, ‘DMB’가 최종 주제어들로 선정되는 경우이다.

이와 같이, 3가지 경우로 제시 가능한 특허회피 설계안은 초기 입력된 핵심 키워드 즉, ‘사물인터넷’, ‘IoT’이 표기되어야 바람직하다 할 것이며, 2차와 3차 필터링을 거쳐 선정된 최종 주제어의 앞과 뒤에 표기되는 문구들이나 어구들은 M개의 정보글 중에서 S에 속한 유사성 통과 주제어들, T에 속한 연관성 통과 주제어들이 기재된 정보글을 가져와 적용하는 것이 바람직할 것이다.



[그림 3-42] 사례2의 최종 주제어 설정

IV. 실험 및 검증

본 논문의 봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구는 프로토타입으로 개발한 봇 V1.0를 실험 도구로 삼아 가상의 실험을 실시하였다. 즉, 봇 V1.0은 고객으로부터 핵심 키워드를 전달받으면 사용자 인터페이스를 통해 특허 키워드, 최종 주제어 및 특허회피 설계안을 화면디스플레이 한다.

4.1 특허회피 설계안 자동생성 봇 구현

본 논문의 특허회피 설계안 자동생성 봇은 특허회피 설계안을 자동 생성할 수 있도록 만든 가상 개인서비스 분야에 속한 지능형 앱 플랫폼으로서, 개발 환경 즉, 자바 스크립트 : 11.10.0 버전, 파이썬 : 3.7.2, 자바 : 8 버전에서 구현되도록 파이썬 프로그램으로 코딩하여 제작한 것이다.

먼저, 봇 V1.0의 시동은 윈도우 파워 셸 화면을 실행한 다음, cd 명령어를 입력해 특허회피 설계안을 자동 생성하는 소스 프로그램이 있는 경로를 입력한다. 봇 V1.0의 시동은 node 명령어를 넣어 자바스크립트 파일인 index.js를 실행시켜 가상의 로컬 서버를 구축한 후, 로컬 호스트 주소인 <http://localhost:4343>에 접속함에 따라, 특허회피 설계안을 자동 생성하는 사용자 인터페이스 화면창을 오픈한다.

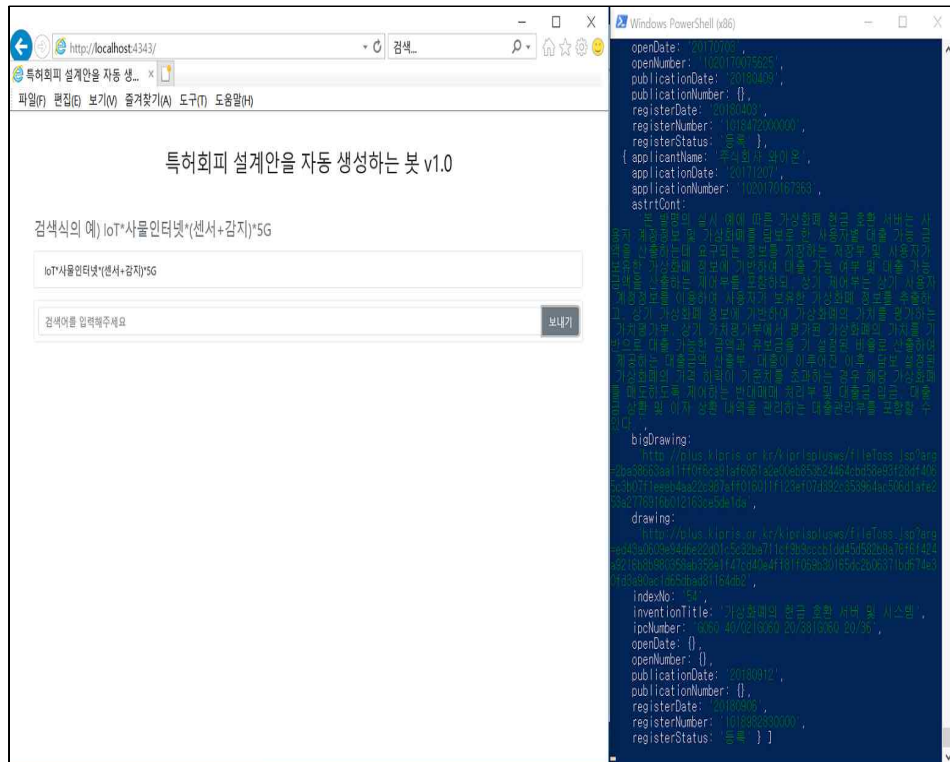
이때부터, 봇 V1.0은 고객이 입력하는 핵심 키워드를 전달받아 키프리스에 접속해 핵심 키워드를 활용한 특허 검색을 실시한다.

봇 V1.0은 윈도우 [그림 4-1]의 파워 셸 화면에서 보여지는 것처럼, 키프리스로부터 검색된 N개의 선행특허에 관한 오픈된 정보들, 말하자면 발명의 명칭, IPC코드, 출원번호, 출원날짜, 공개날짜, 공개번호 등록번호, 등록날짜, 출원인정보, 요약 등을 제공한다.

일예로, 봇 V1.0은 고객이 핵심 키워드인 “IoT*사물인터넷*(센서+감

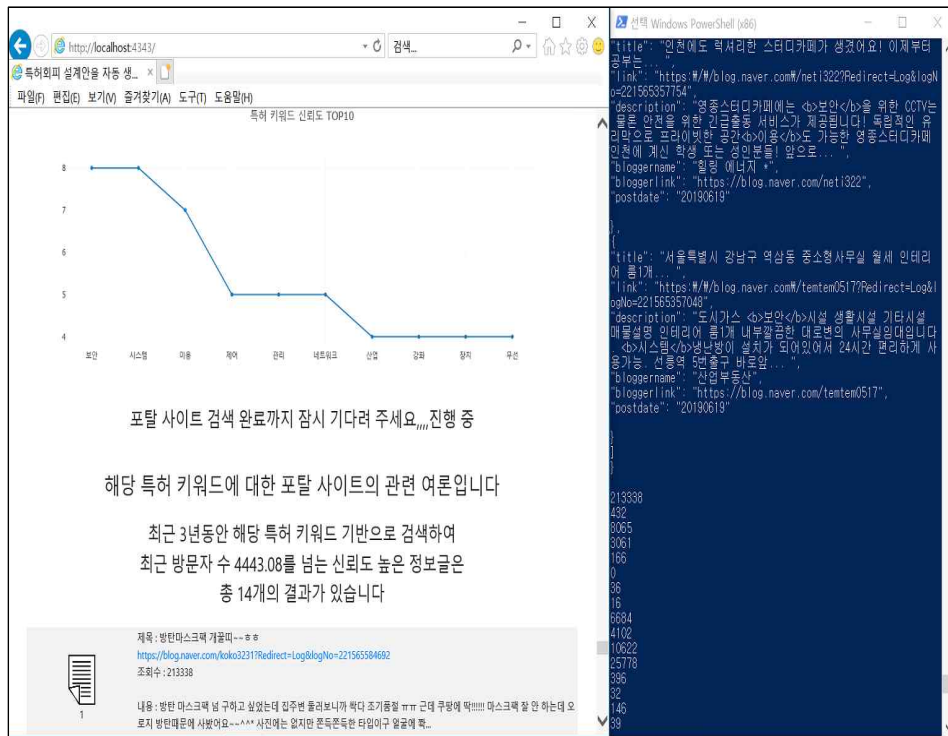
지)*5G”를 전달받아 키프리스에 접속해 핵심 키워드로부터 검색된 54개의 선행특허를 찾아낸다.

봇 V1.0은 고객이 +, *연산자로 다수의 핵심 키워드를 조합한 검색식을 입력하더라도 이를 인지해 찾고자 하는 N개의 선행특허를 검색할 수 있다.



[그림 4-1] 특허회피 설계안 자동생성 봇의 핵심 키워드 입력

봇 V1.0은 [그림 4-2]에서 보여지는 바와 같이, 키프리스로부터 최근 3년 이내 핵심 키워드와 관련된 상위 54개의 등록특허를 찾아내면서 그에 관한 리스트 정보를 사용자 인터페이스 화면에 제공한다.

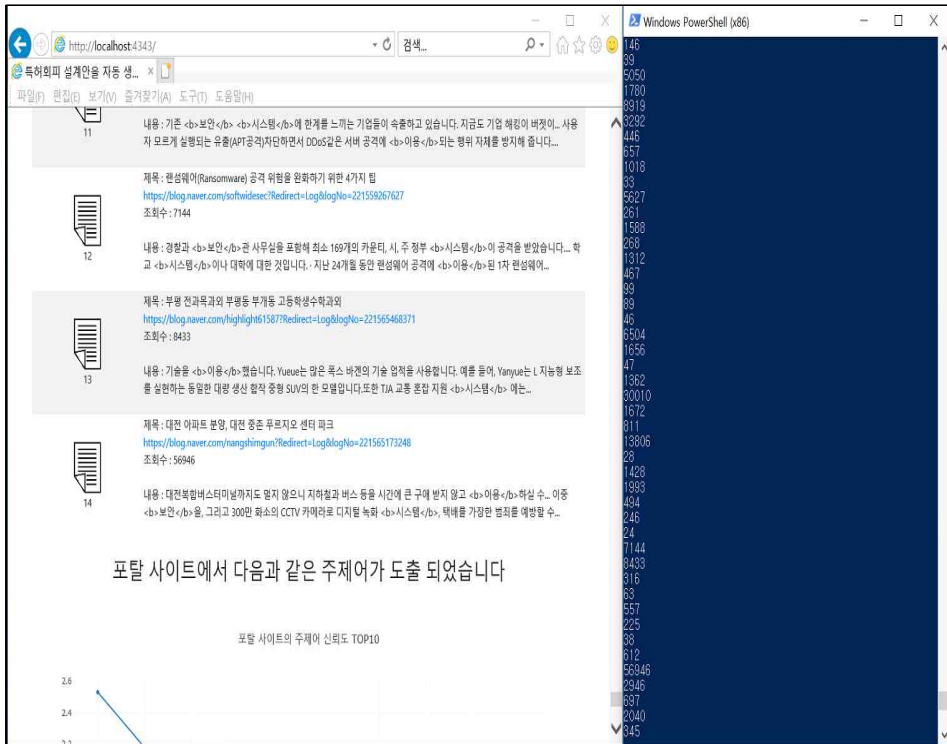


[그림 4-3] 특허회피 설계안 자동생성 봇의 특허 키워드 색출

봇 V1.0은 [그림 4-4]에서 보여지는 바와 같이, 14개의 유용한 정보글을 사용자 인터페이스 화면을 통해 1번부터 14번까지 순차적으로 나열하여 고객에게 제공해 준다.

봇 V1.0은 사용자 인터페이스 화면과 윈도우 파워 셸 화면에서도 알 수 있듯이, 12번째 정보글의 조회수가 평균 방문자 수인 4443.08 보다 큰 7144임으로 특허회피 설계안에 사용할 만한 정보 가치가 있다고 판단해 이를 데이터베이스에 저장한다.

마찬가지로, 13번째 정보글의 조회수 역시도 평균 방문자 수 4443.08 수보다 큰 8433임으로 특허회피 설계안에 사용할 만한 정보 가치가 있다고 판단해 이를 데이터베이스에 저장한다. 봇은 1번부터 14번까지 나타난 14개의 유용한 정보글을 특허회피 설계안을 제시하는 빅데이터로 활용하기 위하여 데이터베이스에 임시 저장한다.

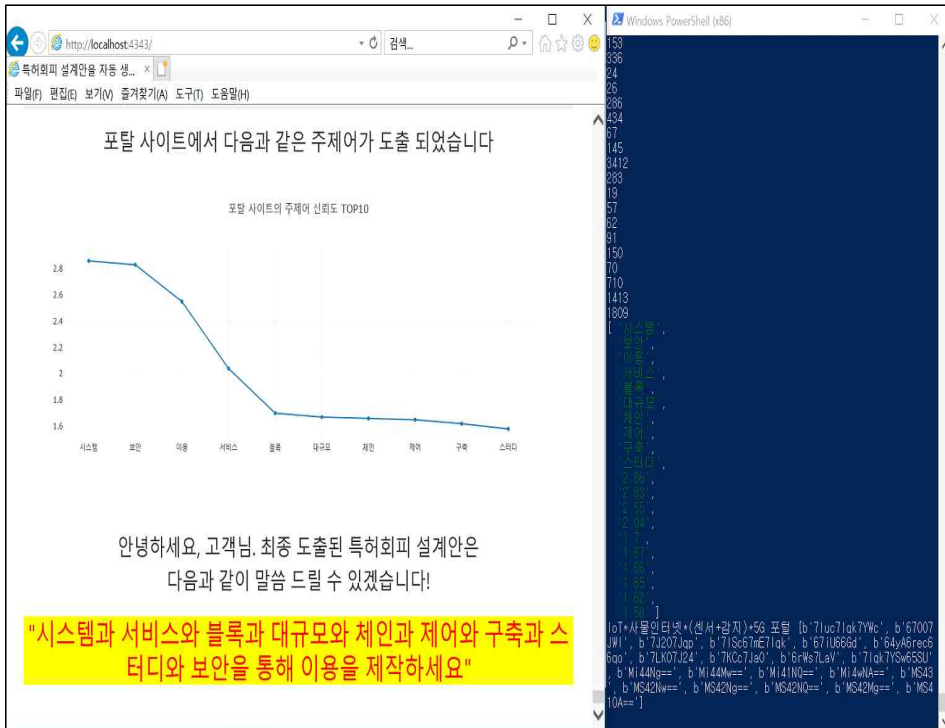


[그림 4-4] 특허회피 설계안 자동생성 봇의 정보를 색출

봇 V1.0은 [그림 4-5]에서 보여지는 바와 같이, 14개의 유용한 정보를 바탕으로 핵심 키워드 또는 특허 키워드와 연관성이 높다고 판단되는 상위 10위 안에 들어갈 만한 최종 주제어들을 선정하며, 특허회피 설계안 제시 시, 최종 주제어들이 표기된 특허회피 설계안을 자동 생성시킨다.

다시 말해, 봇 V1.0은 자연어 처리를 하기 위한 목적으로 이미 700개 이상의 사전 정의된 카테고리 문서를 분류해 놓음에 따라, 봇 V1.0에 의해 선정된 최종 주제어들은 핵심 키워드 또는 특허 키워드와 당위성 면에서 연관성이 있다.

봇 V1.0은 상위 10위 안에 들어갈 만한 최종 주제어들 단어 조합 또는 문장 조합해 특허회피 설계안을 자동 생성시킨 후 이를 고객에게 제시해주는 동작을 한다.

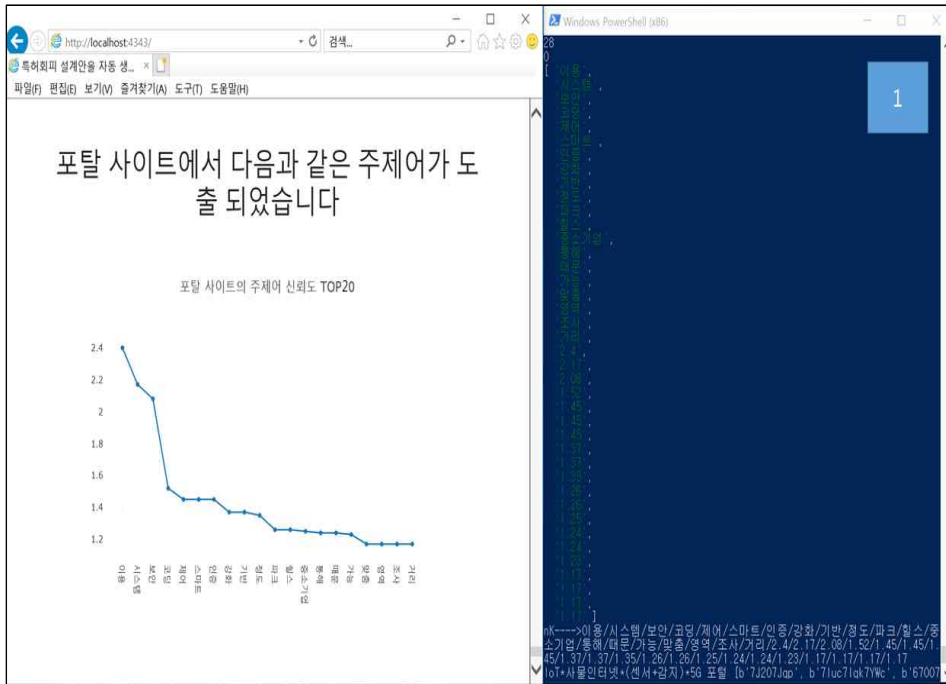


[그림 4-5] 특허회피 설계안 자동생성 봇의 최종 주제어 선정

4.2 봇의 최종 주제어 선정 실험

봇 V1.0이 최종 주제어로 특허회피 설계안을 자동 생성함에 있어, 최종 주제어가 과연 특허회피 설계안을 도출해 내는데 필요한 가치정보일까(설문조사2의 2번 문항 참조)를 검증하기 위해, 4.1장에서 결과로 보여지는 최종 주제어의 수를 10개 범주에서 20개로 늘려 이를 더 고찰해 보았다.

[그림 4-6]에서 보듯이, 봇 V1.0은 고객의 관심분야인 핵심 키워드인 “IoT*사물인터넷*(센서+감지)*5G”를 입력받음에 따라 “시스템, 보안, 이용 관리, 분석, 클라우드, 관리제어, 데이터, 코딩, 인증”이 포함되는 20개의 최종 주제어들을 추출해 낸다.



[그림 4-6] 확장된 최종 주제어 추출

마찬가지 방법으로, 다른 관심 분야에 해당하는 다른 핵심 키워드인 “자율주행*차*(사고+단속)”가 입력되면, 봇은 “시스템, 작업, 중고차, 공간, 원인, 자원, 안전, 자세, 거리”가 포함되는 20개의 최종 주제어들을 다르게 추출해 낸다.

또 다른 핵심 키워드인 “가족여행*추천지역*관광”가 입력되면, 봇은 “기상, 매력, 환경, 체계, 관측, 태풍, 활용, 주민, 실시간”로 조합을 이루는 20개의 최종 주제어들을 또 다르게 추출해 낸다.

인지해야 할 사항으로, 20개의 최종 주제어들의 조합은 키프리스로부터 N개의 특허를 제공받고 포털 사이트로부터 M개의 정보글을 실시간으로 제공받음으로 시간에 따라 계속 변동될 수밖에 없다는 점을 감안해야 할 것이다.

이렇게 동일 방법으로, 각기 다른 분야로 10개의 카테고리가 마련되어 각 카테고리에 해당하는 핵심 키워드가 봇에 입력되면, 봇은 [표 4-1]에서 보여지는 바와 같이, 20개의 조합으로 형성된 각기 다른 최종 주제어 그룹

을 10개의 카테고리별로 만들어 낸다.

[표 4-1] 최종 주제어 제시

	핵심키워드	최종 주제어(20개 제시)
1	IoT*사물인터넷* (센서+감지)*5G	시스템, 보안, 이용관리, 분석, 클라우드, 관리제어, 데이터, 스마트, 강화, 맞춤, 파크, 힐스, 중소기업, 조사, 거리, 영역, 기반, 가능, +코딩, +인증
2	(취업+진로)*상담	시스템, 서비스, 최근, 사회, 국가, 한국, 개발, 역사, 법률, 개방, 경우, 담당, 하락, 지령, 중요, 보고, 가장, 지금, +정신건강, +현금
3	혼합현실*(증강현 실+가상현실)	시스템, 이용, 판단, 기능, 차단, 자동, 기계, 음성, 고도, 기업, 미세, 설립, 의료기기, 국산, 기술, 단위, 인식, 센서, +바이오, +예측
4	헬스케어*(진료+ 의사	시스템, 제중, 자격, 조건, 보험, 혈압, 흡연, 정보, 장치, 카메라, 저장, 복구, 하드, 대형, 개폐, 전동, 동기화, 면세점, +바디, +컨트롤
5	자율주행*무인* 자동차*(사고+단 속)	시스템, 작업, 중고차, 공간, 원인, 자원, 안전, 제어, 메모리, 드라이버, 프로, 가격, 운전, 보조, 방지, 충돌, 동승, 벤츠, +자세, +거리
6	자연어처리*(형태 소+오피니언)	시스템, 사례, 감성, 온라인, 창업, 개선, 소개, 교육, 인공, 색상, 디자인, 지능, 기업, 이진, 주비스, 뉴트, 정리, 완성, +컨설팅, +고객
7	AI*딥러닝*신경 망*유전자	시스템, 러닝, 머신, 기술, 처리, 인공, 분석, 출시, 기계, 개발, 정보, 회수, 현실, 감수, 컴퓨터, 알츠하이머, 자체, 조기, +순위, +인간
8	식품*(빵+발효)* 친환경	조성, 제조, 발효, 과정, 환경, 원료, 문화, 주간, 국내, 대표, 이용, 성분, 지원, 콜린, 투자, 당시, 산패, 모나, +미생물, +바아오
9	노인*치매*치료* 4차산업	억제제, 용도, 시간, 부작용, 마약, 치료, 보충, 제조, 조성, 대한민국, 산업, 자유, 위협, 시행, 일본, 전문, 해양, 쓰레기, +전문, +연구
10	가족여행*추천지 역*명소	기상, 매력, 환경, 체계, 관측, 태풍, 활용, 제공, 시스템, 사업, 정보, 자동, 겨울, 체계, 악취, 정보, 규제, 터치, +주민, +실시간

특허회피 설계안 자동생성 봇은 이렇게 자동 생성된 20개 조합으로 구
성된 최종 주제어들을 자연어 처리하여 고객이 인지 가능한 자국어, 영어
로 표현되는 특허회피 설계안을 자동 생성한다.

4.3 실험 결과 검증

본 논문은 특허회피 설계안 자동생성 봇으로부터 표출된 실험 결과인 최종 주제어에 대한 검증을 받기 위해 미리 준비해 둔 설문조사로 전문가들의 다양한 의견을 받아보았다.

설문조사는 특허법인, 특허사무소 등에서 근무하고 계신 변리사, 통계분석 및 기술경영 컨설팅학 강의 및 연구하고 계신 교수, 또는 특허분쟁 및 대응 예방 전략과제를 실무적으로 담당하고 계신 15년 이상의 지식재산권 경력을 가진 분들을 포함하여 총 12명을 대상으로 실시하였다.

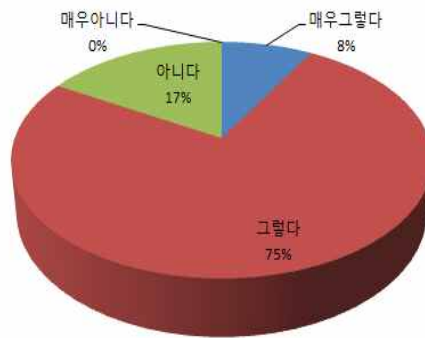
설문조사의 취지는 다음과 같다. 즉, “특허회피 설계안을 자동 생성하는 봇은 출원인의 관심분야를 대표하는 핵심 키워드를 제공받아 키프리스로부터 검색된 특허 명세서의 특허청구범위로부터 특허 키워드를 추출한 다음, 이 특허 키워드를 포털 사이트에 입력해 포털 사이트로부터 기존 장벽 특허들을 회피할 수 있을 만한 유용한 정보글들을 뽑아내고 유용한 정보글들로부터 색출된 후보 주제어들을 대상으로 유사성, 연관성, 적합성 여부를 판단해 최종 주제어들을 선정하는 AI 플랫폼 중 하나라 하겠고 더 나아가 최종 주제어들을 자연어 처리해 특허회피 설계안을 자동 생성하는 가상 개인비서 서비스를 지향하는 지능형 앱 솔루션 중 하나라 말씀드릴 수 있겠습니다.” 라는 취지를 담고 있다.

설문조사는 이러한 과정을 통해 선정된 최종 주제어가 특허회피 설계안을 도출해 내는데 있어 도움이 되는지를 질의한 것으로, 총 5문항에 걸쳐 설문 요청하였고, 그에 관한 전문가들의 설문 응답을 제공받았다.<부록 참조>

다음은 특허회피 설계안 자동생성 봇에 관한 설문조사 응답에 따른 결과분석 내용이다.

1번 문항은 특허회피 설계안 자동생성 봇으로부터 표출된 최종 주제어에 대한 활용도를 묻는 질의로, 이에 대한 설문응답은 ②‘그렇다’라는 반응이 75%로 [그림 4-7]처럼 나타났다. 이는 봇 기반 특허회피 설계안 자동생성 봇이 선정한 최종 주제어는 특허회피 설계안을 제시하는데 있어 도

움이 되며, 비교적 만족스러운 결과임을 표시하였다.



[그림 4-7] 1번 질의-응답결과

2번 문항은 특허회피 설계안 자동생성 봇에 의해 선정된 최종 주제어에 대한 검증을 묻는 문항으로, 전문가가 최종 주제어를 활용해 어떤 내용의 특허회피 설계안을 제시하는지를 파악해 보고, 과연 최종 주제어가 특허회피 설계안을 제시하는데 있어 도움이 되는지를 검증받고자 하는 질의이다.

이에 따른 설문 응답으로 [표 4-2]에서 알 수 있듯이, 전문가가 1부터 10번까지의 각기 다른 관심 분야를 대표하는 핵심 키워드로부터 나온 최종 주제어를 활용해 특허회피 설계안을 제시한 좋은 사례는 최종 주제어에 대한 정체성 여부를 타진해 보는데 있어 충분한 답변이 성립되었다 할 것이다.

즉, 10가지의 예 중 5번째로 제시된 예를 들면, 핵심 키워드는 “자율주행*차*(사고+단속)”이고, 최종 주제어는 특허회피 설계안 자동생성 봇에 의해 “시스템, 작업, 중고차, 공간, 원인, 자원, 안전, +자세, +거리”로 선정된 경우, 전문가는 특허회피 설계안을 “자율주행 기능을 가진 중고차의 고장진단 기술”이라고 제시하였다.

이 안은 핵심 키워드에 기재된 ‘자율주행’을 대상으로 최종 주제어에서 나온 ‘중고차’와 결부시켜 최종적으로 자율주행 기능을 가진 중고차를 대상으로 고장진단을 해주는 기술을 특허회피 설계안으로 제시해주면 어떻까라는 내용을 담고 있다.

이 안은 핵심 키워드와 최종 주제를 모두 가져오면서 새롭게 틈새시장을 노려보는데 제시될만한 용어인 ‘고장진단’과 핵심 키워드와 최종 주제를 연결시켜 주는 용어인 ‘기능’을 조합해 만들어 낸 모범적인 특허회피 설계안이라 할 것이다. 이렇게 조합 구성된 모범적인 특허회피 설계안은 공백기술을 찾아낸 안이자, 개선(개량)특허의 출원 및 시제품 제작을 준비하는데 도움이 될 만한 유용한 가치정보라 할 것이다.

여기서, +자세, +거리는 최종 주제어 선정요건에 관한 제한 사항을 벗어나기 위한 방안으로 최종 주제의 선정범위를 10개에서 20개로 늘렸을 때 확장범위(11위~20위)에서 추가적으로 나온 단어이다.

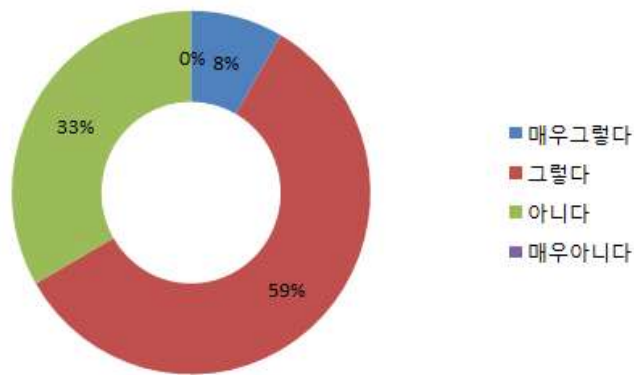
[표 4-2] 특허회피 설계안의 모범 예

	핵심키워드	최종 주제어 (20개 제시)	특허회피 설계안의 모범 예
1	IoT*사물인터넷*(센서+감지)*5G	시스템, 보안, 이용 관리, 분석, 클라우드, 관리제어, 데이터, +코딩, +인증	클라우드 환경에서 Iot를 이 용한 보안시스템의 관리제어 시스템
2	(취업+진로) *상담	시스템, 서비스, 최근, 사회, 국가, 개발, 역사, +정신건강, +현금	최근의 사회 트렌드를 반영 한 취업 및 진로상담 시스템
3	혼합현실*(증강현실+ 가상현실)	시스템, 이용, 판단, 기능, 차단, 자동, 기계, +바이오, +예측	혼합현실 기반 기계고장 진 단기술
4	헬스케어*(진료+의사	시스템, 체중, 자격, 조건, 보험, 혈압, 흡연, +바디, +컨트롤	흡연과 혈압과의 상관관계에 기반한 헬스케어 시스템
5	자율주행* 무인*자동 차*(사고+ 단속)	시스템, 작업, 중고차, 공간, 원인, 자원, 안전, +자세, +거리	자율주행 기능을 가진 중고 차의 고장진단 기술
6	자연어처리 *(형태소+	시스템, 사례, 감성, 온라인, 창업, 개선,	자연어처리 기술 기반의 텍 스트 마이닝을 이용한 창업

	오피니언)	소개, +컨설팅, +고객	트렌드 키워드분석 기술
7	Ai*딥러닝* 신경망*유전자	시스템, 러닝, 머신, 기술, 처리, 인공, 분석, +순위, +인간	딥러닝 기반의 유전자 분석 알고리즘
8	식품*(빵+ 발효)*친환 경	조성, 제조, 발효, 과정, 환경, 원료, 문화, +미생물, +바아오	친환경 발효 기술을 이용한 바이오 에너지 기술
9	노인*치매* 치료*4차산 업	억제제, 용도, 시간, 부작용, 마약, 치료, 보충, +전문, +연구	마약성분 에서 추출한 치매 치료제의 부작용 억제제
10	가족여행* 추천지역* 명소	기상, 매력, 환경, 체계, 관측, 태풍, 활용, +주민, +실시간	환경과 기상 데이터에 기반 한 가족 여행지 추천 시스템

3번 문항은 기존의 특허회피 설계 방법 대비 특허회피 설계안 자동생성 붓이 최종 주제어를 활용해 특허회피 설계안을 자동 생성하는 방법과 어떤 차이점이 있는지를 묻는 질의로 이에 대한 설문응답은 기존방식과 대비 내용면에서 큰 차이가 있을 것이라고 의사 표시한 ②‘그렇다’가 59%를 차지하였고, 그 반대 입장인 큰 차이가 없을 것이라고 의사 표시한 ④‘아니다’라는 반응이 33%로 나타났다.

전문가들의 의견은 “향후 특허회피 설계안 자동생성 붓의 출현이 특허회피 전략을 원하는 고객에게 도움을 줄 것이다”라는 긍정적인 전망을 내린 것에 입각해 기존의 특허회피 설계시 대두되었던 여러 가지 문제점을 해결하고 특허회피 전략을 보다 지능적으로 제시하는 앱 플랫폼 개발의 지속적인 연구가 필요함을 간접 표시하였다.<그림 4-8 참조>



[그림 4-8] 3번 질의-응답결과

4번 문항은 향후 특허회피 설계안 자동생성 봇의 보완점을 묻는 주관식 질의로, 이에 대한 설문응답 중 3가지의 보완점이 지적되었다.

첫째는 봇이 자동 생성한 특허회피 설계안이 장벽특허와 비교해 균등범위에 속하는지 여부도 검토해 주고, 특허 출원 후 심사과정에서 변동되는 특허청구범위 변동에 관한 해석도 가능하길 희망한다고 언급하였다.

둘째는 키워드를 추출하여 조합하는 방법 외에도 기술별 특허동향을 파악해 그로부터 공백기술을 도출하는 방식을 제안하면서 특허의 원천기술에 대한 과학적, 논리적 분석도 가능한지에 대한 필요성도 강조하였다.

셋째는 특허회피 설계안을 비즈니스적으로 연결해 볼 수 있고, 특허선점 기업이 확보한 장벽특허를 공유할 수 있는 계기가 마련되어 특허분쟁 대응이나 특허분쟁 예방하는데 유용할 것 같다는 긍정적 평가 의견을 주었다.

5번 문항은 특허회피 설계안 자동생성 봇에 대한 유용성 여부를 OX로 묻는 7개의 단답형 질의로 이에 대한 설문응답은 [표 4-3]와 같이 나왔다.

[표 4-3] 5번 질의-응답결과

	유용성 여부에 관한 질의	그렇다	아니다
1	수익을 증대시키는데 동반 상승 역할을 해줄 것이라 기대된다	33.3%	66.7%
2	출원인이 개량 발명을 새롭게 고안해 내는데 오히려 혼돈을 줄 것이라 본다	58.3%	41.7%
3	법률 특히, 지식재산권 서비스를 대중화시키는데 이바지할 것이라 본다.	91.7%	8.3%
4	출원인이 독점권을 유지하는데 오히려 방해요소로 작용할 꺼라 본다	66.7%	33.3%
5	특허회피 설계안이 나온다면, 출원인의 독점권을 강화시키려는 전략도 나올 수 있다고 본다	100%	0%
6	특허회피 설계안은 특허 컨설팅을 통해 전문가가 출원인에게 제공해줄 수 있는 것이지, AI가 전문가가 할 수 있는 역할을 대신할 순 없다고 본다	75%	25%
7	AI를 이용해 법률, 특히 지식재산권 서비스를 제공하는 플랫폼은 앞으로 4차 산업 및 인류에게 도움이 되는 미래 산업 기술 및 연구 분야라 본다	91.7%	8.3%

즉, 5번 문항의 1번 OX 질의에는 33.3%가 O를 66.7%가 X를 표기하여 향후 인공지능으로 인한 일자리 난이 지식재산권 서비스 분야에도 있을꺼 라는 우려 표시를 하였다. 5번 문항의 2번 OX 질의에는 58.3%가 O, 41.7%가 X를 표기하여 특허회피 설계안 자동생성 봇의 영향력에 대한 기대반, 우려반으로 나뉜 것으로 확인되었다. 5번 문항의 3번 OX 질의에는 91.7%가 O, 8.3%가 X를 표기하여 특허회피 설계안 자동생성 봇 도입에 대해 비교적 긍정적 평가를 하였다.

또한, 5번 문항의 4번 OX 질의에는 66.7%가 O를 33.3%가 X를 표기하여 특허회피 설계의 활성화와 대비되는 반대급부 측면에서의 본 독점권 행사도 중요시 여겨야 함을 간접 표명하였다. 5번 문항의 5번 OX 질의에는 모두 O를 표기하여 창과 방패와 같은 전략적 수평관계에 대한 중요성을 재차 확인할 수 있었다. 5번 문항의 6번 OX 질의에는 75%가 O를 25%가 X를 표기하여 전문가로서의 책임감과 그 역할에 대한 필요성을 더욱 강조하였다.

마지막으로, 5번 문항의 7번 OX 질의에는 91.7%가 O를 8.3%가 X를 표기하여 특허회피 설계안 자동생성 붓의 이용 가능성에 대해 비교적 긍정적인 반응을 내비쳤다. 이로써, 본 설문조사는 전문가들에게 본 논문의 지속적인 연구에 도움이 될 총 5문항의 설문을 12명의 전문가에게 요청하여 그에 대한 설문 응답을 받아보았다.

V. 결 론

본 논문에서는 특허청 정보검색 사이트와 포털 사이트에서 제공되는 다양한 정보를 이용해 중소기업이나 개인 고객들이 시제품을 양산하기 전에 반드시 이행할 필요가 있는 특허회피 설계안을 봇을 통해 다양하게 자동 생성해 주는 방법을 제안하였다.

기업이나 개인 고객들은 시제품 출시 전 미리 기획 단계 또는 본격적인 연구 개발하기 전에 미리 시제품과 동일 유사한 제품이 시장에 이미 출시되어 있는지 여부를 알고 싶은 요구사항이 많았으나 시장 조사 및 선행 조사로 밖에는 시행되지 못하였다.

기존의 시장 조사 및 선행 조사와는 달리 특허회피 설계전략은 지식재산권이 출원 및 등록되어 있는지 여부를 사전에 조사해 줄 뿐만 아니라 이미 확보된 특허권의 보호범위를 벗어나 공백특허 또는 틈새시장까지 알려주는 방식을 취해 왔다.

그러나, 기존의 특허회피 설계전략 역시도 특허회피 전략이 특허 전문가에 따라 각자 다르게 의견이 나올 수 있고, 각자 다른 견해를 갖는 특허 전문가의 의견이 과연 기업이나 개인 고객들에게 유리하도록 설계된 특허회피 전략일지 확실치 않은 문제점이 있었다.

이를 위해, 본 논문은 고객으로부터 입력된 핵심 키워드를 입력받아 키워드로부터 검색된 다수의 선행 특허로부터 특허 키워드를 색출하고, 특허 키워드를 이용해 포털 사이트로부터 다수의 정보글을 추출한 다음, 다수의 정보글로부터 후보 주제어들을 색출하였다.

또한, 본 논문은 후보 주제어들을 대상으로 유사성 판단, 연관성 판단, 적합성 판단을 실시해 봇을 통해 다양한 특허회피 설계안을 사용될 만한 최종 주제어들을 선정하였다.

즉, 본 논문은 특허 정보검색 사이트와 포털 사이트를 통해 유용한 정보를 색출하고, 그로부터 색출된 후보 주제어들을 1, 2, 3 필터링을 통해 선정하고, 최종적으로 선정된 최종 주제어들을 대화형 구문형식에 맞도록 단

어 조합해 특허회피 설계안을 제시하는 봇 기반 특허회피 설계안 자동 생성 방법을 연구 및 설계 구현하였다.

본 논문은 기업 및 개인 고객들이 아이디어 개발, 시제품이나 시제품 개발과 연관된 연구 및 노력을 충분히 기울였음에도 불구하고 특허회피 전략에 대한 중요성을 소홀히 하여 앞으로 벌어질 만한 우려 즉, 지식재산권으로 인한 인적, 물적 피해를 입지 않도록 방향성 및 길을 제시해 주는 나침반 역할을 해 줄 봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구이다.

향후, 진행하고자 하는 앞으로의 연구로는 고객이 만족스러워하는 양질의 특허회피 설계안을 다양하게 제시해 줄 수 있도록 봇의 딥러닝(강화학습, 비강화학습)을 더욱 강화시켜 빠르고 정확하면서 신뢰성이 높은 정보를 자동 생성시켜 주는 봇에 관한 연구를 새롭게 제안해 보고자 한다.

또한, 앞으로의 다른 연구로는 최종 주제어를 다양한 자연어처리 기법(데이터 마이닝, 텍스트 마이닝, 온톨로지, 메타, 시멘틱, 오피니언 마이닝 등)에 적용하여 적합한 문구나 어구를 찾아 정밀 조합해 일반 대중이 충분히 알아들을 수 있는 자국어, 영어 형태로 특허회피 설계안을 제시해 주는 봇 기반 특허회피 설계안 자동 생성 방법에 관한 연구도 새롭게 제안해 보고자 한다.

참 고 문 헌

1. 국내문헌

- 강아름. (2012). “온라인 게임 봇 탐지를 위한 사용자 행위 분석”. 고려대학교 석사학위논문.
- 강훈, 유지상, 이재진. (2001). 『확률, 랜덤변수 및 랜덤신호 원리』. 경기 : 한국맥그로힐 출판사.
- 강희주, 김승인. (2017). 메신저 기반의 모바일 챗봇 서비스 사용자 경험 평가 -구글(Allo)과 페이스북(M messenger)을 중심으로-. 『한국융합학회』, 8(9), 271-276.
- 강민정. (2018). 사용목적에 따라 선호하는 챗봇의 성격에 관한 연구. 『한국콘텐츠학회』, 18(5), 319-329.
- 김건우, 김태민, 이창환, 김지은, 홍정희. (2017). 이미지 인식과 대화를 통한 의약품 정보제공 챗봇. 『한국시뮬레이션학회 학술대회』, 21-23.
- 김광범. (2002). “화학발명에 관한 회피설계 사례 연구”. 연세대학교 석사학위논문.
- 김동현. (2015). 『팔리는 상품 끌리는 브랜드』. 서울 : 매일경제신문사.
- 김무진. (2016). “특허 기반의 기술 트렌드 모니터링을 위한 의미론적 텍스트 마이닝 접근 방법”. 건국대학교 석사학위논문.
- 김민영. (2018). 당뇨병환자의 자가관리를 위한 챗봇 서비스. 『대한전자공학회 학술대회』, 1711-1714.
- 김방룡, 황성현. (2009). 특허 정보를 활용한 IT 유망기술 도출에 관한 연구. 『한국통신학회』, 34(10), 1021-1030.
- 김보경, 김현아, 심채인, 정진우. (2017). 소셜 챗봇을 활용한 클라우드 소셜 기반 질의응답 시스템의 설계 및 구현. 『한국정보기술학회』,

15(11), 125-138.

- 김영인. (2017). 『국방 인공지능(AI) 활용방안 연구』. 서울 : 한국융합안보연구원.
- 김유정, 한상규, 윤종묵, 허은영, 김정훈, 이중식. (2017). 챗봇 에이전트 정체성에 따른 사용자의 인식 및 행동 차이에 대한 연구. 『한국 HCI학회』, 12(4), 45-55.
- 김윤정. (2016). 인공지능 기술 발전이 가져올 미래 사회 변화, 『한국과학기술기획평가원』, 12.
- 김용석. (2008). “산학협력을 통한 중소기업애로기술 해결에 관한 연구”. 울산대학교 박사학위논문.
- 김익태. (2015). “전통의학 정보포탈 논문검색을 위한 연관검색어 추출 알고리즘에 대한 연구”. 한밭대학교 석사학위논문.
- 김태영. (2009). 『인터넷 판매자가 꼭 알아야 할 상표권과 상표권침해』. 경기 : 도서출판 비비컴.
- 김재국. (2002). 특허권 침해의 유형. 『한국기업법학회』, 11, 249-281.
- 김현우, 김종찬, 이준혁, 박상성, 장동식. (2015). 핵심 기술 및 특허 추출을 위한 IP 마이닝에 관한 연구. 『한국지능시스템학회』, 25(4), 392-397.
- 나도성. (2017). 『한국 컨설팅시장의 이론과 실제3』. 서울 ; 한성대학교출판부.
- 남호현. (2019). 『태양 아래 모든 것이 특허대상이다』. 서울 : 도서출판 예가.
- 마이클 네그네트스키. (2017). 『인공지능 개론』, 서울 : 한빛아카데미.
- 박동아. (2017). 인공지능 기반 대화형 공공 행정 챗봇 서비스에 관한 연구. 『한국멀티미디어학회』, 20(8), 1347-1356.
- 박대진. (1996). 『특허청구범위의 작성과 회피설계』. 서울 : 한빛지적소유권센터.
- 박윤희. (2004). 『돈 버는 아이디어 엿보기』. 서울 : 한국생산성본부.

- 박정호. (2019). 『챗봇서비스 구현을 위한 기술로드맵』. 대전 : 중소기업 기술정보진흥원.
- 박종진. (2018). 자연어 처리를 이용한 감정 스트레스 인지 및 관리 챗봇 개발. 『대한전기학회』, (67)7, 954-961.
- 박우중. (2013). “중소기업의 실효적인 지식재산 경영방안에 대한 연구”. 홍익대학교 석사학위논문.
- 박현우, 성태웅, 변정은, 유연우. (2018). 『기술전략과 R&D 기획』. 서울 : 한성대학교출판부.
- 서성현, 김승천, 노광현. (2018). 특허회피 설계를 위한 챗봇 및 알고리즘 연구. 『대한전자공학회 학술대회』, 1085-1088.
- 송근원. (2003). 조사 결과의 평가점수 부여에 대한 새로운 접근. 『사회과학연구』, 19(1), 179-197.
- 송지준. (2015). 『SPSS/AMOS 통계분석방법』. 경기 : 도서출판 : 21세기 사.
- 오병석. (2009). 『특허가치전략』. 서울 : 도서출판 페이퍼하우스.
- 오세욱. (2016). 미디어로서의 봇(bot) : 뉴스 챗봇에 대한 시론적 논의. 『한국언론정보학회』, 79(10), 70-103.
- 윤병수. (2017). “기계학습을 이용한 핵심 특허문헌 추출시스템에 관한 연구”. 한국해양대학교 박사학위논문.
- 윤수진. (2016). “뉴 실버세대의 감성 커뮤니케이션과 맞춤형 가상비서 (Virtual Assistant) 콘텐츠 연구”. 경희대학교 박사학위논문.
- 윤여선. (2007). “미국특허회피 설계 연구”. 연세대학교 석사학위논문.
- 이기은. (2019). “중소기업의 혁신을 위한 적합기술 탐색 방법론의 개발과 응용 연구”. 동국대학교 박사학위논문.
- 이대근, 나승유. (2018). 중소상인을 위한 AI 챗봇 플랫폼의 설계 및 구현. 『한국전자통신학회』, 13(3), 561-570.
- 이민정, 송창현, 김연배. (2018). 국제특허분류(IPC) 공동출현 네트워크 분석을 통해 살펴본 기업의 지식융합 특성이 혁신성과에 미치는 영향. 『지식재산연구』, 13(1), 245-284.

- 이순근. (2016). “특허문서의 키워드 추출을 위한 텍스트 마이닝 기반의 관계성 그래프 모델”. 강릉원주대학교 박사학위논문.
- 이시윤. (2014). 『신민사소송법 제8판』. 경기 : 도서출판 박영사.
- 이지현. (2014). “특허회피설계 방법론과 그에 따른 발명의 특허성 확보 방안에 관한 연구”. 연세대학교 석사학위논문.
- 이용규. (2019). “Deep Learning을 활용한 3D 모델링 디자인권의 지식재산권 침해 예측 평가 연구”. 호서대학교 박사학위논문.
- 이일현. (2016). 『EasyFlow 회귀분석』. 서울 : 한나래 아카데미.
- 임병웅. (2007). 『특허법 제6판』. 서울 : 한빛지적소유권센터.
- 장인규. (2016). “허가특허연계제도 시행 이후 제네릭 의약품 특허회피전략 분석”. 성균관대학교 석사학위논문.
- 장준희. (2016). 『인공지능 기반의 ‘챗봇(ChatBot)’ 서비스 등장과 발전 동향』. 대구 : NIA 한국정보화진흥원.
- 장한수, 황성하, 김유빈. (2015). 특허분석을 통한 신기술의 잠재산업 발굴 - 핵융합 실증로 기획을 중심으로. 『한국기술혁신학회』, 206-211.
- 전기억. (2001). “신기술 영역에서 특허보호와 관련된 한국의 기술경쟁력”. 서울대학교 박사학위논문.
- 전원표, 김학수. (2011). 음성 기반 챗봇을 위한 중복 발화 회피 기법. 『한국정보과학회 학술대회』, 225-227.
- 정병기, 고남욱, 경진영, 최덕용, 윤장혁. (2017). 장벽특허의 무효화 분석을 위한 선행특허기술 발굴시스템 개발. 『엔트루정보기술저널지』, 16(1), 97-108.
- 정은희, 이병관. (2015). 오피니언 마이닝 기반 SNS 감성 정보 분석 전략 설계. 『한국정보전자통신기술학회』, 8(6), 544-550.
- 정소라. (2017). “특허 검색을 통한 인공지능 기반 음성인식 유망 기술 분석”. 고려대학교 석사학위논문.
- 정지원. (2016). “QFD 기법을 활용한 전기자동차 무선충전기의 이물질 검출 기술에 관한 회피설계 전략”. 성균관대학교 박사학위논문.

- 조경선, 임재용. (2014). 『지식재산금융과 가치평가실무』. 서울 : 한국금융연수원.
- 주종호. (1998). 『미국 특허출원서 쓰는 방법』. 대전 : 한지재권 연구회.
- 최수민, 최용순. (2017). 모바일 메신저기반 인공지능 챗봇의 상품주문 결제 인터페이스 연구. 『한국 HCI학회 학술대회』, 237-240.
- 최옥희, 유명식. (2017). 특허 포토폴리오 구축을 위한 응용 제품에 대한 신규 아이디어 도출 알고리즘 연구. 『한국통신학회 학술대회』, 455-456.
- 챗봇 관련 홈페이지(<http://www.Chatbots.org>), (2019년 11월 11일자 검색).
- 커넥팅립. (2019). 『지금 우리에게 5G란 무엇인가』. 서울 : 미래의 창.
- KB금융지주연구소, (2016). 인공지능 기술의 발달과 가상 개인비서 서비스의 진화, 『KB지식비타민』, 16(53).
- 황금하, 권오욱, 이경순, 김영길. (2017). 목적지향 대화 시스템을 위한 챗봇 연구. 『정보처리학회』, 6(11), 499-506.

2. 국외문헌

- Abbas, A., Zhang, L., and Khan, S. U.. (2014). “A literature review on the state-of-the-art in patent analysis”, *World Patent Information*, 37, 3–13.
- Ahuja, G., and Lampert, C. M.. (2001). “Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions”, *Strategic management journal*, 22, 6–7.
- Benedikt, F. C.. (2013). “The future of employment: How susceptible are jobs to computerisation?”, *Oxford*.
- Candello, H., Pinhanez, C., and Figueiredo, F.. (2017). “Typefaces and the Perception of Humanness in Natural Language Chatbots”, *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 3476–3487.
- Curran, C., and Leker, J.. (2011). “Patent indicators for monitoring convergence—examples from NFF and ICT”, *Technological Forecasting and Social Change*, 78(2).
- Das, T. K., and Teng, B. S.. (2003). “Partner analysis and alliance performance”, *Scandinavian journal of management*, 19(3), 279–308.
- David, T.. (2015). “In 10 years, your job might not exist. Here’s how to make sure you’re still employable.”, *The Washington Post*.
- Ernst, H.. (2003). “Patent information for strategic technology management”, *World Patent Information*, 25(3).
- Fontana, R., et al.. (2008). “Mapping technological trajectories as patent citation networks: An application to data communication standards”, *Economics of Innovation and New Technology*, 18(4).

- Geum, Y, et al.. (2012). “Technological convergence of IT and BT: Evidence from patent analysis”, *ETRI Journal*, 34(3).
- Harhoff, D., Scherer, F. M., and Vopel, K.. (2003). “Citations, Family Size, Opposition and the value of Patent Rights”, *Research Policy*, 32(8), 1343–1363.
- Harold, S., Michael, Z., and Justin, R.. (2015). “The Robotics Revolution”, *BCG*.
- Harrigan, K. R., and DiGuardo, M. C.. (2017). “Sustainability of patent-based competitive advantage in the US communications services industry”, *The Journal of Technology Transfer*, 42(6), 1334–1361.
- Hill, J., Ford, W. R., and Farreras, I. G.. (2015). “Real conversations with artificial intelligence: A comparison between human – human online conversations and human – chatbot conversations”, *Computers in Human Behavior*, 49, 245–250.
- Hinton, G. E., et al.. (2012). “Deep neural networks for acoustic modeling in speech recognition : The shared views of four research groups”, *IEEE Signal Processing Magazine*, 82–97.
- Hu, A. G., and Jaffe, A. B.. (2003). “Patent citations and international knowledge flow: the cases of Korea and Taiwan”, *International Journal of Industrial Organization*, 21(6).
- Ioana, C., Bot, R. I., and Wanka, G.. (2004). “Patent document classification based on mutual information feature selection”, *Italian journal of pure and applied mathematics*, 23, 121–136.
- Jing, H., et al.. (2013). “Semantic naïve bayes classifier for document classification”, *International Joint Conference on Natural Language Processing*, 1117–1123.
- Kim, G., et al.. (2014). “Technology Analysis from Patent Data Using Latent Dirichlet Allocation”, *Soft Computing in Big Data*

- Processing, Springer International Publishing*, 271, 71–80.
- Kim, G. J., Park, S. S., and Jang, D. S.. (2015). “Technology forecasting using topic-based patent analysis”, *Journal of Scientific and Industrial Research*, 74(5), 265–270.
- Khattak, A. S., and Heyer, G.. (2011). “Significance of low frequent terms in patent classification using IPC hierarchy”, *Proceedings of the 6th International Multi-Conference on Computing in the Global Information Technology*, 239–250.
- Lai, K. K., and Wu, S. J.. (2005). “Using the patent co-citation approach to establish a new patent classification system”, *Information Processing & Management*, 41(2).
- Lin, Y. S., Jiang, J. Y., and Lee S. J.. (2014). “A similarity measure for text classification and clustering”, *IEEE Transactions on Knowledge and Data Engineering*, 26(7), 1575–1590.
- Magdy, W., and Jones, G. J.. (2011). “A study on query expansion methods for patent retrieval”, *Paper presented at the Proceedings of the 4th workshop on Patent information retrieval*.
- Metra, M.. (2013). “Positive Impact of Industrial Robots on Employment”, *IFR*.
- Michael, C., James, M., and Mehdi, M.. (2015). “Four fundamentals of workplace automation”, *McKinsey*.
- Oriol, V., and Quoc, V. L.. (2015). “A neural conversational model”, *Proceedings of the 31 st International Conference on Machine Learning*, Deep Learning Workshop.
- Park, C. J., Seong, D. S., and Lee, K. B.. (2012). “Automatic IPC classification for patent documents using machine learning”, *Journal of Korean Institute of Information Technology*, 10(4), 119–128.

- Park, C. J., et al.. (1014). “Automatic IPC classification of patent documents using the term clustering”, *Journal of Korean Institute of Information Technology*, 12(9), 127–139.
- Park, H., et al.. (2013). “A patent intelligence system for strategic technology planning”, *Expert Systems with Applications*, 40(7), 2373–2390.
- Park, Y., et al.. (2015). “A patent recommendation method for technology opportunity exploration using firms own patents: the case of 3D printing”, *The Journal of Intellectual Property*, 10(1), 169–200.
- Rui, Y., Yiping, S., and Hua, W.. (2016). “Learning to respond with deep neural networks for retrieval-based human-computer conversational system”, in *Proceedings of SIGIR’16*, Pisa, Italy.
- Song, Y., and Hou, L.. (2012). “A methodology for text classification based on feature clustering”, *Proceedings of the International Conference on Automatic Control and Artificial Intelligence*, 119–124.
- Seo, S. H., Ro, K. H., and Lee, H.. (2018). “Chatbot Engine and Algorithm for Patent Avoidance Design”, *International Journal of Engineering & Technology(IJET)*, 7(2.33), 1999–1204. [SCOPUS, published on August 6, 2018, ISSN: 2227–524X].
- Tang, J., et al.. (2012). “Patent miner : Topic-driven patent analysis and mining”, *Proceedings of the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1366–1374.
- Tong, L. H., He, C., and Shen, L.. (2006). “Automatic classification of patent documents for TRIZ users”, *Proceedings of the World Patent Information*, 28(1), 6–13.
- Trappey, A. J., et al.. (2006). “Development of a patent document

- classification and search platform using a backpropagation network”, *Journal of Expert Systems wiith Applications*, 31(4), 755–765.
- Wu, C. H., Yun, K., and Tao, H.. (2010). “Patent classification system using a new hybrid genetic algorithm support vector machine”, *Journal of Applied Soft Computing*, 10,(4), 1164–1177.
- Yoon, B., and Park, Y.. (2004). “A text-mining based patent network: Analytical tool for high-technology trend”, *Journal of High Technology Management Research*, 15(1).
- Yoon, J., et al.. (2013). “Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis”, *Scientometrics*, 94(1): 313–331.
- Yuan, W., Guan, D., Lee, S., and Lee, Y. K.. (2007). “The role of trust in ubiquitous healthcare”, *Application and Services, 2007 9th International Conference on e-Health Networking*, IEEE, 312–315.
- Hu, Z., H., Fang, S., and Liang, T.. (2014). “Empirical study of constructing a knowledge organization system of patent documents using topic modeling”, *Scientometrics*, 100(3): 787–799.

부 록

<특허회피 설계안 고찰에 관한 설문지>

1-3. 귀하는 변리사업 혹은 지식재산 관련 업무에 근무하였거나 현재 근무하고 계십니까?

- ☐ 네, 과거, 현재 근무하고 있습니다.
- ☐ 네, 과거에는 변리사업(지식재산 관련 업무)에서 근무하였으나, 현재에는 다른 업종에서 일하고 있습니다.
- ☐ 네, 과거에는 다른 업종에서 일을 하였으나, 현재 변리사업(지식재산 관련 업무)에 종사하고 있습니다.
- ☐ 아니요, 전혀 없습니다.

1-4. 귀하의 경력연수는 어떻게 되십니까?

- ☐ 지체권 분야에서 3년 이하
- ☐ 지체권 분야에서 10년 이하
- ☐ 지체권 분야에서 10년 이상
- ☐ 다른 산업권 분야에서 3년 이하
- ☐ 다른 산업권 분야에서 10년 이하
- ☐ 다른 산업권 분야에서 10년 이상

2. 지식재산권 및 특허회피 설계에 관하여 귀하의 의견도 여쭙고 싶습니다.

2-1. 지식재산권의 획득이 어떤 대응도 필요하다고 보십니까?

- ☐ 시장에서의 독점과 이익을 확보하기 위해
- ☐ 특허권장 및 권리보호를 받기 위해
- ☐ R&D 투자비용 및 향후 추가기술 개발에 대한 리스크 관리 측면을 위해
- ☐ 정부의 각종 정책자금 및 다양한 세제 혜택을 계층별기 위해

2-2. 지식재산권(특허권이 어떤 점에서 도움이 되었는지요?

- ☐ 정부 R&D 사업 과제 추진 및 수행으로 기업 가치와 성장성에 도움이 됨
- ☐ 정부출연(보조)자금 또는 정부대출과 같은 자본금 확충에 도움이 됨
- ☐ 우수제품 개발 및 품질 향상을 위한 시장 진입에 도움이 됨
- ☐ 제품 마케팅의 성공으로 기업 수익성 향상에 도움이 됨

2-3. 기업이 지식재산권의 확보에 하는 이유가 있다면 무엇이라 생각하십니까?

- ☐ 시장에서의 독점적 지위를 확보할 수 있음으로
- ☐ 특허발명의 예방 및 권리 보호를 받을 수 있음으로
- ☐ 연구개발 투자비용의 회수 수단 및 추가 기술개발의 원인이 됨
- ☐ 연구개발의 방향성 설정과 연구기간 및 비용 절감에 유리함으로

2-4. 지식재산권의 확보를 위해 당사는 특허 회피설계를 수행 및 경험한 적이 계십니까?

- ☐ 네, 계속 현장에서 수행하거나 경험하고 있습니다.
- ☐ 아니요, 전에는 수행 및 경험하였으나 현재는 다른 업종에 있습니다.
- ☐ 네, 과거에는 없었으나, 현재 수행 및 경험하고 있습니다.

#지식재산권 분야 중 특허회피 설계안 고찰에 관한 설문조사#

안녕하세요~
최근 지식재산권 분야에서 국내외적으로 필수 불가결하게 일어나고 있는 특허회피로 인한 법적 분쟁에 대한 도전을 완충하기 위해 도움을 주며, 기업의 미래 성장동력을 발굴 육성해 주는 데 나 공헌 역할을 하고 있는 특허회피 설계안들을 더 바르게 알수 및 고찰하고자 다음과 같은 설문 조사를 하였습니다.

하여, 귀하의 대표 등, 특허사무소 변리사 등 이와 지식재산권 분야에 많은 관심을 갖고 계시는 자사직원 대상으로 간단한 설문조사를 실시하고자 합니다.

본 설문을 작성하기 위한 소요시간은 10분 정도로 예상됩니다. 비번시타면도 잠시만 시간을 내 시어 설문지 응해 주시면 대단히 감사하겠습니다. 수집된 개인정보는 외부로 절대 공개되지 않으며 조사 결과는 한정대학교 스마틴홀딩스 컨설팅학과 연구과제를 위한 내부 기초 자료로만 사용될 것입니다.

본 설문과 관련하여 궁금하신 점은 특허법인(대한) 사상의 이서(010-3982-8134, sghyun11@naver.com)으로 당직은 문자해 주시면 되겠습니다.

응답기간 : 2019.03.20(수) - 2019.04.07(월)

*는 필수항목입니다.

1. 먼저 자료 분류를 위해 귀하의 개인적 사항에 대해서 몇 가지 여쭙겠습니다.

1-1. 귀하의 성별은 무엇입니까?

- ☐ 남성
- ☐ 여성

1-2. 귀하의 연령대는 어떻게 되십니까?

- ☐ 20대
- ☐ 30대
- ☐ 40대
- ☐ 50대
- ☐ 60대
- ☐ 70대

○ 아니요, 전혀 없습니다.

2-5. 당신은 특허회피 설계안 또는 특허회피 전략 보고서를 받아 보았거나 향후 받아 볼
당, 도움이 되었거나 도움이 될까라 또는지요?

- ☐ 네, 많은 도움이 되었거나 될까라 본다.
☐ 아니요, 전혀 도움이 없었거나, 전혀 없 될까라 본다.
☐ 네, 일부 도움이 되었거나 조금 도움이 될까라 본다.
☐ 아니요, 전혀 도움이 없었으나, 필요성은 있다고 본다.

2-6. 2-5번 문항에 대한 답변을 왜 그리 선택하셨는지 이유를 간단히 적어주세요.

2-7. 특허회피 설계안이 왜 필요하다고 보는지요(복수선택)

- ☐ 특허 등록
☐ 기술 사업화
☐ 틈새시장 발굴을 통한 시장 선점
☐ 자사기술 확보 및 보호장벽 마련
☐ 기타

2-8. 다음 제시된 특허 회피설계를 하기 위한 방법들 중 가장 중요시 여겨야 한다고 보는
핵심사항은 무엇이라 보는지요?

- ☐ 발명자의 직무발명신고서
☐ 키워드리스를 활용한 선행기술조사
☐ 포털 사이트에 올라온 검색정보 활용
☐ 외부 전문가의 표본트 및 방향제시

2-9. 특허회피 설계안을 보고 받았다고 가정했을 시, 그 후이 특허성 또는 진보성이 있다
고 판단하는 기준은 무엇이라 보는지요(복수선택)

- ☐ 특허 전문가인 변호사의 우수한 경험지
☐ 특허회피 설계안에서 실시되는 선행기술 조사
☐ 포털 사이트에서 제공하는 다양한 검색정보 적용
☐ 발명자의 우수한 기술개발 능력 및 성과경험
☐ 기타

2-10. 특허회피 설계안 제시하기 전에 먼저 이루어져야 할 사항은 무엇일지요(복수선택)

- ☐ 발명기술의 유지 파악
☐ 발명기술과 유사한 선행기술 인지
☐ 발명기술 관련 시장조사
☐ 발명기술 관련 사업성평가
☐ 기타

2-11. 특허사무소를 통해 제시되었던 종래 잘못된 특허회피 전략 보고서에서 느꼈던 문제
점이 있었으면 무엇이었을지요(복수선택)

- ☐ 주관적 관점의 추 제시로 일부 편향된 사업 모델로 인한 미래 방향성 상실
☐ 제시된 보고서에 따라 줄임되었던 특허의 등록거절
☐ 줄임 후에 등록받은 특허의 효용성 및 가치성 문제
☐ 등록된 특허의 제용화 추진 후, 시장의 반응미비
☐ 기타

3. 기업발전에 도움을 줄 목적으로 특허회피 전략을 제시해 주는 AI 플랫폼에 관한 설문
리들을 다양하게 설계 및 연구하기 위해, 이하 여러가지 다양한 의견을 여쭙보고자 합니다.
영역 신물리이드에서 167 무인자동차를 인벤키에 의한 키워드와 도합사이드를 통해 얻어진
주제어를 활용하여 특허회피 설계안을 제시해 주는 것을 V0.1로 불리고 있는 모델을 보실 수 있으
실 것입니다.



3-1. 특허회피 전략을 제시해 주는 AI 플랫폼이 시장에 출시되어 사용했을 경우, 판매 할
고 지내던 변호사님들 통해 제언받았던 특허회피 설계안과 대비해 어떤 점에서 차이를
발견할 수 있을지요(복수선택)

- ☐ 버리사의 컴퓨팅 비용 vs 저렴한 수수료
- ☐ 다양한 특허회피 설계안을 제공
- ☐ 인제 이디스는 믿고 사용할 수 있음
- ☐ 계속적인 AI의 지능화로 결과지 신뢰
- ☐ 기타

3-2. 향후, 특허회피 전략을 제시해 주는 AI 플랫폼이 출시된다면, AI 플랫폼이 기본적으로 갖춰야 하는 핵심 요소가 있다면 무엇이든 모든지(가)꼭수신해)

- ☐ 유저와 AI 플랫폼 간의 밀접한 소통
- ☐ 다양한 특허회피 설계안 제시
- ☐ 사용자 인터페이스의 조작 편리성 및 이해 속도
- ☐ 핵심 키워드로 검색되는 관련 기술동향 및 산업트렌드 제공
- ☐ 기타

3-3. 특허회피 전략을 제시해 주는 AI 플랫폼의 설계 역역의 일부는 기업들이 영위하고 있는 산업군에 해당하는 관심 분야를 특허정보 검색사이트(키프리스)에 입력하여 선행 특허를 추출하려 합니다. 이때 추출된 선행 특허에 기재되어 있는 특허청구범위(청구항)를 포괄도 스크랩할 수 있게 구현해 보고자 합니다. 이때 고려할 점이 있다면 무엇이 있을지요?

- ☐ 정보의 다양화를 높이기 위해 특허청구범위 외 다른 구분영역도 스크랩할 필요가 있다
- ☐ 정보의 대표성을 높이기 위해 특허청구범위 중 독립항을 중심으로 스크랩할 필요가 있다
- ☐ 정보의 표준화를 높이기 위해 특허청구범위의 전체를 스크랩할 필요가 있다고 볼
- ☐ 정보의 간단 명료성을 높이기 위해 부호의 활용, 요약시ופן 스크랩해도 충분하리라

3-4. 특허회피 전략을 제시해 주는 AI 플랫폼의 설계 역역의 일부는 이미 추출된 선행 특허로부터 유추성을 갖는 단어 즉, 특허 키워드를 뽑아낸 다음 포괄 사어드에 입력해 다수의 정보(청구 특허회피 전략을 제시하는데 유용하게 사용되는 백데이터)를 추출해 나고자 합니다. 이때 고려할 점이 있다면 무엇이 있을지요?

- ☐ 유저의 산업군과 추출된 정보들 간의 연관성 판단
- ☐ 추출된 정보들간의 공통분모를 찾아 카테고리별 분류
- ☐ 유저의 산업군에 속한 현재 관심도가 높은 정보를 보여주
- ☐ 과거 데이터와 되었던 검색어가 기재된 정보를 보여주

3-5. 특허회피 전략을 제시해 주는 AI 플랫폼의 설계 역역의 일부는 다수의 정보로부터 특허어들을 추출한 다음, IPC(국제특허분류)를 활용해 기업의 산업군과 특허어들 간의 연관성에 따라 4개의 카테고리로 분류해 내는 1차 필터링을 실시하고자 합니다. 이때 고려할 점이 있다면 무엇이 있을지요?

- ☐ IPC의 부정확성에 대한 분류반식 채택 권유
- ☐ IPC의 특허분류코드를 부여받은 유저의 산업군과 특허어들 간의 비교분석
- ☐ 4개의 카테고리도 분류된 특허어들의 활용도 모색

- ☐ 코드간 비유사로 저가중지 집단으로 구분된 특허어들의 분류상태에 대한 제원중

3-6. 특허회피 전략을 제시해 주는 AI 플랫폼의 설계 역역의 일부는 1차 필터링된 특허어들의 IPC를 각각의 기준(기업의 산업군과)에 속한 관심분야에 해당하는 IPC 코드의 기준(전문가)에 해당하는 5급 특허어부터 인지되는 IPC 코드도 각기 비례해, 1차 필터링된 특허어들(1차 필터링된 특허어)과 관련된 청구(제) 기술(제)를 추출하고, 도출된 백데이터들(제)에는 각각의 설계(제)에 대한 IPC 코드를, 2차 필터링된 특허어들의 사어에 대해, 관련 IPC 코드를, 3차 필터링을 실시하고자 합니다. 이때 고려할 점이 있다면 무엇이 있을지요?

- ☐ 05를 기준으로 한 피어는 상관계수의 일관적 접근에 대한 용이성 판단
- ☐ 거중지 값 부여에 따른 객관적인 기준 설정 모색
- ☐ 전문도가 반영한 5급 특허의 제각 기준에 대한 고시 의무화
- ☐ 상관성으로 구한 피어는 상관계수 값으로 알아낸 최적점수로 제거한 특허어들 이외의 제각

3-7. 특허회피 전략을 제시해 주는 AI 플랫폼의 설계 역역의 일부는 1차 필터링 또는 2차 필터링된 특허어들을 확장된 오픈미인 마인스 기법(EMOSE)으로 검색해가 <고려 및중점적점도 조사>를 실시해 사용자에게 추천된 내장적인 특허어를 골라냄에 따라 향후 특허회피 전략을 제시하는데 내장적하게 사용될 특허어를 최종 선정하는 3차 필터링을 실시하고자 합니다. 이때 고려할 점이 있다면 무엇이 있을지요?

- ☐ 추출된 정보들간에 표현된 감정, 부정, 중립 어부에 대한 판단기준 모색
- ☐ 특허어에 대한 최종 선정 기준을 상호 평표지 이외 다른 기준으로 채택
- ☐ 기존 오픈미인 마인스 기법에 정용사들 더 추가시킨 이유에 대한 적합한 설명
- ☐ 최종 선정된 특허어들의 자어어처리(NLP) 및 자국어로 표현된 특허회피 전략의 언어 구성력

3-8. 특허회피 전략을 제시해 주는 AI 플랫폼의 중시를 받고 필요성을 느껴 사용하게 되었을 시, 기능적으로 혹은 사업적으로 어떤 점들을 더 보완하거나 강화하면 좋으시 조원 부탁드립니다.

3-9. 특허정보 검색사이트(www.kipris.or.kr)를 통해 검색된 특허 건수가 몇 건 정도냐여야 나의 관심분야에 사용되는 특허건수가 몇 건 정도냐여야

- ☐ 10건 이하
- ☐ 10건 내지 30건
- ☐ 30건 내지 70건
- ☐ 70건 내지 100건
- ☐ 기타

3-10. 특허정보 검색사이트(www.kipris.or.kr)를 통해 검색된 특허 건수 중 몇 건 정도가 특허회피 전략을 제시해 주는 AI 플랫폼에 적용되어야 내가 원하는 특허회피 설계안을 중

본히 제시해 줄 수 있는 것이라 보는지요?

- 1건 내지 2건
○ 3건 내지 5건
○ 6건 내지 10건
○ 10건 이상
○ 기타

3-11. 3-10의 권역에 대한 선택을 그리 정한 이유가 있다면 간단히 말씀 부탁드립니다.*

3-12. 포털 사이트를 통해 주중에 년 정보글들 중 최고점(100%)의 클릭율을 차지해 본 경
험자가 몇 회 정도이면, 특허회피 권역에 활용을 한한 정보글로 채택될만 할까요?(3년 기
준)

- 2회 이하
○ 3회 내지 5회
○ 6회 내지 10회
○ 10회 이상
○ 기타

3-13. 포털 사이트를 통해 주중에 년 정보글들 중 최고점(100%)의 클릭율을 차지해 본 경
험자가 해당 산업군에 따라 빈번하게 나타나거나 혹은 드물게 나타나는 것입니다. 다
음 표기 중 클릭률의 빈도가 매우 심하게 떨어진 경험은 있을 것이라 예측되는 산업군은
무엇이라 보는지요?(3년 기준)

- 의학, 식품, 헬스케어, 성화학 분야
○ ICT, IoT, AI, 가전, 로봇, 정보처리 분야
○ 기계, 메카트로닉스, 자동차, SoC 분야
○ SW, 법률(금융, 복지), 플랫폼 서비스, 문화예술 분야
○ 농수산, 유통물류, 조선해양, 항공우주 분야
○ 기타

3-14. 특허회피 권역을 제시해 주는 AI 플랫폼이 특허회피 전략을 최초 1회 제시함에 있
어, 포털 사이트를 통해 주중에 년 정보글로부터 떨어진 주제어들이 어느 정도나되어 적절
하다고 보는지요?

- 10개 미만
○ 10개 내지 30개 미만

- 30개 내지 50개 미만
○ 50개 이상

3-15. 특허회피 권역을 제시해 주는 AI 플랫폼이 특허회피 전략을 제시함에 있어, 기업들
에게 도움이 될 정보글의 선정에 무엇보다도 필요한 것인지는 어떤가요? 필요하면 포털 사이트로
는 어디가 좋을지요?

- 민간포털사이트
○ 정부포털사이트
○ 언론매체사이트
○ 소셜미디어사이트
○ 기타

3-16. 특허회피 권역을 제시해 주는 AI 플랫폼이 특허회피 전략을 제시함에 있어, 기업들
에게 도움이 될 정보글의 선정에 무엇보다도 필요한 것인지는 어떤가요? 필요하면 포털 사이트를
하회피 전략으로 활용하는데 가장 중점을 두어야 할 것이 있다면 무엇일까요?

- 과거부터 현재까지 핫이슈로 상위 검색순위에 여러차례 올라오는 정보글이어야 함 (10
년 기준)
○ 현재 핫이슈로, 2~3회 정도 상위 검색순위에 올라온 정보글이어야 함 (10년 기준)
○ 과거 핫이슈로, 상위 검색순위에 여러차례 올라온 정보글이어야 함 (10년 기준)
○ 과거 핫이슈로, 2~3회 정도에 한해 상위 검색순위에 올라왔던 정보글이어야 함 (10년
기준)

3-17. 저문으로 제시된 다양한 플랫폼 중 특허회피 설계안을 제시하는 플랫폼 타원은 어
떤 형태가 적절할 것이라 보는지요?

- 가게(store) 비즈니스 모델(iOS와 앱스토어를 기반으로 앱과 콘텐츠를 공유할 수 있도
록 한 모델)
○ 사찰(chain) 비즈니스 모델(대형 유통사에 제품 주문량, 주문시간을 전체 관리할 수 있
도록 맡기는 모델)
○ 공유(share) 비즈니스 모델(소상공인을 통해 공급자와 수요자가 비즈니스 활동을
공동으로 하는 모델)
○ 대리인(agent) 비즈니스 모델(대리인을 선임해 간접 매개가 이루어지도록 한 모델)
○ 펠리비스(felivis) 비즈니스 모델(각종 원격 컴퓨터 서비스를 활성화시켜 비즈니스
활동으로 이어갈 수 있도록 한 모델)
○ 게임(game) 비즈니스 모델(게임 방식을 통해 만난 공급자와 수요자가 서로 비즈니스
활동을 할 수 있게 한 모델)

3-18. 특허회피 권역을 제시해 주는 AI 플랫폼은 특허회피 전략을 자극어로 이끌어 줄 수 있
도록 하기 위해 다양한 자연어 처리기법을 적용시켜야 한다고 판단됩니다. 따라서, 다음
보기에 서술된 자연어 처리기법 중 향후 특허회피 전략의 고급화를 위해 채택을 필요로 하
고 보여지는 핵심기법이 있다면 무엇이라 보는지요?(복수선택)

- ☐ 오피니언 마이닝(특정 주제에 대한 여론이나 정보(댓글이나 게시글)를 수집, 분석해 평
판을 도출하는 기법)
☐ 시맨틱 분석(컴퓨터가 정보자원의 뜻을 이해하고, 논리적 추론까지 할 수 있는 기법)
☐ 텍스트 마이닝(한정형/비정형 텍스트 데이터를 정제하고, 특정 추출 및 의미있는 정
보를 발견해 내는 기법)
☐ 단어 임베딩(단어와 단어 간의 유사성, 유사한 의미를 가진 단어들을 발견하여, 미래에
사용될 단어 정보들을 추출해내고, 단어 간의 연결(이동)을 가능하게 하는 기법)
☐ 문맥포지셔닝(단어 간 관계 정의를 통해 문장의 의미를 파악할 수 있는 기법)

- ☐ 메타분석(동일하거나 유사한 연구 주제로 실시된 많은 통계적 연구를 다시 통계적으로 통합하고 종합하는 기법)
- ☐ 기타

3-19. 설문 후기~<자유로운 평가를 부탁드립니다.>*

<특허회피 설계안 자동생성 봇에 관한 설문지>

#특허회피 설계안 자동 생성용 봇에 관한 설문 조사#

본 설문조사는 특허회피 설계안 자동 생성용 봇을 통해 나온 결과에 대한 특허 전문가들의 의견을 여쭙봄을 통해 향후 개선점 및 진행 방향을 고찰해 볼 기회로 삼고자 다음과 같이 요청하게 되었습니다. 하여, 본 설문조사는 특허법인, 특허사무소 등에서 근무하고 계신 변리사, 통계분석 및 기술경영 컨설팅학 강의 및 연구하고 계신 교수님, 또는 특허분쟁 및 대응 예방 전략과제를 실무적으로 담당하고 계신 15년 이상의 지식재산권 경력을 가진 자분들을 대상으로 실시하고자 합니다.

본 설문을 작성하기 위한 소요시간은 10분 정도로 예상됩니다. 바쁘시더라도 잠시만 시간을 내시어 설문에 응해 주시면 대단히 감사하겠습니다. 수집된 개인정보는 외부로 절대 공개되지 않으며 조사 결과는 한성대학교 스마트융합 컨설팅학과 연구과제를 위한 내부 기초 자료로만 사용될 것입니다.

본 설문조사는 5개 문항이다. 조사기간은 2019년도 11월 15일부터 2019 12월 5일까지이다.

1. 최종 주제는 봇이 특허회피 설계안을 제시하는데 쓰일 Start Point라 할 수 있습니다. 출원인의 관심분야로 검색하신 핵심 키워드(ex : IoT*사물인터넷*(센서+감지)*5G)를 특허회피 설계안 자동 생성용 봇에 넣고 검색한 결과, [시스템,보안,이용 관리, 분석,클라우드,관리제어,데이터]단어들이 최종 주제로 선정되어 나왔습니다. 전문가님이 보셨을 때, 특허회피 설계안 자동 생성용 봇으로부터 표출된 최종 주제가과연 특허회피 설계안을 제시하는데 도움이 될만한 단어들이 나왔다고 볼 수있을까요?

①매우 그렇다 ②그렇다 ③보통이다 ④아니다 ⑤매우아니다

2. 특허회피 설계안을 자동 생성하는 봇에 출원인의 관심 분야라 할 수 있는 ①~⑩ 검색하신 핵심 키워드를 넣고 검색한 결과, 최종 주제가 이하 표처럼 나왔을 시에, 특허 전문가님은 이 최종 주제를 가지고 어떠한 특허회피 설계안을 제시해 보실 수 있을까요?(예시참조)

예시) ①5G 분산 클라우드 시스템에서 빅 데이터의 로그를 분석하여 장애를 관리하는 틸레 기술을 제안해 봅니다(+이하에 표 시된 단어는 필수적으로 들어가야 하는 주제어라고 봄)

	핵심 키워드	최종 주제어	특허회피 설계안
1	IoT*사물인터넷*(센서+감지)*5G	시스템,보안,이용관리, 분석,클라우드, 관리제어,데이터(+코딩, +인증)	
2	(취업+진로)*상담	시스템,서비스,최근,사회,국가,개발,역사(+정신건강, +현금)	
3	혼합현실*(증강현실+가상현실)	시스템,이용,판단,기능,차단,자동,기계(+바이오,+예측)	
4	헬스케어*(진료+의사)	시스템, 체중, 자격, 조건, 보험, 혈압, 흡연(+바디,+컨트롤)	
5	자율주행*차*(사고+단속)	시스템,작업,중고차,공간,원인,자원,안전(+자세,+거리)	
6	자연어처리*(형태소+오피니언)	시스템,사례,감성,온라인,창업,개선,소개(+컨설팅,+고객)	
7	AI*딥러닝*신경망*유전자	시스템,러닝,머신,기술,처리,인공,분석(+순위,+인간)	

8	식품*(빵+발효)*친환경	조성,제조,발효,과정,환경,원료,문화(+미생물,+바아오)	
9	노인*치매*치료*성인병	억제제,용도,시간,부작용,마약,치료,보충(+전문,+연구)	
10	가족여행*추천지역*관광	기상,매력,환경,체계,관측,태풍,활용(+주민,+실시간)	

3. 기존에 전문가님이 출원인의 관심분야를 파악해 특허회피 설계안을 제시한 결과와, 향후 특허회피 설계안 자동 생성하는 봇이 최종 주제를 이용해 특허회피 설계안을 제시한 결과 대비시, 내용면에서 큰 차이가 있을 거라 보시는지요?

①매우 그렇다 ②그렇다 ③보통이다 ④아니다 ⑤매우아니다

4. 특허회피 설계안 자동 생성용 봇이 제시하는 특허회피 설계안과, 실제 전문가가 제시하는 특허회피 설계안은 절대 내용적이거나 질적으로도 큰 차이가 있으며 AI가 아무리 뛰어나더라도 사람의 생각보단 못하다고 보신다면, 앞으로 특허회피 설계안 자동 생성용 봇이 보완해야 할 사항은 어떤 점들이 있을지 간단히 기재 요청드립니다. (주관식)

5. 특허회피 설계안 자동 생성용 봇이 대중화된다면, 특허 전문가 입장에서 바라봤을 때 현업에 근무하는 종사자 또는 출원인(고객)에게 좋은 영향을 끼칠지 나쁜 영향을 끼칠지에 대한 필요성을 여쭙습니다.

필요성에 대한 다양한 질의	그렇다	아니다
수익을 증대시키는데 동반 상승 역할을 해줄 것이라 기대된다.		
출원인이 개량 발명을 새롭게 고안해 내는데 오히려 혼돈을 줄 것이라 본다.		
법률 특허, 지식재산권 서비스를 대중화시키는데 이바지할 것이라 본다.		
출원인이 독점권을 유지하는데 오히려 방해요소로 작용할 꺼라 본다.		
특허회피 설계안이 나온다면, 출원인의 독점권을 강화시키려는 전략도 나올 수 있다고 본다.		
특허회피 설계안은 특허 컨설팅을 통해 전문가가 출원인에게 제공해줄 수 있는 것이지, AI가 전문가가 할 수 있는 역할을 대신할 순 없다고 본다.		
AI를 이용해 법률, 특허 지식재산권 서비스를 제공하는 플랫폼은 앞으로 4차 산업 및 인류에게 도움이 되는 미래 산업 기술 및 연구 분야라 본다.		

ABSTRACT

A Study on a Bot-based Automatic Generation Method of Patent Avoidance Design

Seo, Sung-Hyun

Major in Smart convergence Product

Dept. of Smart Convergence Consulting

The Graduate School

Hansung University

The patent avoidance strategy is mainly used to find ways to secure quality patents aimed at blank technology and niche markets in advance for companies or individual customers who feel the need for market analysis and prior investigation before launching a prototype or entering research and development.

Currently, efforts continue to be made to escape the already preeminent barrier patents and the various evils occurring in the imperfect competitive market, but there are concerns that different opinions may arise depending on the patent experts when presenting the patent avoidance strategy, and the opinions of experts in particular who have different views are not sure whether they are designed to benefit the company or individual customers.

The patent avoidance strategy provides strategies to minimize human

and physical waste and avoid barriers to future research and development, thereby helping the president of the monopolist to prevent disputes over the intellectual property rights of the other party in advance, and promoting the technology of the business that further enhances and promotes the activation of products or related technologies that customers are interested in.

This paper is about a study of suggesting bot-based automatic generation method of a patent avoidance design for customers who need a patent avoidance design plan, and searches for patent case by applying a bot-based automatic generation program developed to select the final subject word that applies to automatic generation of patent avoidance design according to the application of prototypes that a company wants to release and core keyword representing its corresponding technology areas.

It will go through the process of finding out the patent keyword by analyzing the scope of patent claims listed in prior patent detected. it accesses the patent keyword to the portal site, retrieves the information article that are essential to presenting the patent avoidance design plan, and extracts the candidate subject word from the information article. As the primary filtering (similarity judgement), secondary filtering (consistency judgement), and tertiary filtering (conformity judgement) are conducted on these extracted candidate subject word, a bot-based automatic generation method of patent avoidance design presents the final subject word to be used in the patent avoidance design plan.

Accordingly, this paper has conducted an expert survey on the final subject word selected by implementing the bot-based automatic generation program to verify the usefulness of the final subject word applied to automatically generate the patent avoidance design plan,

and as a result, it proposes a bot-based automatic generation method of patent avoidance design to confirm the importance of the information that the final subject word has in using it as a patent avoidance strategy.

In other words, the study on a bot-based automatic generation method of patent avoidance design proposed in this paper could provide a variety of patent avoidance design plan that would help secure space patents and pioneer new markets, as well as ensure that ideas or superior inventions from company or individual customers are not affected by the intellectual property rights of the other party.

【Keyword】 bot, patent avoidance design, core keyword, patent keyword, information article, candidate subject word, final subject word