

제 16회 한국대학생산업공학프로젝트경진대회

사회적 매체분석을 통한 자동차 Supplier 기술 경쟁력 평가

김도연 김민지 오희경 이수영 이채림



INDEX

01. 프로젝트 배경

- ▶ 프로젝트 배경 및 필요성
- ▶ 현황분석
- ▶ 문제정의

02. 프로젝트 목표

- ▶ 프로젝트 목표
- ▶ 프로젝트 범위

03. 프로젝트 프로세스

- ▶ Conception
- ▶ 데이터 분석 프로세스
- ▶ 평가요소별 분석 프로세스

04. 결과물

- ▶ AHP 분석
- ▶ 최종 평가 지표
- ▶ 전문가 검증

05. 결론

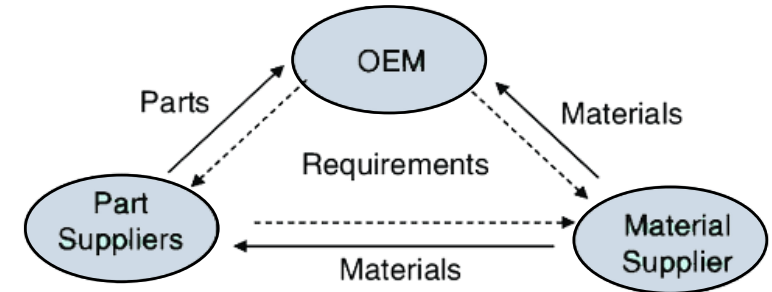
- ▶ AS IS - TO BE
- ▶ 프로젝트 의의 및 파급효과와 한계점



프로젝트 배경 및 필요성

Supplier 기술경쟁력의 중요성

- 대부분의 자동차 생산 방식은 주문자 상표 부착 생산(OEM) 시스템으로 이루어짐
→ Supplier의 기술경쟁력은 곧 OEM 기업의 경쟁력으로 연결됨
- 따라서 OEM 기업에게는 기술력이 우수한 Supplier를 찾아내는 것이 중요함



[그림 1.1] 자동차 생산구조

데이터 분석에 대한 니즈

- 경쟁사 분석 및 Supplier의 기술역량 파악을 위한 객관적인 근거가 필요함
- 중복 기사로 인한 데이터 분석의 어려움이 존재함
- 유망 기술과 소비자 니즈에 대해 지속적으로 파악하는 것이 필요함

**자동차 산업 분야에서, Supplier의 기술경쟁력 평가는 매우 중요하며,
현업에서 지속적인 데이터 분석에 대한 니즈가 존재함**

현황분석 : 기존의 기술경쟁력 평가 방법의 한계

기존의 기술경쟁력 평가의 특징

- 기술경쟁력 평가는 주로 특허/논문 분석을 통해 다음과 같은 요소를 기준으로 평가되고 있음

ex) 기술 생애 주기, 기술적 성숙도, 시장 지배력 ,기술적 영향력, 기술 집중도 등

[표 1.1] 기존의 매체분석을 통한 기술경쟁력 평가 사례 및 한계

	특허 분석	논문 분석
사례	<ul style="list-style-type: none"> · 김민수, 임선영, & 송영화. (2019). 국내 연료전지 기술경쟁력 분석에 관한 연구: 특허 분석을 중심으로. 한국기술혁신학회 학술대회, 2096-2101. · 오새새, 최의현, & 최병헌. (2017). 중국 전기자동차 산업의 발전과 기술경쟁력 비교 분석-특허 통계를 중심으로. 중소연구, 41(3), 45-79. 	<ul style="list-style-type: none"> · 김태영, & 박희성. (2015). 도로교통 기술수준 평가. 대한토목학회논문집, 35(6), 1395-1404. · Christou, C., Hadjipaschalis, I., & Poullikkas, A. (2008). Assessment of integrated gasification combined cycle technology competitiveness. Renewable and Sustainable Energy Reviews, 12(9), 2459-2471.
한계	<ul style="list-style-type: none"> · 특허출원부터 등록 및 거절까지의 절차에 약 18~20개월이 소요됨 · 기술분야, 국가, 기업, 국가적 상황에 따라 특허의 중요도가 떨어짐 	<ul style="list-style-type: none"> · 논문 출판까지 평균적으로 약 100일이 소요됨 · 신생기업 파악이 어려움

특허/논문 분석을 통한 기술경쟁력 평가에는 다양한 한계점이 존재함

문제정의 : 특허/논문의 기술경쟁력 평가 한계점 및 새로운 평가요소 도출

특허/논문을 이용한 기술경쟁력 평가의 한계점

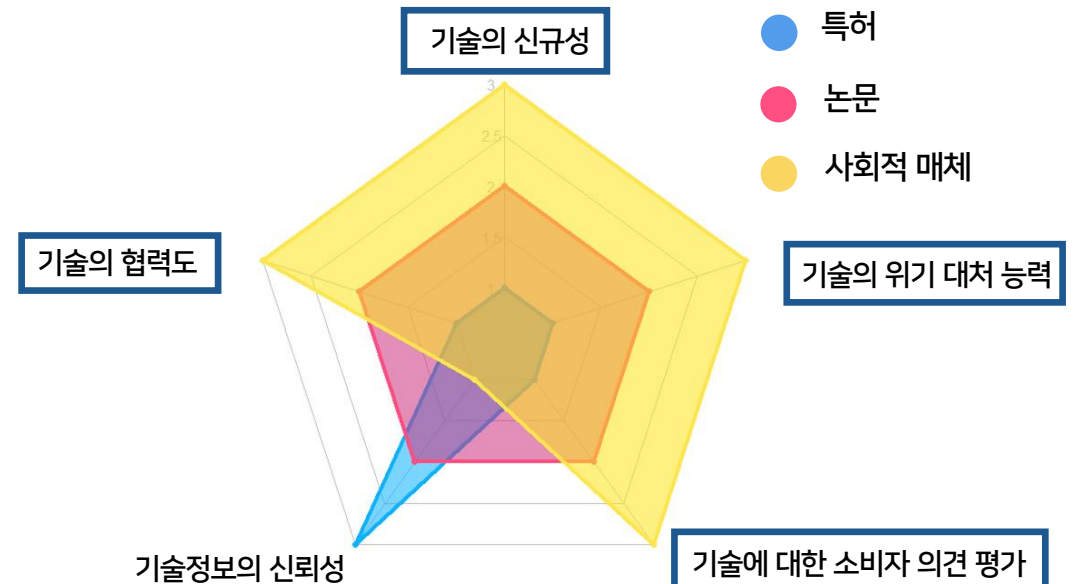
○ 신속한 정보 파악 불가, 신생기업 파악 어려움, 소비자의 의견을 평가하기 어려움

→ 사회적 매체(뉴미디어) 분석을 통해 특허/논문을 통한 기술경쟁력 평가의 한계점을 보완할 수 있는 기술경쟁력 평가요소 도출

기존 기술경쟁력 평가의 한계점을 보완한 새로운 평가요소 도출

[표 1.2] 새로운 기술경쟁력 평가요소에 대한 매체 평가 방법

평가요소	매체 평가 방법
기술의 신규성	해당 매체에서 특정 기술의 개발/양산 시기를 얼마나 신속하게 파악할 수 있는가
기술의 협력도	해당 매체에서 기업간의 협약 관계를 파악할 수 있는가
기술의 위기 대처 능력	해당 매체에서 기술적 위기 이슈에 대한 조치를 취하는지 알 수 있는가
기술에 대한 소비자 의견 평가	해당 매체에서 특정 기술에 대한 소비자의 의견 파악이 가능한가



[그림 1.2] 새로운 기술경쟁력 평가요소에 대한 매체별 평가 가능성

프로젝트 목표

프로젝트 목표

사회적 매체 분석을 통해 자동차 Supplier 기술 경쟁력 평가 지표 개발

→ 특허/논문 매체를 이용한 기술력 평가 지표의 한계점 보완

프로젝트 세부 목표

- 특허와 논문의 한계점(Missing Area)을 극복한 분석이 이루어질 수 있도록 함
 - Missing Area를 평가 요소화 함 : 기술의 신규성, 기술의 협력도, 기술의 위기대처능력, 기술에 대한 소비자 의견평가
- 사회적 매체가 가지는 한계점인 '매체의 편향성'문제를 극복할 수 있어야 함
- 평가지표를 이용하고자 하는 목적에 따라 평가요소별 가중치를 부여해 기술경쟁력을 평가할 수 있도록 함
- 기술 레벨 (Level1~3)에 따라 기술경쟁력 평가 결과를 확인할 수 있도록 함

사회적 매체 분석을 통해 기존 특허/논문 매체기반 기술경쟁력 평가의 한계를 보완할 수 있는, 새로운 기술경쟁력 평가 지표 개발

프로젝트 범위

▶ 프로젝트 범위

- 현대자동차의 주력 기술 분야인 * MECA 중 E(Electrification)의 기술 분야를 분석해 Tech tree를 작성함 [표2.1]

▶ 사회적 매체 내 데이터 분석 범위

- 자동차 OEM의 최신 동향·기술 키워드 수집을 위한 자동차 OEM 사이트 (Toyota, Mini 외 12곳)
- 자동차 Supplier의 최신 동향·기술 이슈를 수집하기 위한 자동차 Supplier 사이트 (AutomotiveWorld, Greencarcongress 외 10곳)
- 기술에 대한 소비자의 의견을 평가하기 위한 사용자 리뷰 사이트 (Consummeraffairs, Nextgreencar 외 6곳)

[표 2.1] 프로젝트 범위 : Tech tree

Level 1 : OEM 주력기술	Level 2 : Level1 세부기술 분야	Level 3 : Level2 세부기술 분야	Level 4 : Level3 세부기술 분야
Electrification	Power control system	Reduction Gear for EV	reduction gear, reduction gear for ev, gear box, etc...
		Inverter	Inverter for EV, Inverter, Rear run inverter, etc ...

	Drive train system	Motor core	Motor core, special molding compound, motor stator, etc ...

Electrification의 기술 분야를 Tech tree 구조로 체계화 하고,
수집한 데이터로부터 기술 경쟁력을 분석하고자 함

*MECA : 현대자동차(OEM) 주력 기술 분야 키워드
M : Mobility E : Electrification C : Connectivity A : Autonomous

Conception

자동차 Supplier 기술 경쟁력 평가지표 개발

사회적 매체를 분석하여
특허/논문과 상호보완 관점의
Supplier 기술경쟁력 평가 지표를 개발

01

Planning 분석기획

- 현업 사례분석 및 한계 도출
- 기존 기술경쟁력 평가방법 선행연구 조사
- 기존 매체 분석의 Missing Area 파악 및 새로운 평가요소 도출

02

Preparing 데이터 수집

- 사회적 매체(웹 기반 뉴스, 소비자 리뷰) 데이터 수집
- 텍스트 데이터 전처리

03

Analyzing 데이터 분석

- OEM-Sup-Tech 매칭 데이터 & 양산/개발·협력·위기·Action 사전 구축
- Doc2Vec 알고리즘 기반의 중복기사 스크리닝
- 평가요소별 분석 프로세스에 따른 분석 수행

04

Developing 시스템 구현

- 평가요소별 Supplier 기업의 Total 합 도출 및 점수 정규화
- 평가요소별 가중치 부여 (AHP분석)
- 평가요소별 가중치에 따른 기술경쟁력 평가지표 도출

05

Deploying 평가 및 전개

- 최종 평가지표의 타당성 검증
- 특허/논문과 상호보완적 결합
- 프로젝트 의의 및 파급효과

데이터 분석 프로세스

데이터 수집 및 전처리

데이터 수집

- 수집 데이터 언어 : 영어
- 데이터 수집 방법 : 웹 크롤링
- 데이터 분석 언어 : Python
- 데이터 수집 결과 : 뉴스 기사(26326개), 리뷰(3429개)

※ 중복기사 스크리닝을 통해 중복 기사 제거

※ 모든 데이터 수집은 5년 내 보도된 기사 데이터로 한정

※ 기사 내 수집 데이터: Date, Title, Contents, Summary

데이터 전처리

- Date Format Integration
- Normalization
- Handling Missing Data
- Tokenization
- Remove Stop-words
- Stemming

데이터 선정

데이터 수집

데이터 전처리

데이터 분석

지표 생성

지표 통합·정규화

결과 도출

데이터 분석

데이터 분석

1. 중복기사 스크리닝

- 목적 : 매체 편향성 완화
- 방법 : Doc2Vec 알고리즘

2. 연도별 기술동향 분석

- 목적 : 연도별 기술동향 분석(기술의 신규성)
- 방법 : LDA 토픽 모델링

3. 리뷰데이터 내 형용사 감성 값 도출

- 목적 : 감성분석(기술에 대한 소비자 의견 평가)
- 방법 : SentiWordnet 기반 형용사 감성분석

4. 각 평가요소별 데이터 분석

데이터 분석 프로세스

중복기사 스크리닝

```
from gensim.models.doc2vec import Doc2Vec, TaggedDocument
from nltk.tokenize import word_tokenize
import pandas as pd
import gensim

df = pd.read_excel('C:/Users/taim/Puhoer/Projects/LBY/02/2020/2020_11_11.xlsx', index=False)

doc = df['Contents']
sentences = []
for i in range(len(df['Contents'])):
    sentences.append(df['Contents'][i])
    #print(sentences)

sentences = [s.lower().strip().split(" ") for s in sentences]

sims_fing_Lover2 = []
sims = []
fims = []
fims_red = []

tagged_documents = []

for i, s in enumerate(sentences):
    tagged_documents.append(gensim.models.doc2vec.TaggedDocument(s, [i]))

Doc2Vec_model = gensim.models.doc2vec.Doc2Vec(vector_size=60, min_count=1)
Doc2Vec_model.build_with_vocab(tagged_documents)
Doc2Vec_model.train(tagged_documents, total_examples=len(tagged_documents), epochs=100)

def calculate_sims(a):
    s1_doc = word_tokenize(df['Contents'][a])
    s1_vec = Doc2Vec_model.docvecs[s1_doc]
    s1_sim = []
    s1_fim = []
    s1_fim_red = []
```

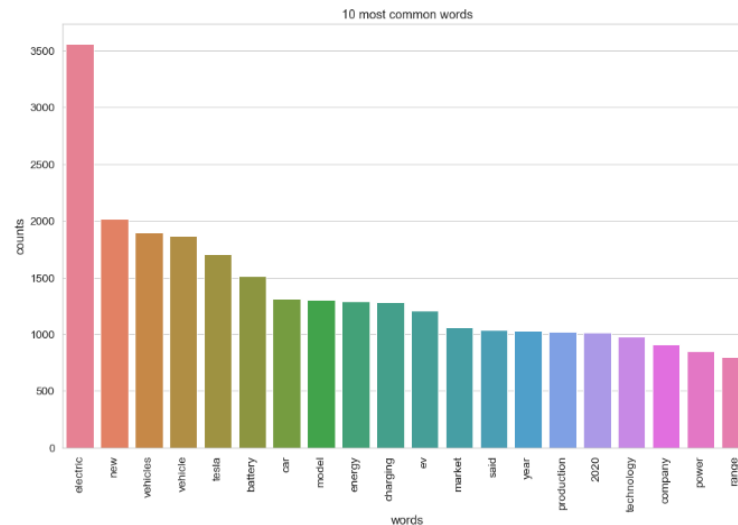
[그림 3.1] 중복기사 스크리닝 소스코드 일부분

조건1. 기간 내(3개월) 유사도가 높은 기사들

조건2. 기사 Contents, Summary에 대한 평균 유사도 값이 0.75 이상인 경우

중복 제거 후 뉴스 데이터 27,775개 → 26,326개

연도별 기술동향 분석



[그림 3.1] 2020년 Top 20 topic 그래프

* LDA 토픽 모델링을 통해,
2015~2020 연도별 Top 20 topic 도출함

리뷰데이터 내 형용사 감성 값 도출

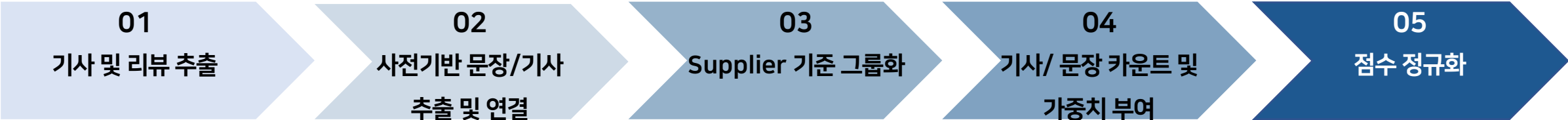
Highlighted Text	Phrase	Phrase Sentiment
...nst the bumper when closing. which was damaging	damaging	-0.40000006
...ned on its own and he had to perform an	emergency	-0.649999976
...o the erdington dealership as he didn t trust them	trust	0.49000001
he refused to send the car back to the erdington d...	refused	-0.5
...ry and they can t tell me why all these warnings are	warnings	-0.588000059
...way. the gear box felt sometimes a bit too long , but	too long	-0.5
... to common belief about toyotas service excellence ,	excellence	0.49000001
... toyotas service excellence, i have had horrible	horrible	-0.800000012
... carried over from the gasoline cr-v. the best cr-v is	the best	0.620000005
...same mechanical all-wheel drive system (real time	real time	0.600587964
...brid uses electric-motor propulsion the vast	vast majority	0.600000024
...ffer. the 5-speed manual is smooth and glitch free,	glitch	-0.49000001
...g roads in goa. kia has worked on the durability	durability	0.389999986
... while many automakers have put forth a great effort	great	0.600000024
i had expected it to respond to steep hills but found	steep hills	0.205369994
...great effort to mask them, the distinct whines of the	whines	-0.518148124
...front has a design variation for better efficiency ,	efficiency	0.5
...one in front has a design variation for better	better	0.300000012
the gear shifts are lightning-quick and precise, with	lightning-quick	0.24000001
...ll, making this powertrain a thoroughly enjoyable	enjoyable	0.842800021

[그림 3.2] 리뷰데이터 형용사 감성 분석 결과

SentiWordnet을 기반으로,
리뷰데이터 내 형용사 감성 값을 도출함

*LDA: 대량의 문자에서 확률에 근거하여 토픽을 추출하는 토픽모델링 기법

평가요소별 분석 프로세스



[표 3.1] 평가요소별 분석 프로세스

	Process 1	Process 2	Process 3	Process 4	Process 5
기술의 신규성	[기술, 양산 or 개발, 기업명] 키워드 동시 출현 기사 추출	양산/개발 키워드 사전 기반 문장 추출	Supplier 기준 Level3 기술 그룹화	LDA 결과 기반 트렌드 가중치 부여 및 연도별 문장 Count	1위 Supplier 점수를 기준으로 Total 값 정규화
기술의 협력도	[기술, 협력, 기업명] 키워드 동시 출현 기사 추출	협력 키워드 사전 기반 문장 추출		협력유형에 따른 가중치 부여	
기술의 위기 대처 능력	[기술, 위기, 기업명] 키워드가 동시 출현 기사 추출	기사 데이터 내 (주어 : 기업, 동사 : Action 키워드) 문장이 포함된 기사만 추출		기사 Count	
기술에 대한 소비자 의견 평가	[기술] 키워드 언급된 문장 추출	OEM-Supplier-Tech 매칭 사전 기반 Supplier-Tech 연결		형용사 감성분석 기반 가중치 부여 및 문장 Count	

AHP 분석 & 최종 평가 지표

AHP 분석

- 평가요소의 상대 비교를 통해 평가요소에 대한 가중치를 부여함
- 설문 대상 : 현대 NGV 기획 지원 부서(5명) & 기술정보조사·분석 부서(5명)
- 쌍대비교 행렬을 이용한 Eigen Vector 계산 결과 도출 및 일관성 검토 결과
*CR < 0.1 인 답변자의 답변을 이용하였음
- 소비자 의견평가 > 신규성 > 위기 대처 능력 > 협력도 순으로 가중치를 부여함

[표 4.1] 전문가 10명의 평가항목별 가중치 평가 결과

	기술의 신규성	기술의 협력도	기술의 위기 대처 능력	기술의 소비자 의견 평가
답변자 1	0.4027	0.1802	0.0856	0.3315
답변자 2	0.1459	0.4717	0.3209	0.0618
...
답변자10	0.1196	0.0711	0.3108	0.4985
평균	0.2498	0.1856	0.1947	0.3700

기술경쟁력 최종 평가 지표

- Level3 : Traction motor에 대한 최종 평가 지표 결과물은 [표4.2]와 같음
- Level2, Level1 관점에서도 각각 기술경쟁력 평가 결과를 확인할 수 있음

[표 4.2] 기술경쟁력 평가 지표(Level3 : Traction Motor)

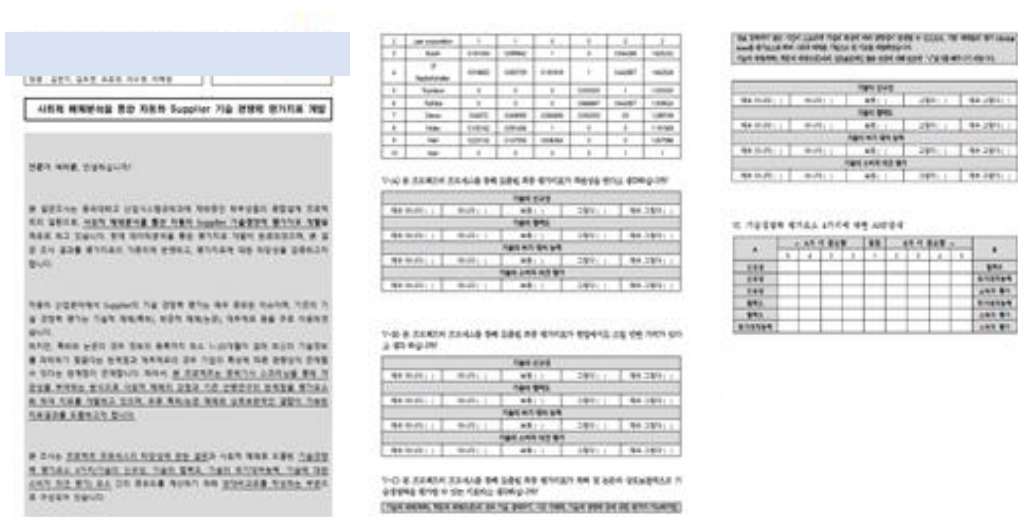
Electrification - Drive train system- Traction Motor						
순위	Supplier	기술의 신규성 (0.2498)	기술의 협력도 (0.1856)	기술의 위기 대처 능력 (0.1947)	기술에 대한 소비자 의견 평가 (0.3700)	Total
1	Bosch	0.1264	1	1	0.6897	0.6670
2	Volkswagen	0.6647	0	0.3334	0.8349	0.5398
3	Hyundai mobis	0.0313	0	1	0.9042	0.5370
4	Zf friedrichshafen	0.0112	0.1819	1	0.7213	0.4981
...
9	Toshiba	0	0	0.6667	0.6856	0.3835
10	Aisin	0	0	0	1	0.3700

* CR(일관성 비율) = $\frac{CI(\text{일관성 지수})}{RI(\text{Random Index})}$

전문가 검증

전문가 검증

○ 전문가 검증을 위한 설문지



[그림 4.1] 전문가 검증을 위한 설문지

전문가 검증 결과

- 설문지 결과 '분석방법의 적절성'에서 가장 높은 평가를 받음
- 적절성과 객관성을 갖춘 지표로 평가 받음

[표 4.3] 설문지 최종 점수

설문지 상위 항목	*최종 평가점수 평균
문제 정의	4.250
평가요소의 적절성	4.075
분석방법의 적절성	4.400
평가 지표의 객관성	4.025
평가 지표의 가치	4.225
다른 매체와의 상호보완성	4.125
총 평균	4.183

제안한 프로세스의 문제 정의 및 평가요소의 선정과 분석방법은 타당하고,
도출된 평가 지표는 객관적이고 가치 있는 평가 지표로 타 매체와의 상호보완성을 띤다고 할 수 있음

· 설문 대상: 현대 NGV 기획 지원 부서(5명) & 기술정보조사 · 분석 부서(5명)

* 최종 평가점수 = 5점 만점

05 결론

AS IS - TO BE

AS IS

중복기사
스크리닝 미수행

“ 특허·논문 중심의
기술경쟁력 평가 ”

특허·논문 분석 시
Missing Area가 존재

지속적인
데이터 분석 니즈 불충족

현업에서 Supplier의 기술
경쟁력 평가는 필수적

데이터 수집

데이터 분석

종합
분석·평가

- 특허, 논문, 시장동향
- 기술·경쟁사별 종합 분석
- 기술평가 및 경쟁력 분석
- 정형·비정형 데이터 분석
- 전략적 의사결정 지원

- √ 특허·논문 중심의 기술경쟁력 평가를 수행, Missing Area 존재
- √ 중복기사 스크리닝 미수행, 지속적인 데이터 분석의 니즈 불충족

TO BE

“ 사회적 매체를 이용한 기술경쟁력 평가 ”

→ 중복기사 스크리닝을 통한 사회적 매체 분석으로

특허와 논문의 Missing Area를 해결한 평가 지표 개발

데이터 수집

데이터 분석

종합
분석·평가

- 사회적 매체
- 새로운 평가요소 도출
- 기술평가 및 경쟁력 분석
- 중복기사 스크리닝
- 기술·경쟁사별 종합 분석
- 합리적 의사결정 지원
- 정형·비정형 데이터 분석

- √ 사회적 매체를 이용해 특허·논문의 Missing Area를 해결
- √ 중복기사 스크리닝을 통해 사회적 매체의 편향성 극복
- √ 사회적 매체를 활용해 지속적인 데이터 분석의 니즈를 충족

프로젝트 의의 및 파급효과와 한계점

산업공학적 의의 및 파급효과

- Trend에 민감하게 반응하는 산업과 사회적 매체 분석이 필요한 분야에 평가 프로세스를 제공함 → 합리적인 의사결정을 지원함
- 중복기사 스크리닝을 통해 현업에서 필요로 하는 데이터 분석의 요구사항을 충족함
- 새로운 평가요소의 도출 및 기존 기술경쟁력 평가 지표와 결합한 상호보완적 형태의 평가 지표를 개발하였음 ([표5.1] 참고)

프로젝트 한계점

- 사회적 매체 내 모든 내용은 사실이라는 가정을 수반해야 함
- 보다 신뢰성 높은 기술경쟁력 평가를 위해서는, 사회적 매체를 통한 다양한 평가요소의 도출이 추가적으로 필요함

[표 5.1] 상호보완적 형태의 평가 지표 예시

Electrification										
순위		기술의 권리적 측면			기술의 경제적 측면		기술의 환경적 측면			Total
	가중치	0.126	0.124	0.130	0.125	0.123	0.127	0.115	0.130	
	Supplier	기술의 신규성 (특허/사회)	기술적 영향력 (논문/특허)	기술적 성숙도 (특허)	기술의 시장력 (특허/논문)	기술의 생애주기 (특허)	기술의 협력도 (사회)	기술의 위기 대처 능력 (사회)	기술에 대한 소비자 의견 평가 (사회)	
1	Volkswagen	0.5681	0.3242	0	0.2481	0	0.7152	0.2871	0.1509	0.9382
2	Denso	0.3152	0	0.6298	0.2281	0.1273	0.4328	0	0.3214	0.8957
3	Zf Friedrichshafen ag	0.6120	0.8731	0	0	0.2415	0.1210	0	0.1539	0.8726

※ 요소별 가중치는 프로젝트의 유형에 따라 사용자 기준에 맞게 변경 가능함

※ 하나의 평가요소에 대해 두 가지 매체를 이용해 분석하는 경우 매체 별 가중치를 설정해 Total 계산