

# 산업공학 인식에 대한 온라인 커뮤니티 데이터 분석

한주희<sup>1</sup> · 유진이<sup>1</sup> · 이영훈<sup>2\*</sup>

<sup>1</sup>서울과학기술대학교 데이터사이언스학과 / <sup>2</sup>서울과학기술대학교 산업공학과

## Analysis of Online Community Data on Industrial Engineering

Juhee Han<sup>1</sup> · Jinyi Yu<sup>1</sup> · Younghoon Lee<sup>2</sup>

<sup>1</sup>Department of Data Science, Seoul National University of Science and Technology

<sup>2</sup>Department of Industrial Engineering, Seoul National University of Science and Technology

This study aims to analyze the perception of the public on industrial engineering with respect to the recent changes driven by the 4th Industrial Revolution. First, we collected posts and comments related to industrial engineering in online communities from 2019 to 2021. We found out keywords and areas of interest through keyword analysis based on the collected data and then clustered the main and sub groups through network-based clustering. Based on the main and sub groups, we studied sentiment polarity and trends during 2019 - 2021 through sentiment analysis and topic modeling. The results of this study can be used as a tool for setting the direction of industrial engineering education, such as curriculum composition and promotion strategies.

**Keywords:** Text Mining, Network Analysis, Clustering, Topic Modeling, Sentiment Analysis, Industrial Engineering

### 1. 서론

산업공학은 ‘공학·과학적 지식에 경영 기법을 접목해 다양한 산업계의 문제를 해결하는 학문’이며, 이러한 특징으로 인해 과거 공대 내에서 문과에 가까운 학문으로 인식되었다. 하지만, 4차 산업혁명으로 인해 전 산업에 AI, 빅데이터의 중요성이 대두되고, 산업 간 융합이 활발해졌다. 이에 따라, 다양한 기술과 산업의 특성을 파악하고 최적화된 형태로 문제를 해결하는 산업공학은 융합, 사회변화 속에서 트렌드를 예측하는 능력이 뛰어나다고 평가 받고 있다. 따라서, 전 산업 분야에서 산업공학과 수요가 증대됨에 따라 산업공학과에 대한 관심과 호감이 높아지고 있다. 4차 산업혁명과 함께 변화하고 있는 산업공학에 대한 인식 및 변화추이를 파악하여 산업공학 교육에 반영하려는 노력이 필요하다.

Kim *et al.*(2016)은 국내 논문 초록 데이터를 바탕으로 산업공학 연구 동향을 Cho *et al.*(2014)은 국내 산업공학 관련 저널들에 게재된 논문들의 주제어들과 초록 내용들을 바탕으로 산업공학의 주요 연구기법들의 동향을 분석하였다. 위와 같이 산업공학 관련 주제어 또는 데이터를 이용한 연구는 대부분

산업공학의 연구 동향과 관련된 연구들이며, 최근 4차산업혁명으로 인한 변화 속에서 실제 사람들의 산업공학에 대한 인식을 분석한 연구는 미비하다. 따라서, 본 논문은 산업공학과 학생들 이외에 타 전공 학생, 수험생을 포함한 모든 학생들 및 일반인들의 산업공학에 대한 인식을 분석하고자 산업공학과에 대한 의견을 알 수 있는 온라인 커뮤니티들의 데이터를 수집하고 분석을 진행하였다.

이후 본 논문의 구성은 다음과 같다. 본 연구의 제2장에서는 온라인 커뮤니티 기반 텍스트 마이닝에 대해 소개하고, 제3장은 키워드 분석, 네트워크 분석, 감성 분석 및 토픽모델링을 수행하기 위한 연구 방법을 기술한다. 제4장~제6장은 연구 방법을 활용한 분석 결과에 해당한다. 각각 제4장은 산업공학과 인식 키워드 분석, 제5장은 산업공학과 인식 네트워크 분석, 제6장은 산업공학과 감성 분석 및 토픽모델링 분석 결과를 설명한다. 마지막으로 제7장에서는 결론을 기술한다.

본 연구의 결과를 이용하면, 산업공학에 관한 관심 키워드 및 분야를 파악하여 커리큘럼 구성 및 홍보 전략 등 산업공학 교육의 방향성 설정에 도움이 되리라 기대한다.

이 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2020R1F1A1067914).

\* 연락저자 : 이영훈 교수, 01811, 서울시 노원구 공릉동 서울과학기술대학교 산업공학과, Tel : 02-970-6467, E-mail : yhoon.lee@seoultech.ac.kr  
2022년 5월 10일 접수; 2022년 5월 24일 수정본 접수; 2022년 5월 30일 게재 확정.

## 2. 온라인 커뮤니티 기반 텍스트 마이닝

온라인 커뮤니티는 다양한 분야 및 영역에서, 공통된 관심사 및 특징을 가진 사람들이 모여서 서로의 정보 및 콘텐츠를 교환하는 가상의 사회적 그룹이다(Park, 2013). 이에 따라 온라인 커뮤니티에서는 다양한 분야 및 영역 집단의 의견이 교환 및 공유된다. 또한, 온라인 커뮤니티 게시글은 특정 집단에 의해 의도된 내용이 아니라, 많은 사람이 자유롭게 작성하는 것이기 때문에 현실을 잘 반영한다는 장점이 있다(Hajli *et al.*, 2014). Park(2013)에 따르면 정보 기술의 발전과 소셜 네트워크의 확산으로 온라인 커뮤니티의 중요성과 영향력은 점점 커지고 있다. 이러한 온라인 커뮤니티의 특성과 더불어 국내 커뮤니티를 활용한 연구 사례는 다음과 같다.

먼저, Yang *et al.*(2014)은 트위터에서 수집한 SNS 데이터를 활용하여 연관 키워드 및 감성 분석을 수행하여 국내 대학에 대한 평판 분석, Lee *et al.*(2020)은 네이버 뉴스 기사 및 카페 글을 활용하여 빈도분석, 워드 클라우드, 잠재 디리클레 할당(LDA)을 활용한 토픽모델링, Ji *et al.*(2021)은 국내 간호사들이 이용하는 온라인 커뮤니티 게시글을 활용한 네트워크 분석과 토픽모델링을 각각 수행하였다. 특히, Ji *et al.*(2021)은 네트워크 분석을 통해 다빈도 키워드 분석을 진행하였고, 토픽모델링을 통해 주요 토픽을 분류하여 국내 임상간호사의 경험과 고충을 폭넓게 탐색하였다. 위와 같이 온라인 커뮤니티 글을 기반으로 키워드 분석, 네트워크 분석, 감성 분석 및 토픽모델링 등을 진행한 연구는 다수 존재한다.

Lee *et al.*(2020), Ji *et al.*(2021)에서 볼 수 있듯이 기존 연구들은 텍스트 데이터를 대상으로 클러스터링 시 토픽모델링 또는 네트워크 분석을 통한 클러스터링 중 하나의 방법만을 활용하였다. 하지만, 네트워크 분석을 통한 클러스터링 및 토픽모델링을 연관 지어 분석한 연구는 미비하다. 본 연구는 원하는 토

픽에 대한 보다 심층적인 분석을 위해 네트워크 클러스터링 결과를 토픽모델링에 활용하였다. 먼저, 분석에 필요한 주요 그룹에 포함된 키워드를 기반으로 데이터를 재추출하여 토픽 모델링을 진행하였다. 이러한 데이터 재추출 방법은 노이즈가 많은 온라인 커뮤니티 데이터 특성상 불필요한 데이터를 제거할 수 있다는 장점이 있다. 또한, 잠재 디리클레 할당(LDA)은 토픽 수에 따라 그 결과가 매우 다르게 나타나므로 토픽 개수의 근거에 대한 정당화 과정이 필요하다(Baek *et al.*, 2018; Kim *et al.*, 2017). 이에 따라 본 연구는 서브 클러스터링 결과 도출된 그룹의 수를 기반으로 최적의 토픽 개수를 설정하였다.

## 3. 연구 방법

### 3.1 연구 프레임워크

본 연구의 연구 프레임워크(figure)는 <Figure 1>과 같이 데이터 수집, 전처리, 키워드 분석, 네트워크 분석, 감성 분석 및 토픽모델링으로 구성된다.

### 3.2 데이터 수집 및 전처리

본 연구에서는 학생들의 의견을 반영할 수 있다고 판단된 여섯 개의 플랫폼을 선정하였다. 최근 산업공학의 전 분야가 데이터 기반 의사결정으로 진화하며 트렌드가 변화함을 확인하였다. 따라서, 본 연구는 산업공학과와 최근 트렌드 변화를 확인하기 위해 2019-2021년 동안 검색 키워드를 기반으로 한 게시글 제목, 본문 및 댓글을 수집하였으며, 총 데이터 개수는 8,186개이다. 분석에 활용한 최종 데이터 형식은 <Table 1>과 같으며, 수집한 데이터의 검색 키워드, 플랫폼 특성 및 수집한 총 데이터 개수는 <Table 2>에 나타내었다.

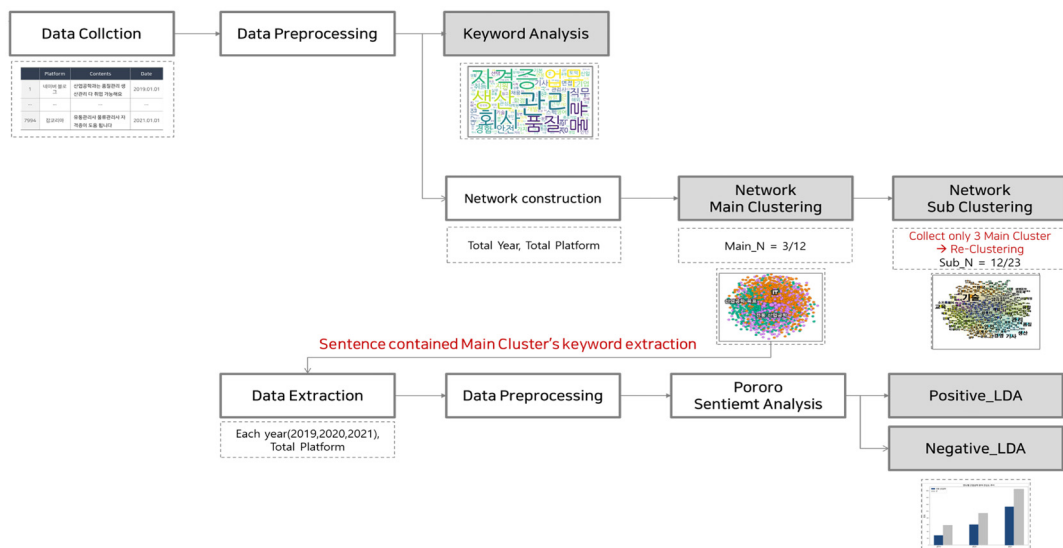


Figure 1. Study Framework

Table 1. Description of Dataset

	Platform	Contents	Date
1	Youtube	Students of the Department of Industrial Engineering can get a job in the field of quality management and production management	2019.01.01
...	...	...	...
8186	Jobkorea	The Distribution manager certificate helps you get a job.	2021.12.31

수집된 텍스트 데이터 전처리 과정은 다음과 같다. Konlpy 패키지의 Okt 형태소 분석기를 사용하였으며, 한국어를 제외한 특수 문자(?,!,@), 영어 등은 제거하였다. 감성 분석 및 토픽 모델링에서는 맞춤법 교정, 띄어쓰기를 추가로 진행하였다. 키워드 분석, 네트워크 분석에서는 ‘명사’만 추출하였고, 불용어 처리를 진행하였다. 불용어 처리 과정은 다음과 같다. 수집된 데이터 내 거의 모든 단어들과 동시출현을 보인 ‘산업’, ‘공학’은 특정 주제를 대표할 수 없어 불용어로 처리하였다. ‘대학교’, ‘학원’ 등 광고글에 포함된 주요 키워드 및 그 외 관련 없는 단어들도 추가 불용어로 처리하였다.

### 3.3 빈도, Text Rank 기반 키워드 분석

Text Rank란 Page Rank에 착안하여 한 문장 내에서의 동시출현 빈도를 통해 중요도가 높은 단어를 추출하는 알고리즘이다. 본 논문에서는, 빈도 및 Text Rank를 기반으로 중요 키워드를 추출하고 이를 Word Cloud로 시각화하였다. 이를 통해 산업공학의 주요 관심 분야를 찾아내고, 플랫폼 특성 및 연도별 핵심 키워드를 비교 분석하였다.

### 3.4 네트워크 구축 및 클러스터링

동시출현 네트워크는 특정 텍스트 단위에서 공동으로 출현한 단어의 집합적 상호 연결을 표현하는 방식이다. 이러한 동시출현 단어분석을 통해 특정 주제 영역이 독립적으로 다루고 있는 분야, 공유하고 있는 분야, 시기에 따른 변화 추이를 분석할 수 있다(Liu *et al.*, 2012). 데이터 전처리 과정을 통해 추출된 명사를 대상으로 동시출현 행렬을 생성하였다. 생성된 동시출

현 단어를 대상으로 Gephi 0.9.2 버전을 사용하여 네트워크 클러스터링 분석 및 시각화 분석을 진행하였다. 먼저, Gephi의 Modularity Class(node)를 활용해 그룹을 생성하였다. 생성된 그룹을 기반으로 네트워크 분석을 진행하였다. 네트워크 클러스터링은 총 두 번 진행하였다. 먼저, 전체 형성된 동시출현 행렬을 기반으로 클러스터링을 진행해 총 12개의 그룹으로 분류하였다. 그중 15% 이하에 해당하며 산업공학과 무관한 내용에 해당하는 그룹을 제외한 나머지 3개 클러스터를 대상으로 네트워크 메인 클러스터링 결과 분석을 진행하였다. 두 번째 클러스터링은 3개의 메인 그룹에 해당하는 키워드만을 대상으로 다시 23개의 그룹으로 분류하였다. 이 중 1% 이하, 광고성 키워드가 대다수인 그룹을 제외한 나머지 12개의 그룹을 대상으로 네트워크 서브 클러스터링 결과 분석을 진행하였다.

### 3.5 감성 분석 및 토픽모델링

‘산업공학 분야’에 대한 세부적인 분석을 위해 감성 분석과 토픽모델링을 진행하였다. 이를 위한 네트워크 클러스터링 결과 기반의 데이터 추출 작업은 다음과 같다. 먼저, 3개의 메인 그룹 중 ‘산업공학 분야’에 해당하는 2개 그룹에 대해서만 분석을 진행하였다. 이를 위해, 연도별로 2개 그룹에 해당하는 키워드들을 포함하는 문장만 추출하였다. 추출된 데이터를 활용한 감성 분석을 통해 산업공학 분야의 긍정/부정 결과의 연도별 추이를 분석하였다. 본 논문에서는, Pororo 감성 분석 모델을 활용하였다. Pororo란, 카카오브레인에서 개발한 요약, 감성 분석 등의 자연어 태스크를 간단하게 수행할 수 있는 라이브러리로, Pororo의 3가지 감성 분석 모델 중 네이버 영화 리뷰 데이터로 fine-tuning시킨 모델을 활용하였다(Heo *et al.*,

Table 2. Description of Platform for Collecting Data

Platform Name	Youtube	Naver Blog	Naver Kin	Naver Cafe (Dokchisa, Specup)	Orbi	Jobkorea	Dcincide
Search Keyword	Department of Industrial Engineering	Department of Industrial Engineering	Department of Industrial Engineering	Department of Industrial Engineering	Prospects and Pathways of Industrial Engineering	Department of Industrial Engineering	Department of Industrial Engineering
Platform Properties	Video Sharing Platform	Information and Opinion Sharing Platform	Information Sharing Platform	Information sharing and communication platform for job seekers	Information sharing and communication platform for examines	Information sharing and communication platform for job seekers	the largest community in the country
Number of data	891	1,779	1,090	212	2,233	435	1,758

2021). 연도별 ‘산업공학 분야’에 대한 긍정/부정 반응 분석을 위해 감성 분석 결과를 기반으로 토픽모델링을 수행하였다. 토픽모델링은 잠재 디리클레 할당(LDA)을 활용하였으며, 잠재 디리클레 할당(LDA)은 적정 토픽의 개수를 직접 결정해야 한다는 특징이 있다.

실험 내용은 다음과 같다. 감성 분석 결과를 바탕으로 연도별 긍정/부정을 총 6개의 카테고리로 분류하여 각각 잠재 디리클레 할당(LDA)을 진행하였다. 적정 토픽의 개수는 서브 클러스터링의 결과를 기반으로 결정하였다. 토픽모델링 결과 선정된 토픽의 주요 키워드 및 문장을 통해 ‘산업공학 분야’에 대한 긍정/부정 반응을 분석하였다.

#### 4. 키워드 분석 결과 및 해석

##### 4.1 수험생 및 취업준비생 키워드 분석

‘오르비’는 수험생들의 입시 관련 정보 공유 및 소통 커뮤니티이며, ‘잡코리아’ 및 ‘네이버 카페(독취사, 스펙업)’는 취업준비생들의 정보 공유 및 소통 커뮤니티이다. 이는 각각 수험생과 취업준비생을 대표하는 플랫폼에 해당한다. 따라서, 해당 플랫폼들의 키워드 분석 결과를 통해 산업공학에 대한 수험생과 취업준비생의 인식을 비교·분석하였다. 분석 결과는 <Figure 2>, <Figure 3>과 같다.

키워드 분석 결과 수험생들의 경우 입시와 관련된 키워드가 산업공학과 관련된 키워드보다 상대적으로 많이 등장하였다. 또한, ‘생산’, ‘품질’ 등의 ‘전통 산업공학’과 관련된 키워드에 비해 ‘IT’와 관련된 키워드의 등장 빈도가 높았다. 따라서, 수험생들의 경우 ‘전통 산업공학’보다는 ‘IT’에 대한 관심도가 더 높았다. 하지만, 취업준비생들의 경우 ‘전통 산업공학 분야’에 대한 관심도가 높았다.

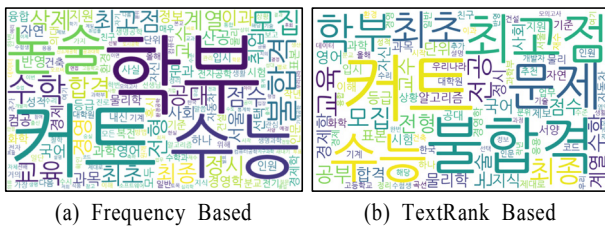


Figure 2. Orbi WordCloud

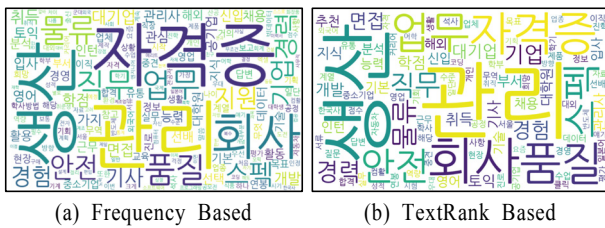


Figure 3. Naver-Cafe and JobKorea WordCloud

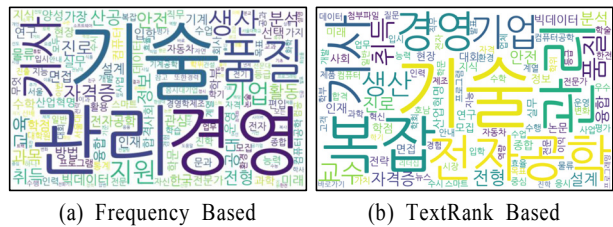


Figure 4. 2019 Word Cloud

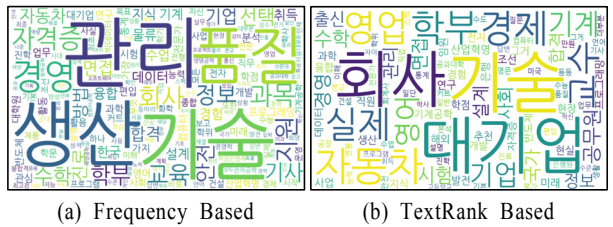


Figure 5. 2020 Word Cloud

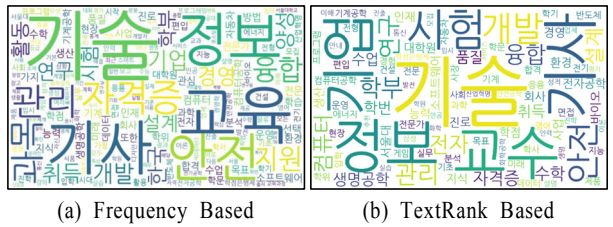


Figure 6. 2021 Word Cloud

##### 4.2 연도별 키워드 분석

연도별 산업공학에 관한 관심 분야 및 키워드 변화 분석을 위해 전체 플랫폼에 대해 연도별 키워드 분석을 진행하였으며 결과는 <Figure 4>, <Figure 5>, <Figure 6>과 같다.

키워드 분석 결과 3개년 모두 ‘전통 산업공학’, ‘IT’에 대한 관심이 지속되었다. 2019년과 2020년에 비해 2021년에는 ‘생산’, ‘품질’ 등 ‘전통 산업공학’과 관련된 키워드의 빈도가 낮아졌다. 또한, 3개년 모두 ‘IT’와 관련된 키워드가 꾸준히 등장하였으며, 2021년에는 ‘인공지능’과 같은 새로운 키워드가 등장하였다.

#### 5. 네트워크 분석을 통한 클러스터링 결과 및 해석

##### 5.1 메인 네트워크 구축 및 클러스터링을 통한 메인 그룹 도출

동시등장 행렬 기반 네트워크 분석은 산업공학의 관심 분야를 파악하는 데 유용하다. 2019년부터 2021년 동안의 네트워크는 419개의 키워드가 총 12,021번 관계를 맺는다. 메인 클러스터링 결과 도출된 그룹 중 최종 선정된 3개의 그룹은 <Figure 7>과 같으며, 각각 ‘IT’, ‘전통 산업공학’, ‘산업공학 목표’ 그룹으로 명명하였다. <Table 3>은 기본적인 노드와 링크의 수의 정보를 담은 네트워크의 통계적 분석 결과이다. 이외에도 가중치가 반영된 노드끼리의 연결 수, 그래프의 밀도 등의 옵션 정보와 그에 따른 값을 보여주고 있다.





Figure 7. Main Group Network

Table 3. Network Statistics Analysis Result

Option	Value
The Number of Keywords	419
The Number of Links	12,021
Weighted Degree	20.784 +
Density	0.137
Diameter	3
Average Path Distance	1.901
Size of Keywords	Pagerank ranking

‘IT’ 그룹에서 PageRank 값 기반 메인 키워드는 데이터, 프로그래밍, 연구, 컴퓨터, 개발, 반도체 등에 해당한다. 언급한 키워드 이외에도 생명공학, 화학공학, 반도체 등 타 학문도 산업공학의 IT 분야와 연계되어 있다는 것을 의미한다. ‘전통 산업공학’ 그룹에서 PageRank 값 기반 메인 키워드는 생산, 관리, 품질, 안전, 기사, 효율, 설계 등에 해당한다. 이와 같은 키워드는 전통 산업공학 분야에 해당하는 것이다. ‘산업공학 목표’ 그룹에서 PageRank 값 기반 메인 키워드는 기술, 산업혁명, 미래, 융합, 인재양성, 교육 등에 해당한다. 이는 키워드 간의 조합을 고려해볼 때, 산업공학이 추구하는 방향 및 목표의 내용을 담고 있다.

5.2 서브 네트워크 구축 및 클러스터링을 통한 서브 그룹 도출

도출된 세 가지 메인 그룹에 대해, 각각 세부적으로 토픽을 나누어 키워드를 분석하고자 서브 네트워크 클러스터링을 진행했다. 2019년부터 2021년 동안의 서브 네트워크는 335개의 키워드가 총 7,623번 관계를 맺는다. 서브 클러스터링 결과 총 12개로 분류한 그룹의 결과는 <Figure 8>이며, 그래프 분석의 수치적 결과는 <Table 4>이다. 12개 그룹의 명칭은 위의 <Table 5>와 같다. 각 그룹 내 가중치가 반영된 중심성이 높은 상위 10개의 키워드를 살펴보면, <Table 6>과 같다.

메인 그룹 내에서 가장 높은 비율을 차지하는 서브 그룹은, IT는 ‘SW 개발’, ‘전통 산업공학’은 ‘생산관리 및 품질’, ‘산업공학 목표’는 ‘기술 발전’에 해당한다. 첫 번째, ‘SW 개발’ 그룹의 가중치 중심성이 높은 키워드는 개발 > 연구 > 전문가 > 컴퓨터 > 소프트웨어 순이다. 키워드 간 조합을 보면, 컴퓨터



Figure 8. Sub-group Network

Table 4. Network Statistics Analysis Result

Option	Value
The Number of Keywords	335
The Number of Links	7,623
Weighted Degree	51.27+
Density	0.136
Diameter	3
Average Path Distance	1.908
Size of Keywords	Pagerank ranking

Table 5. Sub Groups

	2.	3.
1. IT	Traditional Industrial Engineering	Industrial Engineering Goals
SW Development	Production and quality Control	Technological Advancement
Data Linkage Studies	Efficiency Design	Economy and Society
Data Analysis	Smart Factory	Future Convergence
4th industrial revolution (Smart City)	Expense Savings	Human Resource Development Education

공학과 관련된 분야가 산업공학에서도 주요 분야로 등장하고 있다는 것을 의미한다. 두 번째, ‘생산관리 및 품질’ 그룹의 가중치 중심성이 높은 키워드는 관리 > 생산 > 안전 > 기사 > 품질 순이다. 키워드 간 조합을 보면, 생산관리, 제조 공정 과정에 대한 관심도가 높다는 것을 의미한다. 세 번째, ‘기술 발전’ 그룹의 가중치 중심성이 높은 키워드는 기술 > 한국 > 발전 > 핵심 > 국가 순이다. 키워드 간 조합을 보면, 한국의 핵심기술 발전을 위한 연구에 관심이 있는 것을 의미한다.

메인 그룹에 속하는 서브 그룹의 비율은 <Figure 9>와 같다. 첫 번째, ‘IT’ 그룹 내 서브 그룹의 높은 비율은 SW 개발 > 데이터 연계 학문 > 데이터 분석 > 4차산업혁명(스마트시티) 순이다. ‘IT’ 그룹은 타 학문과의 연계를 통한 개발 및 분석이 주요 키워드로 구성되었다.

두 번째, ‘전통 산업공학’ 그룹 내 서브 그룹의 높은 비율은

생산관리 및 품질 > 효율 설계 > 스마트팩토리 > 비용 절감 순이다. ‘전통 산업공학’ 그룹은 생산 및 품질관리에 효율적인 프로세스 설계를 통한 비용 절감을 목표로 하는 주요 키워드로 구성되었다. 세 번째 ‘산업공학 목표’ 그룹 내 서브 그룹의 높은 비율은 기술 발전 > 경제·사회 > 미래 융합 > 인재양성 교육이다. ‘산업공학 목표’ 그룹에서는 4차산업혁명, 융합, 혁신 등을 목표로 인재양성 교육을 통한 기술, 사회의 발전 추구를 목표로 하는 주요 키워드로 구성되었다.

메인, 서브 클러스터의 결과를 종합했을 때, 3개의 메인 그룹 중에서 ‘IT’, ‘전통 산업공학’ 그룹은 데이터, 개발, 분석, 생산 및 품질관리 등 산업공학의 교육 및 진출 분야에 해당한다. 따라서, 본 연구에서는 두 그룹을 ‘산업공학 분야’로 새롭게 설정하였으며, 감성 분석과 토픽모델링을 활용해 ‘산업공학 분야’에 대한 구체적인 의견에 대해 분석하였다.

Table 6. Top 5 Keywords in Sub Group

Subgroup Name	Top 5 Keywords with Weight Centrality in Subgroup
SW Development	Development > Research > Experts > Computers > Software
Data Linkage Studies	Production Management and Quality > Efficiency Design > Smart Factory > Reduce Expense
Data Analysis	Technological advancement > Economy, Society > Future convergence > Human Resource Development Education
4th industrial revolution (Smart City)	Environment > Energy > Work > Resources > Power
Production and quality Control	Management > Production > Safety > Articles > Quality
Efficiency Design	Design > Operations > Efficiency > Business > Framework
Smart Factory	Smart > Factory > Manufacturing > Drone > Factory
Expense Savings	Occurring > Costs > Cause > Burden > Price
Technological Advancement	Technology > Korea > Development > Core > Country
Economy and Society	Society > Us > Economy > Life > Humans
Future Convergence	Convergence > Future > Times > Bio > Change
Human Resource Development Education	Development > Research > Experts > Computers > Software

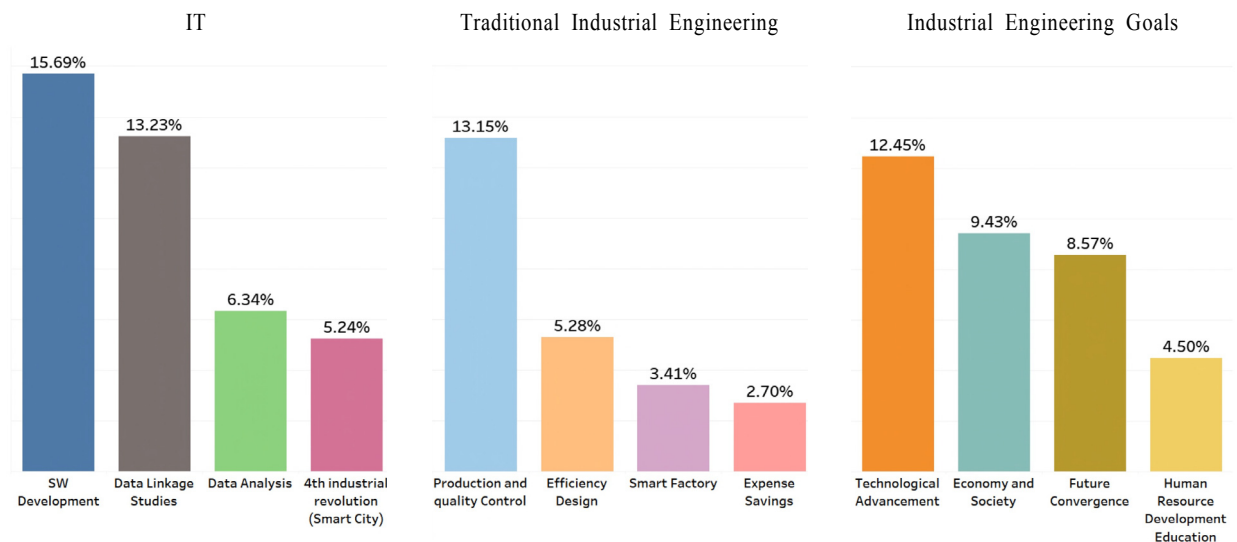


Figure 9. Sub Group Ratio in Main Group

## 6. 감성 분석 및 토픽모델링 결과 및 해석

### 6.1 감성 분석 결과

감성 분석에 활용한 연도별 데이터의 총수는 2021, 2020, 2019 순으로 많았으며, 분석 결과는 <Figure 10>과 같다. 3개년 모두 부정보다 긍정의 빈도가 높았으며, 2021년의 경우, 부정에 비해 긍정 반응의 빈도가 2배 이상 높았다. 이는, 과거에 비해 2021년에 대해 ‘산업공학 분야’에 대한 긍정적인 인식이 증가했음을 의미한다.

### 6.2 토픽모델링 결과

연도별 긍부정 데이터에 대해 토픽모델링을 진행하였으며, 모델링 결과를 바탕으로 토픽별 주요 키워드와 문장을 분석하여 ‘산업공학 분야’에 해당하는 주요 토픽을 선정하였다. 전체 토픽 중 해당 ‘산업공학 분야’가 차지하는 비율을 통해 파악한 연도별 ‘산업공학 분야’에 대한 긍부정 반응의 변화율은 <Figure 11>과 같다. ‘전통 산업공학’에 대한 긍정/부정 반응은 3년간 거의 비슷하게 나타났다. 이에 반해, ‘IT’의 경우 3개년 동안 긍정 반응이 5배 이상 증가했다.

<Figure 12>는 연도별 긍부정 반응 빈도의 총합을 통해 파악

한 연도별 ‘산업공학 분야’에 대한 관심도 추이이다. 3개년 모두 ‘전통 산업공학’에 비해 ‘IT’의 관심도가 높은 것을 확인했다. 또한, 앞서 키워드 분석 결과를 통해 2021년에 ‘인공지능’과 같은 새로운 키워드가 등장하는 것을 확인했다. 따라서, 꾸준히 산업공학의 ‘IT’에 대한 관심도는 높으며, 매년 이슈로 떠오르는 키워드는 다르다고 할 수 있다.

토픽모델링 결과 ‘산업공학 분야’에 해당하는 주요 토픽의 주제는 <Table 7>과 같다. 토픽의 주제는 토픽별 주요 키워드 및 문장 분석을 통해 선정하였으며, 연도 및 긍부정 반응에 따라 총 12개의 카테고리로 분류하였다.

먼저, ‘산업공학의 진출 분야’와 관련된 주요 토픽 내용은 다음과 같다. 2019년, 2020년에는 ‘전통 산업공학’이 ‘산업공학의 주요 진출 분야’ 인식이 뚜렷했지만, 2021년에는 ‘IT’가 ‘산업공학의 주요 진출 분야’라는 인식이 등장하여 변화를 보였다. 이를 통해, 최근 3년 동안 ‘IT’가 산업공학의 주요 분야로 떠오르고 있음을 파악했다.

‘산업공학의 취업’과 관련된 주요 토픽 내용으로는, ‘전통 산업공학’의 경우 ‘취업이 잘 안 되는 분야’, ‘취업 시 관련 자격증의 중요성’이 3개년 모두 부정 반응으로 등장했다. 반면, ‘IT’의 경우 ‘석사 이상의 학력 필요’에 대한 인식이 3개년 연속 등장하여 IT 관련 직업의 대부분은 석사 이상의 전문지식을 갖춘 인재를 요구하고 있다는 것을 알 수 있다. 이처럼 ‘전통

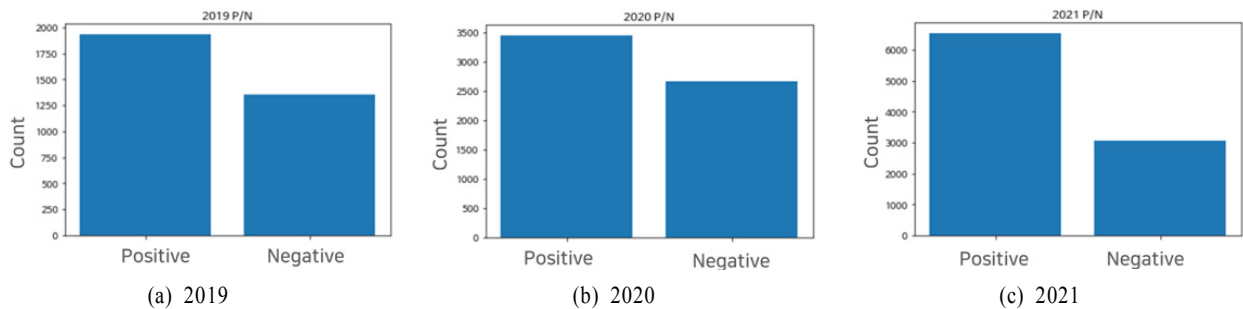


Figure 10. Result of sentiment analysis

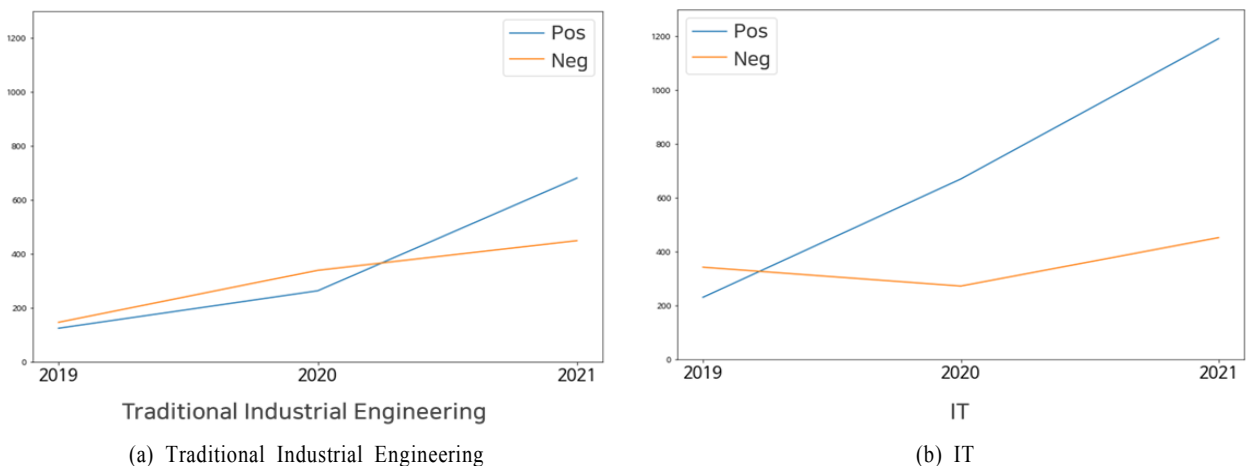


Figure 11. Change of Sentiment Rate of the Industrial Engineering Areas

산업공학'과 'IT' 분야는 취업을 위한 역량에서 관점의 차이를 보이고 있다. '산업공학 학문의 특성'과 관련된 주요 토픽 내용은

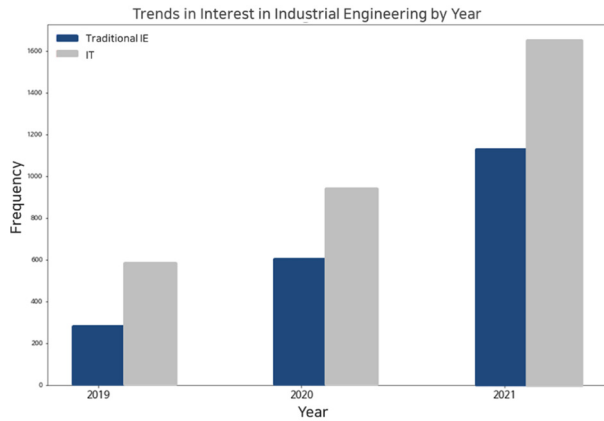


Figure 12. Change of rate of the Industrial Engineering areas

다음과 같다. 최근 산업공학의 'IT 분야'는 경영학 뿐만 아니라 기계, 화학 등 다양한 분야와 연계될 수 있는 융합 학문으로 자리 잡고 있으며, '데이터 사이언스, 인공지능 학과와의 연관성'도 높아지고 있다. 또한, 여러 분야를 배우는 산업공학의 특성에 대해 'IT'의 경우 '전문성이 떨어지는 분야'라는 부정적인 인식으로 전문성을 주장할 수 없지만, 동시에 '프로그래밍 능력의 중요성', '데이터 분석의 중요성'도 함께 등장한 것으로 보아 IT 관련 역량에 대한 전문성을 중요하게 인식하고 있음을 알 수 있다. 하지만, '전통 산업공학'의 경우 '다양한 분야 학습 가능'하다는 긍정적인 인식으로 등장하였으며, 2020년, 2021년에는 '학문의 추상성'이라는 인식이 등장하였다. 이는 산업공학 진학을 희망하는 학생들이 산업공학과의 진출 분야 및 학습 과목에 대한 인식이 부족하다고 할 수 있다.

마지막으로, <Table 8>을 통해 최근 3년간 'IT 분야'에 대한 관점 변화를 살펴볼 수 있다. 이를 통해, 산업공학의 'IT 분야'가

Table 7. Main Topics of Industrial Engineering Areas (2019-2021)

		Traditional Industrial Engineering		IT	
		Positive	Negative	Positive	Negative
2019	Main Topic 1	The Relationship between Management and Industrial Engineering	A relevant qualification is required for employment	Data, Artificial Intelligence and Industrial Engineering Association	A master's degree or higher is required
	Main Topic 2	Recognition that production and quality are major areas of industrial engineering	the perception that employment is a poor area	A wide range of learning areas	A less professional field
2020	Main Topic 1	Recognition that production and quality are major areas of industrial engineering	A relevant qualification is required for employment	Importance of programming literacy	A master's degree or higher is required
	Main Topic 2	-	The perception that employment is a poor area	Importance of Data Analytics	The Relationship between Data Science and Industrial Engineering
	Main Topic 3	-	The abstraction of learning	Recognize as a convergence major	-
2021	Main Topic 1	A wide range of learning areas	The abstraction of learning	A master's degree or higher is required	A master's degree or higher is required
	Main Topic 2	Synergistic effects of related certificates exist in employment	the perception that employment is a poor area	Recognition of the Major Field of Industrial Engineering	A less professional field

Table 8. Changes of perception of the IT areas (2019-2021)

Year	Perspective	Contents
2019	Increasing interest in IT	The recognition that there is a connection between data and industrial engineering, and the recognition that it is a 'diverse field that can be learned', have increased interest and interest in industrial engineering
2020	Increasing the concrete and necessary of IT	The importance of 'programming literacy', 'data analysis', and 'recognition as convergence disciplines' are many opinions, and the need for learning IT and other related disciplines is increasing
2021	Increasing the importance of IT	The importance of the IT field in industrial engineering has increased with the emergence of 'Master's or higher education required' and 'the recognition that it is a major field of industrial engineering'



시간의 흐름에 따라 빠르게 변화하고, 산업공학에서 차지하는 비율이 높아지고 있음을 파악하였다.

## 7. 결론

본 논문은 산업공학에 대한 온라인 커뮤니티의 반응을 기반으로 키워드 분석, 네트워크 분석, 감성 분석 및 토픽모델링을 진행하였다. 산업공학과의 최근 트렌드 변화를 확인하기 위해 2019 - 2021년의 온라인 커뮤니티 데이터를 활용한 분석을 진행하였다. 연구 목표는 산업공학에 대한 학생들 및 일반인들의 실제 관심 분야 및 인식 파악을 통한 산업공학 교육의 방향성을 설정하는 것이다.

연구 주요 결과는 다음과 같다. 첫째, 2019년과 2020년에는 전통 산업공학 분야가 산업공학의 주요 진출 분야라는 인식이 존재했지만, 2021년에는 IT분야가 주요 진출 분야라는 인식으로 등장하였다. 이처럼 전통 산업공학 분야와 더불어 IT분야가 산업공학의 새로운 진출 분야로 떠오르고 있음을 시사한다. 둘째, 최근 산업공학 IT분야는 경영학뿐만 아니라 기계, 화학 등 다양한 분야와 연계될 수 있는 융합 학문으로 자리 잡았다. 이는, 산업공학의 IT분야가 시대의 흐름에 적응하여 신속하게 변화하여 적합한 자리매김을 하고 있음을 의미한다. 셋째, 산업공학 진학을 희망하는 수험생들의 경우 산업공학과의 진출 분야 및 학습 과목에 대한 정보가 부족함을 확인하였다. 이를 통해, 산업공학과 진학 시 학습 과목 및 구체적인 커리큘럼을 홍보하는 것이 필요함을 확인하였다. 본 연구는, 3개년의 데이터를 대상으로 분석을 진행하였으나, 데이터 수집 기간을 늘린다면 더 장기적인 변화의 추이를 살펴볼 수 있을 것이다. 또한, 데이터 전처리 시 광고 데이터가 완벽히 제거된다면, 더욱 정확한 분석 결과를 도출할 수 있을 것이다. 하지만, 본 연구의 분석 결과를 산업공학과 커리큘럼 구성 및 홍보 전략 등 산업공학 교육의 방향성 설정에 활용하여 산업공학의 위상 강화에 이바지할 수 있을 것이다.

## 참고문헌

- Baek, Y. (2018), *Text-Mining using R*, Hanul, Seoul.
- Cho, G. H., Lim, S. Y., and Hur, S. (2014), An Analysis of the Research Methodologies and Techniques in the Industrial Engineering Using Text Mining, *Journal of the Korean Institute of Industrial Engineers*, **40**(1), 52-59.
- Hajli, M. N., Sims, J., Featherman, M., and Love, P. E. (2015), Credibility of Information in Online Communities, *Journal of Strategic Marketing*, **23**(3), 238-253.

- Heo, H., Ko, H., Kim, S., Han, G., Park, J., and Park, G. (2021), PORORO: Platform Of neuRal mOdels for natuRal language prOcessing, Available from: <https://github.com/kakaobrain/pororo>.
- Ji, H., Lim, A., and Lee, S. E. (2021), Exploring Nurses' Experience and Grievance: Network Analysis and Topic Modeling Using a Social Networking Service, *Journal of Korean Academy of Nursing Administration - KoreaMed*, **27**(3), 169-180.
- Kim, N., Lee, D., Choi, H., and Wong, W. X. S. (2017), Investigations on techniques and applications of text analytics, *The Journal of Korean Institute of Communications and Information Sciences*, **42**(2), 471-492.
- Kim, S. and Jang, S. Y. (2016), A Study on the Research Trends in Domestic Industrial and Management Engineering Using Topic Modeling, *Journal of the Korea Management Engineers Society*, **21**(3), 71-95.
- Lee, S. M., Ryu, S. E., and Ahn, S. (2020), Mass Media and Social Media Agenda Analysis Using Text Mining: Focused on '5-day Rotation Mask Distribution System', *Journal of the Korea Contents Association*, **20**(6), 460-469.
- Liu, G. Y., Hu, J. M., and Wang, H. L. (2012), A Co-word Analysis of Digital Library Field in China, *Scientometrics*, **91**(1), 203-217.
- Park, D. H. (2014), The Effect of Online Community, Members, and Personal Characteristics on Lurking Behavior: Why do people only consume rather than create contents?, *Journal of Internet Computing and Services*, **15**(1), 73-88.
- Yang, M. H., Jung, I. S., Kim, Y. T., and Cho, W. S. (2014), An Awareness Identification and Preference Analysis for Domestic University Using SNS Data, *Journal of The Korea Big Data Service Society*, **1**(1), 1-13.

## 저자소개

**한주희**: 서울과학기술대학교 산업공학과에서 2020년 학사학위를 취득하고 일반대학원에서 데이터사이언스학과 석사과정에 재학 중이다. 연구분야는 데이터 마이닝, 인공 신경망이다.

**유진이**: 서울과학기술대학교 산업공학과에서 2021년 학사학위를 취득하고 일반대학원에서 데이터사이언스학과 석사과정에 재학 중이다. 연구분야는 자연어처리, 인공신경망, 텍스트마이닝, 지식 그래프 임베딩이다.

**이영훈**: 서울대학교 산업공학과에서 2007년 학사, 2009년 석사, 2019년 박사학위를 취득하였다. 현대자동차 빅데이터실 책임연구원, LG전자 UX연구소/선행디자인연구소 선임연구원을 역임하고, 2019년부터 서울과학기술대학교 산업공학과 교수로 재직하고 있다. 연구분야는 데이터마이닝, 사용자경험(UX) 디자인이다.