

# 산업공학에서의 데이터과학: 과거와 현재, 그리고 미래

강석호<sup>1</sup> · 강필성<sup>2</sup> · 조성준<sup>2\*</sup>

<sup>1</sup>성균관대학교 산업공학과 / <sup>2</sup>서울대학교 산업공학과

## Data Science in Industrial Engineering: Past, Present, and Future

Seokho Kang<sup>1</sup> · Pilsung Kang<sup>2</sup> · Sungzoon Cho<sup>2</sup>

<sup>1</sup>Department of Industrial Engineering, Sungkyunkwan University

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

Data science has become increasingly important in modern industrial engineering, serving as a versatile tool for addressing real-world industrial challenges. In this paper, we analyze the past, present, and future of data science in industrial engineering. To understand the past and present, we examine data science-related papers published in the Journal of the Korean Institute of Industrial Engineers. To outlook the future, we conduct a survey among industrial engineers who specialize in data science. We discuss the results and provide insights into future prospects.

**Keywords:** Data Science, Data Mining, Big Data, Artificial Intelligence, Machine Learning, Deep Learning

### 1. 서론

데이터과학(Data Science)은 데이터 속에 숨겨진 패턴이나 의미를 찾아내어 가치를 창출하는 것을 목표로 하는 다학제적인 융합 학문 분야이다(Cao, 2017). 일반적으로 수학, 통계학, 컴퓨터과학, 정보공학, 인공지능, 데이터마이닝, 데이터베이스 등 데이터와 관련된 모든 분야를 아우르는 포괄적 용어로 사용된다. 최근, 다양한 데이터 소스로부터 수집되는 데이터의 양이 급증함에 따라 데이터과학은 제조, 금융, 에너지, 보안 모든 산업에서 현실 문제를 해결하기 위한 중요한 도구로서 그 활용 범위도 점차 넓어지고 있다(Wu *et al.*, 2014).

산업공학은 공학적 방법론을 활용하여 산업 시스템의 설계, 운영, 개선을 다루는 학문으로 설명된다(Bailey *et al.*, 2005; Biles, 1991). 전통적으로는 확률, 통계, 최적화 등을 공학적 방법론으로 많이 활용해 왔으며, 데이터과학 또한 산업공학에서 중요한 방법론으로 자리매김하고 있다(Cho *et al.*, 2014; Jeong and Lee, 2016; Cho and Kim, 2012). 데이터과학에 대한 높은 산업의 수요와 함께, 현대의 산업공학에서 데이터과학의 중요성

과 활용 빈도는 점차 높아지고 있다. 국내의 대부분의 산업공학과 교육과정에서 데이터과학 관련 교과목 개설이 늘어나고 있으며, 데이터과학을 연구 주제로 다루는 연구실의 비중도 높아지고 있다. 데이터과학의 지속적인 발전은 산업공학에 있어 새로운 기회를 제공해 줄 것으로 기대된다.

본 논문에서는 대한산업공학회 창립 50주년을 맞아 산업공학에서의 데이터과학의 과거와 현재, 그리고 미래를 분석하였다. 산업공학에서의 데이터과학의 과거와 현재를 살펴보기 위해, 대한산업공학회지에서 1975년부터 현재까지 출판된 논문들에 대한 메타분석을 통해 데이터과학의 비중과 트렌드를 확인하였다. 산업공학에서의 데이터과학의 미래를 전망하기 위해, 각 계에서 데이터과학 분야에 활발히 활동하고 있는 산업공학 전공자들을 대상으로 설문한 결과를 정리하여 제시한다.

### 2. 대한산업공학회지 메타분석으로 보는 과거와 현재

산업공학에서의 데이터과학의 과거와 현재를 보기 위해 대한

\* 연락처: 조성준 교수, 08826 서울시 관악구 관악로 1 서울대학교 공과대학 산업공학과, Tel : 02-880-6275, Fax : 02-889-8560, E-mail : zoon@snu.ac.kr

2024년 8월 22일 접수; 2024년 9월 25일 게재 확정.

산업공학회지에서 출간된 관련 논문에 대한 메타분석을 수행하였다. 대한산업공학회지(Journal of the Korean Institute of Industrial Engineers)는 대한산업공학회에서 발행하는 학술지로 1975년에 창간되어 현재까지 산업공학의 이론 및 응용에 관한 다수의 연구 논문이 발표되고 있다. 1976년 6월부터 2024년 6월까지 출판된 총 1,856편의 논문들 중 데이터과학에 관련된 논문의 비중을 확인하고, 시간의 흐름에 따른 연구 동향 및 주요 키워드를 분석하였다.

<Table 1>은 기간 별 대한산업공학회지에 출판된 논문 수 및 데이터과학에 관련된 논문 비율의 추이를 보여준다. 초기에는 5%에 미치지 못했던 데이터과학 관련 논문의 비율이 시간이 지남에 따라 점차 증가하며, 특히 2010년도 이후 빠른 증가 속도를 보여주었다. 특히 2021년부터 현재까지 출판된 논문들 중에서는 절반이 넘는 61.1%가 데이터과학 관련 논문인 것으로 확인되었다. 대한산업공학회지에 출판된 논문들은 주로 데이터과학 이론 연구보다는 제조 등 다양한 산업분야의 현실 문제 해결을 위해 데이터과학을 도구로 사용하는 응용 논문들이 주로 출판되었다.

**Table 1.** Trends of Data Science-Related Publications in the Journal of the Korean Institute of Industrial Engineers

Period	No. Published Papers	No. Papers Related to Data Science	Proportion. Papers Related to Data Science
1975 ~ 1990	298	12	4.00%
1991 ~ 2000	451	47	10.40%
2001 ~ 2005	207	26	12.60%
2006 ~ 2010	195	28	14.40%
2011 ~ 2015	275	65	23.60%
2016 ~ 2020	255	108	42.40%
2021 ~ present	175	107	61.10%
Total	1,856	393	21.20%

<Table 2>는 기간별 데이터과학 관련 논문에서 다루어진 주요 키워드를 나타낸 것이며, <Figure 1>은 전체 기간에 대한 데이터과학 관련 논문 키워드들의 워드 클라우드를 나타내고 있다. 전체 기간 관점에서는 방법론 측면에서 머신러닝(Machine Learning), 데이터마이닝(Data Mining), 딥러닝(Deep Learning), 군집화(Clustering)와 같은 일반적인 키워드 뿐만 아니라, 유전알고리즘(Genetic Algorithm), 토픽모델링(Topic Modeling), 네트워크분석(Network Analysis) 등과 같은 구체적인 분석 방법론들이 자주 사용된 것을 알 수 있다. 이에 더하여 텍스트마이닝(Text Mining), 감성분석(Sentiment Analysis), 자연어처리(Natural Language Processing)와 같은 자연어 및 텍스트 데이터를 분석하는 연구가 많이 수행되었다는 것 또한 확인할 수 있다. 데이터과학 방법론이 적용된 응용 분야/산업 관점으로는 반도체 제조업과 특히 분석이 가장 높은 비중을 차지하고 있

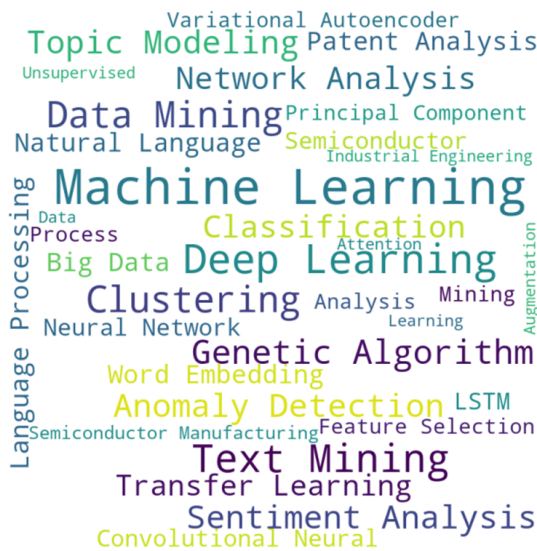
는 것으로 나타났다. 주요 키워드를 통해 파악할 수 있는 기간 별 트렌드는 다음과 같다.

- **[1975~1990: 기초 구축기]** 이 기간 동안 연구는 주로 기초 알고리즘과 최적화 기법에 중점을 두고 있었던 것으로 보이며 유전알고리즘(Genetic Algorithm), 입자군집최적화(Particle Swarm Optimization), 군집화(Clustering) 등의 방법론이 주로 연구되었다.
- **[1991~2000: 비지도 학습 및 자연어 처리의 출현]** 1990년대는 비지도학습(Unsupervised Learning), 링크예측(Link Prediction), 자연어처리(Natural Language Processing)와 같은 다양한 응용 분야에 데이터과학이 적용되기 시작하였다. 1990년대 이전까지의 기초 방법론 연구에서 데이터 기반 문제 해결을 중점적으로 다루는 전환점이 되는 시기로 볼 수 있다.
- **[2006~2010: 네트워크 분석 및 공공 데이터로의 다각화]** 2006년부터 2010년까지는 머신러닝(Machine Learning)을 비롯하여 텍스트마이닝(Text Mining), 네트워크분석(Network Analysis)과 같이 연구 주제가 다양해진 시기이다. 이에 더하여 기술융합(Technology Convergence) 및 공공데이터(Public Data)로 대변되는 관심사의 다각화가 두드러지며, 이는 다양한 데이터 소스의 가용성이 증가함에 따라 복잡한 네트워크 분석과 공공 데이터 활용에 대한 관심이 증가했기 때문으로 볼 수 있다.
- **[2011~2015: 딥러닝 및 합성곱신경망의 등장]** 이 기간 동안 머신러닝(Machine Learning), 유전알고리즘(Genetic Algorithm), 딥러닝(Deep Learning) 키워드의 빈도가 크게 증가하였다. 합성곱신경망(Convolutional Neural Network; CNN)과 추천시스템(Recommendation Systems)도 주요 키워드로 등장하였으며, AlexNet의 등장 이후 폭발적으로 증가한 CNN 및 딥러닝에 대한 관심이 본 지에도 반영된 것으로 볼 수 있다.
- **[2016~2020: 스마트 팩토리 관련 연구 증가]** 이 시기에는 유전알고리즘(Genetic Algorithm), 머신러닝(Machine Learning), 이상탐지(Anomaly Detection) 등과 같은 방법론이 스마트팩토리(Smart Factory), 품질예측(Quality Prediction) 등, 스마트팩토리와 관련된 분야에 많이 적용된 것을 볼 수 있다. 실제로 이 시기는 4차 산업혁명에 대한 관심이 매우 높았던 시기이며, 산업공학 측면에서 4차 산업혁명의 주요 분야인 스마트팩토리에 대한 연구가 활발히 수행되었다.
- **[2021~현재: 딥러닝의 확장 및 데이터과학의 범용성 증가]** 이 시기는 딥러닝(Deep Learning), 합성곱신경망(CNN)으로 대변되는 딥러닝 방법론 뿐만 아니라 전이학습(Transfer Learning)과 같은 방법론의 확장이 주요 키워드로 나타났다. 이에 더하여 지속적으로 자주 등장한 머신러닝(Machine Learning)과 유전알고리즘(Genetic Algorithm), 데이터마이닝(Data Mining), 텍스트마이닝(Text Mining)과 같은 키워

**Table 2.** Trends of Major Keywords in Data Science-Related Publications in the Journal of the Korean Institute of Industrial Engineers

Period	Major Keywords (Frequency)
Overall	Machine Learning (45), Genetic Algorithm (33), Neural Network (24), Deep Learning (24), Data Mining (23), Text Mining (20), Anomaly Detection (19), Network Analysis (15), Clustering (14), Time Series (13)
1975 ~ 1990	Genetic Algorithm (2), Particle Swarm Optimization (2), Clustering Problem (2)
1991 ~ 2000	Unsupervised Learning (4), Link Prediction (4), Natural Language Processing (4), Text Mining (3), Anomaly Detection (3), Time Series (3), Domain Adaptation (3)
2001 ~ 2005	Reinforcement Learning (5), Data Analysis (3)
2006 ~ 2010	Machine Learning (6), Text Mining (5), Network Analysis (4), Genetic Algorithm (3), Anomaly Detection (3), Technology Convergence (3), Public Data (3), Social Network Analysis (3)
2011 ~ 2015	Machine Learning (18), Genetic Algorithm (8), Deep Learning (8), Data Mining (8), Feature Selection (4), Convolutional Neural Network (3), Recommendation System (3)
2016 ~ 2020	Genetic Algorithm (12), Machine Learning (11), Anomaly Detection (8), Smart Factory (7), Sentiment Analysis (6), Quality Prediction (6)
2021 ~ present	Deep Learning (10), Convolutional Neural Network (8), Transfer Learning (8) Genetic Algorithm (7), Data Mining (7), Text Mining (7)

드의 절대적인 빈도가 증가함으로써 데이터과학이 산업공학의 다양한 문제를 해결하는 핵심 도구로 활용되고 있다는 점을 확인하였다.



**Figure 1.** Word Cloud of Keywords from Data Science-Related Publications in the Journal of the Korean Institute of Industrial Engineers

<Table 3>은 논문의 제목과 저자 키워드를 이용하여 토픽모델링을 수행한 결과이다. 토픽모델링 방법으로는 잠재 디리클레 할당(Latent Dirichlet Allocation; LDA)(Blei *et al.*, 2003)을 사용하였고, 토픽의 수는 10, 15, 20개의 후보 토픽 중에서 토픽별 고빈도 단어들을 확인 후 정성적으로 10개로 선정하였다. 토픽모델링 결과 산업공학회지에 게재된 논문의 주요 토픽으로는 시스템 최적화(System Optimization), 의사결정지원

시스템(Decision Support System)과 같은 전통적인 산업공학의 주요 분야가 추출되었으며, 방법론 측면에서는 머신러닝 알고리즘(Machine Learning Algorithms), 학습 전략 및 이미지 분석(Learning Strategy & Image Analytics), 텍스트 마이닝 및 네트워크 분석(Text Mining & Network Analysis), 토픽모델링(Topic Modeling) 등이 추출되었다. 또한, 주요 응용 분야로는 제조(Manufacturing) 및 물류(Logistics)와 같은 전통적인 산업공학의 활용 분야가 활발히 연구되는 것으로 나타났으며, 특히 주목할만한 점은 제조 중에서도 국내의 큰 비중을 차지하는 반도체 관련 연구들이 별도의 토픽으로 추출될 만큼 활발하게 연구가 되었다는 점이다. 이 외에도 최근의 트렌드를 반영한 스마트 팩토리(Smart Factory)도 주요 연구 토픽으로 추출되었다. 산업공학에서 활발하게 연구되고 있는 인간공학 및 인간-컴퓨터 상호 작용(Human-Computer Interaction; HCI) 분야에도 데이터과학 기법을 접목한 연구가 활발하게 이루어지고 있으나, 관련 논문들이 해당 분야에 특화된 학술지에 주로 게재되고 있어 대한산업공학회지 메타분석 과정에는 잘 드러나지 않은 것으로 보인다.

<Figure 2>는 각 토픽의 기간별 비중 추이를 나타낸 것이다. 절대적인 비중 관점에서는 학습 전략 및 이미지 분석에 관한 연구가 가장 높은 비중을 차지하며 꾸준히 연구되었고, 그 뒤를 이어 반도체 응용과 텍스트 마이닝 및 네트워크 분석이 높은 비중을 차지하고 있다. 열 가지의 토픽 중에서 기간에 따라 유의미하게 증가하는 토픽은 머신러닝 기법, 토픽모델링, 물류 응용으로 나타났다. 토픽모델링의 경우 토픽모델링의 방법론을 개발한다기보다는 본고의 목적과 유사하게 토픽모델링을 활용하여 특허 분석이나 논문 분석 등 대량의 문서 데이터에 대한 분석 도구로써 사용이 된 것으로 보이며, 이를 활용하여 유망기술 분석, 고객 반응 분석, 특정 분야의 핵심 연구 동향 분석 등을 수행한 연구들이 증가했다고 볼 수 있다.

**Table 3.** Topic Modeling Results Using Paper Titles and Keywords

Topic	Frequent Words
System Optimization	optimization, analysis, time, system, technique, model, method, selection, machine, study, mining, tree, decision, problem, classification, prediction, feature, multivariate, learning, clustering, application, run, text, classification, touch, information, design, series, process
Decision Support System	model, learning, technique, prediction, network, research, detection, data, decision, machine, development, system, analysis, process, text, algorithm, mining, self, data mining, study, method, case, clustering, price, customer, industry, trend, problem, item, programming
Machine Learning Algorithms	model, data, prediction, machine learning, analysis, learning, process, intelligence, quality, development, deep learning, machine, classification, detection, text, technique, method, business, manufacturing, anomaly detection, weight, problem, time, system, text mining, data, correlation, case, technology, factor
Application: Semiconductor	technique, method, data, pattern, network, wafer, detection, system, extraction, learning, structure, Semiconductor, monitoring, diagnosis, search, model, vehicle, feature extraction, one-class classification, feature, estimation, fault detection, classifier, development, information, semiconductor, recommendation, study, defect, self
Learning Strategies & Image Analytics	time, datum, classification, research, system, model, analysis, genetic algorithm, performance, technique, methodology, processing, Transfer Learning, learning, network, study, image, quality, production, prediction, site, determination, modeling, wave, change, case, value, content, defect, convolutional neural network
Text Mining & Network Analysis	analysis, network, learning, model, system, detection, prediction, classification, sentiment analysis, mining, process, study, word embedding, market, word, index, case, development, data, machine learning, classification, opportunity, clustering, stock work, reinforcement learning, property, component, calculation
Applications: Manufacturing	model, analysis, detection, sequence, line, input, study, process, selection, datum, learning, number, welding, diffusion, data, assembly, behavior, shipping, network, recognition, end, Machine Learning, size, process mining, trace clustering, correlation, variable, clustering, transfer learning, performance
Topic Modeling	analysis, model, datum, learning, technology, prediction, research, product, detection, patent, topic modeling, method, network, machine learning, deep learning, review, power, level, factor, generation, study, process, estimation, clustering, life, sentiment analysis, problem, anomaly detection, classification
Application: Smart Factory	analysis, data, method, model, prediction, development, classification, traffic, similarity, estimation, variable, semiconductor, service, measurement, test, evaluation, learning, process, framework, data clustering, manufacturing, factory wafer, smart factory, reliability, characteristic network analysis, research, item, member
Application: Logistics	analysis, research, model, learning, algorithm, problem, performance, process, vehicle, development, network, relationship, stage, clustering, technique, text mining, delivery, technology, patent, route, prediction, product, routing, production, detection, design, optimization, level, topic modeling, tree



**Figure 2.** Trends in the Proportion of Each Topic Over Time

Period	1st	2nd	3rd	4th	5th
1975 ~ 1990	Learning Strategies & Image Analytics (39.70%)	Text Mining & Network Analysis (20.89%)	Application: Semiconductor (17.96%)	Decision Support System (7.45%)	Application: Manufacturing (4.66%)
1991 ~ 2000	Learning Strategies & Image Analytics (23.53%)	Text Mining & Network Analysis (15.90%)	Application: Semiconductor (13.39%)	Topic Modeling (11.93%)	Decision Support System (8.74%)
2001 ~ 2005	Learning Strategies & Image Analytics (33.93%)	Text Mining & Network Analysis (14.34%)	Application: Semiconductor (10.54%)	Topic Modeling (9.60%)	System Optimization (8.05%)
2006 ~ 2010	Learning Strategies & Image Analytics (34.85%)	Text Mining & Network Analysis (21.17%)	Application: Semiconductor (11.78%)	Topic Modeling (7.94%)	Machine Learning Algorithm (5.83%)
2011 ~ 2015	Learning Strategies & Image Analytics (20.13%)	Application: Semiconductor (14.44%)	Topic Modeling (13.76%)	Decision Support System (13.64%)	Text Mining & Network Analysis (13.26%)
2016 ~ 2020	Learning Strategies & Image Analytics (22.60%)	Application: Semiconductor (16.89%)	Text Mining & Network Analysis (14.96%)	Topic Modeling (11.66%)	Decision Support System (8.00%)
2021 ~ present	Learning Strategies & Image Analytics (18.73%)	Topic Modeling (16.92%)	Application: Semiconductor (14.75%)	Text Mining & Network Analysis (11.63%)	Application: Logistics (10.99%)

Figure 3. Top-5 Topics in Each Period

<Figure 3>은 각 기간별로 비중이 가장 높은 상위 5개의 토픽을 나타낸 것이다. 데이터과학 분야의 연구이므로 가장 방법론 연구의 기본이라 볼 수 있는 학습 전략 및 이미지 분석이 전 기간에 걸쳐 가장 높은 비중을 차지하고 있는 것을 알 수 있다. 반도체 응용 또한 꾸준히 상위권에 위치하고 있으나, 텍스트 마이닝과 네트워크 분석은 최근 비중이 줄어드는 경향을 여기서도 확인할 수 있다. 반면, 토픽모델링의 경우 초창기에는 상위권에 위치하지 않았으나 최근 들어 더욱 활발하게 연구에 활용되고 있는 것으로 확인되었다.

이와 같이 본 장에서는 50년간 산업공학회지에 출판된 데이터과학 관련 논문들에 대한 메타분석을 수행하여 지금까지의 연구 흐름을 살펴보았다. 이를 통해 산업공학에서 데이터과학을 도구로 사용하여 활발한 연구가 수행되어 왔다는 것을 확인할 수 있었으며, 제조 및 물류 등과 같은 전통적인 산업공학의 적용 분야에 대한 연구가 큰 비중을 차지했다는 것 또한 확인할 수 있었다. 마지막으로 최근 인공지능 발전과 함께 방법론으로는 딥러닝을 포함한 최신 알고리즘들을 적극적으로 활용하며 데이터 관점에서는 정형 데이터뿐만 아니라 비정형 데이터인 텍스트와 이미지를 활용한 연구 또한 활발히 수행되어 왔다는 것을 알 수 있다.

### 3. 산업공학 전공자에게 묻는 미래

산업공학에서의 데이터과학의 미래를 전망하기 위해서, 대학교 또는 대학원에서 산업공학을 전공하고 데이터과학 관련 업무에 종사하고 있거나 종사할 예정인 사람들을 대상으로 설문을 진행하였다. 설문 문항은 산업공학에서의 데이터과학의 미래 및 나아가야 할 방향에 대한 총 3개의 질문으로 구성된다.

설문에는 교수 9인, 대학원생 18인, 그리고 산업체 임직원 8인으로 총 35명이 응답하였다. 항목별 응답자의 의견을 분석하여 요약한 내용을 제시한다.

#### 3.1 미래의 산업공학에서 데이터과학의 중요도와 역할에 대한 전망

- **[산업공학에서의 융합과 혁신의 원동력]** 산업공학은 공학적 방법론의 현실 활용을 통해 제조 및 다양한 산업 분야의 복잡한 문제 해결에 기여해 왔다. 대부분의 응답자가 데이터과학의 기법들은 산업 문제 해결을 위한 범용 도구로서 역할이 더 커질 것으로 전망하였다. 데이터과학이 전통적인 산업공학 분야와 융합되어 새로운 혁신을 이끌어낼 수 있을 것으로 예상된다. 특히, 빅데이터 분석과 인공지능 기술의 도입이 확대되면서, 산업공학의 중요한 도구로 자리 잡아 산업의 지속 가능한 발전을 촉진하는 핵심 역할을 할 것으로 예상된다.
- **[산업 현장의 실용적 응용 도구로서의 발전]** 산업공학의 관점에서 데이터과학의 활용이 다양한 산업분야에서 효율성 개선과 의사결정 지원을 이끄는 핵심 요소로서 중요한 역할을 할 것으로 전망하였다. 산업 현장의 최적화, 예측모델링, 실시간 모니터링 등에 데이터과학 기법의 활용이 크게 기여할 것으로 예상되며, 이는 산업 현장의 생산성 향상, 비용 절감, 전략적 우위 확보로 이어질 수 있다.
- **[산업공학 교육 커리큘럼의 확장]** 산업공학 교육은 프로그래밍, 데이터마이닝, 머신러닝, 딥러닝 등 데이터과학 관련 전공 수업으로 확장되어 오고 있으며, 그 비중이 점차 증가하고 있다. 앞으로도 학문의 변화 및 산업 수요를 반영하여 산업공학에서 데이터과학 관련 교육이 강화될

것으로 예상된다.

### 3.2 미래의 산업공학에서 데이터과학이 제공할 기회

- **[기존 산업공학 연구분야와 데이터과학의 융합]** 기존 산업공학에서 활발하게 연구가 이루어져 왔던 생산 및 품질 관리, 물류 관리, 최적화, 인간공학 등의 문제에 데이터과학을 융합하는 새로운 학문적 접근법을 탐색하는 노력이 꾸준히 이어질 것으로 예상된다. 다수의 응답자가 이러한 연구 개발은 산업공학의 위상과 중요도를 강화하는 기회가 될 것으로 전망하였다. 이는 산업공학이 타 학문분야와 차별화된 데이터과학의 활용을 모색하는 데도 기여할 수 있다.
- **[산업 혁신 및 새로운 비즈니스 모델 창출]** 산업공학은 역사적으로 각 시대 별 산업의 요구에 부응하는 연구개발을 선도해 왔으며, 미래의 산업공학에서 데이터과학의 활용은 다양한 산업 분야에서 데이터를 기반으로 한 산업 혁신과 새로운 비즈니스 모델 창출의 기회를 제공할 것으로 전망된다. 산업공학이 주로 다루어왔던 산업분야와 문제뿐 아니라 신산업을 포함한 산업 분야 전반으로 산업공학의 역할이 확대될 것으로 예상된다.
- **[산업공학 전공자의 커리어 기회 확장]** 데이터과학 역량을 갖춘 산업공학 전문 인력의 중요성이 더욱 부각될 것으로 전망하였다. 데이터과학을 접목한 연구개발은 산업공학의 학문적 발전뿐만 아니라 산업공학 전공자의 직업 및 직무 범위를 넓힐 기회를 제공할 것으로 기대된다.

### 3.3 데이터과학에 관하여 산업공학이 나아가야 할 방향

데이터과학에서의 산업공학의 역할 컴퓨터공학, 통계학, 경영학 등 다른 학문 분야에서의 데이터과학과 차별화되는 산업공학의 역할을 정의할 필요가 있다는 의견이 제시되었다. 산업공학에서의 데이터과학은 단순히 입력과 출력을 다루는 계산 문제를 넘어서, 시스템적 및 비즈니스 관점에서 데이터를 분석하는 목적이 무엇인지, 도출된 모델 또는 분석 결과가 실제 산업 현장에서 어떻게 사용될지, 그리고 투자 대비 기대 수익과 효과는 어떻게 되는지에 대해서 고민하고 분석하는 데 강점을 가진다. 이러한 강점을 바탕으로 데이터과학에서의 산업공학의 정체성을 명확히 하고, 산업공학이 잘할 수 있는 데이터과학의 세부 영역 및 응용 분야를 지속적으로 발굴하고 개척하는 노력이 필요하다는 의견도 제시되었다.

- **[실용적 산업 활용 사례 확대]** 다수의 응답자가 전통적으로 산업공학에서 많이 다루어온 산업 분야뿐만 아니라 다양한 신산업 분야와의 협력을 통해 데이터과학의 적용 범위를 넓혀 나가야 함을 강조하였다. 이는 산업공학이 기술적 변화를 선도하고 산업적 활용 가치를 실현하는 데 중요하다. 일부 응답자는 산업계와 학계 간의 교류를 촉

진하고 실용성 있는 기술을 개발하는 것 또한 중요한 도전과제이며, 이를 가속화하기 위해 산업공학 내 데이터과학 관련 실무 및 학문적 의견을 공유하는 공론장을 활성화해야 한다는 의견을 제시하였다. 또한, 데이터과학의 산업 적용을 어렵게 하는 데이터 윤리와 보안 이슈에 관련한 연구개발도 중요하게 다루어져야 한다는 의견이 제시되었다.

- **[최신 인공지능 연구개발 트렌드 접목]** 최신 인공지능 기술의 급속한 발전이 산업공학의 문제 해결에 효과적인 해결책을 제공하고 있다. 인공지능 기술의 도입이 제조 및 다양한 산업 분야에서 현실 문제를 해결하는 데 큰 기여를 하고 있으며, 산업공학에서 최신 인공지능 기술을 접목하고 응용하는 산업 활용 사례를 늘려나가는 것이 중요한 도전 과제로 전망하였다.
- **[산업공학 전공자의 데이터 리터러시 강화]** 산업공학 전공자들이 데이터과학 분야에서 더 많은 연구와 실제 적용을 통해 가치를 창출할 수 있어야 한다. 산업공학 전공자들이 데이터를 읽고 관리하며 활용할 수 있는 '데이터 리터러시(Data Literacy)' 능력을 강화하는 것이 중요하다. 이는 산업 현장에서 직접적인 데이터 활용과 문제 해결에 기여하는 것을 넘어, 데이터 기반 의사결정 문화를 확립하고 데이터과학의 중요성을 전파해 나가는 역할을 할 수 있을 것으로 예상된다.

## 4. 결론 및 제언

본 논문에서는 현대 산업공학에서 중요한 비중을 차지하는 데이터과학의 역할을 분석하였다. 이를 위해 산업공학에서 데이터과학이 어떻게 발전해왔는지, 그리고 앞으로 어떤 방향으로 나아갈지를 과거, 현재, 그리고 미래의 관점에서 고찰하였다. 과거와 현재를 살펴보기 위해 대한산업공학회지에 게재된 논문을 분석한 결과, 데이터과학과 관련된 논문의 비중이 시간이 흐름에 따라 점차 증가하고 있으며, 최근 3년간은 과반의 논문이 데이터과학과 연관된 주제를 다루고 있음을 확인하였다. 또한, 인공지능 분야, 특히 딥러닝의 발전이 산업공학 내에서 데이터과학의 중요성을 높이는 데 크게 기여했음을 확인하였다. 미래를 전망하기 위해 산업공학을 전공한 데이터과학 관련 전문가들을 대상으로 설문조사 실시한 결과, 응답자들은 데이터과학을 적극적으로 활용하여 산업공학의 전통적 강점을 현대적 도전과 기회에 맞게 조율하고, 이를 통해 산업 현장에서 실질적인 변화와 혁신을 이끌어내야 한다고 강조하였다. 이러한 방향성은 산업공학이 데이터 중심의 현대 사회에서 지속적으로 가치를 창출할 수 있는 기반이 될 것으로 기대된다.

산업공학은 시대의 흐름에 따라 변화하는 다양한 산업의 요구에 부응하며 지속적으로 발전해왔다. 산업공학의 과거와 현재를 돌아보면, 빅데이터 시대의 흐름에 맞추어 데이터과학

기법들을 도입함으로써 산업공학에 큰 변화를 가져왔으며, 이러한 변화는 산업 응용 및 확산을 통해 새로운 산업적 가치를 창출하고, 산업 전반의 지속적인 발전을 촉진하는 데 기여해 왔다. 향후 데이터과학의 발전과 함께 새로운 산업의 수요에 발맞춰 지속적으로 발전하는 미래의 산업공학의 위상을 기대해 본다.

## 참고문헌

- Bailey, D. E. and Barley, S. R. (2005), Return to Work: Toward Post-industrial Engineering, *IIE Transactions*, **37**(8), 737-752.
- Biles, W. E. (1991). The Engineering Science of Industrial Engineering: A Viewpoint of the Industrial Engineering Curriculum. *IIE Transactions*, **23**(3), 208-214.
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003), Latent Dirichlet Allocation, *Journal of Machine Learning Research*, **3**, 993-1022.
- Cao, L. (2017), Data Science: A Comprehensive Overview, *ACM Computing Surveys*, **50**(3), 1-42.
- 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.
- Cho, S. G. and Kim, S. B. (2012), Finding Meaningful Pattern of Key Words in IIE Transactions Using Text Mining, *Journal of the Korean Institute of Industrial Engineers*, **38**(1), 67-73.
- Jeong, B. and Lee, H. (2016), Research Topics in Industrial Engineering 2001~2015, *Journal of the Korean Institute of Industrial Engineers*, **42**(6), 421-431.
- Wu, X., Zhu, X., Wu, G. Q., and Ding, W. (2014), Data Mining with Big Data, *IEEE Transactions on Knowledge and Data Engineering*, **26**(1), 97-107.

## 저자소개

**강석호:** 서울대학교 산업공학과에서 2011년 학사, 2015년 박사 학위를 취득하였으며 현재 성균관대학교에 재직 중이다. 산업 문제의 현실적 제약 하에서 데이터마이닝 및 머신러닝 방법론의 효과적인 활용을 위한 연구를 수행해 오고 있다.

**강필성:** 서울대학교 산업공학과에서 2003년 학사, 2010년 박사 학위를 취득하였으며 현재는 서울대학교에 재직 중이다. 시계열, 자연어, 이미지 데이터 등의 멀티모달 데이터를 활용한 산업 인공지능 연구를 수행해 오고 있다.

**조성준:** 서울대학교 산업공학과 학사, 석사학위 취득 후, 미국 워싱턴대학교 컴퓨터사이언스 학과에서 인공지능 석사학위 및 메릴랜드대학교 컴퓨터사이언스 학과에서 뉴럴네트워크, 머신러닝 분야로 박사학위를 받았다. 현재 서울대학교 산업공학과 교수와 빅데이터AI센터장으로 재직하고 있다. 국가데이터정책위원회 위원을 맡고 있다. 공공데이터전략위원장, 정부3.0추진위원회 빅데이터전문위원장과 한국BI데이터마이닝학회 회장을 역임하였다. 인공지능, 뉴럴네트워크를 시작으로, 머신러닝, 데이터마이닝을 연구하여 왔고, 최근에는 딥러닝, 텍스트마이닝 등 빅데이터와 AI를 연구하고 있다. 이러한 방법론을 바탕으로 제조, 금융, 마케팅, 인사 분야의 IoT 센서 데이터, 텍스트 데이터, 거래 데이터로부터 인사이트를 도출하고 있다. 경험이나 감에 의존하는 주관적 의사결정을 넘어 데이터에 기반한 객관적 의사결정을 일상화하기 위해 필요한 프로세스 변화, 조직 변화, 임직원 교육에도 관심 있다.