

KAST 한국과학기술한림원
The Korean Academy of Science and Technology



ISBN 979-11-86795-92-7

이 사업은 복권기금 및 과학기술진흥기금의 지원을 통한 사업으로 우리나라의 사회적 가치 증진에 기여하고 있습니다.

2023년 석학 커리어 디시전스 과학기술정책제언

통계학: 학문의 길과 산업응용의 길

박성현 서울대학교 명예교수

KAST 한국과학기술한림원
The Korean Academy of Science and Technology

2023년 석학 커리어 디시전스
과학기술정책제언

통계학: 학문의 길과 산업응용의 길

박성현 서울대학교 명예교수



KAST 한국과학기술한림원
The Korean Academy of Science and Technology

발행처

한국과학기술한림원
031) 726-7900
kast@kast.or.kr

발행인

유옥준

발행일

2023년 12월

홈페이지

<http://www.kast.or.kr>

정책제언 및 집필

- 정책제언: 박성현 서울대학교 명예교수
- 집필: 전승민 과학기술분야 전문기자

디자인/인쇄

경성문화사



이 보고서의 모든 저작권은 한국과학기술 한림원에 있습니다.



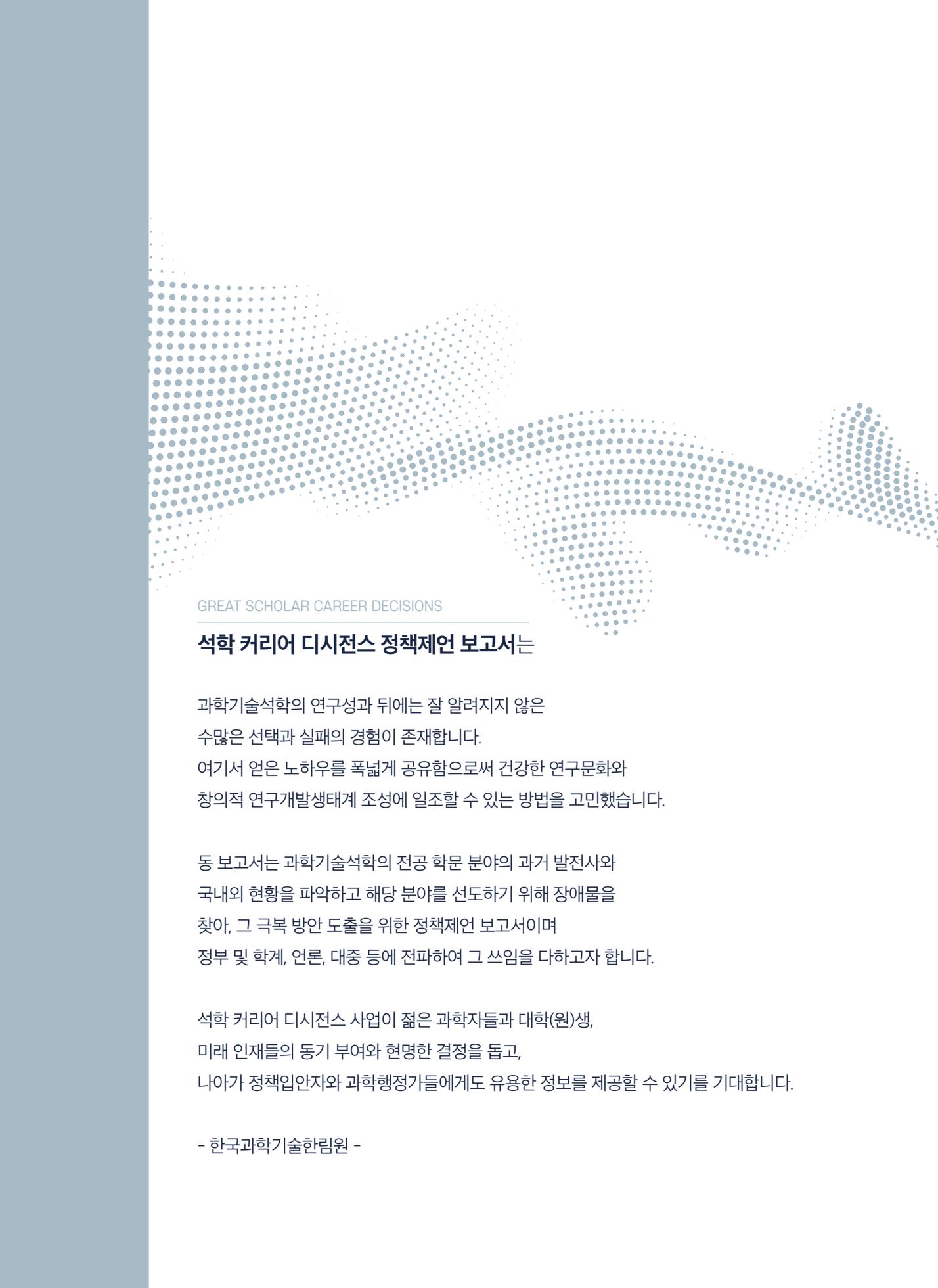
한국과학기술한림원은

대한민국 과학기술분야를 대표하는 석학단체로서 1994년 설립되었습니다.
우리나라 과학기술계를 대표하는 석학집단으로서,
1,000여명의 과학기술분야 석학들이 한국과학기술한림원의 회원이며,
대한민국이 과학기술 선진국으로 도약하여
글로벌 난제를 해결할 수 있는 최첨단 과학기술 업적을 창출할 수 있도록
창의와 도전의 연구개발 생태계를 만드는 데 기여하고자 합니다.

한국과학기술한림원 더 알아보기

- 🌐 홈페이지 www.kast.or.kr
- 📖 블로그 kast.tistory.com
- ✉️ 포스트 post.naver.com/kast1994
- 📘 페이스북 www.facebook.com/kastnews





GREAT SCHOLAR CAREER DECISIONS

석학 커리어 디시전스 정책제언 보고서는

과학기술석학의 연구성과 뒤에는 잘 알려지지 않은
수많은 선택과 실패의 경험이 존재합니다.
여기서 얻은 노하우를 폭넓게 공유함으로써 건강한 연구문화와
창의적 연구개발생태계 구성에 일조할 수 있는 방법을 고민했습니다.

동 보고서는 과학기술석학의 전공 학문 분야의 과거 발전사와
국내외 현황을 파악하고 해당 분야를 선도하기 위해 장애물을
찾아, 그 극복 방안 도출을 위한 정책제언 보고서이며
정부 및 학계, 언론, 대중 등에 전파하여 그 쓰임을 다하고자 합니다.

석학 커리어 디시전스 사업이 젊은 과학자들과 대학(원)생,
미래 인재들의 동기 부여와 현명한 결정을 돕고,
나아가 정책입안자와 과학행정가들에게도 유용한 정보를 제공할 수 있기를 기대합니다.

- 한국과학기술한림원 -



박 성 현
서울대학교 명예교수
(한국과학기술한림원 종신회원)

‘대한민국 산업통계의 선구자’로 불리는 박성현 교수는 기업체 품질관리 및 생산성 향상 교육과 컨설팅 등을 하며 통계학의 산업 적용에 크게 이바지했다. 쌍용정유(현 에쓰오일), SK하이닉스, 삼성전기 등 다수의 국내 대기업에 기업컨설팅을 하며 품질 및 생산성 혁신에 기여한 공로로 고(故) 김대중 대통령으로부터 근정훈장 홍조장을 받기도 했다.

박성현 교수는 현재 대한민국학술원 자연과학부 회장, 한국시니어과학기술인협회 회장, 그리고 사회적책임경영품질원 회장을 겸임하고 있다. 한국과학기술한림원 7대 원장을 지냈으며, 건국대학교 경영대학 기술경영학과 석좌교수를 지냈다. 미국에서 공부하고 미시시피주립대학교 경영대학 조교수를 역임한 후 서울대학교 통계학과 부교수, 교수를 거치며 자연과학대학 학장을 역임했다. 2010년부터 2년간 한국연구재단 기초연구본부장을 지내며 국가 과학기술 기초연구 분야 행정을 총괄한 바 있다. 1995~1997년에 한국통계학회 회장, 1999~2001년에 한국품질경영학회 회장을 지냈다.



통계학: 학문의 길과 산업응용의 길

1. ‘산업기술 강국’ 대한민국과 통계학	04
2. 과학의 산업화, 어떻게 이를 것인가?	12
3. 정책제언: 데이터 산업 육성 및 진흥 전략	28
4. 시대에 걸맞은 ‘산업형 과학기술인’ 육성시스템 마련해야	36
5. 통계학자로서의 철학과 삶	42

1

‘산업기술 강국’ 대한민국과 통계학

• 통계학의 정의, 발전사

'산업기술 강국' 대한민국과 통계학

인간은 왜 역사를 공부할까. '과거'를 되돌아보는 것은 어떤 의미가 있을까. 우리는 왜 '위인'들의 성공 사례를 되짚어 보는 것일까. 한 분야에 탁월한 업적과 성과를 창출한 상징적 인물들의 지식과 경험, 노하우를 되살펴 보는 까닭은 과연 무엇일까.

그것은 성공을 '우연'이라 말할 수 없기 때문이다. 거기에는 '노력'이나 '재능'이라는 한마디 말로 설명할 수 없는 다양한 노하우가 수없이 녹아 있다. 어떤 조직, 혹은 개인이 그 노하우를 충분히 배우고 익혀 실천할 수 있다면, 그 미래는 밝다고 이야기할 수 있지 않을까.

과학기술계, 그리고 대한민국의 미래 역시 마찬가지이다. 다행히 대한민국은 많은 고난과 역경에도 불구하고 비교적 성공적 행보를 걸어왔다. 1953년 한국의 1인당 국민총소득(GNI)은 76달러였으나, 2022년 1인당 GNI는 3만 2,661달러로 나타나 한국전쟁 이후 430배 이상 급성장했다. 그 이면에는 과학기술계의 부단한 노력과 뒷받침이 있었다는 점에 이견을 제시할 사람은 그리 많지 않을 것이다. 우리에게 남은 숙제는 이런 성공을 앞으로도 계속해 나갈 수 있느냐 하는 점이다.

'석학 커리어 디지전스' 정책보고서는 이런 생각에서 출발한다. 대한민국의 성공사를 함께 써 온 과학기술 주역들의 성공 비결을 듣고, 그들의 지식과 경험, 노하우를 최대한 정리하여 새로운 시대를 만들어 나갈 이 땅의 후임들에게 전하려는 것이 그 목적이다. 한국과학기술한림원은 이를 책자 형태로 정리하여, 과학기술 석학의 세부 연구 분야의 태동과 발달, 국내 발전사 및 현황 등을 정리하고 향후 발전 전략을 도출하도록 함으로써 유용한 정책적 자료이자 과학사적 사료로 활용할 수 있도록 하고자 한다. 그리고 더 나아가 동료 연구자 및 미래 인재들은 물론이고, 정책입안자들에게 유용한 자료로 제공, 이러한 정보가 미래 과학기술계, 나아가 국가 발전에 더없이 큰 무형의 자산이 되기를 기대한다.

이번 정책보고서는 대상 석학의 철학과 생각, 노하우와 철학 등을 가감 없이, 최대한 알기 쉽게 전달하는데 주안점을 뒀고, 목차나 제목 등의 구성도 기존의 보고서 형식에 맞추기보다는 내용의 전달 그 자체에 목적을 두어 고민하고 집필했다.

'과학의 산업화' 편의 제언자로 선정된 '박성현 서울대학교 명예교수(이하 박성현 교수)'의 업적과 삶, 경험은 과학기술에 대한 깊은 관심과 연구가 산업에 얼마나 긍정적 영향을 미칠 수 있는지를 보여준다. 박성현 교수 본인도 "과학기술의 발전이 결국 사회의 발전, 나아가 행복한 사회를 만드는 밑바탕이 되며 본인의 지식과 경험이 대한민국의 미래에 작은 도움이나마 되길 바란다"는 의지를 내비쳤다.

박성현 교수는 통계학을 공업적으로 적용하는 방법을 책과 논문, 교육을 통해 전파하여 대한민국 공업기술의 수준을 끌어올리는 데 크게 이바지한 인물이다. 이른바 '기초과학'을 뿌리 깊이 이해하고, 이를 산업에 접목하는데 앞장선 진정한 '기초과학자이면서 동시에 산업형 과학기술인'이라고 구분할 수 있을 것이다. 박성현 교수는 우리나라의 산업화가 한창이던 1980~1990년대 시절, 산업계의 부족한 기초과학의 지식을 제공해 주었던, 이른바 '산업을 도운 과학지식인'으로서 그 가치가 큰 인물이라 정의할 수 있다.

그 과정에서 박성현 교수가 이룬 업적은 크게 두 가지로 나뉜다. 첫째는 그의 전공 분야인 통계학, 그중에서도 회귀분석(regression analysis), 실험계획법(design of experiments), 통계적 품질관리(statistical quality control) 분야에서 일생 연구한 학문을 다수의 책과 논문으로 정리하여 학계와 대중에 전파한 점이다. 둘째는 산업 분야 실무자들을 위해 현장 교육, 그리고 기업컨설팅에 팔을 걷어붙이고 나섰다는 점을 들 수 있다.

일생 그가 저술한 책은 93편, 논문은 178편에 이른다. 이런 책 중 상당수는 기업에서 교재로 쓰일 만큼 실용성이 컸다. 산업체의 생산과정에 문제가 발생하면 직접 팔을 걷어붙이고 달려가, 전공인 통계학을 이용해 문제를 해결해 주는 '해결사' 역할도 마다하지 않았다. 그의 학문적 업적 대다수는 산업통계에 관한 것들이며, 베스트셀러 반열에 오른 다양한 통계학 관련 저서들도 모두 산업통계에 관한 내용을 담고 있다. 그는 국내에 산업통계를 전파한 창시자인 셈이며, 그의 업적은 대한민국이 '산업기술 강국'으로 거듭나는데 적지 않은 기여를 했다. 즉 이런 박성현 교수의 업적과 제언은 보고서로서 정리해 둘 충분한 가치가 있다 여겨진다.

통계학의 정의, 발전사

박성현 교수는 통계학과 관련, 국내 과학 및 산업기술 발전에 큰 업적을 남겼다. 그의 업적과 제언을 온전히 이해하기 위해서는 우선 통계학이란 어떤 것인지 미리 알아볼 필요가 있다.

통계학이란 학문은 유럽에서 18세기에 영국과 독일을 중심으로 태동한 학문이나, 그 학문적 기초는 오래전부터 존재했다. 고대 이집트, 메소포타미아, 중국 등에 있던 국가들에서도 이미 인구, 납세, 농지 등의 수량적 조사나 관찰을 통해 국가의 운영을 위한 정보를 통계의 형태로 구해 왔기 때문이다. 그러나 18세기 전까지는 나라의 운영을 위하여 필요한 통계를 생산, 활용해 온 것뿐으로, 통계학이란 학문이 따로 존재하지는 않았다.

통계학(統計學, Statistics)이란 단어가 공식적으로 통용된 것은 1797년에 대영백과사전(Encyclopedia Britannica)에 처음으로 등재된 후부터이다. 이 단어는 엄밀하게 어떤 뜻일까. 우선 어원을 살펴보자. 폴크(Folks)¹에 따르면 라틴어인 status(국가를 의미하는 state)와 statista(정치인을 의미하는 statesman)에서 온 것이라고 한다. 특히 statista는 국가의 업무를 관장하는 사람을 일컫는 말이었다.

즉 당시 Statistics의 의미는 정치인들이 '국가의 살림을 꾸려 나가는 데 필요한 통계자료를 체계적으로 산출해 내는 정치 산술'로 정의됐다. 초기의 통계학 영역은 오늘날 통계청에서 생산하는 인구주택총조사, 산업총조사 등이나, 한국은행에서 생산하는 국내총생산(GDP) 조사 등이 만들어 내는 국가통계에 관한 것으로, 정치 산술학 성격이 강했다.

통계학이 영국에서 정치산술학으로서 발생했다면, 이를 한 단계 발전시킨 것은 독일 등 유럽 여러 나라에서 발전한 '확률론'이다. 19세기 초에 초기의 통계학이 확률론을 만나면서 통계학 개념이 확대되어 수리통계학(mathematical statistics)의 모습을 띠게 됐다. 확률론을 기반으로 하는 수리통계학은 차츰 추측통계학(inferential statistics)으로 발전되고, 20세기 초에 영국을 중심으로 '현대통계학'이 싹트게 된다. 현대통계학은 정치산술학과 추측통계학을 포함하면서, 표본 데이터에서 모집단의 성격을 파악하려는 오늘날의 통계학으로 발전하게 됐다.

현대통계학의 정의는 다양하지만, 학자 사이에서 통용되는 명백한 기준은 있다. 보고서의 제언

자인 박성현 교수에 따르면, 현대통계학은 '사회, 자연, 공업 및 인간 생활의 온갖 현상을 연구하기 위하여 불확실성(uncertainty)이 내포된 데이터(data)의 수집, 분석, 추정 및 검정을 통하여 의사결정(decision-making)에 필요한 정보(information)의 획득과 처리 방법을 연구하는 학문'²이라고 한다. 즉, 통계학은 '데이터로부터 효과적으로 정보를 추출하는 방법을 연구하는 학문'으로 볼 수 있다.

현대에 기초과학이 가장 발전한 나라로 단연 미국을 꼽는다. 통계학도 20세기 중반에 그 중심이 영국에서 미국으로 옮겨졌다. 이를 중심으로 현대통계학은 급성장하고 있다. 통계학 자체의 이론 연구는 물론, 다른 분야에 도움을 주는 보조 학문으로도 급성장하고 있으며, 다른 학문과 통계학이 만나 새로운 학문 분야의 이름이 생겨나고 있다. 예를 들면, 생물학 분야는 생물통계학(Biostatistics, Biology+Statistics), 경제학 분야는 계량경제학(Econometrics, Economics+Statistics), 사회학 분야는 계량사회학(혹은 사회통계학(Sociometrics, Sociology+Statistics), 공업 분야는 공업통계학 혹은 산업통계학(Industrial Statistics)으로 발전하고 있다.

우리나라에 현대통계학이 도입되기 시작한 것은 1960년대 서울의 5개 사립대(고려대(1963), 동국대(1963), 성균관대(1964), 중앙대(1964), 연세대(1966))에서 통계학과가 설치되면서부터라고 할 수 있다. 동국대는 이과대학에 설치되었으나, 기타 4개 대학은 모두 경상 계열에 설치하여 그 당시 경제·사회 발전을 위한 통계학의 기여를 강조하였음을 알 수 있다. 초기 통계학과들을 세운 학자들은 대부분 계량적인 분석 등의 응용통계에 많은 관심이 있었기 때문에 이 당시 통계학 연구는 주로 계량경제 등을 중심으로 하는 응용통계 분야였다. 1971년에 한국통계학회 창립됐고, 이 학회는 오늘날 통계학 연구와 교류의 중심지가 됐다.

1970년대 이후 미국 등을 중심으로 외국에서 통계학의 이론적인 내용을 공부한 학자들이 귀국함으로써 통계적 추론, 의사결정론 등 이론통계학이 본격적으로 연구되기 시작했다. 그 중심에는 서울대학교 자연과학대학에 1975년에 설치된 계산통계(計算統計, Computational Science and Statistics)학과이다. 이 학과는 계산학과 통계학을 같이 교육하는 학과로, 통계학 분야에서는 이론통계를 중시하여 국내에서 현대통계학 연구의 뿌리를 튼튼하게 하는 역할을 감당했다. 박성현 교

¹ Folks, J. I. 1981. Ideas of Statistics. NY: John Wiley & Sons.

² 박성현. 2003. 현대실험계획법, 개정판. 서울: 민영사, 612쪽.

수는 서울대 계산통계학과에 1977년에 부임한 선구자적 인물로, 대한민국의 현대통계학의 발전에 크게 기여했다고 말할 수 있다. 특히 통계학 전문 분야의 서적이 거의 없던 시절에 각 분야의 전문 서적으로 '회귀분석', '현대실험계획법', '통계적품질관리'란 책을 출판하여 국내 현대통계학의 발전에 큰 몫을 담당했다.

1980년대에 국내 많은 대학에서 통계학과를 설치하면서 오늘날에는 전국에 70여 개의 대학에 통계학 관련 학과가 설치되어 있다. 이들은 초기의 통계학과처럼 경상계열에 있어 통계의 경제·사회 분야 통계응용을 주로 교육하는 곳도 있고, 자연계열에 소속되어 있어 이론통계학을 중심으로 교육하는 곳도 있고, 컴퓨터 분야와 어우러져 '데이터 과학(data science)' 목적에서 주로 교육하는 곳도 있다. 통계학 연구의 다양성을 보여주고 있는 현상이라 할 수 있다.



2

과학의 산업화, 어떻게 이룰 것인가?

- 각국의 과학기술인 산업 참여 사례
- 한국의 사례: 통계학으로 이끈 산업화
- 데이터과학의 발전과 산업 활용
- 데이터 산업의 현황과 발전 전략

과학의 산업화, 어떻게 이를 것인가?

“기초과학 연구 성과가 산업으로 이어지려면 적어도 30년은 걸린다”는 이야기를 들어본 사람이 많을 것이다. 이 말을 들은 사람들은 어떤 반응을 보일까. 표면적으로는 ‘과학기술 투자는 최소 30년 이상 꾸준히 해야 한다’는 이야기로 해석한다. 그런데 현실에서는 이야기가 달라진다. 실제로 과학기술 투자를 진행하는 입장에서 ‘시간만 오래 걸리고 효율은 떨어지니, 기초과학 연구 성과는 해외의 우수한 것을 빌려 오고, 우리는 산업기술 위주로 투자를 진행해야 한다’고 판단하는 경우가 많다. 실제로 국내 과학기술 투자 정책은 이런 기초를 따르는 경우가 많았고, 산업기술의 빠른 확보가 필요한 국내 현실에서 이는 주효한 정책이었다. 그 결과 대한민국은 세계적 산업 국가로서 발돋움할 수 있게 됐고, 현재 세계 국내총생산(GDP) 순위 12위¹를 기록할 만큼 성장했다.

이 때문에 한국은 기초과학 분야를 등한시하고 있다고 착각하는 경우가 적지 않다. 그러나 시점을 달리하면 한국의 기초과학 수준은 의외로 높다고 할 수 있는데, 기초과학을 완전히 등한시한 상태에서 고도의 산업화는 절대 가능하지 않기 때문이다. 산업이란 과학의 토대 위에서 성립한다. 기초과학을 연구하고 발전시키는 것, 인류가 쌓아 올린 기초과학의 토대를 받아들여 그 위에서 산업을 발전시키는 것, 이 두 가지 모두 기초과학에 대한 이해가 필수 조건일 수밖에 없는데, 상당수의 과학기술자도 ‘산업 발전에 집중하면 기초과학을 홀대할 것’이라 착각하는 경향이 있다. 이 점을 이해하지 못하면 과학 및 산업정책을 집행하는 데 있어 적잖은 시행착오를 겪게 될 수 있다.

¹ 2023년 4월 추정치(IMF 기준)

이는 해외라고 해서 예외가 아니다. 기초과학 전문가가 나서 산업기술의 성공을 견인한 성공 사례는 의외로 자주 발견된다. 이는 기초과학기술이 이미 발전해 있고, 이를 전공한 전문가가 산업 기술 분야에 뛰어들거나, 혹은 산업계 프로세스 전반에 대해 조연함으로써 산업의 발전에 크게 이바지한 사례를 이야기한다.

각국의 과학자 산업 참여 사례

01 코로나19 백신 개발 과학자 우그르 사힌

대표적인 사례가 ‘신종 코로나바이러스 감염증(코로나19) 백신을 개발한 ‘우구르 사힌(Uğur Şahin)’이다. 잘 알려진 바와 같이 미국 화이자 사는 이 코로나19의 백신을 최초로 개발, 판매하는데 성공했는데, 내막을 살펴보면 실제로 백신을 만든 회사는 화이자가 아니다. 백신은 화이자와 손잡은 기업 ‘바이오엔테크(bioNtech)’가 개발했고, 화이자가 생산만을 맡고 있다. 그 바이오엔테크 창업자가 우구르 사힌이다. 그는 사업가이기 이전에 의학자로 알려져 있다. 터키 출신 독일 이민자로, 내과 의사 출신으로 다양한 기초의학 분야 연구자였다. 이런 그가 산업계에 뛰어들고, 의학 지식의 근간을 제공하면서 기업은 산업계에서 두각을 나타내기에 이른다. 그가 창업을 결심한 것 자체가 자신이 연구하던 면역요법(immunotherapy)이 암 환자에게 도움이 되겠다는 생각에서였다. 그는 이렇게 만들었던 회사를 일본 제약회사에 팔아 자금을 마련한 후, 그 자금을 바탕으로 다시 바이오엔테크를 설립했다. 우구르 사힌이 면역 분야 연구를 위해 쌓아왔던 기초과학 지식은 그대로 전 세계를 펜데믹에서 구하는 희망이 됐다. 면역요법 연구를 위해 mRNA를 연구해오던 바이오엔테크는 코로나19 대유행 시기에 mRNA 기술을 백신 만드는데 응용했고 그렇게 코로나19 백신을 개발에 성공했다.



‘우구르 사힌’,
바이오엔테크 홈페이지.

02 P&G사의 세탁용 세제 개발

화학 기업 'P&G(PROCTER&GAMBLE)' 사의 혁신 사례도 이 같은 기초과학기술인의 협력에서 출발한다. P&G사는 50여 년 전 세탁용 세제의 성능을 끌어 올리기 위해 고심하고 있었는데, 새롭게 등장한 단백질분해(proteases) 등의 방법이 세탁물의 다양한 얼룩을 제거하는 데 효과적이라는 사실을 알아냈다. 그러나 알레르기 반응 등 실제 사용하는데 적잖은 문제를 일으켰고, 세계 산업계는 1974년에 안전성을 확인할 수 있을 때까지 관련 기술이 적용된 세제를 사용하지 않기로 했다. 이에 P&G사는 유명 대학의 주요 학과로부터 관련 분야 기초 지식을 갖춘 과학기술인들을 영입, 소규모의 생명기술 그룹을 만들었고, 10여 년 동안 본격적인 연구개발을 진행, 모든 문제를 유전공학적으로 해결한 효소를 개발하는 데 성공했다. 기초연구조직과 상품개발그룹과의 효과적인 연계를 통하여 성공적이고 안정적으로 전 세계에서 사용하고 있는 세제를 개발한 것이다. 기업에 기초과학 연구진을 수혈한 결과 얻어낸 값진 결과라고 할 수 있다. 이처럼 과학기술인이 산업의 발전에 직접 기여하는 것은 과학과 산업이 고루 발전한 선진국일수록 두드러진다. 미국의 경우 이론 물리, 첨단재료 및 분자생물학의 근본적 문제를 탐구하는 기초연구자들조차 그들의 작업이 새로운 칩의 개발, 다양한 종류의 비행기 혹은 암 치료 약 개발에 관련될 것이라 기대하고 연구하고 있다.²

03 노벨상을 수상한 일본의 기초과학자들

생산기술 기반 산업이 주가 되는 일본의 경우 '기업의 과학기술인 활용'이 더욱 두드러진다. 본래 일본기업의 경우 기초연구를 중시하는 풍조가 강했다. 심지어 기업이 기초연구를 직접 지원하는 경우도 많았다. 예를 들어 2019년 리튬이온전지 발명으로 노벨화학상을 수상한 요시노아키라(吉野彰) 아사히카세이(旭化成) 명예펠로는 1981년부터 관련 연구를 시작했다. 본래는 전기가 통하는 플라스틱을 소재로, 이를 활용할 수 있는 방법에 관한 긴 연구를 진행한 결과 노벨상을 받았다. 2002년 질량분석 기술로 노벨화학상을 받은 '다나카고이치(田中耕一) 시마즈세이사쿠소(島津製作所) 시니어펠로는 1987년에 발표한 단백질 질량 분석기술의 개발로 노벨상을 받게 됐

² 한국과학재단. 2004. 미국의 기초연구-기초연구를 통한 번영-. 한국과학재단 보고서.

다. 2014년 청색발광다이오드로 노벨물리학상을 받은 3인 중 하나인 아카사키이사무(赤崎勇)는 1981년까지 마쓰시타 덴키(松下電器, 현재의 파나소닉)에 재직하고 있었다. 이처럼 과거 1980년대에 기업연구소에서 이루어진 기초연구들이 2000년대 이후 노벨상 수상으로 이어지고 있다. 기업의 기초연구가 왕성하던 시절, 기업연구소의 1인당 연구비는 2,000만~3,000만 엔 대학의 1인당 연구비 200만~300만 엔의 10배에 달했다. 그러나 일본도 거품경제 붕괴 이후 기업의 경영성과에 직접 관련되는 제품화 연구를 중시하는 방향으로 전환되면서 한국과 같이 '과학지식의 발 빠른 산업응용'에 관심을 두는 경향이 높아지고 있다. 문부과학성이 매년 발표하는 '대학 등의 산학연계 등 실시상황에 대하여'라는 통계에 따르면, 2013년을 전후로 대학과 민간의 산학협력이 증가하는 추세에 있음이 확인된다. '민간기업으로부터의 연구자금 등 수입액(공동연구, 수탁연구, 지식재산활동 등을 포괄하는 넓은 범위의 수입액)'은 2018년 약 1,075억 엔으로, 2003년 조사 개시 이후 최초로 1,000억 엔을 돌파했다. 2012년의 약 600억 엔에 비해서도 상당한 증가세를 보인다. 그중에서도 공동연구가 683억 엔으로 63.5%를 차지하여 산학협력의 중심적 역할을 하고 있으며, 그 추세 또한 공동 연구의 증가에 기인하고 있음이 확인된다.³



노벨물리학상 수상자 아카사키 이사무. Meijo University 홈페이지.

04 품질관리의 아버지, 미국 데밍 박사

수학, 물리, 통계학 분야의 기초과학자로서, 산업현장의 품질관리와 품질경영에 기여한 해외 학자는 누가 있었을까? 대표적인 인물로 에드워드 데밍(Edwards Deming, 1900~1993) 박사를 꼽을 수 있을 것이다. 그는 학사, 석사, 박사 학위를 각각 전기공학, 수학, 물리학으로 받고, 졸업 후에는 자연과 산업의 현상을 수치로 규명하는 통계학에 심취하여 데이터를 활용하는 통계적 품질관리(statistical quality control)와 품질경영(quality management) 방법을 개척하고 이 분

³ 정성춘. 2020. 일본의 개방형 혁신전략: 산학협력을 중심으로. 대외경제정책연구원.

야의 선구자가 됐다. 미국 정부 통계국(the Census Bureau)에 근무할 때 맥아더 장군의 요청으로 일본 정부의 관리자들과 기업인들을 상대로 강연하였는데, 그의 강연 내용인 통계적공정관리(statistical process control)와 품질경영 이론은 전후 일본 산업의 재건에 큰 활력이 됐다. 일본은 데밍 박사의 공로를 기념하기 위해 1951년부터 데밍상(Deming Award)을 제정하여 매년 최우수 기업에게 수상하고 있다. 데밍 박사는 아직까지도 일본기업들의 품질경영에 크게 기여하고 있는 셈이다.



에드워드 데밍. 위키피디아.

데밍 박사는 공무원을 그만둔 후 뉴욕대학(New York University)의 교수가 되어 후임을 양성하는 한편, 집필활동에도 주력해 통계학, 샘플링 기법, 품질경영 기법 등에 관한 수없이 많은 저서를 출판했다. 이런 업적으로 데밍 박사는 데이터를 활용한 통계적 품질관리의 아버지라고 불리우고 있다. 그가 제안한 품질관리의 기본 사이클인 PDCA(Plan-Do-Check-Act)은 지금도 모든 산업에서 애용되고 있다. 이 보고서의 제언자 박성현 교수 역시 데밍 박사를 제일 존경한다고 하며, 그 이유 역시 '기초과학자로서 그의 지식을 산업화 도구로 변환시킨 탁월한 리더이기 때문'이라고 밝히고 있다. 데밍 박사의 대표적인 저서로는 아래에 정리한 6권이 주로 꼽히는데, 처음 3권은 그가 직접 저술한 것이고, 나중 3권은 그의 업적을 기리기 위해 그의 평소의 업적을 모아 후학들이 만들어낸 저서이다.

에드워드 데밍 박사의 대표적 저서

- Statistical Adjustment of Data, Dover, 1964.
- Some Theory of Sampling, Dover, 1966.
- Quality, Productivity, and Competitive Position, MIT Press, 1982.
- The New Economics: For Industry, Government and Education, MIT Press, 2000.
- The Essential Deming: Leadership Principles from the Father of Quality, McGraw Hill Professional, 2012.
- Out of Crisis, MIT Press, 2018.

한국의 사례: 통계학으로 이끈 산업화



국내 학계 및 산업계에도 산업의 발전에 필요한 기초과학을 제공하는 전문가가 존재한다. 과학 기술자로서 산업 현장의 품질관리와 품질경영에 가장 크게 기여한 이는 역시 이 보고서의 제언자인 박성현 교수 사례가 대표적인 것이다.

박성현 교수는 서울대를 졸업하고, 미국 노스캐롤라이나 주립대학교에서 산업공학석사와 통계학 이학박사 학위를 취득했다. 이후 미시시피주립대학교 조교수를 지내다 모교인 서울대학교 자연과학대 교수로 부임, 2010년 정년 때까지 학계를 지켰다. 이런 이력을 인정받아 2010년부터 2년간 대한민국 기초과학 분야 연구 개발 투자를 총괄하는 '한국연구재단 기초연구본부장' 직을 차질없이 수행해 내기도 했다. 2013년부터 3년 동안은 국내 석학들의 모임인 '한국과학기술한림원' 원장을 지냈다.

그는 특히 '통계학의 산업응용'에 관심을 가지고 다방면으로 활동했다. 학계에서도 그가 '산업 특화형 과학자'라는 점을 인정하고 있다. 2012년부터 3년 동안 건국대 경영대 기술경영학과 석좌교수를 지내기도 했으며, 한국통계학회 회장, 한국품질경영학회 회장 등을 지내며 국내 산업계에 다양한 통계 응용기법을 전수하기 위해 노력했다. 그는 통계학 분야와 품질경영 분야의 업적을 인정받아 미국 통계학회와 품질학회로부터 영예스러운 석좌회원(fellow)으로 선정되기도 했다.

그는 기업들이 응용할 수 있는 통계 응용기법 관련 책을 저술한 것도 수십여 권에 이른다. 총 저서는 93권으로, 11권의 영문 저서, 국문 저서로는 28권의 단독 저서와 54권의 공동저서 등이 그의 손끝에서 나왔다. 통계 관련 학술 논문은 총 178편에 달한다. 이런 노력으로 국내 산업계에선

그를 “통계 기법을 응용한 품질관리 분야에서 국내 최고의 실력자”라고 불렀다. 2000년에는 ‘품질경영’ 유공자로 근정훈장 홍조장(대통령상)을 수상했으며, 2002년엔 한국품질경영학회 학회장상을 수상하기도 했다. 그를 ‘산업을 위한 과학자’로서 국내 대표적인 인물로 꼽는 것은 이런 점 때문이다. 그는 모든 조직의 사회적 책임(social responsibility)에도 큰 관심을 갖고, 2016년부터 사회적책임경영품질원의 회장직을 수행하면서 기업, 산업체 등의 사회적 책임경영은 물론 ESG(환경, 사회, 지배구조) 경영의 보급에도 노력하고 있다.

박성현 교수의 품질 분석 기법의 핵심은 지금으로 보면 ‘빅데이터 분석에 의한 품질 향상’기술이라고 볼 수 있다. 그 당시에는 빅데이터 프로그램이 아직 나오지 않은 상황이었지만 기존의 통계 패키지인 SAS, SPSS 등으로 상당 부분 해결이 가능한 부분이 적지 않았다. 박성현 교수는 이런 데이터 기반 품질·생산성 향상 분석 기법을 ‘데이터 기술(DT, data technology)’이라고 작명하고, 이 기법을 널리 전수하기 위해서 노력했다. 오늘날의 빅데이터 기술의 초기 버전이라고 볼 수 있으며, 박성현 교수는 DT의 제안을 자신도 매우 의미있게 생각하고 있다. 그와 관련해 논문으로 발표된 박성현 교수의 연구 성과만 보더라도 상당수에 이른다. 아래에 정리해 둔다.

박성현 교수의 데이터 기술(DT) 관련 논문

- (1) 지식기반 사회에서의 통계학 패러다임의 변화와 데이터 기술의 발전, 경영정보논총, 서울대학교 경영대학 발행, 제11권, p.53-59, 2001. 12.
- (2) Visions of Data Technology and e-Statistics with their roles in Industry and Government, 경영정보논총, 서울대학교 경영대학 발행, 13권, pp. 59-68, 2003. 8.
- (3) Data technology and knowledge-based Six Sigma, Asian Journal on Quality, 4(1), p. 40-45, 2003.
- (4) Sung H. Park and Moon W. Suh (2008), Data Technology as a new discipline for broader application of statistics, Journal of Data Science, Vol. 6, No. 3, pp. 357-368, July 2008 issue.
- (5) Pasquale Erto, Giuliana Pallotta and Sung H. Park (2008), An example of data technology product: a control chart for Weibull processes, International Statistical Review, Vol. 76, No. 2, pp. 157-166.

데이터과학의 발전과 산업 활용



과학기술인의 산업 참여는 물론 중요한 주제지만, 다가올 미래에 ‘어떤 분야의 과학기술인이 산업계에서 주목받을 것인가?’ 하는 점 역시 결코 간과할 수 없는 주제라 할 수 있다.

지금은 ‘데이터의 시대’다. 인공지능(AI), 사물인터넷(IoT) 등 이른바 4차 산업혁명의 시대의 핵심 기술은 예외 없이 데이터를 기반으로 움직인다. 이 시대에는 데이터로부터 신속하고 정확하게 정보를 창출하고 활용하는 과학이 중요해지고 있다. 이런 학문을 데이터과학(data science) 혹은 데이터 사이언스라고 부른다. 보고서의 저언자 박성현 교수는 “지금이야말로 통계학과 관련된 다양한 학문의 발전에 관심을 가져야 하는 시기”라고 지적한다.

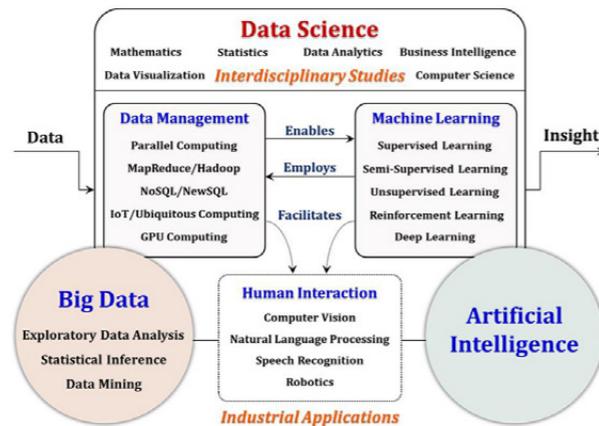
데이터과학이란, 데이터의 수집과 저장에 필요한 데이터 프로세싱 기술과 데이터 분석에 관한 지식(통계학, 데이터 마이닝, 기계학습 등)을 기반으로 다량의 데이터로부터 패턴을 찾아내고, 통계적 추정, 예측 모델링 등을 통하여 필요한 정보를 창출하고, 이를 실제로 활용하는 것을 연구하는 융합과학(convergence science)⁴이다.

넓은 의미의 데이터과학은 아래 그림에서 보는 바와 같이 데이터로부터 빅데이터와 AI 기술을 통하여 통찰력(insight) 있는 정보를 만들어내는 21세기에 태어난 새로운 학문이다. 좁은 의미의 데이터과학은 데이터로부터 통계학의 탐색적 데이터 분석(exploratory data analysis), 통계적

⁴ 박성현, 오진호, 권순선. 2016. 통계로 풀어나가는 빅데이터. 한국표준협회미디어.

추론(statistical inference), 데이터 마이닝(data mining) 등을 통하여 빅데이터 기술을 활용함으로써 정보를 창출하는 학문이다. 그러나 최근에는 넓은 의미의 데이터과학이 더 많이 통용되는 추세라고 볼 수 있다.

데이터과학의 밑바탕이 되는 기존의 학문은 수학, 통계학, 컴퓨터과학, 데이터 분석학(analytics), 데이터 시각화(data visualization), 비즈니스 인텔리전스(business intelligence) 등이다. 데이터과학은 명실공히 다학제간 연구 학문이라 할 수 있다. 데이터과학은 빅데이터를 처리, 분석, 저장할 수 있도록 물리적 환경을 조성해 주는 데이터관리(data management) 분야와 데이터로부터 모델을 만들어 시로 구현할 수 있도록 해주는 기계학습(machine learning) 분야로 나누어진다.



데이터 과학의 전체 조감도. 한국과학기술한림원⁵

데이터관리에 사용되는 기본적인 도구는 병렬 컴퓨팅과 맵리듀스(Map Reduce)와 하둡(Hadoop), 그래픽 처리장치(GPU, Graphic Processing Unit) 컴퓨팅 기술 등을 포함한다. 기계 학습은 학습하는 방법과 분석 모델의 특성에 따라 지도학습(supervised learning), 비지도 학습(unsupervised learning), 준지도 학습(semi-supervised learning), 강화학습(reinforcement learning), 심층학습(deep learning) 등이 있다.

데이터과학의 산업 활용에 대표적인 것이 바로 빅데이터와 AI이다. 오늘날을 ‘데이터·AI 경제 시대’라고도 부른다. 지금 시대엔 데이터 과학이 모든 산업의 핵심적인 역할을 하고 있는 것이다. 1990년대 중반부터는 AI 연구가 통계 기반 기계학습(machine learning)이 도입되면서 급진전이 이루어졌다. 1997년에는 IBM의 AI 슈퍼컴퓨터 딥블루(Deep Blue)가 세계 체스 챔피언에게 승리했고, 1998년에는 소니 최초의 AI 로봇 애완동물 아이보(Aibo)가 개발됐다. 2010년대부터는 딥러닝(deep learning) 기법이 도입되면서 AI의 획기적이 발전이 이루어지기 시작했다. 2011년에는 IBM이 개발한 슈퍼컴퓨터 왓슨(Watson)이 미국의 유명한 퀴즈쇼 ‘제퍼디(Jeopardy)’에서 우승하는 사건이 발생하였고, 또한 애플 AI 음성 인식 소프트웨어 시리(Siri)가 발표됐다. 2014년에는 페이스북(Facebook)에서 얼굴인식 기술인 딥페이스(Deep Face)를 공개하여 세상을 놀라게 한 바 있다.

2022년 하반기부터 마이크로소프트(MS)의 지원을 받은 ‘Open AI’사에서 개발한 ChatGPT라는 생성형 AI가 출시되어 선풍적인 인기를 끌고있다. ChatGPT는 대화형(chat) 생성형 사전학습변환기(generative pre-trained transformer)라고 볼 수 있으며, 대화식으로 하면서도 매우 빠르게 질문자에 성실하게 답해 주는 능력을 가진 AI 소프트웨어이다. 이 소프트웨어는 데이터 과학의 산물이라고 볼 수 있으며, 앞으로 모든 분야(산업, 과학기술, 교육, 문화 등)에 심대한 영향력을 행사할 것이 틀림없다. 최근에는 미국의 기타 대형 IT 기업들도 ChatGPT와 유사한 소프트웨어 개발에 나서고 있고, 한국에서도 삼성, 네이버, KT 등도 참여하고 있어 앞으로의 경쟁이 주목된다.

5 이영조, 박성현, 장중순, 김선, 박태성, 신현정, 이재길. 2019. 빅데이터·인공지능 산업 진흥을 위한 데이터 과학의 발전 전략 연구. 한국과학기술한림원 연구보고서 130

데이터 산업의 현황과 발전 전략

빅데이터와 관련된 모든 비즈니스(보급, 지도, 분석, 유통 등)는 데이터 산업(data industry)에 속한다. '데이터 산업'에 대해 아래에서 조금 짚고 넘어가기로 하자.

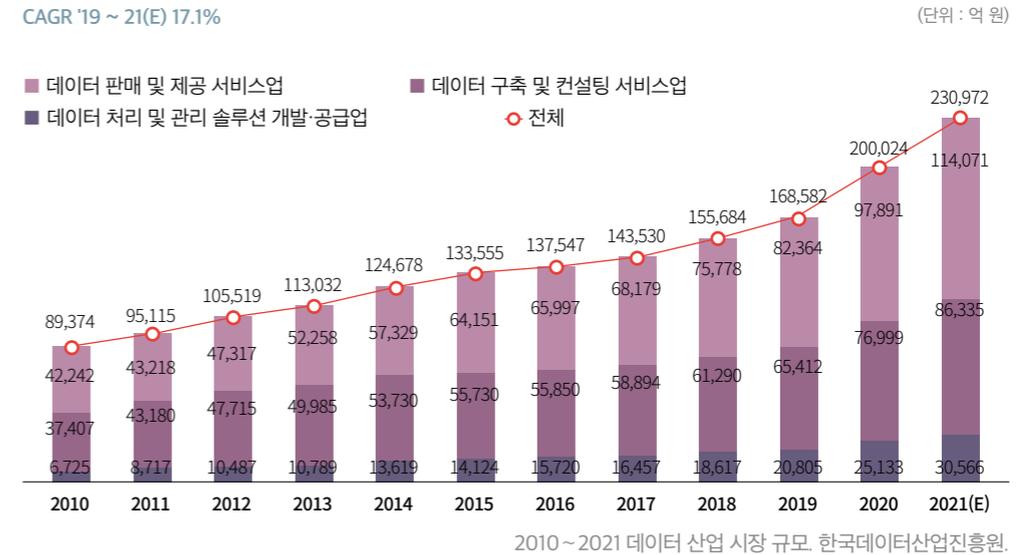
2022년에 한국데이터산업진흥원에서 발표한 「2021 데이터 산업 현황조사」에 의하면, 데이터 산업(data industry)은 '데이터의 생산, 수집, 처리, 분석, 유통, 활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업'으로 정의한다. 데이터의 가치 사슬상에 나타난 데이터와 관련된 제반 활동을 포함해 데이터로부터 가치가 창출되는 일련의 모든 과정, 이와 연관된 활동을 포함한다.

데이터 산업의 정의에 따라 데이터 산업의 대분류는 3가지로 한다. 이들은 첫째로, 데이터와 관련한 제품을 판매하거나 기술을 제공하는 데이터 처리 및 관리 솔루션 개발·공급업(빅데이터 포함), 두 번째로 데이터 구축 및 컨설팅 서비스업, 그리고 세 번째로 데이터를 판매하거나 이를 기반으로 정보제공 및 분석 서비스(클라우드 사업 포함)를 제공하는 데이터 판매 및 제공 서비스업이다.

01 국내 데이터 산업의 시장 규모와 인력

「2021 데이터 산업 현황조사」에 따르면, 데이터 산업의 시장규모는 다음 페이지의 도표에 나타나 있듯이 지난 12년간 매년 급성장해 왔다. 더구나 이런 추세는 앞으로도 계속 이어질 것으로 보인다. 특히 최근 3년간(2019~2021년)의 연평균 성장률(CAGR)이 17.1%에 달할 정도로 매우 유망한 산업 분야다. 2021년 기준 데이터 산업의 부문별 시장 규모를 보면 데이터 판매 및 제공 서비스업이 11조 4,071억 원(49.4%), 데이터 구축 및 컨설팅 서비스업이 8조 6,335억 원(37.4%), 그리고 데이터 처리 및 관리 솔루션 개발·공급업이 3조 566억 원(13.2%)를 차지하고 있다.

데이터 산업에 종사하는 인력은 2021년 기준 39만 4,709명으로, 이 중 데이터 직무인력이 12만 2,431명이고, 데이터 직무 외 인력이 27만 2,278명이다. 여기서 데이터 직무인력이란 데이터 산업에 종사하는 데이터 개발자, 데이터베이스 관리자, 데이터 엔지니어, 데이터 분석가, 데이터 과학자, 데이터 기획자, 데이터 컨설턴트, 데이터 아키텍트 등 데이터 관련 업무를 수행하는 전문 인력을 뜻한다. 데이터 직무 외 인력이란 데이터 산업에 종사하는 인력 중에서 데이터 직무인력을 제외한 인력을 이야기한다.



데이터 산업 전체 인력은 2015년 28만 323명이었으나 2021년에 39만 4,709명으로 40.8% 증가했다. 무엇보다 데이터 관련 직무인력의 증가세가 놀랍다. 2015년에 7만 338명에서 2021년에 12만 2,431명으로, 불과 수년 사이 74.1%라는 경이적인 증가율을 보였다. 2020년과 2021년을 비교해도 20.1%나 증가했으니 놀라운 증가율이 아닐 수 없다. 일자리 창출에서도 데이터 산업이 우리나라의 유망산업이라는 증거는 이처럼 데이터로 볼 수 있다.

「2021 데이터 산업 현황조사」에 따르면 2020년도에 데이터 산업은 아니나 '데이터 산업 관련 업종'의 시장 규모는 106조 4,337억 원이다. 2020년 데이터 산업 시장 규모(20조 24억 원)보다 약 5.3배 정도 크며, 이들은 데이터 산업의 진흥에 큰 역할을 담당하고 있다. 데이터 산업 관련 업종으로 아래와 같은 3가지가 있으며, 이들의 2020년도 시장 규모도 매우 크다.

- SW 개발 및 공급업 시장 규모	59조 6,891억 원
- 컴퓨터 프로그래밍, 시스템 통합 및 관리업 시장 규모	25조 9,277억 원
- 정보서비스업 시장 규모	20조 8,169억 원
데이터 산업 관련 업종 총 시장 규모 합계	106조 4,337억 원

2021 데이터 산업 현황조사. 한국데이터산업진흥원

2020년도에 데이터 산업의 시장 규모(20조 24억 원)와 데이터 산업 관련 업종의 시장 규모(106조 4,337억 원)가 전체 우리나라 산업의 시장 규모(4,116조 9,953억 원)에 비하여 아직은 그 비중이 각각 0.49%, 2.59%로 작으나, 앞으로 그 비중이 더욱 커질 것은 자명한 일이다. 이들 산업은 4차 산업혁명 기술들의 핵심적인 기초 기술에 해당하므로, 데이터산업과 그 관련 산업의 발전은 매우 유망하다 할 수 있다.

02 데이터 산업의 국제 현황

그렇다면 데이터 산업의 국제 현황은 어떨까. 아래 표는 주요국(미국, EU, 일본, 브라질, 중국)의 2016~2021년 사이 데이터 산업 시장 규모를 보여주고 있다. 이 자료는 「2021 데이터 산업 현황조사」에서 발췌한 것으로, 원자료는 EU집행위원회가 발간한 「유럽 데이터 시장 연구(European Data Market Study)」에서 얻은 것이다.

표 1 2016~2021년 사이 주요국 데이터 산업 시장 규모(Billion Euro)

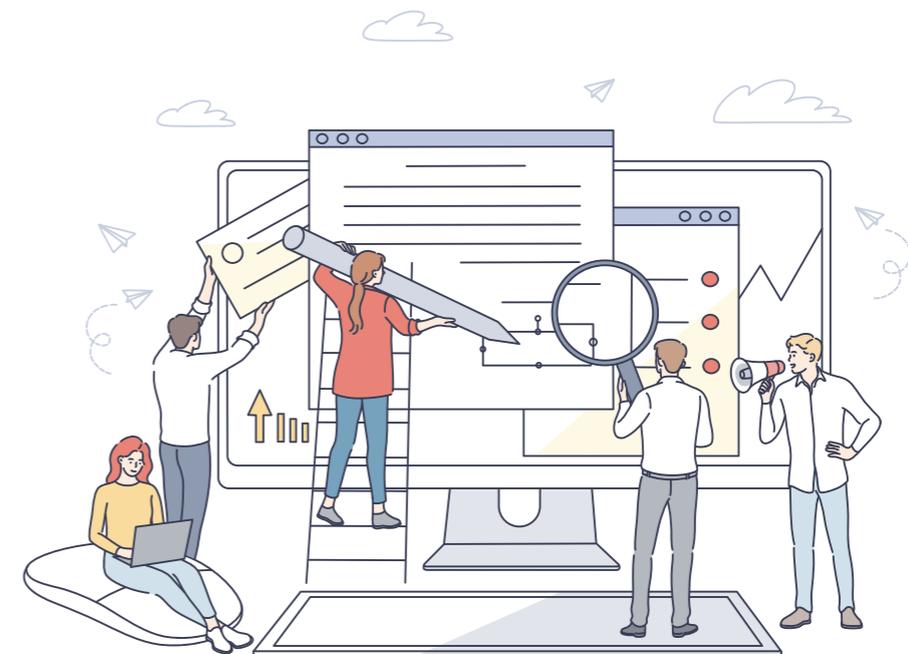
구분	2016	2017	2018	2019	2020	2021	5년 평균 성장률(%)
미국	129.2	147.0	164.0	184.9	213.5	240.0	13.2
EU	61.8	68.8	77.3	83.5	91.3	94.4	8.8
일본	25.5	26.7	29.8	32.9	36.6	40.0	9.4
브라질	6.0	7.0	7.4	7.9	8.4	8.9	8.1
중국	-	-	-	24.2	27.5	31.7	-
합계	222.5	249.5	278.5	309.2	377.3	414.8	13.3

EU집행위원회. 2022. The European Data Market Study 2021~2023. First Report on Facts and Figures IDC & The Lisbon Council.

이 자료에 따르면, 미국의 데이터 산업 시장 규모는 2021년 2,400억(추정치) 유로로 2020년 대비 12.4% 증가했으며, 5년(2017년~2021년)간 13.2%의 고속 성장이 이루어졌다. EU의 경우 시장 규모는 미국보다 작지만, 상황은 크게 다르지 않다. 2021년 기준 미국의 약 39.3% 수준의 시장 규모를 유지하고 있으나, 5년 평균 성장률이 8.8% 정도의 고성장을 유지하고 있다. 2021년 한국의 데이터 산업 규모(236조 972억 원, 약 1,760억 유로)는 미국의 데이터 산업 시장 규모 대비 약

7.3% 수준이고, 유럽의 데이터 산업 시장 규모 대비 약 18.6%인 것으로 산출된다. 국내 데이터 산업 역시 급성장하고 있으므로 조만간 미국에 비해 10% 수준까지는 상승하리라 여겨진다.

아쉬운 점은 적극적인 '글로벌화'가 부족하다는 점이다. 한국의 데이터 산업은 해외 진출이 적으며 주로 국내용이다. 한국 데이터 산업 중에서 해외 시장에 진출한 경험이 있는 기업은 6.9%에 불과하다. 해외 국가별 매출 발생 비중은 일본(36.6%), 베트남(32.4%)이 가장 크고, 다음으로 중국(28.2%), 미국(21.1%) 등의 순으로 나타나 아시아 국가들의 비중이 높다. 앞으로 데이터 산업에서 더욱 적극적인 해외 진출을 모색해야 할 것이다.



3

정책제언: 데이터 산업 육성 및 진흥 전략

- 빅데이터 도입률 제고
- 데이터 산업 인력 정예화
- 플랫폼 간 표준화의 필요성과 데이터 품질관리
- 데이터 과학·AI 국제협력센터 구축 서둘러야
- 데이터·AI 연구 융합 클러스터 조성 필요

데이터 산업 육성 및 진흥 전략



대한민국은 이제 선진국 반열에 올랐다. 명실공히 세계 경제에서 큰 비중을 차지하는 경제 대국으로 꼽힌다. 반만년 역사 속에서 대한민국이 이처럼 인구가 많고 잘 살았던 시기는 없었다. 우리나라보다 앞서 국민소득 3만 달러 이상 인구 5,000만 명 이상을 달성한 나라는 미국, 일본, 독일, 프랑스, 영국, 이탈리아뿐이다.

그러나 우리 한국은 지금 저출산, 과학기술계 투자 위축, 이공계 기피 등 다양한 불안 요소를 안고 있어 더 이상의 성장동력을 찾기 어렵다는 지적이 적지 않다. 지속적 발전을 통해 세계적 선진국으로 거듭날지, 다시금 개발도상국의 위치로 물러설지를 결정하는 '기로운 서 있다'고 보아도 과언은 아닌 형국이다. 이제 우리는 어느 방향을 향해 나아가야 할까.

과학기술적 관점에서 미래 사회의 기술적 방향은 상당 부분 예견돼 있다. AI와 메타버스, 로봇 기술이 하나로 합쳐진 4차 산업혁명 시대를 코앞에 두고 있다는 점이 그것이다. 따라서 앞으로는 철저한 데이터 중심사회로 나아갈 것이라는 점이 자명하다. 통계학을 무기로 한국의 산업화를 도왔던 박성현 교수 역시 이점에 집중한다. 이에 박성현 교수는 국내 데이터 산업의 진흥 전략을 아래 다섯 가지로 나누어 제언했다.

빅데이터 도입률 제고

미래는 AI의 시대라는데 이의를 제기할 사람은 그리 많지 않을 것이다. 그리고 AI의 근간은 통계학이다. 실제로 신경망 회로에서 다양한 연산을 할 때, AI의 판단은 통계 및 확률에 근거해 이뤄진다. 그리고 AI는 데이터에 기반해 학습하고 결론을 내린다. 따라서 '빅데이터'의 활용능력이 곧 경쟁력인 시대다.

그러나 국내 산업계의 빅데이터 도입은 아직 부족하여 제고가 필요한 실정이다. 2022년 한국 데이터 산업진흥원에서 실시한 「2021 데이터 산업 현황조사」에 의하면, 2021년 국내 기업의 빅데이터 도입률은 15.9%에 불과하다. 업종별로는 공공의 도입률이 55.4%로 가장 높게 나타났고, 민간 기업에서는 금융(40.0%), 통신·미디어(31.0%) 등의 순으로 높게 나타났다. 의료, 금융, 교육 등 도입 추진 중이거나 고려 중인 비율이 높게 나타났다. 종사자 규모별(공공 제외)은 100~299명 8.4%, 300~999명 16.4%, 1,000명 이상 53.8% 수준으로 기업체 규모가 클수록 도입률이 높은 경향을 보였다. 매출 규모별로는 1,000억 미만인 7.3%이며, 1,000억 이상 기업은 38.9% 도입하고 있는 것으로 조사됐다. 즉 규모가 큰 기업을 중심으로 빅데이터 도입률이 높아지고 있으나, 소규모 기업의 경우 상대적으로 도입이 더딘 것으로 나타나 불균형이 큰 상황이다.

전반적인 흐름을 볼 때, 공공, 그리고 대기업 이상의 빅데이터 도입률은 다소 높게 나타나고 있으며, 향후 국내 기업 빅데이터 도입률은 차츰 높아질 것으로 예상된다. 그러나 이러한 자연적인 증가세로는 현재 충분하지 않다. 도입률을 앞으로 크게 제고할 필요가 있다. 우리 기업들에서 빅데이터 도입률을 올리고, 그 활용률을 제고하는 것은 데이터 산업의 진흥에 크게 기여할 것이 자명한 만큼, 국가적으로 그 도입률을 제고할 필요가 있다. 중소기업의 도입률 역시 높일 수 있도록 다양한 지원 방안을 마련할 필요도 있다.



데이터 산업 인력 정예화

‘데이터 전문 인력’의 양성은 국가적 과제가 되고 있다. 이미 인력 부족 문제가 현실로 나타나고 있다. 「2021 데이터 산업 현황조사」에 의하면 데이터 산업에 종사하는 인력은 크게 부족한 것으로 나타났다. 향후 5년까지(2026년) 데이터 산업에서 추가로 필요한 데이터 직무인력은 1만 6,984명이며, 일반산업을 포함하는 전 산업에서 추가로 필요한 데이터 직무인력은 총 2만 4,409명으로 조사됐다. 전체적으로 현재의 12.2%가 부족한 셈이다. 향후 5년까지 데이터 직무별 필요 인력 비중이 가장 큰 직무는 데이터 개발자이며, 인력 부족률이 가장 큰 직무는 데이터 과학자로 나타났다. 즉 데이터 산업의 필수 인력이라 할 수 있는 데이터 과학자(36.5% 부족), 데이터 개발자(16.0% 부족), 데이터 분석가(15.6% 부족) 등이 심각한 인력 부족에 시달리고 있는 것이다. 심지어 이 문제는 해마다 심각해 지고 있다. 한 해 전 조사인 「2021 데이터 산업 현황조사」와 비교해 데이터 과학자(31.4% 부족), 데이터 개발자(14.5% 부족), 데이터 분석가(14.5% 부족) 등의 부족 수치가 모두 증가했다. 데이터 과학자 등을 충분히 확보하여 인력 정예화를 시키고 생산성을 제고 하는 것은 데이터 산업 진흥의 첩경인 형국이다.

표 2 데이터직무 인력 수요

(단위: 명)

구분	데이터산업				전 산업*			
	2022년 필요인력	향후 5년까지(2026년) 필요인력		인력부족률	2022년 필요인력	향후 5년까지(2026년) 필요인력		인력부족률
인력수	비중	인력부족률	인력수		비중	인력부족률		
데이터 아키텍트	140	446	2.6%	4.6%	325	824	3.4%	5.4%
데이터 개발자	4,201	8,035	47.3%	16.0%	4,761	9,247	37.9%	15.2%
데이터 엔지니어	1,021	2,131	12.5%	10.1%	1,305	2,864	11.7%	11.1%
데이터 분석가	596	1,744	10.3%	15.6%	1,170	2,821	11.6%	14.7%
데이터베이스관리자	392	825	4.9%	4.5%	1,827	3,765	15.4%	7.7%
데이터 과학자	483	1,738	10.2%	36.5%	630	1,989	8.1%	33.3%
데이터 컨설턴트	474	1,062	6.3%	10.6%	570	1,302	5.3%	11.1%
데이터 기획자	424	1,004	5.9%	7.3%	688	1,598	6.5%	8.9%
전체	7,731	16,984	100.0%	12.2%	11,274	24,409	100.0%	11.9%

출처: 데이터 직무 인력수요. 한국데이터산업진흥원.

플랫폼 간 표준화의 필요성과 데이터 품질관리

정보란 흐름이다. 따라서 그 통로를 확보하지 않으면 의미가 없어진다. 이에 플랫폼 간 메타 데이터(meta data, 데이터를 효율적으로 이용하기 위해 구조화한 데이터)를 공유·활용할 수 있도록 표준화할 필요성이 생겨났다. 이 과정에서 데이터의 가치를 높이기 위한 품질관리(quality control) 역시 필요해졌다. 산업 각 분야별로 많은 종류의 빅데이터가 작성되어 있으나 이들 간에 데이터 표준화가 미비해 통합 작업이 어렵고, 정보 교환에 애로가 많은 것이 현실이다. 누구나 쉽게 플랫폼들에 축적된 데이터를 찾고 상호 비교와 활용이 가능하도록 할 필요가 있다. 양질의 데이터를 확보하는 것은 데이터 산업 진흥에 필수적인 요소이며, 그 과정에서 플랫폼 간 표준화와 데이터 품질관리는 필수적이다.

데이터의 품질은 정확성, 시의성, 비교성, 접근성, 편의성 등에 의해 결정된다. 이런 품질의 가치를 유지하기 위해서는 데이터에 대한 지속적인 품질관리가 요구된다. 통계에서는 ‘쓰레기 데이터가 들어가면 쓰레기 정보만이 나온다(garbage data in, garbage information out)’란 말이 있다. 품질관리가 안 된 데이터가 들어가면 좋은 정보 창출이 될 수 없다는 의미이다. 따라서 데이터를 생산하는 조직은 데이터의 전 생명주기에서 데이터의 품질관리에 대한 책임을 수행해야 할 필요가 있다.



데이터 과학·AI 국제협력센터 구축 서둘러야



AI의 근간은 데이터다. 따라서 데이터과학의 발전 없이는 AI 분야 역시 성장하기 어렵다. 한국이 국제적 AI 중심국가로 거듭나기 위해서는 ‘인재가 몰려드는 나라’를 만들어야 한다. 국가적 관점에서 전문가, 특히 과학기술인들의 이주를 많이 받는 등 적극적인 두뇌 유치 전략을 펼쳐야 하고, 더 나아가 한국의 과학·산업계가 ‘데이터과학 및 AI 분야 과학기술인들의 사랑방’ 역할을 해야 한다.

그만한 투자 역시 필요하다. 우선 데이터과학 및 AI 분야의 국제적인 저명한 전문가들이 한국에 단기간이라도 체류하면서 국내의 전문가들과 교류가 이루어지도록 할 필요가 있다. 이런 교류를 국가에서 파격적으로 지원하는 가칭 ‘데이터 과학·AI 국제협력센터(Data Science·AI International Cooperation Center)’를 만들 필요가 있다. 이 센터에서 전문가들 간에 국제교류가 이루어지도록 하고, 저명 외국인 전문가들이 국내 대학, 연구소, 기업들과 협력 교류를 할 수 있으면 국내 데이터과학·AI의 발전은 물론 데이터 산업 육성에도 촉매제 역할을 다 할 수 있으리라 예상된다.

데이터·AI 연구 융합 클러스터 조성 필요



산업이란 과학기술의 토대 위에 성립한다. 이는 분야를 가리지 않는다. 앞으로 데이터 과학 및 AI 분야는 최근 우리나라의 급성장 분야인 방위산업, 우주항공산업, 문화산업, 원전산업 등 모든 분야를 뒷받침하는 기반이 될 것이다.

이 과정에서 데이터과학과 AI를 전문으로 연구할 ‘클러스터’의 필요성 역시 대두된다. 기업, 대학, 연구소, 공공기관 등이 한 곳에 모여 데이터·AI 관련 연구 개발 및 인력 양성, 창업 등을 지원할 수 있는 가칭 ‘데이터·AI 연구 융합 클러스터’를 조성하자는 것이다.

이를 통해 데이터의 수집, 연구, 활용 등이 활발히 일어나도록 하고, 그 성과를 AI 연구, 더 나아가 국내 모든 과학기술 및 산업과 접목하는 방안을 연구할 필요가 있다. 실제로 미국의 경우에는 국립과학재단(NSF)의 지원으로 2017년부터 4개 빅데이터 지역혁신허브(Big Data Hub)를 만들어 산업계, 학계, 지방정부 등이 적극적인 협업으로 빅데이터 허브를 만들어 운영하고 있다. 이를 벤치마킹하여 우리나라도 대략 4개권(수도·강원권, 충청권, 경상권, 전라·제주권)으로 나누어 빅데이터·AI 지역혁신허브를 구축하는 것이 바람직하다. 2019년에 정부에서 발표한 ‘데이터·AI 경제 활성화 계획’을 보면, 2023년까지 전국에 빅데이터 센터 100개 구축을 목표로 하고 있다. 전국에 4개 ‘빅데이터·AI 연구 융합 클러스터’를 두고, 이들 클러스터가 지역별로 다수의 빅데이터 센터를 관리하게 하는 것이 좋은 방안이 될 것이다.

4

시대에 걸맞은 '산업형 과학기술인'
육성시스템 마련해야

GREAT SCHOLAR CAREER DECISIONS

시대에 걸맞은 '산업형 과학기술인' 육성시스템 마련해야



현대를 이른바 '과학기술 패권시대'라고 부른다. 기술을 매개로 새로운 국제질서와 동맹이 만들어지는 시대를 일컫는다. 기초과학을 더욱더 깊이 있게 발전시키는 것은, 마땅히 인류의 발전과 존속에 있어 필수 불가결한 요소이다. 그러나 산업의 발전에 있어 당면의 요소는, 현재 인류가 현재까지 확보한 과학의 영역을 충분히 이해하고, 이를 산업에 적용할 수 있도록 돕는 전문가의 역량이다. 이 과정에서 산업 분야 기술 확보를 위한 과학기술 전문가의 확보는 중차대한 국가 전략으로서 관리될 필요가 있다.

산업의 성공 사례를 꼽을 때 과학기술인이 활약한 사례는 이루 헤아리기 어려울 정도로 많다. 기초과학의 토대를 받아들여 그 위에서 산업을 발전시키는 것은 실로 기초과학에 대한 깊은 이해와 응용이 필수적이기 때문이다. '코로나19 백신 개발' 과학자 우그르 사힌의 사례, P&G사의 세탁용 세제 개발 사례, 일본의 기업 과학기술인들의 기초과학연구 사례 등을 살펴볼 때 이런 점은 일목요연하다고 할 수 있다. 수학과 통계학을 산업에 적용한 대표적 인물로는 미국의 데밍 박사와 한국의 박성현 교수를 꼽을 수 있다. 이 두 사람은 각각 미국과 한국에서 통계적 품질관리와 품질경영 방법을 보급해 양 국가의 산업화를 이끈 공로자이다.

앞으로는 통계학 분야가 더욱 발전한 '데이터과학' 분야로 크게 주목받을 것이 자명하다. AI, IoT, 로봇 등 이른바 4차 산업혁명의 시대 기술의 핵심은 바로 데이터 처리에 있기 때문이다. 즉 우리나라의 미래를 대비하기 위해서는 통계학의 발전된 형태인 '데이터과학' 전문가를 집중적으로 육성하고, 이들이 산업계에서 활약하도록 그 길을 열어줄 필요가 있다.

데이터 산업은 전 세계적으로 5년 평균 성장률 13.3%에 달할 정도로 급속도로 발전하고 있다. 데이터 산업의 시장규모는 지난 12년간 매년 급성장해 왔으며, 더구나 이런 추세는 앞으로도 계속 이어질 것이 확실시되고 있다. 따라서 데이터 산업에 종사할 인력 수요 역시 크게 늘어나고 있다. 반면 전문 인력은 대단히 부족하며, 그 과정에서 과학적 지식을 제공할 전문 과학기술인의 중요도는 이루 말할 수 없이 높다고 할 수 있다. 데이터 개발자, 데이터베이스 관리자, 데이터 엔지니어, 데이터 분석가, 데이터 기획자, 데이터 컨설턴트, 데이터 아키텍트 등의 전문인력 수요는 크게 늘어날 것으로 보이는데, 그 근간이라 할 수 있는 데이터 과학자가 부족해질 경우 산업의 뿌리가 흔들릴 수 있음은 자명한 일이다.

이미 인력 부족이 현실로 나타나고 있다. 2021년 전 산업에서 추가로 필요한 데이터 직무 인력은 총 2만 4,409명으로 조사돼, 현재 12.2%의 부족률을 보였다. 더구나 핵심 인력이라 할 수 있는 데이터 과학자의 경우 부족률이 36.5%에 달해 대안 마련이 시급한 것으로 나타났다.

이에 대해 보고서의 제안자인 박성현 교수는 '데이터 산업 육성 및 진흥 전략'으로서 크게 5가지 대응책을 제시했다. 먼저 국내 산업계의 **△빅데이터 도입률 제고**를 꼽았다. 전체적으로 도입률이 미비하다(15.9%). 매출 규모별로는 1,000억 미만 중소기업의 도입률이 미진한 점(7.3%) 등의 불균형 해소 등도 필요하다고 지적했다. 두 번째로 **△데이터 산업 인력 정예화**를 꼽았다. 데이터 과학자 등을 충분히 확보하여 인력을 정예화 시키고 생산성을 제고 하는 것은 데이터 산업 진흥의 첩경이라고 지적하고 있다. 세 번째로 꼽은 것이 **△플랫폼 간 표준화의 필요성과 데이터 품질 관리**이다. 플랫폼 간 메타데이터의 공유·활용이 가능하도록 표준화하고, 데이터의 가치를 높이기 위한 품질관리 방안을 마련하자는 주장이다. 네 번째로 **△데이터 과학·AI 국제협력센터 구축**해 국내외 전문가들과 교류가 활발히 이뤄지도록 할 필요 역시 있다고 제안했다. 다섯 번째로 기업, 대학, 연구소, 공공기관 등이 한 곳에 모일 수 있도록 **△데이터·AI 연구 융합 클러스터 조성**해야 한다고 했다. 데이터·AI 관련 연구 개발 및 인력 양성, 창업 등을 지원할 수 있는 '지역혁신허브'를 구축하자는 것이다.

박성현 교수는 이 보고서를 빌어 자신이 일생을 바쳐 온 '통계학' 분야 후배들에게도 조언을 남

겼다. 우선 통계학을 전공하는 학생들에게는 '학문의 중요성이 점점 높아지고 있으니, 더더욱 연구에 몰두해 주길 바란다'고 했다. 최근 데이터과학이 발전하면서 통계학의 영역이 대단히 넓어지고 있으니, 박사과정 이후의 학생들은 앞으로 데이터 과학자로서 활약할 수 있도록 '빅데이터' 분석능력 역시 갖추도록 조언했다. 과거의 통계학자들은 빅데이터 분석능력에 대해 엄두를 내지 못했으나 앞으로는 유능한 통계학자가 유능한 데이터 과학자가 되어야 하며, 그렇게 되기 위해서는 빅데이터 분석 능력이 필수라고 강조했다. 통계학 분야 교수와 교육자들에게는 "좀 더 다방면으로 연구하라"는 말을 남겼다. 그는 "시대가 시대이니만큼 통계학을 다방면에 활용하는 방법을 연구하고, 이를 가르치는데 거리낌이 없어야 한다"며 "수리통계학 등 분야를 학문적으로 깊이 파고드는 것은 물론 필요하지만, 통계학을 비롯해 모든 학문의 의미는, 배우고 익힌 것을 현실에 응용해 새로운 가치를 창출하는 데 있다는 점을 잊어서는 안 된다"고 했다.

방법론 면에서는 앞으로 모두의 고민이 필요하다 여겨진다. 과학기술계의 지식을 산업계로 빠르게 전달할 수 있는 효율적인 체계를 만들어 나갈 필요가 있다. 교육 단계에서부터 산-학-연 공동 시스템이 확고히 작동할 수 있도록 길을 닦아 주어야 한다. 나아가 주니어 단계의 연구자 때부터 산-학-연 간 교류에 활발히 참여할 수 있도록 하여야 한다. 미래 세대에 적합한 '데이터 과학기술인'들이 자연스럽게 산업계로 유입될 수 있는 시스템 마련에 총력을 기울여야 할 시점이다.

이 과정에서 해외 인재의 국내 유입을 유도하는 것 역시 고려해야 한다. 개발도상국 인재의 한국 이주를 적극적으로 유도해야 하며, 이들이 국내에서 활약할 때는 산-학-연 연계시스템 내에서 활약할 수 있도록 우선하여 유도할 필요가 있다. 이 같은 시스템은 현재 문제가 되고 있는 두뇌 유출(Brain Drain) 문제에도 다소나마 해결책이 될 수 있을 것이다.



5

통계학자로서의 철학과 삶

- 학문적 업적
- 산업교육과 연구
- 과학기술계 활동

통계학자로서의 철학과 삶



박성현 교수는 삶은 '중용(中庸)'이라는 한 단어로 정리할 수 있지 않을까 여겨진다. 기초과학의 바탕이랄 수 있는 통계학을 공부했지만 이에 치우치지 않고 산업계에도 눈을 돌려 그 역량을 십분 발휘하였고, 국가통계의 발전과 과학기술 정책 활동 역시 왕성하게 수행했다. 그의 이 같은 '삶의 이야기'를 보고서의 말미에 정리해 두는 것은 국가적 석학의 업적과 역사를 남겨 둔다는 점에서 큰 의미가 있다고 여겨진다.

그가 이룬 업적은 크게 세 가지로 나뉜다. 첫째는 일생 연구한 학문을 다수의 책과 논문으로 정리하여 대중에 전파한 점이다. 그의 저서는 통계학 관련 후배 양성의 토양이 됐고, 나아가 국내 산업계 전문가들이 통계학을 산업에 적용하는 밑바탕이 됐다. 두 번째는 산업 분야 실무자들을 위해 통계학의 산업화 현장 교육, 그리고 기업컨설팅에 팔을 걷어붙이고 나섰다는 점이다. 세 번째로 정년 후에는 과학기술 정책에 관심을 가지고 수없이 많은 기고를 통해서 과학기술 정책 입안에 기여했다.

학문적 업적

박성현 교수의 저술 활동은 1977년 그가 미시시피주립대학교 경영대 조교수를 마치고 서울대학교 통계학과 조교수로 처음 부임하던 첫해부터 시작됐다. 당시 그가 귀국하여 국내에 발간된 통계학 교재들을 살펴보니 대학에서 쓸만한 좋은 책이 거의 없었다. 통계학이 국내에 도입된 초창기 인지라, 통계학 개론 책은 조금 있었지만, 통계학 분야의 전공 서적이라 할 수 있는 책이 거의 없어 영어로 된 원서를 사용할 수밖에 없는 실정이었다. 이에 '당장 급한 대로 후배들을 가르칠 수 있는 통계학책을 써야겠다'고 결심하였고, 1979년 1월부터 전공 분야인 '회귀분석'과 관련된 책을 쓰기로 마음먹었다. 그는 통계학 분야 중에서도 △회귀분석 △실험계획법 △통계적 품질관리 3가지 분야를 선택하였는데, 이는 그의 전공 분야와도 가깝고, 국내에 전문서적이 없었기 때문이다. 그는 이 3가지 분야에 대해서 만큼은 제대로 된 책을 국내에 남겨야겠다고 결심했다.

당시는 컴퓨터나 워드프로세서가 없어 원고지에 펜으로 꼭꼭 눌러 써 집필을 하던 시절이었다. 박성현 교수는 막연히 책을 써야 한다는 생각에, 출판사와 계약도 하지 않고 자기 스스로 일 년 반 정도 집필활동을 계속해 1980년 6월에 200자 원고지 2,500매에 달하는 두꺼운 회귀분석 책 원고를 탈고하는 데 성공한다. 이 책을 자연과학 서적 출판으로 유명한 '박영사'에 연락해 출판을 부탁했는데, 출판사 담당 부장이 원고지 문치를 가지고 갔다가 한 달 만에 돌아와서 한다는 말이 '너무 전문적인 책이라 잘 팔리지 않을 것 같다'는 부정적인 답변이었다. 책을 출간하지 않겠다는 소리였다. 당연히 박성현 교수의 실망도 이만저만한 것이 아니었다. 그런데 당시 실망하고 있던 박성현 교수에게 한 청년이 찾아와 '그 책을 내가 출간하게 해 달라'고 요청해 오기에 이른다. 그 청년은 박영사의 신입사원이었는데, 그 책이 자기가 보기엔 잘 팔릴 것 같으니, 출판사를 새로 만들어서라도 출간을 하겠다고 나선 것이다. 그 출판사가 지금은 통계학책을 많이 내기로 유명한 '민영사'이다. 그 신입사원의 예상은 기가 막히게 맞아 들어가 박성현 교수의 '회귀분석' 책은 날개 돋친 듯 팔려나갔다. 당시 많은 대학에 통계학과가 설치되기 시작했는데, 교재로 쓰일 책이 전혀 없었던 탓이다. 그 책은 회귀분석 분야의 스테디셀러로 아직도 꾸준히 판매되고 있다.

이후 박성현 교수는 실험계획법과 관련된 책을 쓰기 시작해 이 책도 1982년 민영사에서 '현대 실험계획법'이란 제목으로 출판하였고, 곧 품질관리에 관한 책도 쓰기 시작해 1984년도에 민영사를 통해 '통계적 품질관리'란 제목으로 출판했다. 이런 책들은 통계학과뿐 아니라 산업공학과 학생들도 공부하는 책이고, 이외의 기업에서 품질관리 기사가 되고 싶은 사람들이 반드시 공부해야 하

는 과목이므로, 2010년대 후반까지 오랫동안 각 분야에서 가장 인기 있는 책이었다. 박성현 교수는 지금도 이 책 3권을 '일생의 역작'으로 생각하고 있다고 했다.

이후 박성현 교수는 끊임없이 책을 저술해 왔다. 그가 단독으로 저술한 책만 28권. 공저, 영문 저서 등을 모두 합하면 총 93권에 달한다. 어지간한 사람은 일생 책 한 권을 쓰기 어렵다는 점을 생각하면 실로 대단한 업적이 아닐 수 없다. 박성현 교수는 무엇보다 이런 책들이 우리나라의 통계학 저변확대 및 산업화 응용과정에 크게 이바지하였을 것이라는 점에서 큰 보람을 느낀다고 했다.

박성현 교수가 저술 활동에만 매달린 것은 아니다. 학자가 해야 할 일의 기본 중 하나가 학문의 정보교류 수단, 이른바 '논문'을 생산하는 일일 것이다. 이 분야에서도 박성현 교수의 업적은 실로 지대하다. 총 논문 수는 178편. 이중 국내 논문은 114편이며 국외 논문이 64편에 달한다. 국제논문 색인(SCI) 급 논문 수도 60여 편에 달해 논문의 절대적 수준 역시 대단히 높다고 할 수 있다.

박성현 교수의 연구 주제를 살펴보면 회귀분석과 실험계획법 외에 반응표면분석, 품질경영, 식스 시그마 등, 통계학을 산업에 응용하는 주제의 논문이 적지 않다. 이 밖에 통계학을 응용해 사회 문제를 객관적으로 풀어나가는 방안을 제시하는 논문이 상당수 눈에 들어온다. 1980~1990년대에 주로 이 같은 연구를 많이 진행했다. 활발히 연구하던 당시에는 일 년에도 수차례씩 학술발표를 위해 해외를 나가고는 했다. 미국통계학회, 미국품질학회, 국제통계학회, 국제품질아카데미 등의 초청을 자주 받았다고 했다. 이런 박성현 교수의 학문적 노력은 사회적으로 대단히 큰 인정을 받아 왔다. 국가를 대표할만한 과학자로 인정받아 다양한 대외활동으로도 이어졌다. 박성현 교수는 미

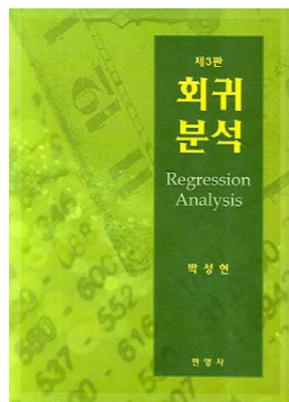
국통계학회와 미국품질학회의 석좌 회원(fellow)이며, 국제품질아카데미의 석좌 회원(academician)이기도 하다.

박성현 교수는 2000년대 들어서부터 '데이터로 보는 세상', '데이터로부터 해결의 실마리 찾기' 등 데이터로부터 정보를 얻고, 이를 이용해 해결방안을 제시하는 방법론 등에 관심을 두고 연구하고 있다. 이런 연구는 이른바 '빅데이터 분석' 기법 등과도 연관이 큰 방식이며, 향후 AI 관련 기법의 발전에도 큰 도움이 될 것으로 박성현 교수는 기대하고 있다.

박성현 교수의 숨은 업적 중 하나로 우리나라의 통계학 기틀을 다진 점 역시 빼 놓을 수 없다. 우리나라 대학에 통계학과가 들어서기 시작한 것은 1960년대로 거슬러 올라간다. 당시 이

미 5개 국내 사립대학에 통계학과가 설치되어 있었으나, 전문적으로 통계학을 전공한 교수진이 활약하기 시작한 것은 1970년대 들어서였으며, 실제로 '한국통계학회'가 운영을 시작한 것도 1973년이다. 서울대학교에 '계산통계학과'라는 명칭으로 국립대학교에 처음 통계학과가 설치된 것이 1975년이라는 것을 생각하면, 대한민국의 통계학 역사는 이제 겨우 50년 남짓 지났을 뿐이다. 박성현 교수가 서울대 교수로 부임한 것이 1977년이니, 그는 한국의 통계학 역사를 두 눈으로 지켜봐 온 '산 증인'이라고 해도 과언은 아니다.

서울대를 거친 이후 박성현 교수 자신도 "나의 학문의 길에서 특별히 의미 있게 기록하고 싶은 것은 바로 13대 한국통계학회 회장직을 수행한 것"이라고 이야기할 정도이다. 당시 박성현 교수는 한국통계학회장을 지내면서 공식통계연구회, 통계계산연구회, 통계교육상담연구회, 조사통계연구회, 생물통계연구회, 공업통계연구회 등 다양한 연구회의 발족 및 활동을 지원했다. 지회 활동도 적극적으로 지원했다. 기존의 대구·경북 지회, 부산·경남 지회, 호남·제주 지회, 강원·경기·인천 지회에 이어 북미주지회 발족을 지원하고 연구회와 지회가 모여 각종 학술발표회, 기념행사, 산학 간 협력 사업 등을 하도록 이끌기도 했다. 한-일 공동 통계 학술회의를 지속해서 개최하도록 지원하기도 했다. 한편으로 통계학 용어집을 발간하도록 해 통계 용어의 표준화를 이끌었다.



박성현 교수의 첫 저서, '회귀분석'. 40년 이상이 지난 지금까지도 꾸준히 팔리고 있는 통계학 분야 스테디셀러이다.



1987년 서울대 계산통계학과 교수진(전산과학 4명, 통계학에 8명)

산업교육과 연구

박성현 교수의 업적 중 결코 빼 놓을 수 없는 부분이 '대외강연'이다. 박성현 교수는 산업 인력들이 통계를 업무에 제대로 적용하지 못하는 것을 크게 안타까워했고, 이를 교육을 통해 바로 잡고자 했다. 이에 박성현 교수는 '품질관리기사(현 품질경영기사)' 제도에 주목했다. 당시 모든 공장에는 규모에 따라 품질관리 기사를 반드시 배치하도록 법으로 정하고 있었고, 실제로 품질관리기사가 되기 위한 시험 과목 중 통계적 품질관리, 실험계획법 등 통계학과 관련된 과목이 포함돼 있었다. 엔지니어들은 품질관리기사가 되기 위해 이런 과목을 공부할 필요가 있었고, 따라서 다양한 교육 수요가 있었다. 주로 품질관리기사 양성과정 등을 운영하는 주요 기관, 즉 한국표준협회 등에서 그를 강사로 초빙하는 경우가 대단히 많았다. 1980년대부터 대학의 수업에 방해가 되지 않는 한 전국을 누비면서 이런 강연에 응하곤 했다. 울산공단, 창원공단, 구미공단 등 엔지니어들이 있는 곳을 찾아다니며 수없이 많은 통계 강의를 하곤 했다고 하니, 그가 대한민국 산업계의 체질 개선에 미친 영향은 대단히 크다고 할 수 있다.

그에게 기업들이 컨설팅을 요청하는 사례도 적지 않게 있었다. 그중에서도 박성현 교수가 대표적 사례로 꼽는 것이 쌍용정유(현 에쓰오일) 품질개선 건이다. 쌍용정유는 당시 석유제품의 품질이 생각만큼 높게 나오지 않아 고심하고 있었다. 이러한 사실을 전해 들은 박성현 교수는 즉시 산업 프로세스 전반을 살펴보았는데, 데이터에 기반을 둔 통계적 품질관리가 전혀 되지 않고 있다는 사실을 알아내고 즉시 컨설팅에 들어갔다. 당시 쌍용정유에서 하던 일은 외국 회사의 품질관리 소프트웨어를 매년 비싼 사용료를 내고 가져다 쓰는 정도였으며, 그 소프트웨어에서 지시하는 값에 맞춰 품질관리를 하고 있었다. 그러나 국내 실정에 맞지 않게 만들어진 소프트웨어라 최종 제품의 품질이 그리 좋지 않은 것은 어찌 보면 당연했다. 이에 박성현 교수는 석유제품의 품질을 결정하는 옥탄가 등 축적된 데이터로부터 그것을 끄집어내 통계적으로 분석하고, 모형을 개발하여 그 값에 따라 쌍용정유에서 컴퓨터 소프트웨어를 자체적으로 개발하도록 이끌었다. 그렇게 개발한 국산 소프트웨어는 그동안 사용하던 외국산 소프트웨어보다 훨씬 정확했고, 쌍용에서 생산하던 석유제품 품질도 월등히 높아질 수 있었다.

비슷한 사례가 또 있다. 박성현 교수는 현대전자(현 SK하이닉스)의 반도체 생산공정에서도 비슷한 기여를 한 바 있다. 당시 현대전자는 반도체 생산공정에서 수율이 경쟁사보다 낮아 고심이 컸다. 이야기를 듣고 간 박성현 교수가 생산공정 전체의 데이터를 분석하고, 수율 부분의 데이터를

통계적으로 분석해 공정개선에 대해 조언했다. 그 결과 현대전자의 반도체 생산 수율은 경쟁사 못지않게 높아질 수 있게 됐다. 이만한 규모 사업체의 수율을 개선한 건 국가적으로도 무시할 수 없는 성과이다. 박성현 교수에 따르면 이런 사례가 적지 않았다고 한다. 쌍용정유와 현대전자 이외에도 대표적인 사례를 들면 삼성전자(반도체 공정관리를 위한 통계 로직 개발), 통계청(통계품질관리 표준 매뉴얼 작성), 삼성코닝(식스 시그마 공정데이터의 해석 패키지 개발), 삼성인력개발원(자기학습 통계프로그램 개발) 등이다.

박성현 교수의 연구 일생을 이야기하면서 '산업통계학' 혹은 '공업통계학'을 빼놓을 수는 없다. 그가 학문 분야에서 일생 가장 큰 관심을 두었던 부분이 '공업통계학의 발전'이다. 통계학은 이론통계학과 응용통계학으로 구분할 수 있다. 이론통계학은 당연히 통계학 그 자체를 고도화하고, 발전시킬 방법을 연구한다. 과학 발전의 기초 중의 기초에 해당하는 학문이라 할 수 있다. 자신은 서울대 자연과학대학 소속으로, 마땅히 이론통계에만 집중하면 되는 입장이었다. 그러나 그는 여기서도 중용의 시각에서 산업에 적용할 수 있는 '응용통계학', 그중에서도 공업통계학의 발전에 관심을 두고 큰 노력을 기울여 왔다.



네팔에서 개최한 ANQ(Asia Network for Quality) 회의 강연 후

응용통계의 분야는 다양하다. 경제통계, 사회통계, 공업통계, 의료통계, 심리통계 등 다양한 분야가 존재하는데, 박성현 교수는 그중에서도 자연스럽게 공업통계를 더욱 중시했었다고 한다. 이는 박성현 교수가 학사 과정에서 화학공학을, 석사과정에서 산업공학을 전공한 점도 큰 영향을 끼쳤지만, 그보다 공업통계를 보급하는 것이 곧 우리나라의 산업사회 발전에 크게 이바지할 수 있을 것이라는 믿음 때문이기도 했다. 우리나라는 제조업을 기반으로 하는 산업국가다. 그리고 공업통계는 제조기업에서 생산과정에서 발생하는 수많은 데이터를 보고 생산, 품질, 공정 등을 관리하는데 요긴하게 사용할 수 있는 강점이 있다.

이런 점에서 박성현 교수는 일생의 많은 부분을 책 저술, 강연, 연구, 논문작성 등으로 공업통계학의 발전에 이바지했다. 그가 쓴 저서 중 절반 이상은 공업통계에 관한 것들이며, 베스트셀러 반열에 오른 다양한 공업통계 관련 저서들도 모두 공업통계에 관한 내용을 담고 있다. 11권에 달하는 영문 저서는 모두 공업통계 관련 저서들이다. 대표적인 저서들을 소개하면 다음과 같다.

| 공업통계 관련 대표 저서 |

- 통계적품질관리(민영사)
- 공업통계학(민영사)
- 응용통계(한국방송통신대학교. 이태림 공저)
- 다구찌방법을 중심으로 한 응용실험계획법(영지문화사)
- 품질공학(민영사)
- 통계적 공정관리(민영사. 박영현, 이명주 공저)
- 품질경영(한국방송통신대학교. 백재욱 공저)
- 6 시그마 이론과 실제(한국표준협회. 이명주, 정목용 공저)
- 6 시그마 설계를 위한 DFSS(한국표준협회. 이명주, 이강군 공저)
- 통계적 품질관리와 6 시그마의 이해(민영사. 박영현, 이재영 공저)

| 공업통계 관련 영문 저서 |

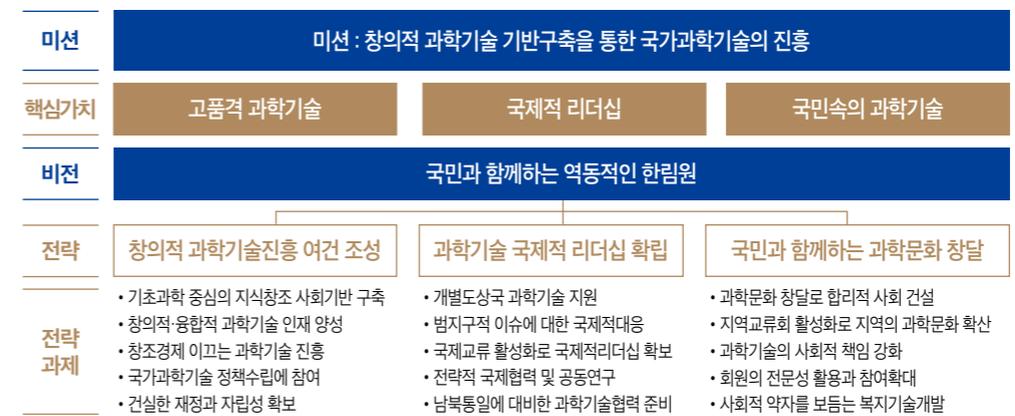
- Robust Design and Analysis for Quality Engineering, Chapman & Hall, 1996.
- Statistical Process Monitoring and Optimization, Marcel Dekker, 1999.
- Six Sigma for Quality and Productivity Promotion, Asian Productivity Organization, 2003.
- Robust Design for Quality Engineering and Six Sigma, World Scientific, 2008.

과학기술계 활동

박성현 교수는 2010년 정년 이후 학문적 연구와 산업 현장 교육 및 컨설팅 일을 줄이면서 과학기술 분야의 연구자들과 국가의 과학기술 정책을 돕는 일에 참여하기 시작했다. 이때 시작한 일이 '한국연구재단 기초연구본부장'이었다. 대학 또는 정부출연연구기관 등에 근무하는 기초연구자들에게 정부의 정부 예산을 지급하는 일이다. 대전은 국가 과학기술의 요람이라 할 수 있다. 한국과학기술원(KAIST)을 중심으로 다수의 대학, 연구기관 등이 한 동네에 모여있는 대한민국 대표 사이언스 타운이다. 이런 상징적 의미 때문에 연구재단 역시 이곳에 자리하고 있다. 따라서 박성현 교수 역시 재임 기간 대전에 내려가 오피스텔에 묵으며 기초연구본부 업무를 수행했다. 기초연구본부장은 행정적으로 막중한 자리이다. 즉 약 1조 원에 달하는 국가 연구비가 결정되는 중요 보직이다.

이후 박성현 교수는 2013년부터 3년간은 '한국과학기술한림원(한림원) 원장' 직을 맡아 수행하였다. 박성현 교수는 이때를 "일생에서 가장 보람 있는 시간 중 하나"라고까지 평하고 있다. 대한민국 과학기술 분야 모든 석학을 대표하는 자리이니 국가 간 협력 자리에 나설 때도 많았다. 미국, 독일, 프랑스, 영국, 스웨덴, 이탈리아 등 많은 주요 선진국을 방문하였고, 한국과의 과학기술 교류 증진을 위해 노력했다. 많은 과학기술포럼을 개최하는 동시에, 국가 과학기술 석학기관으로서 정부에 국가의 과학기술 정책 방향을 제시하기도 했다. 그는 한림원 원장으로 있으면서, 한림원의 비전 및 전략 체계를 처음으로 만들어 기관의 지속 가능한 발전 토대를 마련한 점이 특히 보람찼던 일이었다고 했다. 이 비전과 전략은 지금도 그대로 남아 한림원 발전의 토대가 되고 있다.

그림 '한국과학기술한림원의 미션-비전-전략' 체계도



하지만 그가 한림원 원장 재임 시절 가장 자랑하는 업적은 따로 있는데, 그것은 ‘과학기술유공자 예우 및 지원에 관한 법률’을 기획하고 입안해 국회에 실제로 제안했던 일이다. 이 법안은 2017년 실제로 통과됐다. 과학기술인들을 우대할 수 있도록 우리나라 법률을 개정하는데 한림원이 밀착 작업을 했고, 그 결과 과학기술인들이 조금이라도 살기 좋은 세상이 만들어지는데, 한림원이 일조하도록 그가 이끈 것이다.

박성현 교수는 2012년 3월부터 3년간 다시 학계 일을 보기도 했다. 3년간 건국대학교 기술경영학과 석좌교수직을 맡아 과학기술 정책, 통계학 분야(통계가 과학기술경영 선진화에 미치는 영향 등)에 대해 주로 강의했다.

2013년 9월부터 박성현 교수는 국가를 위해 중요한 일을 할 기회를 얻게 된다. 대통령 직속 과학기술자문회의 자문위원으로 2년간 근무하게 된 것이다. 특히 과학기술기반 분과 의장을 맡아 국가 발전을 위해 다양한 과학기술 정책 방향을 건의했다. 이 밖에 박성현 교수는 다양한 보직을 맡아



2021년 5월 부산 만덕고등학교에서 가진 '한림원석학과의 만남' 강연



강연 중인 박성현 교수

역임해 왔다. 한국과학기술단체총연합회 주관 국가발전포럼 운영위원장직, ‘국회 4차 산업혁명 포럼’에서 ‘기초과학 및 R&D 혁신위원회 위원장’, 국가과학기술연구회 산하 ‘남북과학기술협력위원회 위원장’ 등이다. 이런 노력으로 박성현 교수는 과학기술훈장 혁신장: 과학기술 분야 훈장 제86호(대통령상, 2015) 등을 수상하기도 했다.

현재도 박성현 교수는 바쁘게 지내고 있다. 맡은 직함만 해도 △사회적책임경영품질원 회장(2016~현재) △대한민국 학술원 자연과학부 회장(2022~현재) △한국시니어과학기술인협회 회장(2022~현재) 등으로 적지 않다. 그 밖에도 산업 및 학술 분야 다양한 대외활동을 끊임없이 이어가고 있다. 최근에는 신문잡지 등에 다양한 글을 기고하는데 많은 시간을 쓰고 있다. ‘미래한국’이란 잡지에 편집위원으로 참여하고 있으며, 매달 ‘데이터로 보는 세상’에 글을 쓰고 있다. 이 밖에 문화일보 고정칼럼니스트로도 활동하고 있다. 과학기술 정책이나 국가 통계정책 등에 대하여 언론 기관(신문, 잡지 등)에 수없이 많은 기고문을 게재하였으며, 이는 2012년 이후에만 182편에 달한다. 학생 및 일반인 대상의 다양한 대중강연 역시 끊임없이 소화하는 등, 지금도 청장년층 못지않게 활발한 활동을 이어 나가고 있다.

