

국내기업의 빅데이터 이용의도에 미치는 영향요인 분석 : 공공데이터 활용여부를 기준으로

정화민, 이상윤*
서강대학교 정보통신대학원

Analysis of Factors Affecting Big Data Use Intention of Korean Companies : Based on public data availability

HwaMin Jeong, SangYun Lee*
Graduate School of Information & Technology / Sogang University

요약 본 연구는 공공데이터 활용 여부에 따른 국내 기업체들의 빅데이터 기술사용 및 서비스 이용 의도에 미치는 영향요인을 파악하기 위한 탐색적 연구이다. 2017년 13,136개 국내 기업의 전산담당자와 총무담당자를 대상으로 정보화 통계 설문조사 데이터를 R 3.5.3 프로그램을 이용하여 연구 모형에 따른 가설검정을 시행하였다. 먼저 빅데이터 기술 및 서비스 이용의도에 영향을 미치는 요인에 대한 로지스틱 회귀분석 결과, 공공데이터를 활용하는 업체에서는 비용 절감이 빅데이터 이용의도에 긍정적 영향을 미치는 요인으로 나타났고, 공공데이터를 활용하지 않는 기업의 경우 고객만족은 긍정적 영향, 의사결정 지원은 빅데이터 이용에 부정적인 영향을 미치는 요인으로 나타났다. 최근 들어 우리 정부는 기업 지원에 있어 4차 산업 혁명의 동력 중 공공데이터 개방과 더불어 빅데이터 활용 부분에 중점을 두고 있다. 그러나 본 연구결과, 아직까지 국내 기업 전반에 공공데이터 및 빅데이터 활용이 미진한 것으로 밝혀졌다. 단, 본 연구를 위한 정보화 통계 데이터가 2017년도 데이터임을 감안할 때 정부의 빅데이터 지원 정책 등 실효적 성과를 분석하기 위해서는 지속적인 빅데이터 활용 및 이에 따른 성과에 대한 연구도 필요하다 하겠다.

Abstract This is an exploratory study to examine factors affecting South Korean companies' intentions to use big data technology and services based on whether the companies use public data or not. This study, using R, conducted chi-squared tests and logistic regression analysis. As a result of the logistic regression analysis, cost reduction had a positive effect on the big data-use intentions in companies that use public data, whereas with companies that do not use public data, customer satisfaction had a positive impact, and support for decision-making had a negative impact on the intention to use big data. Recently, the South Korean government has focused on improving the utilization of public data and big data. However, as a result of this study, the use of public data and big data in South Korea is still insufficient. Yet, considering that the data utilized for this study was created in 2017, additional study using public data and big data is also required.

Keywords : Bid Data, Public Data, Korean Company, Use Intention of Big Data, R Program.

*Corresponding Author : Sang-Yoon Lee(Sogang Univ.)

email: yunpha123@naver.com

Received June 10, 2019

Accepted October 4, 2019

Revised July 29, 2019

Published October 31, 2019

1. 서론

미국, 일본, 중국 등 주요국 중심으로 공공데이터가 국민 공유의 재산 및 국가 경쟁력의 핵심자원으로 인식되면서 공공데이터 활용을 위한 오픈 데이터(Open Data) 정책이 대두되고 있다[1]. 이러한 경쟁 흐름에 발맞춰 우리나라도 2013년 '정부 3.0'을 통해 공공데이터 개방을 국정과제로 제시하며 데이터 활용도 제고를 위한 정부 차원의 지원을 시작했다[1].

공공데이터 포털 자료에 의하면 2013년 공공데이터 개방 건수는 5,272건에서 2017년(10월 말 기준) 23,084건으로 2013년 대비 약 4.3배, 활용 건수도 2013년(10월 말 기준) 13923건에서 2017년 3,505,731건으로 약 252배 증가한 것으로 나타났다[2]. 2015년의 경우 OECD 공공데이터 개방 지수에서 평가 대상 30개 회원국 중 1위를 차지했다[3]. 그러나 2017년에 발표된 스위스 국제 경영개발 대학원(IMD)의 국가경쟁력 평가 내 디지털 경쟁력 순위에서는 우리나라의 빅데이터 사용 및 활용능력 수준이 63개의 평가대상 국가 중 56위에 그치는 것으로 나타났다[4]. 2013년 11월 제28차 경제관계 장관회의 이후 빅데이터로써 활용가치가 높은 공공데이터에 대한 국가적 관심이 크게 늘었으나[5], 아직까지 공공데이터 활용에 대한 인식 부족[6], 개인 정보보호 관련 법제와 데이터 연계 및 공유하는 시스템이 미비해 공공데이터와 빅데이터 관련 정책 수립이 체계적으로 이루어지지 않고 있다[3,4].

무엇보다 통계적 지식과 데이터마이닝, 도메인 지식과 같이 데이터 관련 비즈니스 경험을 가진 전문 인력이 부족해 공공데이터를 빅데이터 관점에서 분석과 활용이 부족한 실정이다[4,7,8].

이에 본 연구에서는 매년 국가통계로 실시하고 있는 정보화 통계조사 Raw Data를 활용해 공공데이터 활용 여부에 따른 국내 기업 빅데이터 이용 의도 영향요인을 파악하고 이를 검증하고자 한다. 또한 분석된 결과를 바탕으로 기업의 담당자와 정책 입안자들에게 빅데이터의 활용 요인에 대한 기초적인 정보와 시사점을 제공하고자 한다.

이를 위해 본 연구에서는 공공데이터와 빅데이터에 대한 개념 및 관련 선행 연구를 살펴보고 구체적인 연구 방법을 제시하였으며, 연구결과를 바탕으로 정책적, 실무적 시사점을 제시하였다.

2. 본론

2.1 이론적 고찰

2.1.1 공공데이터 개념 및 활용

공공데이터는 공공기관에서 수집한 정보를 전자적 방식으로 처리된 자료를 의미한다. 주로 대용량이고, 다양한 데이터 세트를 구성하는 빅데이터의 영역에서 다루어지고 있다[5]. 또한 일반 데이터보다 왜곡이나 조작이 발생할 가능성이 적고 매시간 국가 및 국민과 관련된 공공업무가 수반되기 때문에 신뢰성과 지속성이 높은 특성을 가지고 있다[9].

우리나라는 '정부 3.0' 정책을 통해 공공데이터를 개방·공유하고, 부처 간 칸막이를 없애 소통·협력하여 국민 맞춤형 서비스 제공과, 민간영역에서 다양한 활용 및 새로운 서비스를 창출하는 데 목적을 두고 있다[9].

2.1.2 빅데이터 기술 및 서비스

빅데이터 기술은 방대하고 다양한 종류의 데이터들을 융합 및 예측해 새로운 가치를 내는 것을 의미하며 수집, 처리 및 저장, 분석, 시각화의 과정으로 구분할 수 있다[11].

수집은 기관 및 기업 등 분산된 데이터들을 수집 및 변환해 정제된 데이터를 확보하는 기술을 의미한다[10]. 데이터 처리 및 저장은 다양한 형식의 데이터를 실시간으로 처리 및 저장할 수 있는 분산 컴퓨팅 기술을 의미하고 Map Reduce, In Memory DB, HDFS 등이 있다[11].

분석의 방법으로는 데이터마이닝, 텍스트 마이닝, 소셜 분석, 평판 분석 등을 들 수 있고, 이는 모두 다양한 데이터로부터 비즈니스 가치를 발견하는 데 목적을 둔다[12,13].

시각화는 비전문가가 데이터 분석을 수행할 수 있는 환경 제공 및 분석 결과 확인이 가능한 기술로, 인포그래픽스, 그래픽 기반 모델링 도구, 정보 시각화 기술 등이 있다[13,14].

빅데이터를 활용한 국내 공공 분야 서비스로는 부정 거래 개선, 국토 안보, 인적 관리, 국정 모니터링 등이 있고, 국내 기업에서는 클라우드, 소셜 네트워크 분석, 검색 등 다양한 서비스들을 제공하고 있다[15].

2.2 선행연구

2017년 정보화 통계조사 설문 문항은 Fig. 1과 같이 공공데이터를 활용하는 국내 기업체의 빅데이터 기술 및 서비스 이용에 미치는 영향 요인을 비용 절감, 생산성 및

정보 공유 증대, 의사결정 지원, 신규 수익원 및 서비스 창출, 고객만족의 다섯 가지로 구분 했으며(정보화 통계 조사 설문 번호 문 53-1~문 53-5번) 관련 선행연구는 다음과 같다.

2017년 정보화통계조사

인행하십니까?
미행창조과학과 한국정보화진흥원에서는 국가정보기본법 제49조 및 통계법 제17조, 제18조에 의거하여 우리나라 사업체들의 정보화 수준 및 현황을 종합적으로 파악하기 위해 매년 전국의 모든 사업체를 모집단으로 하는 표본조사인 정보화통계조사를 실시하고 있습니다.
바른표준을 이어나 정부와 효과적인 정보화정책 수립에 도움이 될 수 있도록 적극적인 협조를 부탁드립니다. 아울러 작성해 주신 자료는 반드시 조사와 연구에 관련된 목적에만 사용될 것이며, 비밀은 철저히 보장될 것임을 약속드립니다.
잠시만 시간을 내어 협조해 주시면 대단히 감사하겠습니다.
설문조사에 응해 주셔서 감사드리며, 귀사업장의 평인과 번창하심을 기원합니다.

미행창조과학부 · 한국정보화진흥원
[조사대상 범위] 미행창조과학부 소속의 사업체 (02-3415-5036), 조인회 대 (02-3415-5210)

지역: ① 서울 ② 부산 ③ 대구 ④ 인천 ⑤ 광주 ⑥ 대전 ⑦ 울산 ⑧ 세종 ⑨ 경기 ⑩ 강원 ⑪ 충북 ⑫ 충남 ⑬ 전북 ⑭ 전남 ⑮ 경북 ⑯ 경남 ⑰ 제주

사업체명 (필수) []

표본 번호 (필수) []

집단 구분 (필수) []

사업 형태 (필수) []

조직 형태 (필수) []

① 응답해 주실 때 꼭 지켜 주십시오
1. 설문지는 각 페이지마다 순서대로 차례대로 작성해 주십시오. 질문 앞에 특별한 언급이 없는 한 모든 질문에 대해 주십시오. (필수사항을 포함해 1-25페이지까지 작성해야 합니다)
2. 응답은 귀 사업장의 현실과 가장 가까운 대로 작성하십시오. (예를 들어, '예'라고 하면 '예'로, '아니오'라고 하면 '아니오'로 작성하십시오.)
3. 질문 앞에 특별한 언급이 없는 한 모든 질문에 대한 답을 작성하십시오. (예를 들어, '예'라고 하면 '예'로, '아니오'라고 하면 '아니오'로 작성하십시오.)
4. 작성하는 표본의 일련번호도 2016년 12월 31일 기준으로 응답해 주십시오. (예) 귀 사업장은 현재 컴퓨터가 있습니까? → 2016년 12월 31일 현재 컴퓨터 보유 여부를 알음에 주시면 됩니다.
※ 최근 1년간(2016년 1월 ~ 12월)을 기준으로 하는 문항은 문항에 별도로 표기가 되어 있습니다.

주관기관: 미행창조과학부 | 응답기관: 한국정보화진흥원 | 조사기관: 미행창조과학부 | 조사대상: NIA 한국정보화진흥원

Fig. 1. Questionnaire for Informatization Survey in 2017 (Example)

2.2.1 비용절감

빅데이터 도입 시 기업 및 기관에서는 보유한 다양한 인프라 자산들에 대한 보수 및 수리가 가능해지고 효율화를 높여 비용절감을 할 수 있다[16,23].

국내 공공데이터를 활용한 비용절감 사례로 국제 유가의 분석을 통해 국내 유가 안정을 위한 판매가 예측을 하는 한국 석유공사와, 산업현장의 에너지 절감을 위해 실시간 건물 에너지 진단 및 유틸리티 운전 최적화 시스템을 구축한 (주)소프트웍스가 있고[17], 서울시는 N 버스 도입을 통해 서울시민들의 교통비 절감 및 승용차 운행 수요를 감소시켜 유틸리티 절감 효과를 보았다[17]. 그러므로 비용절감은 빅데이터 기술 및 서비스 이용의도에 긍정적인 영향을 줄 수 있는 요인으로 보았다.

2.2.2 의사결정지원

기업의 경영자는 빅데이터 기술 활용을 통해 광범위한 데이터를 효과적으로 분석해 더 합리적인 의사결정을 내

릴 수 있고 사전 이슈 인지 및 대응 능력을 향상시킬 수 있어 기업의 제품 경쟁력을 확보할 수 있다[18,19]. 또한 고객의 소비패턴, 선호 제품 및 서비스 등을 파악할 수 있는 정보를 제공해 준다[19]. 국내 기업의 활용 사례로 시민들의 요구 및 민원 과제들을 해결하기 위해 민원 관리 시스템을 구축한 포항시와 포트 홀로 인한 교통사고 대응 및 도로 유지 보수 점검을 위해 인공지능 기반의 도로포장 파손 실시간 탐지시스템을 개발한 한국도로공사가 있으며[17], 의사결정 지원 관점에서의 빅데이터 이용에 대한 선행연구와 사례를 바탕으로, 본 연구에서는 의사결정지원을 빅데이터 기술 및 서비스 이용의도에 긍정적인 영향을 미치는 요인으로 가정하였다.

2.2.3 생산성 및 정보공유증대

빅데이터 활용 시, 외부로부터 데이터들을 활용해 이슈 감지 및 대응과 기업 내 데이터들을 연계해 의미 있는 정보를 산출함으로써 경쟁력을 확보할 수도 있다[20,21]. 이와 관련한 국내 기업과 공공기관 사례로, 공정시간과 비용 절감 및 품질을 향상시킨 이미지 데이터를 분석업체 (주)씨마이너와 실시간 생산 시설 분석 및 모니터링 등으로 고객사들의 생산성 향상을 지원하는 삼성SDS의 '브라이텍스'가 있다[22]. 농촌진흥청의 경우 기상청 자료를 바탕으로 생산량 예측 소프트웨어를 개발해 작물의 생육상태 점검 및 농산물 생산성을 예측할 수 있게 되었다[23]. 그러므로 생산성 및 정보 공유 증대는 빅데이터 기술 및 서비스 이용의도에 영향을 주는 요인으로 보았다.

2.2.4 고객만족

빅데이터는 고객의 행동과 소비패턴을 분석하여 맞춤형 서비스를 제공하고 불만사항을 실시간으로 파악해 사전에 이탈 가능성이 높은 잠재 고객을 예측하는데도 사용될 수 있다[24,25].

국내 주요 사례로 고객 목소리 분석 시스템으로 고객 만족도를 높인 한국도로공사와 '찾아가는 서비스'를 도입해 재해 결정 시간을 단축시키고 작업 복구율을 향상시킨 근로복지공단이 있다[26]. 이와 같은 사례를 볼 때, 고객만족은 빅데이터 기술 및 서비스 이용의도에 긍정적인 요인으로 작용한다고 판단할 수 있다.

2.2.5 신규수익원 및 서비스 창출

한국 정보화 진흥원(2013)에서는 빅데이터 활용은 경제적 가치 창출, 사회적 문제 해결과 더불어 새로운 비즈니스를 창출하게 될 것이라 주장했다[2,23]. 국내 주요

사례로 내·외국인 관광소비패턴 및 트렌드 분석을 통해 개인 맞춤형 관광정보 제공과 추가 관광지를 개발한 미래창조과학부, 청장년층의 취업난과 실업 등에 맞춘 일자리 지표를 마련한 통계청이 있다[17]. 따라서 주요 선행연구의 실제 사례를 토대로 확인했을 때 신규 수익원 및 서비스 창출은 빅데이터 기술 및 서비스 이용 의도에 긍정적 영향을 줄 수 있다고 볼 수 있다.

2.2.6 빅데이터 이용 의도

빅데이터 이용의도는 기업 혹은 기관에서 빅데이터를 이용 중이거나 계획이 있는 정도를 의미한다.

기술 수용 모델(TAM)을 활용하여 '빅데이터 기술 및 서비스 이용 의도'의 영향요인에 대한 연구는 다수 진행되었다[23,26]. 본 연구에서는 검증된 선행연구 사례들과 2017년 정보화 통계 설문 자료를 바탕으로 '빅데이터 기술을 이용하는 국내 기업의 이용의도'를 종속변수로 활용하여 국내 기업의 빅데이터 기술 및 서비스 이용의도의 영향요인을 분석하였다.

3. 연구방법

3.1 참여기업 특성

2017년 전국의 기업체를 대상으로 실시했던 정보화 통계 설문조사 데이터를 바탕으로 한 참여기업에 대한 분석은 다음과 같다.

지역별 공공데이터 활용 여부에 따른 분포의 차이는 아래 Table 1과 같이 카이제곱 값이 68.470, 유의 확률 값은 0.000으로 통계적으로 유의한 결과로 나타난 바, 수도권(12.7%)과 비수도권(9.4%) 간의 공공데이터의 활용 차이를 확인할 수 있다. 또한 공공데이터 미활용(78%)이 활용(22%)보다 높은 것으로 보아 대부분 기업체에서 공공데이터를 활용을 못하고 있음을 나타낸다.

Table 1. Use of public data by region

division		Public Data		Total	Chi-square /p
		Use	Unused		
Region	Non capital	1233	5253	6486	68.470 / .000***
		9.4%	40.0%	49.4%	
	Capital	1662	4987	6649	
		12.7%	38.0%	50.6%	
Total		2895	10240	13135	
		22.0%	78.0%	100.0%	

*:p<0.05, **:p<0.0

지역별 빅데이터 기술 및 서비스 이용에 대한 분포 차이는 다음과 같다. 아래 Table 2에서 카이제곱은 21.639, 유의 확률 값은 0.000으로 통계적으로 유의 수준에서 지역별로 빅데이터 기술 및 서비스 이용의 차이가 있음을 보여준다. 13,136개의 조사 기업체 중 445(3.4%) 개 기업이 빅데이터 기술 및 서비스를 이용 중이고, 12,691(96.6%) 개의 기업체는 미 이용 중인 것으로 나타나 아직까지 현장에서 빅데이터 기술 및 서비스를 이용하는 분포가 매우 낮은 것으로 나타났다.

Table 2. Use of local big data

Region	Big Data		Chi square/p
	No(%)	Yes(%)	
Seoul	2853(21.7)	153(1.2)	285.783 0.000***
Busan	910(6.9)	26(0.2)	
Daegu	553(4.2)	8(0.1)	
Incheon	602(4.6)	56(0.4)	
Gwangju	345(2.6)	10(0.1)	
Daejeon	288(2.2)	48(0.4)	
Ulsan	285(2.2)	5(0.0)	
Saejong	470(4)	4(0.0)	
Gyeonggi	2921(22.2)	65(0.5)	
Gangwon	319(2.4)	17(0.1)	
Chungbuk	393(3.0)	5(0.0)	
Chungnam	485(3.7)	14(0.1)	
Jeonbuk	431(3.3)	0(0.0)	
Jeonnam	531(4.0)	11(0.1)	
Gyeongbuk	699(5.3)	8(0.1)	
Gyeongnam	877(6.7)	7(0.1)	
Jeju	152(1.2)	8(0.1)	
Total	12,691(96.6)	445(3.4)	

*:p<0.05, **:p<0.01, ***:p<0.001

공공데이터 활용하지 못하는 이유로는 아래 Table 3과 같이 '빅데이터 확보 방법을 모름'(73.20%)이 가장 높은 빈도로 나타났고, 데이터의 품질 수준 이하(9.80%), 신뢰성 부족(8.67%), 현행화 미흡(8.33%)으로 나타났다.

Table 3. Reasons for Not Using Public Data

division	Frequency	Percentage
Securing method unknown	1,934	73.20
Public data lacking in trust	229	8.67
Less than public data quality level	259	9.80
Public data synchronizing insufficient	220	8.33
Total	2,642	100

3.2 연구모형

본 연구의 연구 모형은 아래 Fig 2 와 같이 독립변수인 비용 절감, 의사결정 지원, 생산성 및 정보 공유 증대, 고객만족, 신규 수익원 및 서비스 창출 확대로 설정하였고 종속변수는 범주형 데이터로 빅데이터 기술 및 서비스 이용을 설정하였다. 가설은 다음과 같다.

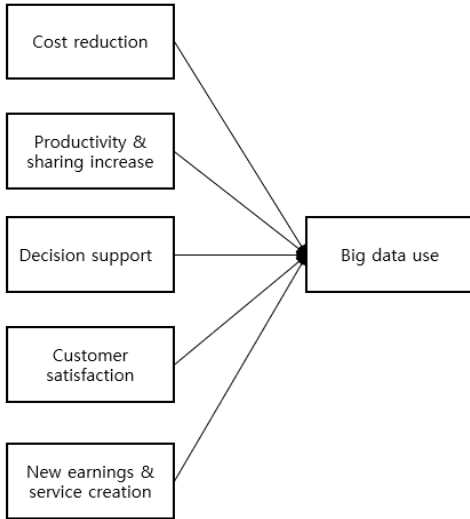


Fig. 2. Research Model

- H1. 비용 절감은 빅데이터 이용의도에 통계적으로 유의한 영향을 미칠 것이다.
- H2. 의사결정 지원은 빅데이터 이용의도에 통계적으로 유의한 영향을 미칠 것이다.
- H3. 생산성 및 정보 공유 증대는 빅데이터 이용의도에 통계적으로 유의한 영향을 미칠 것이다.
- H4. 고객만족은 빅데이터 이용의도에 통계적으로 유의한 영향을 미칠 것이다.
- H5. 신규 수익원 및 서비스 창출 확대는 빅데이터 이용 의도에 통계적으로 유의한 영향을 미칠 것이다.

본 연구의 변수별 조작적 정의는 다음 Table 4와 같이 변수 구분, 변수 설명, 관련 연구의 형태로 나타내었다.

Table 4. Operational definition by variable

dimension		description	Related research
Independent variable	Cost reduction	Cost reduction through price comparison and prediction	[16],[23]
	Productivity & sharing increase	Productivity and service information sharing increase using data in major tasks such as production marketing	[20],[21]
	Decision making support	Support rational decision making through collection analysis of data	[18],[19]
	Customer satisfaction	Increase satisfaction with fast and precise customer response through customer and people dissatisfaction factor analysis	[24],[25]
	New earnings & service creation	Create new earnings and service through collection and analysis of data acquired from big data	[2],[23]
Dependent variable	Big data use	Whether big data technology and service used	[23],[26]

4. 분석결과

4.1 신뢰성 검정 및 변수별 상관분석

가설검정 전 데이터의 타당성 검정을 위한 요인분석은 본 공공데이터의 설문 구조상 설계되지 않아 본 연구에서는 생략하였다. 독립변수 5개(비용 절감, 의사결정 지원, 생산성 정보 및 공유 증대, 고객만족, 신규 수익원 및 서비스 창출 확대)에 대한 신뢰성 검정 결과, Table 5와 같이 Cronbach's α 값이 0.771로 나타나 가설검정을 위한 독립변수들의 신뢰성이 있다고 하겠다.

Table 5. Reliability Test

division	Mean	Standard Deviation	Cronbach's α Deleted Item	Cronbach's α
Cost reduction	3.59	0.775	0.756	.771
Decision support	3.91	0.723	0.730	
Productivity & sharing increase	3.88	0.717	0.717	
Customer satisfaction	3.82	0.747	0.726	
New earnings & service creation	3.70	0.775	0.717	

종속변수(이분형 데이터)를 제외한 독립변수들 간의 상관분석 결과는 아래 Table 6과 같이 나타났고, 독립 변수들 간의 상관관계가 $r=0.8(80\%)$ 이상인 관계가 없어 다중 공선성의 문제는 없다고 할 수 있다.

Table 6. Correlation analysis

division	1	2	3	4	5
1. Cost reduction	1				
2. Decision support	.361**	1			
3. Productivity & sharing increase	.352**	.480**	1		
4. Customer satisfaction	.320**	.364**	.465**	1	
5. New earnings & service creation	.384**	.407**	.415**	.488**	1

*, $p<0.05$, **, $p<0.01$

4.2 가설 검증

가설검증을 위하여 기업의 비용 절감, 의사결정 지원, 생산성 정보 및 공유 증대, 고객만족, 신규 수익원 및 서비스 창출 확대를 독립변수로 빅데이터 이용을 종속변수로 한 이분형 로지스틱 회귀분석(Binomial Logistic Regression Analysis)을 실시하였다. 분석은 공공데이터를 사용하는 집단과 공공데이터를 사용하지 않는 집단을 나누어 아래 Fig3와 같이 R 프로그램 3.5.3 버전을 활용하여 분석하였다.

```
rm(list=ls())
library("psych")
data <- read.table("data.txt", sep=" ", header=F)
names(data) <- c(1:249)#header add
View(data)
datacron <- data[,c('190', '191', '192', '193', '194')]
alpha(datacron)
library(PerformanceAnalytics)
datacorr <- data[,c('190', '191', '192', '193', '194')]
chart.Correlation(datacorr, histogram=TRUE, pch=19)

data1 <- data[data$'58'==1,]
data1$'178' <- ifelse(data1$'178'!=1, 0, 1)
data1$'178' <- as.factor(data1$'178')
model11 <- glm(formula = data1$'178' ~ , family = "binomial")
summary(model11)
exp(coef(model11))

data2 <- data[data$'58'==2,]
data2$'178' <- ifelse(data2$'178'!=1, 0, 1)
data2$'178' <- as.factor(data2$'178')
model12 <- glm(formula = data2$'178' ~ , family = "binomial")
summary(model12)
exp(coef(model12))
```

Fig. 3. R coding: Binomial Logistic Regression

분석 결과, 공공데이터를 활용하는 기업체를 대상으로 가설검정 한 결과는 Table 7과 같이 나타내었다. 가설의 채택은 유의 확률 값(P)이 0.05 미만을 기준으로 하였고 공공데이터를 활용하는 기업체 중 비용 절감만이 빅데이

터 이용에 통계적으로 유의한 영향을 미치는 요인으로 나타났다. 구체적으로 비용 절감이 1 변할 때, Exp(B)를 기준으로 빅데이터를 이용하는 확률이 이용하지 않을 확률보다 약 1.59배 증가하는 것으로 나타났다. 이 외에 고객만족, 의사결정 지원, 생산성 정보 공유, 신규 수익원 및 서비스 창출은 통계적으로 유의한 영향을 주고 있지 못하고 있는 것으로 나타났다.

Table 7. Binomial Logistic Regression Analysis (Using public data)

division	B	S.E	Significance probability	Exp(B)
Cost reduction	0.470	0.136	0.001*	1.599
Decision support	-0.308	0.159	0.053	0.735
Productivity & sharing increase	-0.167	0.160	0.297	0.846
Customer satisfaction	-0.163	0.158	0.300	0.849
New earnings & service creation	-0.009	0.144	0.951	0.991

*, $p<0.05$, **, $p<0.01$, ***, $p<0.001$

(Dependent variable : Big Data use)

또한, 공공데이터 활용하지 않는 기업체를 대상으로 한 분석 결과는 Table 8과 같다. 공공데이터를 활용하는 기업의 분석 결과와 달리 의사결정 지원과 고객만족만이 빅데이터 이용의도에 유의한 영향을 주는 것으로 나타났다. 구체적으로 고객만족이 1 변할 때, Exp(B)를 기준으로 빅데이터를 이용하는 확률이 약 1.57배 증가하는 것으로 나타났고, 의사결정 지원은 베타값이 -0.46이므로 1 변할 때, Exp(B) 기준으로 빅데이터를 이용하는 확률이 0.63배 낮아져 부정적인 영향이 있는 것으로 나타났다. 이 외에 비용 절감, 생산성 및 정보 공유, 신규 수익원 및 서비스 창출은 통계적으로 유의한 영향을 주고 있지 못하고 있는 것으로 나타났다.

Table 8. Binomial Logistic Regression Analysis (Un used public data)

division	B	S.E	Significance probability	Exp(B)
Cost reduction	-0.025	0.162	0.876	0.975
Decision support	-0.460	0.174	0.008*	0.631
Productivity & sharing increase	0.279	0.189	0.141	1.321
Customer satisfaction	0.451	0.171	0.008**	1.570
New earnings & service creation	-0.281	0.173	0.104	0.755

*, $p<0.05$, **, $p<0.01$, ***, $p<0.001$

(Dependent variable : Big Data use)

5. 결론

본 연구는 2017년 정보화 통계 조사 Raw data를 활용하여 13,136개 국내 기업의 공공데이터 활용 여부에 따라 빅데이터 기술 및 서비스 이용 영향요인을 분석하였다. 분석 결과, 공공데이터를 활용하는 국내 기업의 경우, 빅데이터 이용의도의 영향을 주는 변수는 비용 절감만이 긍정적인 영향을 주는 것으로 나타났으며, 이는 선행연구[16,23]들의 연구결과와 유사한 결과로 나타났다. 또한 공공데이터를 활용하지 않는 국내 기업체의 경우, 고객만족은 빅데이터 이용의도에 긍정적인 영향을 주는 변수이나, 의사결정 지원은 빅데이터 이용의도에 부정적인 영향을 주는 변수로 나타났다.

본 연구 결과를 바탕으로 실무적, 정책적 시사점을 제시하면 다음과 같다.

첫째, 기업체에서 데이터 확보를 쉽게 할 수 있는 방법과 데이터의 품질을 검증하고 신뢰성을 확보할 수 있는 공공기관의 설립과 데이터 검증 시스템의 개발도 필요하다 하겠다.

둘째, 국내 기업들은 필요한 공공데이터의 확보와 처리를 위한 인프라 구축을 위해 지속적으로 관련 산업인력 양성 및 투자와 지원을 해야 할 것이다.

셋째, 공공데이터를 사용하고 있는 기업체에서는 빅데이터 이용의도에 영향을 주는 요인이 비용절감만으로 빅데이터 활용이 제한적이고 빅데이터를 사용하지 않는 기업에서는 빅데이터에 대한 확보 방법, 품질, 데이터 신뢰성에 대한 인식이 부족함에 대한 향후 추가적인 연구가 필요하다.

최근 정부에서 대통령 직속 4차 산업 위원회의 설치와 운영, 지능화 혁신 프로젝트 추진, 성장 동력 기술력 확보, 산업 인프라 생태계 조성 등을 통해 4차 산업 활성화를 위한 투자와 노력을 하고 있으며 빅데이터 분야도 많은 관심과 투자를 받고 있다. 이러한 지원과 노력이 성과로 이어지기 위해선 무엇보다 산업계, 학계, 정부 기관의 밀접한 협업도 필요할 것이다.

References

- [1] Seok Hyun Song, Jai Yong Lee, "Comparative Analysis of National Policies for Open Data Government Ecosystem," Journal of Society of Korea Industrial and Systems Engineering, vol. 41, no. 1, pp. 128-139, Mar. 2018.
- [2] In Taek Jung, Kyu-soo Chong, Development of Information Technology Infrastructures through Construction of Big Data Platform for Road Driving Environment Analysis, Korea Academy Industrial Cooperation Society, 19(3), pp.669-678.
DOI: <https://doi.org/10.11627/jkise.2018.41.1.128>
- [3] Seok Hyun Song, Samyool Lee, Yeol Shin, Jai Yong Lee, "A Study on the Effectiveness of Korea's Open Government Data Policy: Ecosystem Perspective", Journal of Korean Society of Information Technology, 20(4), pp.1-34.
DOI: <http://dx.doi.org/10.22313/reik.2018.16.2.427>
- [4] Do Young Jung, Min Chang Kim, Jae Hwan Kim, "Big Data Policy Promotion Status Usage Improvement Plan", The legal policy report, NARS, Korea, pp.2-22. 2018.
- [5] Bong Choi, An Analysis of Economic Effect on Utilization of Big Data and Information Disclosure on Public Sector, Policy Research, The Seoul Institute, Korea, pp.25.
- [6] Samsung Economic Research Institute (2012), "Big Data: Cause of Industrial Crustal Change", CEO Information, Vol.851, pp.3-25.
- [7] Jong Hyeon Kim, "Examples of Big Data Use in Domestic and Foreign Financial Sectors Pre-assignment for activation of introduction", Financial Report, Woori Finance Research Institute, Korea, pp.23-28.
- [8] Seung Beom Choi, "A Study on the Conditions for the Vitalization of Big-data Use in Local Governments", Social enterprise & policy studies, 6(1), pp.177-205.
- [9] Young Il Cha, Sung-Kyou Choi, Kyeong-Seok Han, "An Empirical Study on the Influence on Public Data Usage in Private Business Sectors", Journal of Digital Convergence, 15(6), 9-17. 2017.
DOI: <https://doi.org/10.14400/JDC.2017.15.6.9>
- [10] Sun Woo Lee, "Research on determinants for big data system adoption in organizations", Ph.D, Sungkyunkwan University, Korea, pp.12-30.
- [11] Hyeon Ji Seo, "A Study on the Customer Behavioral Analysis Using Big Data of Distribution Industry", Master's thesis, Gachon University, Korea, pp.12-25.
- [12] Jung Sun Kim, "A Study on Initial Characteristics of the Acceptance of Innovative Big Data Technology in Korea", Ph.D, Ewha Womans University, Korea, pp.2-11, 2015.
- [13] Yeong Nan Yu, "Analysis on technology level of Big Data for Smart Manufacturing in South Korea - Focused on Domestic Patent Trends-", Ph.D, Yonsei University, Korea, pp.7, 2016.
- [14] Ji Young Kim, "Protection Measures for Sensitive Information and Personal Information In Big Data Environment", Ph.D, Sungshin Womans University, Korea, pp.6-7, 2014.

- [15] Jin Hee Song, Jung Sook Kim, Analysis of the best practices big data services, The Korea Contents Society, 12(1), pp.32-37, 2014.
DOI: <http://dx.doi.org.libproxy.sogang.ac.kr/10.20924/CCTHBL.2014.12.1.032>
- [16] Hee Jeong Yoon, "A Study on the Effect of the Organizational, Technical and Environmental Recognition and Utilization of Big Data", Business Management Research ,Vol.7 No.2, Korea ,pp.153-161, Dec. 2014.
- [17] Ji Hye Lee, "Public Big Data Best Practice Collection", Public Data Policy Division, Feb 2018.
- [18] JKa. Hoi Kwang, Jin Soo Kim, "An Empirical Study on the Influencing Factors for Big Data Intended Adoption: Focusing on Strategic Value Recognition and TOE Framework. Asia Pacific Journal of Information Systems" Business Administration, Asia Pacific Journal of Information Systems, Seoul, Korea, 24(4),pp.443-472 ,Feb. 2014.
DOI: <https://doi.org/10.14329/apjis.2014.24.4.443>
- [19] Eun Jung Kim, Jae Woong Lee, Hee in Yang, The Actual Condition of the Information Divide in Korea and its Policy Implications: Focusing on the 2007 Information Gap Survey, journal of the Korean Association of Local Information Science, Vol.1, pp. 75-101, Nov. 2008.
- [20] Hee Jung An, "A Business Performance Study of Data Quality Management for Big Data Adoption : Focused on Corporate Data Quality Management Process", Ph.D, Kookmin University, Seoul, Korea, pp.2-18, Nov. 2016.
- [21] Choon Yeol Lee, A Study on Synchronization Effect of A Multi-dimensional Event Database for Big Data Information Sharing. Journal of Digital Convergence, 15(10), pp.243-251, 2017.
DOI: <https://doi.org/10.14400/JDC.2017.15.10.243>
- [22] KAIT, "Big Data forums, domestic market leverage Big Data for the first activation, competition analysis ceremony held", News Release, 2013.
- [23] Ji Sook Kim, "Study on utilizing big data and analyzing techniques", Master's thesis, Korea University, Seoul, Korea, pp.4-10, 2013.
- [24] KISA, "Internet Secret Issue", Technical Report, Korea, pp11-12, Feb 2012.
- [25] KAIT, "Big Data forums, domestic market leverage Big Data for the first activation, competition analysis ceremony held", News Release, 2013.
- [26] Young Ki Jung, Myung gun Suk, Chang Jae Kim, "A study on the success factors of Big Data through an analysis of introduction effect of Big Data". Journal of Digital Convergence, 12(11), pp.241-248, 2014.
DOI: <https://doi.org/10.14400/JDC.2014.12.11.241>

정 화 민(Hwa-Min Jeong)

[정회원]



- 2009년 8월 : 서강대학교 대학원 MIS전공 (경영정보학박사)
- 2016년 1월 ~ 현재 : 한국정보기술연구원 빅데이터 연구교수
- 2018년 9월 ~ 현재 : 서강대학교 정보통신대학원 데이터사이언스 (인공지능)전공 겸임교수
- 2019년 2월 ~ 현재 : 타우데이터 주식회사 대표이사

<관심분야>

빅데이터, 통계분석, 머신러닝, 인공지능, 국가인적자원개발, ICT교육/창업, 정보보안, e-Business 등

이 상 윤(Sang-Yun Lee)

[정회원]



- 2011년 3월 ~ 2018 2월 : 부천대학교 전자공학 (공학학사)
- 2018년 3월 ~ 현재 : 서강대학교 정보통신대학원 데이터사이언스 전공 (공학석사 재학 중)
- 2015년 3월 ~ 현재 : SK 오앤에스 SKT 검증센터 관리(재직 중)

<관심분야>

정보통신, 빅데이터, 통계분석, 머신러닝, 인공지능