

## 활용 주체별 빅데이터 수용 인식 차이에 관한 연구: 활용 목적, 조직 규모, 업종 특성을 중심으로

이 영 주\*, 양 현 철\*\*

### 요약

빅데이터 관련 기술이 점차 성숙해지고 있고 공공부문을 중심으로 초기 성공사례들이 발표되고 있으나 실질적인 가치 창출에 대한 확신 부족과 개인정보 유출에 대한 여전한 우려로 산업 전반으로의 확산이 더딘 실정이다. 본 연구에서는 빅데이터 도입 과정에서 다양한 이해관계자 집단 별로 긍정적, 부정적 인식이 도입 활성화에 어떠한 차이를 보이는지 탐색적으로 규명하고자 한다. 먼저 기술수용모형(TAM)과 업무기술적합성(TTF) 모형, 프라이버시 계산이론을 바탕으로 긍정적인 평가요인과 부정적인 평가요인을 통합하여 독립변수를 개발하고 빅데이터 이용의도를 종속변수로 하는 연구 모형을 개발하였다. 국내 빅데이터 실사용자 또는 잠재적 사용자를 대상으로 실증 분석한 결과 선행요인 중 빅데이터의 인지된 유용성, 업무기술적합성, 개인정보침해위험이 이용의도에 유의미한 영향을 주고, 사용용이성은 유의성이 없는 것으로 나타났다. 또한 각 요인 별로 빅데이터 활용 주체에 따른 집단 간 변수의 평균 차이를 부분적으로 발견할 수 있었고, 독립변수와 종속변수와의 인과관계에서도 일부 조절 효과를 발견하였다. 본 연구의 시사점을 통해 향후 빅데이터의 산업 활성화를 위해서 이해관계자 별 차별화된 정책 개발이 필요한 시점이다.

주제어: 빅데이터, 사용의도, 프라이버시, 수요자, 이용활성화

## Intention to Use and Group Difference in Adopting Big Data: Towards a Comprehensive View

Lee, Young-Joo, Yang, Hyun-Cheol

### Abstract

Despite the early success story, the pan-industry diffusion of big data has been slow mostly due to lack of confidence of the value creation and privacy-related concerns. The problem leads us to the need to a stakeholder analysis on the adoption process of big data. The present study combines technology acceptance model, task-technology fit theory, and privacy calculus theory to integrate the positive and negative factors on the big data adoption. The empirical analysis was performed based on the survey from the current and potential big data users. Results revealed perceived usefulness, task-technology fit, and privacy concern are significant antecedents to the intention to use big data. Furthermore, there are significant differences in the perceptions of each constructs among groups divided by the types of big data use, with several exceptions. And the control effect was found in the magnitude of the relation between independent variables and dependent variable. The theoretical and politic implications of the analysis are discussed as to the promotion of big data industry.

Keywords: big data, adoption, privacy concern, user, big data policy

2016년 9월 13일 접수, 2016년 9월 28일 심사, 2017년 3월 11일 게재확정

\* 한국정보화진흥원 미래전략센터 수석연구원(lj@nia.or.kr)

\*\* 미래창조과학부 사무관(yhc98@korea.kr)

## I. 서론

디지털 시대에 이용자들이 스스로 만들어 내는 정보의 양은 우리의 예상을 뛰어 넘는다. 매일 20억 개의 인터넷 검색이 구글을 통해 이루어지고 있으며, 300만개의 트윗 대화가 만들어지고 있다(Terence, et al., 2011). 특히 소셜미디어를 통해 생성되는 빅데이터는 기존 고객 유지와 새로운 고객 유치, 새로운 상품이나 서비스의 개발 등 기업의 비즈니스 전략 달성을 위한 기본적인 정보제공에서부터 가치 창출의 원천으로 활용이 확산되고 있다(Chui, et al., 2013).

빅데이터 산업은 데이터의 양적 폭발이 일어났던 빅데이터 1.0시대를 지나 빅데이터로 부터 실질적 가치를 창출하는 빅데이터 2.0시대로 진입하고 있다. 공급 측면에서는 빅데이터 기술이 나날이 정교해지면서 전문적인 데이터 기반 서비스를 제공하는 빅데이터 전문 기업이 부상하고 있으며, 수요측면에서는 미디어·금융·서비스 산업을 중심으로 빅데이터 도입이 빠르게 시도되고 있다(김규남, 2014). 공공부문 또한 데이터 증거에 기반한 과학적 행정과 대민서비스 개선을 위한 빅데이터의 초기 성공 사례들이 발표되고 있다(미래창조과학부 외, 2014). 이렇듯 빅데이터는 민간과 공공 영역에서 광범위한 활용이 기대되고 있으나 실질적 도입 및 확산은 미미한 편이다. 대한상공회의소(2014)의 「빅데이터 활용현황 및 정책과제 연구」에 따르면 조사 대상 기업 4백여 개 업체 가운데 불과 18.4%만이 빅데이터를 이미 활용 중(7.5%)이거나 활용할 계획(10.0%)이 있고, 나머지 81.6%는 아직 활용할 계획이 없는 것으로 나타나고 있다.

빅데이터의 활성화를 막는 배경에는 온라인에서 급격하게 증가되는 개인정보의 유출과 프라이버 침해 가능성에 대한 우려와 연관이 있다(Kim, et al., 2012; 이환수 외, 2013). 따라서 빅데이터 위험 가운데 가장 정책적으로 대응이 필요한 영역은 기술적 위험 중 해킹과 사이버테러, 법제도적 위험 중 개인정보(프라이버시) 침해인 것으로 조사되었다(윤상오, 2013). 한편

사용자들을 위한 개인정보보호와 프라이버시 관련 정책 및 법규의 개선 요구는 빅데이터 산업 생태계의 변화와 기술발전의 흐름에 따라가지 못하고 있는 상황이다(손상영, 2013; 장영훈 외, 2013).

따라서 빅데이터 산업의 활성화 정책은 개인정보의 오남용과 프라이버시 침해 위험을 경감시키면서 동시에 산업적 활용을 억제하지 않는 적절한 균형점을 찾아 나가야 하는 도전에 직면해 있다. 이를 위해서는 빅데이터의 활용을 둘러싼 다양한 이해관계자들이 인지하고 있는 활성화 요인과 저해요인을 종합적으로 살펴볼 필요가 있다. 그러나 국내 빅데이터 도입과 관련된 초기 연구에서는 주로 환경적 특성과 조직 관리적 특성에 초점을 맞추어 왔고 이해관계자 특성을 고려한 연구는 부족한 상황이다.

본 연구는 이러한 문제의식에서 출발하여 빅데이터 기술의 도입과 관련된 다양한 활용주체들의 인식의 차이를 탐색적으로 규명하여 빅데이터 산업 활성화를 위한 정책 방향을 모색하는데 필요한 기초자료로 제공하고자 한다. 이를 위해 빅데이터를 포함한 정보기술 수용에 관한 이론적 고찰을 통해 빅데이터 도입 및 활용과 관련된 선행변수와 연구모형을 도출하고, 이를 기반으로 국내 잠재적 또는 실제 빅데이터 사용자들을 대상으로 실증 분석하였다.

## II. 문헌 검토

### 1. 빅데이터 관련 연구 동향

빅데이터에 대한 초기 개념은 ‘기존의 기술로 감당할 수 없는 범위의 대규모 데이터’로, 다소 협의의 정의에서 출발하였다가 이후 비즈니스적 가치 또는 공익적 가치까지 모두 포괄하는 방향으로 변화하고 있다(채승병, 2011). 초기 빅데이터를 설명하는 특성으로 이른바 3V(Volume, Velocity, Variety)가 제시되고 있으나(Gartner, 2012; McAfee, et al., 2012) 최근에는 가치(Value)와 정확성(Veracity)까지 그 확대해

〈표 1〉 주요기관의 빅데이터 정의

기관	정의
맥킨지 (Chui, 2011)	일반적 데이터베이스체계가 저장, 관리, 분석할 수 있는 범위를 초과하는 규모의 데이터
IDC (2011)	다양한 종류의 대규모 데이터로부터 가치를 추출하고 데이터의 초고속, 수집, 발굴, 분석을 지원하도록 고안된 차세대 기술
가트너 (Gartner, 2012)	향상된 시사점과 더 나은 의사 결정을 위해 사용되는 비용 효율이 높고, 혁신적이며, 대용량, 고속 및 다양성의 특성을 가진 정보 자산
삼성경제연구소 (2012)	기존의 관리 및 분석 체계로는 감당할 수 없을 정도의 거대한 데이터의 집합을 지칭하며, 대규모 데이터와 관계된 기술 및 도구도 포함

서 특징을 이해하고 있다(White, 2012). 빅데이터의 개념에 대해 합의된 정의가 없는 상황에서 본 연구에서는 연구 목적에 따라 데이터나 기술 중 하나에 초점을 맞추지 않고 전체적인(Holistic) 접근법을 취하고자 한다(Boyd, et al., 2012). 이에 본 연구에서는 빅데이터의 개념을 ‘다양한 데이터를 수집, 보관, 정제, 분석하고 가치창출을 이끌어낼 수 있는 IT 도구와 기술 수준’으로 정의한다(Wamba, et al., 2015).

빅데이터의 개념적 범위가 넓은 만큼 관련 학계와 연구 주제 또한 다양하게 펼쳐져 있다. 해외에서는 빅데이터 연구가 크게 디지털 정보의 수집 및 관리 방법에 관한 연구(Information), 구현 기술에 대한 연구(Technology), 분석 방법론 연구(Method), 그리고 사회경제적 가치 창출에 관한 연구(Value) 등 네 가지 분류로 연구가 진행되고 있다(De Mauro, et al., 2015). 이중 네 번째 유형에서는 주로 다양한 산업 분야에서 빅데이터의 활용 사례를 중심으로 그 중요성과 가치를 확인하는 내용이 주를 이루고 있다(McAfee, et al., 2012; Pearson, et al., 2013).

국내 사회과학 분야에서 빅데이터 연구는 초기 빅데이터 개념과 해외 성공사례가 알려지기 시작한 2012년부터 본격적으로 등장하고 있다. 초기에는 빅데이터의 개념과 기술적 특성을 평가하고 어떻게 활용될 것인지에 대한 논의로 시작되었다가(예, 강만모 외, 2012; 안창원 외, 2012), 빅데이터 분석 방법론의 적

용 가능성에 대한 연구(예, 이병엽 외, 2013; 최병정 외, 2012)와 산업 각 분야에의 업무에 적용하는 사례 연구로 확대되었다(예, 김관호 외, 2013; 도해용 외, 2013; 김유신 외, 2012). 빅데이터 도입 사례들이 조금씩 증가한 이후에는 빅데이터의 수용과 관련된 조직 차원의 변인을 규명하는 연구가 시작되고 있다(김은영 외, 2013; 이선우 외, 2014; 김정선 외, 2014).

한편 빅데이터의 보안에 대한 연구는 매우 초기 단계로서, 관련 개념과 기술적 기법에 대한 이해를 해나가는 과정에 있다(박서기 외, 2016). 특히 개인정보 유출과 프라이버시 위협에 대한 기술적 조치에 대한 연구(최대선 외, 2013; 성민경 외, 2014)와 법제도적 조치에 대한 연구(예, 오규철, 2013; 오길영, 2015; 최혜민, 2014; 김선남 외, 2014)가 활발하게 진행되고 있다. 조직 차원의 연구에서도 개인정보 침해 위협 등이 사용자와 조직의 수용에 어떠한 영향을 끼치는지에 대한 개별 변인 차원의 연구는 시도되고 있으나(이환수 외, 2013; 장성호 외, 2014; 윤상오, 2013 등), 개인정보 보호 이슈와 기존의 조직 차원의 변인들과의 통합적인 수용 모형을 탐색하는 연구는 부족한 상황이다. 이에 따라 다음 장에서 빅데이터를 포함한 신기술에 대한 조직의 수용과 관련된 기존의 이론들과 개인정보보호와 관련된 선행요인을 통합하기 위한 논의를 이어가고자 한다.

## 2. 빅데이터 수용 관련 이론

### 1) 기술수용모형(TAM)

Davis(1989)의 기술수용모형(Technology Acceptance Model: TAM)은 정보기술의 수용에 관한 사용자의 행동을 설명하는 보편적인 이론적 틀로 인정되어 왔다. 기술의 수용 과정을 인간의 합리적 행동이론(Fishbein, et al., 1977)에 근거하여 사용자가 새로운 정보기술에 대해 유용성(Usefulness)과 사용용이성(Ease of Use)을 인지할 경우 기술에 대한 긍정적 태도(Attitude)와 지속적 사용의도를 형성하는 과정으로 설명하고 있다. 이후 기술수용모형은 구성변인들이 정보기술에 대한 개인의 태도를 중심으로 구성되어 있는 한계를 극복하기 위해 수용 과정 주변에서 영향을 주는 조직적, 사회적 환경요인들이 추가되면서 확장된 모형들이 제시되고 있다.

빅데이터에 대한 문헌상의 정의를 보았을 때 단순히 대용량의 데이터 자체를 의미하는 것이 아니라 데이터를 처리하고 활용하며, 조직의 업무에 적용하여 활용하는 기술에 대해 집합적으로 정의하고 있기 때문에(IDC, 2011; 국가정보화전략위원회, 2011 등), 빅데이터를 광의의 정보기술로 보고 기술수용모형에 의한 설명이 가능하다. 기술수용모형은 공공부문의 IT 정책형성과정의 연구에서도 자주 활용되고 있으며(권문주 외, 2009), 국내 빅데이터 초기 연구에서도 도입 및 활성화 요인을 파악하기 위해 기술수용모형을 적용하고 있다. 이선우 외(2014)는 빅데이터 시스템의 도입 과정을 설명하기 위해 기술-조직-환경 연계모형(TOE), 혁신 확산모형(DOI), 통합된 기술수용모형(UTAUT)을 통합한 연구모형을 제시하였다. 김은영 외(2013)는 수용조직의 환경요인으로 조직의 혁신성, 여유자원, 인프라 성숙도를, 빅데이터 시스템의 인지된 혜택을 기술적 특성 요인으로 선정하여 기술수용모형의 인지된 유용성과 사용용이성 변수와의 관계를 고찰하였다. 마찬가지로 김정선 외(2014)는 조직의 혁신 성향, 주관적 규범(Subjective Norm), 업무기술적합성(Task-

Technology Fit)이라는 변수를 기술수용모형의 인지된 유용성과 사용용이성 변수와 연계하였다. 또한 이러한 모형을 기술수용의 목적성에 따른 사용자 특성집단으로 나누어서 조절효과를 검증하였다.

### 2) 업무-기술적합모형(Task-Technology Fit Model)

업무-기술적합(Task-Technology Fit: TTF)은 기술, 업무, 사용자의 특성이 서로 맞아야(Fit) 개인의 성과를 이끌어낼 수 있다는 주장이다. Goodhue, et al.(1995)은 특히 정보기술의 역할에 초점을 맞추어 기술이 업무수행에 적절한 정보를 제공해 주고 업무처리를 원만하도록 하며 결과적으로 최적의 의사결정과정과 성과를 이끌어낼 것이라 주장하였다. 업무-기술적합모형은 기술수용모형처럼 사용자의 태도와 의도가 정보기술의 수용을 결정할 것이라는 입장 보다는 업무와 기술이 적절하게 연계되지만 한다면 사용자의 태도와 관계없이 수용을 결정할 수도 있다는 입장이다(Dishaw, et al., 2002). 이러한 근거에 따라 기술수용모형의 단점을 보완할 수 있기 때문에 두 모형을 통합하려는 연구가 시도되었고, 개별 모델을 적용한 것보다 설명력이 높다는 연구 결과가 나타나고 있다(Dishaw, et al., 1999).

빅데이터 사용 맥락에서는 빅데이터의 기술적 특성상 데이터의 수집 및 처리, 분석 단계에서 각각 활용하는 방식과 내용이 달라질 수 있다. 따라서 활용 주체별로 빅데이터 기술이 개인의 업무에 어떻게 맞는지(Fit)에 대하여 인식이 달라질 수 있을 것이다. 대표적으로 서비스 제공자가 느끼는 업무-기술 적합의 의미와 서비스 사용자가 인지하는 적합성에 그 차이가 있을 수 있다. 또한 빅데이터 기술은 아직 성숙기에 도달하지 않은 상태로 기존의 일반적인 정보시스템 보다 업무활용성의 범위가 아직 정해지지 않았고 미지의 영역이 남아 있다는 점에서 업무-기술적합 모형만을 적용하기에는 한계가 있을 것이다. 이러한 측면에서 업무-기술적합모형과 기술수용모형 등 다른 이론과 결합한 통합된 연구모형의 도출이 필요한 상황이다.

### 3) 프라이버시 계산 이론

인간이 어떤 행위를 할 때 그 결과로 초래될 이득과 손실을 평가하는 과정은 여러 이론들에서 검토가 되어 왔다. 사회적 교환이론에서는 다른 사람들과의 관계가 초래할 잠재적 보상과 벌을 식별한 후 전체 보상과 전체 벌을 계산하여 보상 쪽이 클 경우에 사회적 상호작용을 하게 된다고 주장하고 있으며(Homans, 1961), 기대 이론(Expectancy Theory)에서는 인간은 긍정적 결과를 최대화하고 부정적 결과를 최소화하는 쪽으로 행동한다고 주장한다(Vroom, 1964). 프라이버시 계산이론은 이러한 개념을 프라이버시와 관련된 맥락으로 옮겨와, 인간이 특정한 이득을 위해 일정 수준의 프라이버시 희생을 감수할 수 있다는 주장이다(Laufer, et al., 1977). 이 이론은 인터넷 환경에서 사용자가 개인정보에 대한 우려를 극복할 것인가를 설명할 수 있는 틀을 제공할 수 있기 때문에 SNS 등의 인터넷 서비스의 사용의도에 대한 연구에 사용되고 있다(김병수, 2012; 민진영 외, 2013; 차훈상, 2012; Dinev, et al., 2006). 또한 조직에서 절차적 정당성을 통해 개인정보에 대한 우려(Privacy Concern)를 해소할 수 있는가에 대한 설명이론으로 적용되기도 하였다(Culnan, et al., 1999).

모바일 어플리케이션이나 소셜 네트워크 서비스 등의 정보 서비스를 이용하면서 생성되는 위치정보, 검색패턴, 로그 등의 그림자 데이터는 개인의 프라이버시를 침해할 수 있는 가능성을 증가시킨다(이환수 외, 2013). 이처럼 빅데이터가 새로운 사업기회가 제공되기도 할 것이지만, 개인정보가 포함된 데이터 관리를 잘 못하는 경우에는 개인정보 또는 프라이버시의 침해에 대한 손해배상 책임의 부담으로 기업은 엄청난 경제적 손실을 입고, 그와 더불어 기업 이미지의 실추에 따라 시장의 경쟁에서 질 수밖에 없을 것이다(손영화, 2014).

따라서 빅데이터 사용 환경에서 서비스 이용자는 개인의 정보를 제공함으로써 얻을 수 있는 이익과 프라이버시의 침해에 따른 위험 비용을 고려하여 합리적

으로 결정할 것이다(박은경 외, 2014). 또한 서비스 공급자들도 개인정보를 공개하거나 사용해야 하는 상황에서 프라이버시 침해에 대한 우려를 평가하고 다양한 요인들에 대한 위험-혜택 분석(Privacy Calculus Model)을 하게 될 것이다(차훈상, 2012).

## Ⅲ. 연구모형 및 가설

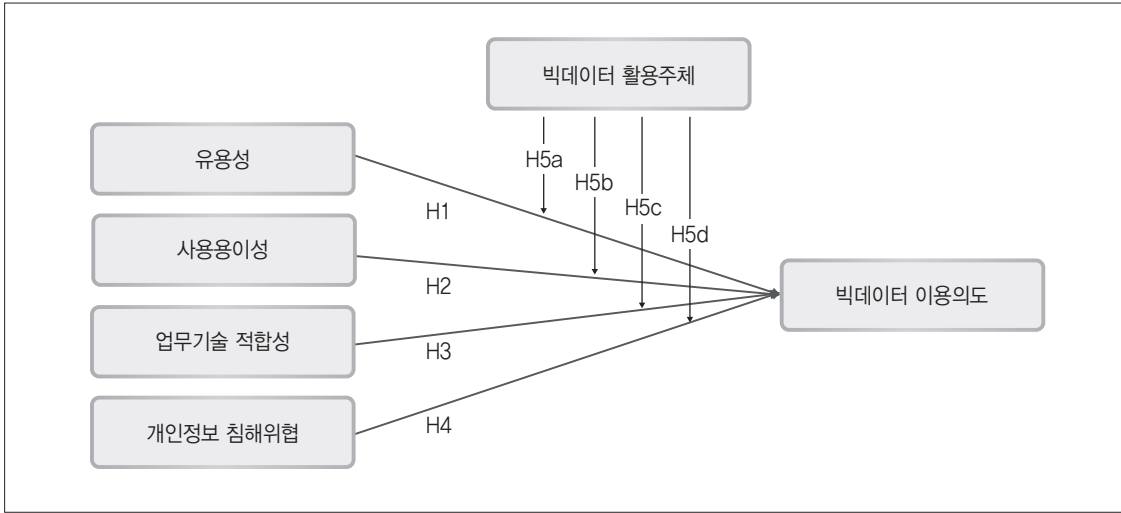
### 1. 연구 모형

앞장의 선행 이론에 대한 검토 결과에 따라 <그림 1>과 같은 연구모형을 수립하였다. 먼저 빅데이터 기술수용모형에서 기본적인 출발점을 삼아 종속변수로 빅데이터 이용의도를 채택하였고, 선행변수로 인지된 유용성과 사용용이성을 채택하였다. 다음으로 업무-기술적합모형과 프라이버시 계산모형을 결합하여 인지된 이득에 해당하는 변수로 업무기술적합성이라는 변수를 채택하였고, 인지된 손해에 해당하는 변수로 인지된 개인정보 침해위험이라는 변수를 채택하였다. 마지막으로 빅데이터 활용 주체는 조절변수로 채택하였다.

### 2. 연구 가설

#### 1) 빅데이터의 인지된 유용성과 사용용이성

Davis(1989)는 인지된 유용성(Perceived Usefulness)을 새로운 기술이 조직의 업무 성과의 달성에 미치는 지각된 효과성으로 정의하였다. 빅데이터 처리 기술과 시각화 기술 등은 복잡한 데이터를 효율적으로 처리하고 의사결정에 필요한 요약 정보를 제공해 줄 수 있기 때문에 조직 내 업무에 유용하게 활용될 수 있다. 인지된 사용용이성(Ease of Use)은 사용자가 많은 노력을 들이지 않고 기술을 사용할 수 있는 정도를 의미한다(Davis, 1989). 날로 발전하는 공개소프트웨어 기반 데이터 처리 알고리즘과 사용자 중심의 솔루션들은 비교적 적은 비용으로 빅데이터를 활용할 수 있는 방법



〈그림 1〉 연구 모형

을 제시하고 있다. 다수의 조직연구에서 기술수용모델(TAM)에서 제시된 인지된 유용성과 사용용이성 정보 기술에 대한 긍정적인 태도와 이용의도에 영향을 미치는 것으로 확인되고 있는 바, 본 연구에서도 이러한 근거에 따라 빅데이터에 대한 인지된 유용성과 사용용이성이 빅데이터 이용의도에 긍정적인 영향을 줄 것이라 가정하고 다음과 같은 두가지 연구가설을 수립하였다.

*H1: 빅데이터에 대한 인지된 유용성은 빅데이터 이용의도에 긍정적인 영향을 줄 것이다.*

*H2: 빅데이터에 대한 인지된 사용용이성은 빅데이터 이용의도에 긍정적인 영향을 줄 것이다.*

2) 업무기술적합성

업무기술적합(Task Technology Fit) 이론에서는 새로운 기술이 특정 과업의 필요조건에 적절한 적합성이 존재한다면 그 기술은 적용될 수 있을 것으로 본다. 정보기술(IT)분야에서는 주로 업무를 수행할 때 필요한 기능을 IT가 얼마나 잘 지원해 주는가의 정도로 정의되고 있다(Goodhue, 1995). 이 이론을 빅데이터 맥

락에 적용할 때도 마찬가지로 빅데이터 수집, 처리, 분석의 목적이 조직의 과업에 대한 수요를 충족시키고 특정 업무에 부합하는지에 대한 사용자의 태도가 형성될 경우 이용의도와 연계될 것이다. 본 연구에서는 업무기술적합 모형 자체의 검증이 아니라 빅데이터 도입 상황에서 업무기술 적합도와 이용활성화 사이의 집단별 차이를 검증하는 것이 주 목적이다. 따라서 연구모형의 간결성을 위해서 해당 개념을 빅데이터의 ‘인지된 업무기술적합성’으로 조작적 정의를 통해 본 연구에 적용하고자 한다. 이에 따라 다음과 같이 연구가설을 수립하였다.

*H3: 빅데이터의 인지된 업무기술적합성이 높을수록 빅데이터 이용의도는 높아질 것이다.*

3) 프라이버시 침해위험

프라이버시 위험은 소비자가 특정 행위를 할 때 주관적으로 지각하는 위험으로써, 의사결정을 위한 목적과 결과를 알 수 없을 때 사용자가 느끼는 불확실성이라고 볼 수 있다(이현중, 2009). 프라이버시 위험은 인터넷 환경에서 발생 가능한 모든 프라이버시 침해

위험에 대한 인지이며(이미나 외, 2009; Bellman, et al., 2004), 특히 제 3자에게 유포 또는 제공될 가능성이 높은 소셜네트워크서비스(SNS)에서 침해 가능성이 더 높아질 수 있다(장성호 외, 2014). 따라서 소셜 네트워크에서 발생하는 빅데이터를 이용하여 마케팅 분석 등 서비스를 제공하는 집단과 이를 사용하는 집단에서는 의도적이든 비의도적이든 개인정보가 공개되는 상황에 대해 위험-혜택 분석(Privacy Calculus Model)을 하게 된다(차훈상, 2012). 즉 개인의 정보를 제공 또는 사용함으로써 얻을 수 있는 이익과 프라이버시의 침해에 따른 위험 비용을 고려하여 수용 의도를 합리적으로 결정할 것이다(박은경 외, 2014). 결과적으로 인지된 프라이버시 침해 위험이 도입에 따른 혜택보다 더 심각하다고 판단할 경우 이용을 거부하려는 태도를 보일 것이다(문혜정 외, 2012). 이러한 근거에 따라 다음과 같이 연구가설을 수립하였다.

*H4: 빅데이터의 인지된 프라이버시 침해 위험이 높을수록 빅데이터 이용의도는 낮아질 것이다.*

#### 4) 빅데이터 활용 주체

빅데이터는 단일한 정보시스템이나 솔루션이 아닌 데이터의 수집부터 전처리, 분석과 활용 과정의 전반적인 기술을 총칭하는 의미로, 빅데이터의 가치 사슬 단계별로 다양한 활용 주체가 참여한다. 백인수(2012)는 빅데이터 생태계를 3개의 구성요소인 서비스 공급업체, 서비스 사용자와 어플리케이션 공급업체로 나누어 구분하였다. 미래창조과학부(2014)는 데이터 제공자, 플랫폼 운영자, 서비스 이용자로 분류하였다. 본 연구에서 가장 관심 있는 활용 주체는 빅데이터 기술 수용 목적에 따른 공급자와 수요자 집단이다. 빅데이터 서비스 공급자는 빅데이터 수집, 저장, 관리, 분석 단계에서 데이터, 서비스, 어플리케이션, 인프라 등을 제공하는 이해관계자를 의미하며, 빅데이터 서비스 수요자는 이를 활용하여 의사결정이나 업무에 적용하여 가치를 창출하고자 하는 이해관계자를 의미한다. 빅데

터 수용의 부정적 요인인 개인정보 침해 위험에 대해 주로 개인정보를 수집하고 재가공하여 서비스를 개발하는 서비스 공급자와 이를 활용하는 수요자 집단 간에는 각각의 업무 특성에 따라 빅데이터에 대한 긍정적 인식(유용성, 사용용이성, 업무기술적합성)과 부정적 인식(프라이버시 침해위험)에 대한 인식의 차이가 클 것으로 예상된다. 마찬가지로 빅데이터 투자 여력과 관련 있는 기업의 규모, 빅데이터 기술에 대한 이해도가 다른 IT업종과 비 IT업종, 정책적 접근의 방식이 다른 공공기관과 민간기업 등 업종에 따른 집단 간에도 인식의 차이가 있을 것으로 예상된다. 이에 따라 각 활용 주체 별 집단 간 인식의 차이와 변수 간 인과관계의 조절효과를 검증하기 위하여 다음과 같이 가설을 수립하였다.

*H5a: 빅데이터 활용 주체에 따라 인지된 유용성이 이용의도에 미치는 영향이 차이가 있다.*

*H5b: 빅데이터 활용 주체에 따라 인지된 용이성이 이용의도에 미치는 영향이 차이가 있다.*

*H5c: 빅데이터 활용 주체에 따라 인지된 업무기술 적합성이 이용의도에 미치는 영향이 차이가 있다.*

*H5d: 빅데이터 활용 주체에 따라 인지된 프라이버시 침해 위험이 이용의도에 미치는 영향이 차이가 있다.*

### 3. 측정도구 개발 및 데이터 수집

#### 1) 연구변수와 측정도구 개발

연구변수들은 아래 <표 2>와 같이 기술수용모형과 업무기술 적합 모형, 프라이버시 계산 모형 등 선행연구에서 제시된 변수를 빅데이터 맥락에 맞추어서 조작적으로 정의하였다. 변수 측정을 위한 설문 항목은 선

〈표 2〉 연구변수의 조작적 정의와 개발된 설문 항목

변수	조작적 정의 / 설문 항목	관련연구
유용성	빅데이터와 관련된 서비스와 기능이 직무에 유용하게 활용될 것이라고 생각하는 정도	Davis, et al.(1989) Venkatesh, et al.(2000)
	1) 빅데이터는 필요한 정보를 빠르게 제공해 줄 것이다	
	2) 빅데이터는 업무성과를 향상 시킬 수 있을 것이다	
	3) 획득한 빅데이터는 유용하게 사용될 것이다	
사용용이성	빅데이터와 관련된 서비스와 기능이 많은 노력을 기울이지 않아도 편리하게 사용할 수 있다고 인지하는 정도	Davis, et al.(1989) Venkatesh, et al.(2000)
	1) 빅데이터는 사용하기 쉽고 이해하기 쉬울 것이다	
	2) 빅데이터는 쉽게 배울 수 있을 것이다	
	3) 빅데이터는 능숙하게 다루기 쉬울 것이다	
업무기술적합성	빅데이터와 관련된 서비스와 기능이 업무를 수행하는데 적합하다고 인지하는 정도	Goodhue, et al.(1995)
	1) 빅데이터는 업무수행에 적합하다	
	2) 빅데이터를 통해 업무처리에 필요한 충분한 정보를 얻을 수 있다	
	3) 빅데이터를 통해 원하는 정보를 쉽고 빠르게 얻을 수 있게 한다	
프라이버시 침해위험	빅데이터 관련 서비스와 기능을 활용할 때 개인정보 침해위험이 높아진다고 인지하는 정도	Malhotra, et al.(2004) Bulgurcu, et al.(2010) Bollier(2010)
	1) 빅데이터 활용 시 개인정보 침해위험을 쉽고 빠르게 대응 할 수 있다(-)	
	2) 빅데이터의 활용 시 개인정보 침해 가능성이 높아진다	
	3) 빅데이터 활용 시 개인정보 침해 방지기술은 효과적이다(-)	
이용의도	빅데이터와 관련된 서비스와 기능을 지속적으로 이용(하게) 할 의도	Davis, et al.(1989) Venkatesh, et al.(2000)
	1) 빅데이터를 지속적으로 이용하게 할 것이다.	
	2) 빅데이터를 업무에 많이 활용될 것이다.	
	3) 빅데이터의 새로운 활용 기회에 이용할 것이다.	
	4) 빅데이터를 계속 사용하게 할 것이다.	

주: 지면의 한계로 문구는 요약하였으며, (-) 표시된 항목은 Reverse Coding 후 적용.

행연구에서 사용된 항목을 대부분을 빅데이터 활용 상황에 맞추어 대폭 보완하였는데, 그 이유는 아직 다수의 기업에서 빅데이터의 도입이 이루어지지 않은 상황에서 응답자의 예상 또는 기대치에 의존할 수 밖에 없기 때문이다. 또한 종속변수인 빅데이터 이용의도의 경우, 본 연구 목적 중 하나인 수요자와 서비스 공급자

의 인식 차이를 확인하기 위해서 공급자 관점에서 빅데이터 이용의도까지 포괄할 수 있도록 설문을 미묘하게 수정하였다. 따라서 수정 및 신규 개발된 설문항목은 별도의 내용타당성 확인 과정이 필요하여, 국내 39명의 빅데이터 산학연 전문가 집단의 예비 검토와 그 결과에 따른 보완 과정을 통해 설문 문항의 모호성을



〈표 3〉 응답자의 인구통계학적 특성

구분		표본		
		빈도	퍼센트	누적(%)
성별	남자	370	71.4	71.4
	여자	148	28.6	100.0
	합계	518	100.0	
나이	18-24	38	7.3	7.3
	25-34	250	48.3	55.6
	35-44	164	31.7	87.3
	45-54	63	12.2	99.4
	55이상	3	0.6	100.0
	합계	518	100.0	
업종	제조업	45	8.7	8.7
	의료	7	1.4	10.0
	학원/교육	41	7.9	18.0
	은행/금융/보험	28	5.4	23.4
	정부기관	11	2.1	25.5
	정보기술	285	55.0	80.5
	공기업	12	2.3	82.8
	기타	89	17.2	100.0
합계	518	100.0		
기업형태	대기업	89	17.2	17.2
	중견기업	68	13.1	30.3
	중소기업	212	40.9	71.2
	소기업	51	9.8	81.1
	개인사업자	4	0.8	81.9
	기타	94	18.1	100.0
	합계	518	100.0	
고용상태	정규직	381	73.6	73.6
	임시직	16	3.1	76.6
	시간제 근무직	4	0.8	77.4
	계약직	40	7.7	85.1
	기타	77	14.9	100.0
	합계	518	100.0	
직무	경영진	44	8.5	8.5
	기술직	144	27.8	36.3
	전문직	96	18.5	54.8
	관리직/사무직	147	28.4	83.2
	기타	87	16.8	100.0
	합계	518	100.0	
이용형태	서비스 공급자	267	51.5	51.5
	서비스 수요자	251	48.5	100.0
	합계	518	100.0	

제거하고 측정 목적에 부합하도록 지표의 적정성을 확인하였다.<sup>1)</sup> 측정 문항별 척도는 모두 리커트 7점 척도를 사용하였다.

2) 표본 데이터 수집

본 연구의 모집단은 빅데이터를 활용할 가능성이 있는 잠재적 사용자 집단이다. 모집단의 특성상 확률적 방법에 기반을 두어 표본을 추출하기 어렵기 때문에 편의적 샘플링 방법을 사용하였다. 한국정보화진흥원의 K-ICT 빅데이터센터의 메일링 리스트에서 1,500여명의 최초 응답자 후보군을 추출하였고 2015년 11월부터 약 1개월 동안 온라인 설문조사를 통해 설문지를 전달하고 응답을 요청한 결과 518명의 유효한 설문을 확보하였다. 응답률은 약 34.5%이며 응답자의 인구통

계학적 분포는 다음 <표 3>과 같다.

3) 측정도구의 타당성 검증

측정도구의 타당성은 먼저 탐색적 요인분석(Exploratory Factor Analysis)을 통해 전반적인 구성타당성과 신뢰도를 평가하였고, 구조방정식 모형 분석에 의한 확인적 요인분석(Confirmatory Factor Analysis)을 통해 집중타당성, 판별타당성을 확인하였다.

요인분석은 주성분 분석(PCA: Principal Component Analysis) 기법을 사용하였고, 직각회전(Varimax Rotation)방식을 사용하여 추출된 요인의 고유값(Eigen Value)이 1 이상인 요인을 최종 선정하였다 그 결과 각 변수와 요인간의 상관관계의 정도를 나타내는

<표 4> 확인적 요인분석 결과(요약)

구분	C.R.	AVE	상관계수					측정항목	표준화 적재값		
			1	2	3	4	5				
유용성	0,824	0,540	-						PU_3	0,851	
									PU_4	0,855	
									PU_2	0,841	
									PU_1	0,811	
사용용이성	0,819	0,601	0,138	-					PE_3	0,935	
									PE_2	0,938	
									PE_1	0,868	
업무기술적합성	0,826	0,613	0,501	0,180	-				TTF_3	0,854	
									TTF_2	0,907	
									TTF_1	0,83	
개인정보 침해위험	0,867	0,621	0,163	0,088	0,166	-				PI_1	0,839
										PI_2	0,892
										PI_3	0,863
										PI_4	0,772
이용의도	0,851	0,587	0,404	0,127	0,410	0,397	-			ITU_1	0,881
										ITU_2	0,891
										ITU_3	0,912
										ITU_4	0,882

1) 각각의 전문가가 “일반적 설명”과 “연구의 목적”이라는 이름의 설문초안을 받았다. 그들은 다음과 같은 세 개의 방식으로 응답하였다. (1) 각각의 문항과 참고자료가 질문에 정확하게 대답하도록 도와주는가를 확인하는 것 (2) 연구에 적합하다고 여겨지는 다른 요인들을 제시할 것 (3) 전체적으로 질문 도구의 구조와 내용물에 대해 평할 것.

요인적재량이 0.5 이상(노형진, 2001)인 요인들을 유의미하게 추출하였고 일부 지표를 제외한 최종 측정항목에서 사전에 정의한 독립변수, 종속변수의 5개 요인이 정상적으로 추출되었고 각 요인의 공통성 분석결과는 모두 0.7이상으로 나타났다. 도출된 요인별 신뢰도 분석 결과 모든 요인항목 별 cronbach  $\alpha$  계수가 0.8 이상으로 나타났다.

다음으로 구조방정식모형을 이용하여 측정항목과 구성변수들에 대한 확인적 요인분석을 실시하였다. 통계패키지 AMOS 21을 이용하여 구조모형 분석을 실시한 결과 <표 4>와 같이 개념 신뢰도확인을 위한 CR (CR: Construct Reliability)값과 집중타당도 확인을 위한 평균분산 추출지수(AVE: Average Variance Extracted)가 각각 일반적인 검증 기준인 0.7과 0.5를 상회하고 있고(Fornell, et al., 1981; 김계수, 2004), 분산추출지수(AVE)의 제곱근과 구성개념 간 상관계수 비교를 통해 판별타당도가 확인되었다.

#### IV. 분석 결과

##### 1. 모형 적합도 및 가설 검증

본 연구모형과 가설의 검증을 위해서 기술통계분석 및 다중회귀분석을 실시하였다. 먼저 연구모형의 적합도 판단을 위해 조사대상 응답자의 인구통계학적 자료 중 연구목적과 연구문제에 적합한 변인인 산업분야(민간-공공, IT-비IT), 활용목적(공급자-수요자), 규모(대기업-중소기업, 100명이상-미만) 등은 각각 0과 1의 값을 갖는 더미(Dummy) 변수로 치환하여 통제변수로 모델에 투입하였다. 다중회귀분석 결과 <표 5>와 같이 통제변수만 투입한 회귀모형보다 독립변수를 추가로 투입한 회귀모형의 설명력이 59.2%로 적합한 수준으로 유의미하게 증가한 것을 확인할 수 있었다.

연구모형에서 설정한 독립변수들의 종속변수에 대한 회귀계수의 통계적 검증 결과 유용성 변수와 업무기술적합성, 개인정보침해위험 변수가 99% 유의수준에서 유의미한 영향을 미치는 것으로 확인되었다. 다만 사용용이성 변수는 종속변수에 유의미한 영향이 없

<표 5> 회귀분석 결과

변수	통제변수(Control)				주효과(Main Effect)			
	B	SE	표준화 계수	t	B	SE	표준화계수	t
(상수)	4.713	.171		27.608	3.814	.242		15.783
공공	.525	.269	.088	1.952 <sup>^</sup>	.195	.176	.033	1.106
직종(IT)	.150	.118	.061	1.269	-.031	.078	-.013	-.404
주체(공급자)	-.075	.123	-.030	-.614	-.021	.080	-.009	-.264
규모(중소기업)	.265	.141	.099	1.873	.058	.092	.022	.631
규모(100명이상)	.328	.131	.132	2.512*	.074	.086	.030	.865
유용성					.265	.036	.298	7.293**
사용용이성					.039	.032	.041	1.213
업무기술적합성					.209	.040	.222	5.221**
개인정보침해위험					-.388	.032	-.396	-12.29**
R2	0.025				0.592			
Adjusted R2	0.015				0.584			
R2 변화량	-				0.567**			

주: <sup>^</sup> p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

〈표 6〉 공급자-수요자 집단간 평균 차이

구성변수		N	평균	표준편차	유의성 (p(0.05))	비고
유용성	공급자	267	4.67	1.422	-	등분산 가정
	수요자	251	4.68	1.343		
사용용이성	공급자	267	4.05	1.248	-	등분산 가정
	수요자	251	4.07	1.332		
업무기술적합성	공급자	267	4.65	1.309	-	등분산 가정
	수요자	251	4.60	1.206		
개인정보침해위험	공급자	267	2.71	1.158	○	등분산 미가정
	수요자	251	3.04	1.331		
이용의도	공급자	267	5.12	1.154	-	등분산 가정
	수요자	251	4.97	1.173		

는 것으로 확인되었다. 따라서 빅데이터에 대한 인지된 유용성은 빅데이터 이용의도에 긍정적인 영향을 줄 것이라는 가설 H1은 채택되었고 빅데이터에 대한 인지된 사용용이성은 빅데이터 이용의도에 긍정적인 영향을 줄 것이라는 가설 H2는 기각되었다. 빅데이터의 인지된 업무기술적합성이 높을수록 빅데이터 이용의도는 높아질 것이라는 가설 H3과 빅데이터의 인지된 프라이버시 침해 위험이 높을수록 빅데이터 이용의도는 낮아질 것이라는 가설 H4는 채택되었다. 각 독립변수의 표준화 회귀계수는 개인정보침해위험이 -.396으로 가장 높게 나타났고 다음으로 인지된 유용성, 업무기술적합성 순으로 나타나 빅데이터의 긍정적인 도입요인 보다 부정적 도입요인의 종속변수에 대한 영향이 큰 것을 알 수 있다.

## 2. 집단간 차이 및 조절효과 분석

활용 주체와 특성에 따른 집단 간 차이 분석을 위해 응답자의 소속 조직 정보를 통해 각 조절변수 별로 집단을 상대적으로 큰 차이를 보이도록 두 개의 집단으로 나누었다. 먼저 집단 별 연구모형의 변수 별 평균값의 차이분석을 통해 전반적인 인식의 차이를 살펴보고, 다음으로 다중회귀모형의 각 회귀계수 별 집단 간 차이 유의성 검정을 통해 조절효과를 확인하였다. 집단 간 평균값의 차이 분석은 t검정을 이용하였고, 집단 별 회귀계수별 차이 검정은 (Paternoster, et al., 1998; Clogg, et al., 1995) 가 제안한 산식<sup>2)</sup>을 통해 z 값의 단측검정을 통해 확인하였다.

### 1) 활용 목적에 따른 차이

빅데이터 서비스 공급자와 수요자 집단별로 각 변수 별 평균의 차이에 대한 t검정을 실시한 결과는 다음 〈표 6〉과 같다. 다섯 개의 연구 변수 중 개인정보침해 위험에 대해서만 유의미한 차이를 보였고(유의수준

2) Brame et a.(1988) 등은 표본이 충분히 클 경우(N>30) z값을 산출하는 전통적인 방법(Cohen, 1983 등)은 표준오차의 편의를 감소시켜 z값이 왜곡될 수 있어 다음과 같이 수정된 산식을 제안하였다.

$$SE_{b-difference} = \sqrt{(SE_{bG1}^2 + SE_{bG2}^2)} \quad \text{and...} \quad Z = \frac{b_{G1} - b_{G2}}{\sqrt{(SE_{bG1}^2 + SE_{bG2}^2)}}$$

(bGi : i번째 집단의 회귀계수, SEbgi : i번째 집단의 회귀계수 표준오차)

〈표 7〉 공급자-수요자 집단간 회귀계수 차이

경로	공급자		수요자		회귀계수 차이
	B	SE	B	SE	
유용성 → 이용의도	.183**	.051	.355**	.051	-.172**
사용용이성 → 이용의도	.022	.046	.041	.045	-.019
업무기술적합성 → 이용의도	.265**	.059	.179**	.054	.086
개인정보침해위험 → 이용의도	-.431**	.048	-.343**	.041	-.088^

주: ^ p < 0.1, \* p < 0.05, \*\* p < 0.01

〈표 8〉 기업규모(대기업/중소기업) 집단 간 평균차이 분석

구성변수		N	평균	표준편차	유의수준 (p(0.05))	비고
유용성	대기업	89	4.89	1.522	○	등분산 가정
	중소기업	212	4.50	1.424		
사용용이성	대기업	89	4.26	1.557	-	등분산 미가정
	중소기업	212	3.93	1.157		
업무기술적합성	대기업	89	4.66	1.476	-	등분산 가정
	중소기업	212	4.53	1.239		
개인정보침해위험	대기업	89	2.90	1.500	-	등분산 미가정
	중소기업	212	2.82	1.194		
이용의도	대기업	89	5.04	1.439	-	미등분산 가정
	중소기업	212	4.98	1.102		

95%), 빅데이터서비스 수요자가 개인정보 침해에 대한 우려가 높은 것을 알 수 있다.

다음으로 두 집단 간 회귀계수의 차이를 검증한 결과 유용성이 이용의도에 미치는 긍정적 영향에 대해서는 공급자, 수요자 집단 모두 유의한 수준이나 수요자 집단이 더 민감하게 인식하고 있었으며, 개인정보 침해 위험이 이용의도에 미치는 부정적 영향에 대해서는 마찬가지로 두 집단 모두 유의한 수준이나 공급자 집단이 더 민감하게 인식하고 있는 것으로 나타났다.

종합해 보면 공급자, 수요자 집단 모두 빅데이터의 유용성을 인식하고 지속적으로 이용할 의도를 보이고 있으나 수요자 측이 그 경향이 강한 한편, 개인정보 침해에 따른 위험성과 그에 따른 빅데이터 이용에 대한 위축 심리는 오히려 공급자 집단이 더 강한 것으로 해

석할 수 있었다.

## 2) 조직 규모에 따른 차이

소속기업의 유형에 대한 응답을 토대로 대기업과 대기업 이외의 기업(중소기업, 1인기업, 벤처 등)을 중소기업으로 분류하여 집단을 나누었다. 대기업과 중소기업의 연구변수 별 평균값의 차이에 대한 t검정을 실시한 결과는 〈표 8〉과 같이 다섯 개의 연구 변수 중 유용성에 대해서만 유의미한 차이를 보였고(유의수준 95%), 대기업 종사자가 유용성에 대해 더 크게 인식하는 것으로 나타났다.

다음으로 두 집단 간 회귀계수의 차이를 검증한 결과 업무기술적합성이 이용의도에 미치는 긍정적 영향에 대해서는 중소기업 집단에서만 유의한 수준으로 나

〈표 9〉 기업규모(대기업/중소기업) 집단 간 평균 차이

경로	대기업		중소기업		회귀계수 차이
	B	SE	B	SE	
유용성 → 이용의도	.290**	.072	.250**	.041	.040
사용용이성 → 이용의도	.075	.058	.026	.039	.049
업무기술적합성 → 이용의도	.104	.081	.262**	.045	-.157*
개인정보침해위험 → 이용의도	-.447**	.061	-.364**	.036	-.083

〈표 10〉 종사자수(100명 미만/100명 이상) 집단 간 평균차이 분석

구성변수		N	평균	표준편차	유의수준 (p(0.05))	비고
유용성	100명 미만	285	4.54	1.360	○	등분산 가정
	100명 이상	231	4.83	1.397		
사용용이성	100명 미만	285	3.95	1.190	○	등분산 미가정
	100명 이상	231	4.19	1.391		
업무기술적합성	100명 미만	285	4.51	1.212	○	등분산 가정
	100명 이상	231	4.76	1.304		
개인정보침해위험	100명 미만	285	2.89	1.199	-	등분산 가정
	100명 이상	231	2.85	1.327		
이용의도	100명 미만	285	4.96	1.111	-	등분산 가정
	100명 이상	231	5.14	1.226		

〈표 11〉 종사자수(100명 미만/100명 이상) 집단 간 회귀계수 차이

경로	100명미만		100명 이상		회귀계수 차이
	B	SE	B	SE	
유용성 → 이용의도	.254**	.048	.268**	.056	-.014
사용용이성 → 이용의도	.029	.043	.051	.050	-.022
업무기술적합성 → 이용의도	.225**	.053	.205**	.062	.020
개인정보침해위험 → 이용의도	-.385**	.041	-.392**	.049	.008

타났다. 개인정보 침해 위험이 이용의도에 미치는 부정적 영향에 대해서는 근소한 차이로 통계적 유의성에 미치지 못하였지만 대기업이 조금 더 민감하게 인식하고 있는 것으로 나타났다.

다음으로 소속기업의 임직원수에 대한 응답을 토대

로 100명 미만과 100명 이상으로 집단을 나누어서 분석한 결과 〈표 10〉과 같이 다섯 개의 연구 변수 중 유용성, 사용용이성, 업무기술적합성에 대해 유의미한 차이를 보였고(유의수준 95%), 전반적으로 100명 이상 기업의 종사자가 높은 응답치를 나타내고 있다. 하

〈표 12〉 IT업종-비IT업종 간 평균 차이

구성변수		N	평균	표준편차	유의수준 (p(0.05))	비 고
유용성	정보통신업	285	4.64	1.304	-	등분산 미가정
	비정보통신업	233	4.73	1.474		
사용용이성	정보통신업	285	4.03	1.164	-	등분산 미가정
	비정보통신업	233	4.09	1.427		
업무기술적합성	정보통신업	285	4.57	1.241	-	등분산 가정
	비정보통신업	233	4.69	1.279		
개인정보침해위험	정보통신업	285	2.86	1.226	-	등분산 가정
	비정보통신업	233	2.89	1.292		
이용의도	정보통신업	285	5.01	1.090	-	등분산 미가정
	비정보통신업	233	5.08	1.251		

〈표 13〉 IT업종-비IT업종 간 회귀계수 차이

경로	IT업종		비 IT업종		회귀계수 차이
	B	SE	B	SE	
유용성 → 이용의도	.230**	.048	.302**	.055	-.072
사용용이성 → 이용의도	.050	.044	.018	.049	.032
업무기술적합성 → 이용의도	.250**	.051	.183*	.063	.068
개인정보침해위험 → 이용의도	-.346**	.041	-.434**	.047	.087^

지만 두 집단 간 회귀계수의 차이를 검증한 결과는 모든 경로에 대해서 유의미한 차이를 나타내지 않았다.

### 3) 업종에 따른 차이

IT업종 종사자 집단과 비IT업종 종사자 집단 간 각 변수별 평균의 차이에 대한 t검정을 실시한 결과는 다음 〈표 12〉와 같이 다섯 개의 연구 변수 모두 집단 간 유의미한 평균 차이는 존재하지 않았다. 다음으로 IT업종과 비IT업종 집단 간 회귀계수의 차이를 검증한 결과 개인정보 침해 위험이 이용의도에 미치는 부정적 영향에 대해서는 마찬가지로 두 집단 모두 유의한 수준

이나 비IT업종이 더 민감하게 인식하고 있는 것으로 나타났다. 나머지 변수들에 대해서는 집단 간 차이를 보이지 않았다. 개인정보 침해에 따른 위험성과 그에 따른 빅데이터 이용에 대한 위축 심리는 비IT업종이 더 강한 것으로 해석할 수 있었다.

다음으로 공공기관과 민간기업 집단 간 변수별 평균의 차이에 대한 t검정을 실시한 결과는 다음 〈표 14〉와 같이 업무기술적합성과 이용의도에서 유의미한 평균 차이를 보였다.<sup>3)</sup> 두 변수 모두 공공기관 종사자 집단에서 더 높은 점수를 보여주고 있다. 다만 두 집단 간 회귀계수의 차이를 검증한 결과는 모든 경로에 대

3) 응답자 중 공공기업 종사자의 표본수가 상대적으로 적기 때문에 제한적으로 해석할 필요가 있다.

〈표 14〉 공공-민간 집단 간 평균 차이

구성변수		N	평균	표준편차	유의수준 (p<0.05)	비고
유용성	공공	23	4.80	1.387	-	등분산 가정
	민간	495	4.67	1.384		
사용용이성	공공	23	4.38	1.649	-	등분산 미가정
	민간	495	4.04	1.269		
업무기술적합성	공공	23	5.22	.935	○	등분산 가정
	민간	495	4.51	1.312		
개인정보침해위험	공공	23	2.51	1.091	-	등분산 가정
	민간	495	2.89	1.260		
이용의도	공공	23	5.64	.962	○	등분산 가정
	민간	495	5.08	1.238		

〈표 15〉 공공-민간 집단 간 회귀계수 차이

경로	공공		민간		회귀계수 차이
	B	SE	B	SE	
유용성 → 이용의도	.248*	.115	.272**	.038	-.024
사용용이성 → 이용의도	.055	.080	.033	.034	.021
업무기술적합성 → 이용의도	.083	.173	.216**	.041	-.133
개인정보침해위험 → 이용의도	-.531**	.123	-.379**	.032	-.152

해서 유의미한 차이를 나타내지 않았다.

## V. 결론

### 1. 연구 결과의 종합적 논의

본 연구는 국내 빅데이터의 활성화에 미치는 긍정적 요인과 부정적 요인의 통합적 이론 개발과 이해관계자 집단에 따른 인식 차이를 탐색하기 위해 수행되었다. 이를 위해 기술 도입 과정을 설명하는 전통적인 기술수용모형과 위험과 혜택에 대한 비교를 통한 도입 의도를 설명하는 업무기술적합성 모형과 프라이버시 계산 모형을 접목한 연구모형을 개발하였고 국내 빅데이터 사용자 또는 잠재적 사용자를 대상으로 실증하였다.

분석 결과를 통해 나타난 이론적, 정책적 시사점은 다음과 같다.

첫째로, 전통적인 기술수용모형은 빅데이터의 이용 상황을 완전히 설명하지 못하고 있다. 주요 선행 변수 중 인지된 유용성은 지속적 이용의도에 유의미한 영향을 미쳤으나 사용용이성은 연관이 없는 것으로 나타났다. 각 지표에 대해서도 응답자는 공통적으로 빅데이터의 유용성보다 사용용이성에 대하여 낮은 점수로 응답하였는데, 이는 빅데이터의 유용성보다 새로운 인프라와 분석 기법 등 새로운 기술을 습득해야 하는 것에 대한 부담감을 가지고 있다고 해석할 수 있다. 한편 업무기술적합 모형과 프라이버시 계산 모형에서 참조한 업무기술적합성과 개인정보 침해위험은 모두 이용의도에 유의미한 영향을 미치고 있으며, 표준화 회귀계



수를 비교해 보았을 때 이중 개인정보 침해위험이 이용의도와 가장 민감하게 영향을 주는 것으로 나타났다. 즉 사용자 입장에서 빅데이터가 아무리 유용하게 업무에 활용할 수 있다 하더라도 위험-혜택 분석에서 아직은 위험(프라이버시 침해)에 대한 우려가 기대된 혜택을 상쇄하지 못하고 있는 상황으로 해석된다. 본 연구 결과를 바탕으로 위험과 혜택을 균형적으로 설명할 있는 이론적 프레임워크의 지속적인 개발을 통해 빅데이터, 인공지능 등 최신 기술이 조직에 도입되는 과정을 더욱 잘 설명할 수 있을 것으로 보인다.

둘째로, 가능한 모든 유형의 활용 주체를 산정하여 각 주체 별로 빅데이터 도입 과정에서 미묘하지만 서로 다른 인식의 차이를 발견하였고, 빅데이터 활용과 산업 활성화 정책 수립을 위해 필요한 기초적인 정보를 제공하였다. 구체적인 내용은 다음과 같다.

먼저 빅데이터의 유용성과 업무기술적합성 등 빅데이터의 긍정적인 특성에 대해서는 상대적으로 큰 조직이, 민간보다는 공공 부문이 더 크게 인식하고 있는 것으로 나타났다. 즉 상대적으로 투자 여력이 있는 대기업과 초기 빅데이터 도입을 주도하고 있는 공공부문에서는 빅데이터 활성화를 위한 초기 여건이 형성되어 있음을 의미한다. 따라서 아직 빅데이터에 대한 인식이 형성되어 있지 않은 민간의 중소기업을 대상으로 맞춤형 컨설팅과 적은 투자비용으로 효과를 볼 수 있는 성공사례의 발굴과 홍보가 필요한 시점이다. 다음으로 빅데이터 서비스 공급자와 수요자 중에 수요자 집단이 빅데이터의 유용성 인지에 따른 이용의도가 높게 형성되어 있는 것으로 나타났는데, 이는 수요자 입장에서 가치에 대한 확신과 성공사례가 지속적으로 확산된다면 앞으로 빅데이터 서비스 수요가 더욱 증가할 것이 예상되는 결과이다.

한편 개인정보 침해 위험에 대해서는 공급자-수요자 집단에서만 인식의 차이를 보이고 있다. 또한 두 집단 모두 개인정보 침해 우려에 따른 이용의도의 위축을 예상할 수 있으나 오히려 공급자 집단이 더 민감한 것으로 나타났다. 이는 개인정보 유출이나 프라이버시

침해의 책임이 서비스 공급자에게 전가될 수 있는 상황을 대별하고 있다. 빅데이터 서비스 개발자 및 공급자들이 개인정보를 악용하는 것에 대한 대비도 필요하지만 익명화, 비식별화 등 정해진 기준과 방법으로 개인정보를 안전하게 서비스 개발하고 공급에 활용할 수 있는 제도적, 기술적 지원 또한 필요한 상황이다. 비슷하게 개인정보보호 관련 기술적 조치에 대하여 배경지식이 있는 IT업종 종사자에 비해 비IT업종 종사자들이 개인정보침해 위험에 대하여 더 민감하게 이용의도의 위축으로 이어지고 있다. 개인정보 보호 관련 제도적, 기술적 조치에 대하여 보다 일반인들이 이해할 수 있는 수준으로 공감대를 형성할 수 있는 정책 아이템의 발굴이 필요한 시점이다.

## 2. 연구의 한계 및 향후 연구 방향

본 연구에서 밝혀진 이론적, 정책적 시사점에도 불구하고 다음과 같은 사항을 고려하여 향후 지속적인 연구가 필요하다.

첫째로 본 연구의 분석모형은 빅데이터의 사용자 또는 잠재적 사용자를 대상으로 한 개인 단위의 분석으로 조직적, 환경적 요인에 대한 추가적인 고려가 필요하다는 점이다. 기존 연구에서 기술수용모형에 추가적으로 도입되고 있는 사회적 규범이나 조직의 혁신성, 데이터 증거에 우선한 의사결정 문화 등 조직적, 환경적 변수와 빅데이터 도입 및 활성화 과정을 연계하여 지속적인 이론 개발이 필요하다.

둘째로 본 연구를 바탕으로 실제적인 빅데이터 도입 과정과 성과에 대한 후속 연구가 필요하다. 앞으로 국내에 빅데이터 기술을 도입하는 조직이 더 많아지면 도입의도 뿐 아니라 도입 성과에 대한 정성적, 정량적 분석이 시도되어야 한다. 빅데이터의 도입 성과를 체계적으로 설명할 수 있는 프레임워크와 성과지표의 개발, 현장의 다양한 사례 연구도 지속될 필요가 있다.

마지막으로 빅데이터의 확산과 함께 지속적으로 증가하고 있는 개인정보 유출, 프라이버시 침해, 사이버

테러 등 빅데이터의 역기능에 대한 제도적 장치에 대한 지속적인 연구가 필요하다. 빅데이터 시대에서의 개인정보 보호의 핵심은 다양한 빅데이터의 유용성 및 기타 장점을 극대화하면서도 동시에 개인정보 보호 관점도 소홀히 하지 않아야 한다는 데 있다(최혜민, 2014). 개인정보 비식별화 기술과 해킹 방어기술 등 기술적 조치에 대한 공학적 연구도 중요하지만 사회과학 연구에서도 빅데이터 산업 활성화를 저해하지 않으면서 빅데이터 제공자와 사용자에 대한 신뢰를 유지할 수 있는 제도적 대안들을 모색할 필요가 있다(윤상오, 2013). 빅데이터 시대를 맞아 데이터 생산자와 사용자 사이에 바람직한 '권력의 견제와 균형'(이경규, 2014)이 형성되지 못한다면 빅데이터가 추구하는 사회적 가치 또한 심각하게 훼손될 수 있으므로, 산업계와 학계의 지속적인 관심과 상호 연계 노력이 필요하다.

■ 참고문헌

가회광·김진수 (2014). “빅데이터 도입의도에 미치는 영향요인에 관한 연구.” 한국경영정보학회 춘계공동학술대회.

강만모·김상락·박상무 (2012). “빅 데이터의 분석과 활용.” 『정보과학회지』, 30(6): 25-32.

권문주·김태웅·김민하 (2008). “국내 공개소프트웨어의 도입 실태 및 활성화 장애요인에 관한 탐색적 연구”, 『정보화정책』, 15(4): 3-21.

김관호·오규협·이영규·정재운 (2013). “스마트카드 빅데이터를 이용한 서울시 지하철 이동패턴 분석.” 『한국전자거래학회지』, 18(3): 211-222.

김규남 (2014). 『빅데이터 2.0시대 주요이슈와 정책적 시사점』, KISDI Premium Report.

김계수 (1997). 『AMOS 구조방정식 모형분석』, 서울: (주)데이터솔루션.

김선남·이환수 (2014). “빅데이터 개인정보보호 가이드라인(안)의 개선 방향에 관한 연구.” 『정보화정책』, 21(4): 20-39.

김정선·송태민 (2014). “빅데이터 기술수용의 초기 특성 연구.” 『한국콘텐츠학회논문지』, 14(9):

538-555.

김병수 (2012). “모바일 소셜네트워크서비스 환경에서 지속 사용 의도의 선행 요인에 관한 연구: 신뢰와 프라이버시 우려의 역할.” 『지식경영연구』, 12(4): 83-100.

김유신·김남규·정승렬 (2012). “뉴스와 주가: 빅데이터 감성분석을 통한 지능형 투자 의사결정모형.” 『지능정보연구』, 18(2): 143-156.

김은영·이정훈·서동욱 (2013). “빅데이터 시스템의 수용의도에 영향을 미치는 수용조직의 환경요인에 관한 연구.” 『Journal of Information Technology Applications & Management』, 20(4): 1-18.

노형진 (2001). 『한글 SPSS 10.0에 의한 조사방법 및 통계분석』, 서울: 형설출판사.

대한상공회의소, “빅데이터 활용현황과 정책과제 연구 보고서”, 2014.

도해용·유기상 (2013). “빅데이터 분석 기법을 이용한 레스토랑 품질과 고객만족이 추천행동에 미치는 영향에 관한 연구.” 『외식경영연구』, 16(2): 233-255.

도해용·이애주·이창원 (2013). “빅데이터를 이용한 레스토랑 품질이 고객만족에 미치는 영향에 관한 연구.” 『외식경영연구』, 16: 49-70.

문혜정·조현석 (2012). “빅 데이터 시대 위험기반의 정책 - 개인정보침해 사례를 중심으로-.” 『정보화정책』, 19(4): 63~82

미래창조과학부 (2014). 『빅데이터 활용 단계별 업무 절차 및 기술 활용 매뉴얼』

미래창조과학부·한국정보화진흥원 (2014). 『2014년도 빅데이터 활용 스마트서비스 사례집』, 서울: 한국정보화진흥원.

민진영·김병수 (2013). “프라이버시 계산 모형을 적용한 SNS 지속 사용 의도에 대한 연구: 페이스북과 카카오톡 사례 중심으로.” 『Information Systems Review』, 15(1): 105-122.

성민경·정연돈 (2014). “데이터 스트림에서 프라이버시 보호를 위한 익명화 기법.” 『정보과학회논문지: 데이터베이스』, 41(1): 8-20.

손민선·문병순 (2012). “빅 데이터 시대의 한국: 갈라파고스가 되지 않으려면.” 『LGERI 리포트』.

손영화 (2014). “빅데이터 시대의 개인정보 보호방

- 안.” 「기업법연구」, 28(3): 355-393.
- 손상영 (2013). 「빅데이터 온라인 마케팅과 프라이버시 보호」, KISDI Premium Report.
- 박은경·최정일·연지영 (2014). “Social Network Service에서의 사용자 저항에 영향을 미치는 요인에 관한 연구.” 「품질경영학회지」, 42(3): 387-405.
- 백인수 (2012). 「빅데이터 시대 : 에코시스템을 둘러싼 시장경쟁과 전략분석」, 서울: 한국정보화진흥원 IT & Future Strategy(4)
- 안창원·황승구 (2012). “빅 데이터 기술과 주요 이슈.” 「정보과학회지」, 30(6): 10-17.
- 오규철 (2013). 「빅데이터 활용 시 개인정보보호 법률이 미치는 영향의 측정 방법론」. 정보통신산업진흥원 IT기획시리즈.
- 오길영 (2015). “데이터 상업화과정으로서의 개인정보 비식별화.” 「민주법학」, 58:179-207.
- 윤상오 (2013). “빅데이터의 위험유형 분류에 관한 연구.” 「한국지역정보화학회지」, 16(2): 93-122.
- 이미나·심재웅 (2012). “성별에 따른 온라인 프라이버시 염려와 프라이버시 보호전략 사용의 차이에 관한연구.” 「미디어 젠더&문화」, 12: 65-190.
- 이병엽·임종태·유재수 (2013). “빅 데이터를 이용한 소셜 미디어 분석 기법의 활용.” 「한국콘텐츠학회논문지」, 13(2): 211-219.
- 이석준·조숙진·박정선·함유근·김용재 (2007). “정보시스템 프로젝트의 위험요인에 대한 현업인력과 서비스제공인력과의 인식도 차이.” 「Journal of Information Technology Applications & Management」, 14(3): 79-94
- 이선우·이희상 (2014). “빅데이터 시스템 도입을 위한 통합모형의 연구.” 「Journal of Information Technology Applications & Management」, 21(4): 463-483.
- 이현중 (2009). 「온라인 쇼핑몰에서의 지각된 위험, 지각된 쇼핑가치, 관계 질, 서비스 충성도간의 구조적 관계」. 남대학교 박사학위논문.
- 이환수·임동원·조향정 (2013). “빅데이터 시대의 개인정보 과잉이 사용자 저항에 미치는 영향.” 「지능정보연구」, 19(1): 125-139.
- 장성호·이기동 (2014). “소셜네트워크서비스 프라이버시 위험이 사용자저항에 미치는 영향.” 「e-비즈니스연구」, 15(3): 323-338.
- 장영훈·오현목·김지민·박명철 (2013). “빅데이터 시대의 개인정보보호정책에 관한 실증적 연구: 효과적인 규제, 프라이버시 통제와 신뢰를 중심으로.” 한국경영정보학회 춘계학술대회.
- 차훈상 (2012). “스마트폰 위치기반 어플리케이션의 이용활성화 의도에 영향을 미치는 요인: 프라이버시 계산 모형의 적용.” 「Asia Pacific Journal of Information Systems」, 22(4): 7-29.
- 채승병 (2011). “정보홍수 속에서 금맥 찾기 : 빅데이터 분석과 활용.” 「SERI경제노트」, 삼성경제연구소.
- 최병정·김혜진·김자호·진서훈 (2012). “빅 데이터 시대의 CRM 을 위한 데이터 분석.” 「Entrue Journal of Information Technology」, 11(1): 19-27.
- 최대선·김석현·조진만·진승헌 (2013). “빅데이터 개인정보 위험 분석 기술.” 「정보보호학회지」, 13(2).
- 최혜민 (2014). “빅데이터 시대와 현행 개인정보보호 법제의 부정합 문제 및 그 해결방안에 대한 연구.” 「IT와 법 연구」, 8: 357-382.
- Bellman, S., Johnson, E. J., Kobrin, S. J. & Lohse, G. L. (2004). “International differences in information privacy concerns: A global survey of consumers.” *The Information Society*, 20(5): 313-324.
- Bollier, D. Rapporteur (2010). “The Promise and Peril of Big Data, The Aspen Institute.” Washington, D.C.
- Bulgurcu, B., Cavusoglu, H. & Benbasat, I. (2010). “Information Security Policy Compliance: An Empirical Study of Rational-based Beliefs and Information Security Awareness.” *MIS Quarterly*, 34(3): 523-548.
- Clogg, C. C., Petkova, E. & Haritou, A. (1995). “Statistical methods for comparing regression coefficients between models.” *American Journal of Sociology*, 1261-

- 1293.
- Chui, M., Manyika, J., Bughin, J., Brown, B., Robers, R., Danielson, J. & Gupta, S. (2013). "Ten IT-enabled Business Trends for the Decade Ahead." *McKinsey Global Institute*, McKinsey & Company.
- Craig, T. & Ludloff, M. E. (2011). "Privacy and Big Data." *O'Reilly Media*
- Culnan, M. J. & Armstrong, P. K. (1999). "Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation." *Organization science*, 10(1): 104-115.
- Davis, F. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly*, 13(3): 318-339.
- De Mauro, A., Greco, M. & Grimaldi, M. (2015). "What is big data? A consensual definition and a review of key research topics." In G. Giannakopoulos, D. P. Sakas, & D. Kyriaki-Manessi (Eds.), *AIP conference proceedings*, 1644(1): 97-104.
- Dinev, T. & P. Hart. (2006). "An Extended Privacy Calculus Model for E-Commerce Transactions." *Information Systems Research*, 17(1): 61-80.
- Dishaw, M. T. & Strong, D. M. (1999). "Extending the technology acceptance model with task-technology fit constructs." *Information & management*, 36(1): 9-21.
- Dishaw, M., Strong, D. & Bandy, D. B. (2002). "Extending the task-technology fit model with self-efficacy constructs." *AMCIS 2002 Proceedings*, 143.
- Fishbein, M. & Ajzen, I. (1977). *Belief, attitude, intention, and behavior: An introduction to theory and research*.
- Fornell, C. & Larcker, D. F. (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error." *Journal of Marketing Research*, 18: 39-50.
- Gartner (2012). "The Importance of 'Big Data': A Definition."
- Goodhue, D. L. & Thompson, R. L. (1995). "Task-technology fit and individual performance." *MIS quarterly*, 213-236.
- Homans, C. G. (1961). "Social Behavior: Its Elementary Forms." New York: Harcourt, Brace and World.
- Kim, H., I. Son & Lee, D. (2012). "The Viral Effect of Online Social Network on New Products Promotion: Investigation Information Diffusion on Twitter." *Journal of Intelligence and Information Systems*, 18(2): 107-130.
- Laufer, R. S., Proshansky, H. M. & Wolfe, M. (1974). "Some Analytic Dimensions of Privacy." In Rikkard Kuller(ed.), *Architectural Psychology*. Proceedings of the Lund Conference, 353-372. Stroudsboung, PA.
- Malhotra, N. K., S. S. Kim & J. Agarwal. (2004). "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model." *Information Systems Research*, 15(4): 336-355.
- McAfee, A. & Brynjolfsson, E. (2012). *Big data: the management revolution*. Harvard business review, 1.
- Paternoster, R., Brame, R., Mazerolle, P. & Piquero, A. (1998). "Using the correct statistical test for the equality of regression coefficients." *Criminology*, 36(4): 859-866.
- Pearson, T. & Wegener, R. (2013). *Big data: the organizational challenge*. Bain Co.
- Venkatesh, V. & Davis, F. (2000). "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies." *Management Science*, 46(2): 186-204.
- Vroom, V. H. (1964). "Work and Motivation."

*John Wiley and Sons, Inc.*

Wamba, S. F., Akter, S., Edwards, A., Chopin, G. & Gnanzou, D. (2015). "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study." *International Journal of Production Economics*, 165: 234-246.

White, M. (2012). Digital workplaces: Vision and reality. *Business Information Review*, 29(4): 205-214.