

데이터 가치에 대한 탐색적 연구: 공공데이터를 중심으로*

이상은** · 이정훈*** · 최현진****

A Study on the Data Value: In Public Data*

Sang Eun Lee** · Jung Hoon Lee*** · Hyun Jin Choi****

■ Abstract ■

The data is a key catalyst for the development of the fourth industry, and has been viewed as an essential element of the new industry, with technology convergence such as artificial intelligence, augmented/virtual reality, self-driving and 5 G. This will determine the price and value of the data as the user uses data in which the data is based on the context of the situation, rather than the data itself of the past supplier-centric data. This study began with, what factors will increase the value of data from a user perspective not a supplier perspective The study was limited to public data and users conducted research on users using data, such as analysis or development based on data. The study was designed to gauge the value of data that was not studied in the user's perspective, and was instrumental in raising the value of data in the jurisdiction of supplying and managing data.

Keyword : Data Value, User Value, Public Data, User

Submitted : November 11, 2021

1st Revision : January 22, 2022

Accepted : February 23, 2022

* 본 연구는 국토교통과학기술진흥원의 지원을 받아 연구되었음(과제번호: 22NSPS-C149910-05).

** 연세대학교 기술경영학 박사과정

*** 연세대학교 정보대학원 교수, 교신저자

**** 연세대학교 정보대학원 석사과정

1. 서 론

과거 산업혁명에서 '전기'와 '석유'를 기반으로 생산성을 개선하여 경제적 발전을 했듯이, 4차산업혁명 시대에는 데이터와 디지털 기술이 새로운 성장 동력으로 주목받고 있다(Satya Narayana Nadella). 특히 데이터는 인공지능(Artificial Intelligence: AI), 사물인터넷(IoT)·자율주행·블록체인 등의 기술과 결합하여 4차산업혁명의 인프라를 만들어내고 있다. 세계경제포럼(WEF)이 조사한 결과에 따르면 2022년까지 중점 투자할 분야로 빅데이터(85%)와 AI(73%)를 꼽았다. 그뿐만 아니라 IDC에 따르면 전 세계 데이터 시장은 2018년 1,660억 달러(191조 8,130억 원)에서 2022년 300조 4,040억 원으로 성장할 것이며, 데이터양은 2020년 44ZB에서, 2025년에는 180ZB까지 기하급수적으로 커질 것으로 예상된다. Reseach & Market 시장 전망보고서에서는 데이터 시장 규모가 연평균 6.8% 성장하여 '25년까지 2,300억 달러에 달할 것으로 예측했고, 국내에서도 지난해 데이터 경제 가속화로, 시장 규모 전년도 대비 14.3%(2.5조) 성장하여 20조를 돌파했다(2020 데이터 산업 현황조사). 이러한 변화에 대응하기 위해 세계 각국은 데이터 개방 활성화를 위한 기술, 제도 강화에 총력을 기울이고 있으며, 데이터와 데이터 가공기술이 곧 경쟁력이 되고 있다(공학저널, 2020).

최근 장기화 된 팬데믹 현상으로 가속화된 재택근무와 전자상거래와 통신망 확대, AI와 딥 러닝 기술의 발전 등으로 디지털 혁신이 강조되고 있는 요즘 데이터 시장은 앞으로 더 빠르게 성장할 것으로 예상된다. 이러한 변화에 대응하기 위해 세계 각국에서도 데이터 개방을 위한 정책 및 기술에 총력을 기울이고 있으며 국내에서도 방대한 예산, 인프라, 데이터 전문인력 등 다양한 시도를 하고 있다. 그러나 국내에서는 데이터 측정 기준 부재와 공급자 관점에서의 데이터 개방, 고품질 데이터 부족 등으로 데이터 활성화가 저조하며, 특히 정책적으로 지속적인 공공데이터 개방 확대에도 불구하고 사용자들이 실제 데이터를 활용할 때, 데이터의 정확성 결여, 중복,

누락 등 품질 또는 형태적 측면에서 낮은 수준으로 지적되었다. 또한 최근 데이터 거래 시장이 조성되면서 데이터 가격 산정 기준의 부재로 인한 불합리한 가격 및 신뢰성 저하도 문제점으로 지적되고 있다. 이처럼 개별 기업 중심의 데이터 유통, 거래제도 미비 등의 양질의 데이터 유통을 제한하고, 데이터가 산업, 사회 혁신 촉매제로 역할이 기대되나 국내 빅데이터 활용 및 분석 수준은 63개국 중 31위이며 국내 기업의 빅데이터 이용률은 10%에 불과하여(IMD, 2018) 활용도 측면에서 산업 전반의 제고력에 한계로 분석되고 있다.

앞으로 데이터는 새로운 제품·서비스·경험·가치 창출의 핵심 자원으로 활용되는 4차산업혁명 시대에서 데이터의 생산·분석·활용은 국가 발전의 기반이 될 것이다(국회입법조사처, 2019). 그러나 국내에서는 고품질 데이터 부족과 불합리한 데이터 가격 산정 등으로 인해 데이터 활용률이 저조함에 따라 데이터 활성화를 위한 방안이 필요하다.

과거에는 데이터는 자체적 품질을 기반으로 정확성에 따라 데이터의 가치를 결정했다. 서로 다른 분야의 데이터 간의 결합을 통해 새로운 인사이트 확보에 필수 요소가 될 것이다. 그러나 데이터가 자산임을 인식하지만, 실질적인 재무 가치 산정은 쉽지 않다. 데이터는 무형 자산으로, 실제 확인이 어렵고, 기업에서는 데이터를 하나의 자산으로 평가하기보다는 관리를 위한 스토리지 도입 및 유지, 애플리케이션, 인건비 등으로 IT 비용을 증가시키는 요소로 평가하고 있다. 그러나 앞으로 데이터도 하나의 자산으로 평가받으며, 자체적인 품질뿐만 아니라 그 데이터로부터 어떤 가치를 창출할 수 있겠느냐 하는 것이 그 데이터의 가격을 결정할 것이다(IBM, 2016). 따라서 데이터를 직접 사용하는 사용자에게 매력적으로 접근해야 한다(성태웅 외, 2016). 앞으로 데이터는 데이터 자체적인 품질뿐만 아니라 사용자가 데이터의 활용이나 재조합, 재사용을 통해 창의적 융합으로 기존에 방식으로는 풀 수 없는 문제를 해결하거나, 가치 없던 데이터를 신기술을 적용하여 가치를 창출하는 등과 같이 데이터가 보유하고

있는 잠재적 가치를 도출하기 위해 얼마나 사용자에게 데이터가 얼마나 편리한 형태인지가 중요한 요소로 작용할 것이다.

따라서 본 연구에서는 최초의 데이터 자체가 아닌 이를 분석하고 개발 등을 통해 데이터를 활용하는 사용자 관점에서 어떠한 데이터의 가치를 추구하고자 하는지 살펴보고자 한다.

연구 목적은 첫째, 데이터를 활용하는 사용자 관점에서 데이터의 가치 요인을 도출하고, 둘째 도출된 가치 요인을 바탕으로 가중치 분석을 통해 사용 목적에 따라 비교한다. 이를 통해 데이터 경제 활성화를 위한 데이터의 생성, 유통, 분석, 적용 단계에서 발생하는 데이터 가치체계를 다루는 시스템의 필요성을 제시하고, 데이터의 가치체계 형성을 통해 데이터 공개(개방)와 공유를 촉진할 수 있는 해결책을 마련하고자 한다.

최근 전 세계적으로 빅데이터가 쟁점이 되면서 민간 분야뿐만 아니라 국내 정부를 포함한 공공분야에서 그 활용을 위해 다각적인 노력과 공공데이터를 민간에게 개방하는 것에 집중하고 있다(삼성경제연구소, 2013). 그뿐만 아니라 국내에서도 '공공데이터의 제공 및 이용 활성화에 관한 법률'이 제정되어 삶의 질 향상을 기대하고 있다(정국환, 2013). 그러나 공공데이터의 활용은 현재 기대만큼 활성화되지 않고 있으며, 데이터 활용 성과들 역시 기대에 미치지 못하는 상황이다. 그 주요한 원인으로 개방 데이터 품질에 대한 가이드라인 부재가 있다(임원준, 2017). 그럼에도 불구하고 이에 관한 연구는 아직 초기 단계 수준이다. 따라서 전 세계적으로 공공데이터를 통해 새로운 일자리 창출과 경제 활성화를 실현하기 위한 경제 성장 동력으로 인식함에 따라 사용자 관점에서 공공데이터의 가치에 미치는 요인에 관한 연구가 요구된다. 따라서 본 연구에서 데이터 생태계를 기반으로 각각 단계별로 사용자를 분류하고, 데이터를 통해 데이터 활용 단계의 비즈니스 수요자 관점에서 도출하고자 한다.

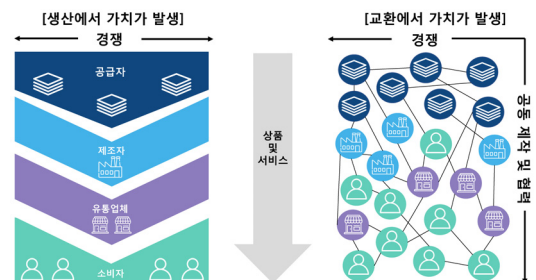
본 연구는 개방형 설문과 메타분석을 통해 데이터 가치 요인을 추출하였고, 이를 FGI를 통해 공공데

이터의 특성 및 사용자 관점에 따라 재정의하였다. 둘째, 최종적으로 도출된 공공데이터 가치 요인을 기반으로 사용자에게 데이터를 사용할 때 고려하는 중요 정도에 대해 5점 척도로 설문을 진행하였으며, 설문 결과를 바탕으로 군집분석을 시행하였다. 셋째 형성된 군집을 기반으로 AHP 분석을 통해 데이터 사용 목적에 따라 가중치를 분석하였다.

2. 이론적 배경

2.1 데이터에 대한 인식 변화

데이터는 인간 또는 컴퓨터를 비롯한 기기에 의해 행해지는 통신 해석 또는 처리로 형식화된 사실, 개념 또는 명령을 표현으로 (한국정보통신기술협회) 데이터 자체의 가치는 크지 않지만, 상황적 맥락에 따라 적재적소에 자본을 투입하여 혁신적인 부가 가치를 창출한다(The Economist, 2017). 양질의 데이터가 정보·지식·상품·서비스로 전환되어 더 나은 의사결정을 지원하고 경제·사회적 편익을 발생시킴으로써 하나의 데이터가 다양한 분야에서 활용하고, 자본과 달리 비경쟁적이어서 무한하게 재이용이 가능한 신개념의 자본으로 부상하고 있다. 이에 따라 데이터는 기존 가치사슬을 더욱 복잡하고 거대한 생태계를 변화시켜 점점 증가하는 데이터 기기에서 생성된 데이터와 서비스의 융합 등을 통해 동적이고 복잡하게 연결되어 새로운 가치사슬로 생태계를 형성할 것이다.



출처: Deloitte analysis, Forbes 재인용.

[그림 1] 데이터로 인한 비즈니스 생태계 창출

과거에는 데이터가 얼마나 정확한지, 얼마나 안전하지 등의 데이터의 자체적인 품질 관점에서 가치를 측정했다. 그러나 4차산업혁명 시대에 빅데이터, 인공지능, IoT 등에서는 데이터의 품질에 대한 새로운 정의와 해석이 필요하게 되었다. 즉, 과거 전통적으로 데이터 자체적인 품질에서 이제는 이 데이터를 통해 어떤 가치를 창출할 수 있는지가 그 데이터의 가치와 가격을 나타낼 것이다. 예를 들어 어떤 데이터들과 융합·분석하여 어떤 가치가 나올 수 있는지, 개인정보는 어느 정도 활용할 수 있는지, 사용자가 분석하기에 얼마나 편리한 형태로 되어 있는지 등이 중요한 요소가 될 수 있다. 그러나 데이터의 가치는 최초의 원시(Raw) 데이터만 보고는 이 데이터를 어떻게 분석하고 가공하여 향후 얼마의 가치를 창출할지를 예측하기는 어렵다. 대표적으로 실시간 교통 정보 자체는 상품성이 적지만 이를 주변의 다른 데이터와 융합하여 내비게이션 서비스를 제공할 수 있다. 따라서 데이터의 가치를 인정하고 평가하지 않는 현재 데이터의 생태계에서는 데이터의 거래가 지속적으로 발전 할 수 있도록 데이터 중심 경제 체제를 중심으로 얻어지는 가치를 나눌 수 있는 대상에 따라 데이터 거래 생태계에서도 필요한 시점이다.

2.2 가치에 대한 개념의 변화

가치란 1950년 초 인간의 가치에 대한 추상적인 개념을 시작으로 마케팅 관점에서 제품/서비스 이용과 연계되어 trade-off 개념으로 경제적 관점에서 경험기반의 가치에 관한 관심이 높아지고 있다. 사용자가 서비스 또는 제품 이용 과정에서 다양한 경험을 하게 되고, 이러한 경험 자체를 가치 있게 여기기 시작했다(Park, 2013). Parsons(1951)는 가치는 기준 역할을 하는 요소로, 다양한 기준 중 상황에 따라 대안을 선택하기 위한 표준이라 하였고, Rokeach(1968)은 사물·행동·목표 등에 내포되어 있거나 주어진다고 생각되는 것이라고 하였다. 이후 1980년대에 마케팅에서는 가치가 발생하기 위해서는 주체와 그 주체와 관계된 대상과의 관계에서 평가를 기반한

다고 했다(Gutman, 1982). 그러나 최근 소비자 관점에서 사용자부터 사용자가 제품을 사용하면서 인지되고, 사용자가 얻을 수 있는 혜택이라고 하였고(Jensen, 2001), 바람직한 존재 상태나 행동의 형태로 정의하여 life value, user value, product/service value 나뉘 분류하였다(Park, 2013). Harvard Business Review(2016) 가치를 기능적, 감정적, 삶의 변화, 사회적 영향의 네 가지 범주로 구성 요소들을 가지고 제품/서비스에 적합한 조합을 만들면 사용자의 충성도와 매출이 증가한다고 하였다.

2.3 사용자의 지각된 가치

사용자가 지닌 가치는 오래전부터 인간 행동에 커다란 영향을 미치는 것으로 인간 행동을 이해하기 위한 하나의 중요한 개념으로 간주하여(김철호, 2013), 광범위한 분야에서 지속해서 연구되어왔다(Dodds and Monroe, 1985; Holbrook, 1994). 최근 사용자는 실제 경험하는 품질 또는 가치에 따라 상품을 구매하고 있다(SIMONSON, ITAMAR, 2015). 아무리 좋은 기능을 가진 제품 또는 서비스라도 사용자에게 어떠한 가치를 제공해주는지가 불명확하다면 사용자에게 외면당할 가능성이 크다(김진우, 2012). 이처럼 사용자(소비자)의 경험 가치에 관한 관심이 높아지고 있으며, 사용자는 전반적인 상품 이용 과정에서 다양한 경험을 하게 되고, 이러한 경험 자체를 가치 있게 여기기도 한다(안은미, 2013). 따라서 소비 대상과 관련된 가치는 소비의 모든 단계를 통해 추구되는 것이기 때문에 중요하게 고려해야 하는 개념이다.

지각된 가치는 사용자가 제품 또는 서비스로부터 얻고자 하는 것과 지급한 비용에 대한 보상으로 사용자가 무엇을 주고 무엇을 받는가에 대한 지각에 근거한 효용에 대한 평가로 정의된다(Zeithaml, 1988). 그뿐만 아니라 사용자들이 지각하는 소비 경험은 문제해결이나 욕구, 만족과 같은 인지적 소비 가치와 즐거움이나 미적 특성과 같은 경험적이고 쾌락적인 소비 가치를 모두 포함하고 있다(Hirschman and Holbrook, 1982).

즉 사용자의 가치는 사용자들의 기분이나 감정 및 흥미 등의 경험적 소비 가치와 사용자의 정보처리 과정에서의 인지적인 소비 가치가 합쳐진 개념이다. 이렇듯 사용자의 지각된 가치에 관한 개념은 경제적, 실용적 가치와 같이 단순한 차원들로 단일의 개념으로 측정하는 방법이 있을 수 있으나, 심리적, 인지적, 정서적 가치와 같은 측정 방법도 있다. 경제적, 실용적 가치는 Zeithaml(1988)의 지각된 가치 개념에 근거하여 Fornell et al.(1996)이 시도한 것으로, 지급한 것과 얻은 것의 차이 또는 비율로 측정하는 것으로 감정적인 부분이 미흡하다. 따라서 심리적, 인지적, 정서적 가치처럼 복잡한 특성을 반영하여 경험의 잠재적 즐거움과 감정적 가치를 반영한 가치가 포함되어야 한다(Babin, 1994). 지각된 가치는 사용자의 행동을 예측하는 데 있어 주요한 지표 중 하나로서 마케팅 분야에서 널리 사용됐다.

2.4 선행연구결과

이론적 배경 및 선행연구 분석 결과 데이터를 단순한 원시 데이터 아닌 4차산업혁명 시대에 핵심 자원으로 제품과 서비스를 위한 투입 요소로써 사용하고 지식, 상품, 서비스 창출을 위한 투입 요소로 인지하고 있다. 그러나 아직 공급자 관점에서 기술 주도적이며 단정적으로 데이터 자체적인 품질 중심에 관한 연구가 이뤄지고 있으며, 공공데이터의 경우 거시적 관점에서 실증적 연구보다는 비 실증적인 연구에 집중하고 있었다. 그뿐만 아니라 데이터의 가치평가는 경제적 가치평가에 집중하여 조건부 평가, 산업연관 분석에 치중되어 있었다. 그러나 가치라는 개념은 과거 경제적 관점에서 경험을 기반한 가치에 관심이 높아짐에 따라 사용자가 제품/서비스 이용 과정에서 다양한 경험을 하면서 이런 경험 자체를 가치로 정의되고 있었다(Zeithaml, 1988). 그뿐만 아니라 선행연구를 기반으로 지각된 가치가 잠재고객들의 행동을 예측하기 위한 지표로써 사용할 수 있다는 것을 뒷받침할 수 있었다.

따라서 데이터를 직접 사용하는 수요자의 다양한

요구에 주목하여 수요자의 요구를 만족시키거나 기대 수준 이상으로 충족시키는 것은 업무 효율 및 성공에 핵심 요소로 사용자의 필요에 부응하여 그 요구를 충족시킬 수 있는 효과성에 초점을 맞춰야 한다. 따라서 본 연구에서는 공공정보의 개방성 제고와 데이터의 가치를 측정하기 위한 데이터 가치에 관한 연구를 진행하고자 한다.

3. 연구 문제 및 방법

3.1 연구문제 및 대상

선행연구 및 이론적인 검토를 통해 데이터의 자체적인 품질을 중점으로 연구되고 있으나, 데이터 품질이 높다고 데이터가 가치가 높은 것이 아니고, 데이터의 가치가 높다고 품질이 좋은 것은 아니다. 이렇듯 데이터를 사용하는 분야와 맥락적 상황에 따라 데이터의 가치가 다르기 때문에 사용자의 관점에서 가치 요인을 도출하고자 한다.

연구 질문 1: 사용자 관점에서의 데이터의 가치를 결정하는 요인은 무엇인가?

연구 질문 2: 데이터 사용 목적 또는 상황별 중요한 공공데이터의 가치는 무엇인가?

본 연구에서 데이터를 가공/처리하지 않은 데이터로 관찰된 객관적인 사실 그 자체인 원시(Raw)로 정의한다(Rowley, 2017). 데이터의 범위를 국내 공공데이터 포털(Data.go.kr)에서 다운로드할 수 있는 공공데이터로 한정하여 사용자가 데이터에 접근하는 환경과 비용 측면을 동일하게 적용했다. 공공데이터의 가치는 사용자가 데이터를 사용할 때, 사용자에게 충족되는 바람직한 상태 또는 모드(Miller, 2008) 또는 사용자가 제품에 대한 요구와 사용자 경험 및 만족에서 출발하여 제품/서비스의 극대화를 목적으로 사용자가 중요하게 인지하는 가치를 반영한 것으로 정의하였다(Zeithaml, 1988).

본 연구에서 사용자를 데이터 중심 체제에서 데이

터로부터 가치를 나눌 수 있는 대상물 [데이터를 최초로 만들어 내는 생산자]-[데이터를 유통하는 유통자-데이터를 분석/개발 등을 통해 활용하는 비즈니스 사용자]-[도출 결과를 사용하는 최종 사용자 (End-user)]로 분류했다(Zuiderwijk, 2012).

〈표 1〉 연구 대상 정의

대상	정의	출처
데이터의 정의	데이터를 가공/처리하지 않은 관찰된 상태. 객관적인 사실 그 자체인 Raw data 정의	Rowley (2007)
데이터 범위	공공데이터 포털(Daga.go.kr)에서 다운로드 받아 사용할 수 있는 공공데이터로 한정함	Harrison (2011), Zuiderwijk (2014)
공공데이터의 가치	사용자가 데이터를 사용할 때, 사용자에게 충족되는 바람직한 존재 상태 또는 모드	Miller(2008), Park(2001)
사용자의 정의	데이터를 분석하고 결과를 활용하는 사용자로 정의	Kim(1986) Doll(1988)
사용자의 범위	공공데이터를 기반으로 데이터 분석가, 데이터 기반 개발자로 분류	Goodhue (1995)

3.2 연구방법

본 연구는 총 3단계에 거쳐 공공데이터 가치 요인을 도출했으며, 단계별 도출 방법은 다음과 같다. 1 단계에서는 개방형 설문, 메타분석, FGI를 사용했고, 2 단계에서는 설문을 통해 도출된 요인들은 SPSS를 활용하여 군집/요인분석을 하여 요인을 축소했다. 마지막 3 단계에서는 축소된 요인들을 대상으로 데이터 사용에 대한 목적에 따라 AHP를 실시하여 상황에 따라 중요하게 생각하는 가치에 대해 도출하였다.

4. 1차 데이터가치 요인도출

4.1 개방형 설문

응답자가 자유롭게 응답할 수 있는 개방형 설문을 통해 사용자 관점에서 데이터의 가치 요인 도출하였

다. 응답의 범위를 설정하지 않고, 단답형으로 진행하였으며, 데이터를 활용하는(예: 개발, 분석) 사용자들을 대상으로 온라인을 통해 진행하였다. 개방형 설문에서는 약 102명의 사용자가 총 304개의 설문 답변을 제출하였고, 제출된 답변에서 중복 및 유사 요인을 제외하여 다음과 같이 재정리하였다. 개방형 설문 응답 내용에서 드러난 각각의 요인의 키워드를 추출한 후, 동일 주제에 따라 분류하였다. 개방형 설문은 데이터의 속성을 자체적 속성, 분석 상황에 맞춘 맥락적 속성으로 분류하였고(Wang, 1996; Lee, 2002), 데이터 사용 단계에 따라 데이터 수집-데이터 이해-데이터 활용으로 분류하였다(David, 2011). 개방형 설문 결과 각각의 요인의 키워드의 빈도를 기준으로 데이터 자체적인 속성 10%, 맥락적인 속성 25%, 데이터 수집 단계에서 16%, 데이터 이해 단계 26%, 데이터 활용 23%가 도출되었다.

4.2 메타분석

메타분석이란 ‘분석들에 대한 분석’을 의미하는 것으로 다량의 관련 연구 결과를 통계적 방법이다. 즉 어느 특정 분야 혹은 주제와 관련하여 누적된 독립적 연구 결과물들을 정량적으로 종합한 뒤, 이를 동일한 측정치로 환산하여 분석하는 것으로 이미 도출된 분석 결과들에 대해 재분석을 하여 총체적인 결론에 도달하는 분석 방법을 말한다(정인숙, 2011). 본 연구에서는 체계적인 연구 결과의 분류작업을 통하여 일반적인 분석 방법에 의해서 파악되기 어려운 데이터 가치에 관한 연구 결과들을 함축적으로 분류하고자 메타분석을 선택하였다(Glass, 1976). ‘데이터 가치’, ‘정보의 가치’, ‘공공데이터의 가치’라는 키워드를 중심으로 연구의 기준 척도로 사용하였으며, 연구 결과의 선정과 평가, 서술에서 오는 리뷰의 주관성을 배제하고 결과의 객관성을 확보하고자 선행연구를 기반으로 데이터의 가치 요인에 대한 일반적 일반적인 특성과 체계적으로 요인을 도출하고자 했다.

본 연구에서는 ‘(공공) 데이터 평가’, ‘데이터 가치’, ‘인지된 가치’의 키워드를 통해 Science Direct 및

각 정부 기관에서 데이터 관리 및 평가 지표를 포함하여 총 143개의 문헌을 대상으로 메타분석을 시행하여 총 118개의 요인을 도출하였다. 메타분석의 자료 수집에 대한 신뢰성과 객관성, 타당성의 확보를 위해 선정된 주제, 즉 국내외 데이터 평가 관련된 연구물들을 대상으로 하였다(정인숙 외, 2011). 본 연구는 데이터(정보) 분류에 대한 실증 분석이 본격적으로 지속하기 시작한 시점, 1967년 Snavely의 연구를 시작으로 현시점(2019.03)까지 진행하였고, Science Direct

를 중심으로 검색하였다. 메타분석에서는 공공데이터 가치뿐만 아니라 데이터 가치까지 포함하여 자료를 수집하였다. 본 연구의 최종 분석 대상인 '데이터의 가치'라는 개념이 아직 표준화되지 않은 개별 연구임에 따라, 선정된 연구물 내 기재된 연구 특성을 체계적으로 분석하기 위한 코딩 매뉴얼의 개발하여 논문 제목, 저자명/기관명, 연도, 출처, 학술지 급, 연구 방법, 독립변수, 연구대상에 대해 수집하였다.

총 143개의 문헌을 기반으로 메타분석을 시행하여 총 118개의 요인을 다음 표와 같이 도출하였고, 도출된 118개의 요인에 대해 문헌에서 언급된 빈도를 기준으로 118개의 요인의 빈도 평균(6.6)을 기준으로 정렬하였다. 그 결과 다음과 같이 59개의 요인을 도출하였다.

〈표 2〉 메타분석 결과

요인명	빈도	요인명	빈도	요인명	빈도
정확성	41	참신함	12	사용자 중심	7
신뢰성	32	해석 가능성	12	상호 운용성	7
완전성	24	현재성	12	소스/출처	7
일관성	24	효율성	12	이식성	7
적합성	23	데이터 내용	11	활용성	7
검증가능성	22	사용자 지원	10	효과성	7
비용	20	시의성	10	Value-Added	7
접근성	19	유용성	10	경쟁력	6
사용빈도/사용량	18	이해 가능성	10	잘못된 표현	6
데이터 오류	17	일치성	10	메타 데이터	6
최신성	16	가용성	9	생산성	6
형식	16	개인정보	9	성공적인 사용	6
적시성	15	복구 가능성	9	사용자 중심	6
데이터양	14	적절성	9	시간감소	6
객관성	13	기밀성	8	유효성	6
반응성	13	무결성	8	융합성	6
보안성	13	재사용	8	조작 용이성	6
데이터 관리	12	표준화	8	창조력	6
범위	12	비교성	7	추적성	6
				투명성	6

4.3 FGI

FGI(Focus Group Interview: 표적집단면접)은 정성적 조사 방법 중(Qualitative Research) 한 가지 방법으로, 진행자(moderator)가 보통 5~8명의 전문가를 대상으로 자연스러운 분위기에서 연구 목적과 관련된 토론을 함으로써 자료를 수집한다. FGI의 목적은 응답자들 간의 자연스러운 대화 과정에서 연구 목적과 관련된 유용한 정보를 얻거나 설문에서 기대하지 못한 결과를 발견하는데 있다(이 고운, 2012). FGI는 주로 기본적 탐색 조사에 사용되며 향후 시장 변화에 대한 예측적 정보 수집하기 위해 사용자 관점에서 전반적인 정보 수집 등을 위해 사용된다.

본 연구에서는 데이터 사용자들이 추구하는 가치를 추가로 파악하기 위하여 FGI(Focus Group Interview: 표적 집단 면접)을 진행했다. 데이터분석 경력 2년 이상 국내에서 제공 중인 공공데이터를 사용해본 사용자들을 대상으로 데이터 기반 서비스 개발자 1명, 서비스 기획자 1명, 프로그래머 1명, 데이터 분석 교육자 1명, 데이터 분석 전공 대학생 1명, 데이터 분석가 1명 공공 데이터 담당자 1명 총 6명으로 구성하였다.

본 FGI는 2019년 5월 26일 오후 2시~3시 30분 동안 중각 WEWORK에서 진행되었고, FGI에서는 도출된 83개의 요인의 중복성 및 재정의와 데이터 품질, 데이터 가치, 사용자 범주에서 자유롭게 진행하였다. 참석한 전문가들은 다음과 같다.

〈표 3〉 FGI 전문가 정보

NO	직무	소속	분야	경력	이름
1	서비스 개발자	스타트업	과학기술	4년	박0재
2	서비스 기획자	대기업	과학기술	3년	정0영
3	개발자	스타트업	과학기술	4년	김0한
4	교육자	스타트업	교육	2년	김0기
5	대학원생	대학교	보건의료	3년	반0민
6	데이터 분석가	대기업	재정금융	2년	김0우
7	공공데이터담당자	공공기관		3년	김0민

FGI에서는 메타분석, 개방형 설문 결과를 기반으로 도출된 데이터 자체적인 품질, 데이터 작업 시 맥락에서 고려해야 할 요구사항, 데이터의 표현, 데이터 접근성 범주에서 진행하였으며 그 외에도 데이터 관리, 내용 범주에서 자유롭게 진행하였다. 그뿐만 아니라 도출된 요인 중 중복적인 개념을 삭제하고, 공공데이터에 대한 특성을 반영하고 도출된 요인들에 대해 사용자 관점에서 요인들을 재정의하였다. 개방형 설문을 기반으로 데이터를 사용하기 전의 과정에서 사용자의 가치 요인을 추가하였고, 메타분석에서 기준(빈도) 미달로 인해 제외되었지만, 공공 데이터 생태계 특성을 반영한 ‘순환성’, ‘수요주도성’, ‘참여성’, ‘지속성’을 추가하고(The World Bank, 2014; Jetzek, 2014) ‘데이터양’, ‘비용’, ‘신뢰성’, ‘접근성’을 제외하였다. 그에 관한 결과는 다음과 같다. 총 58개의 공공데이터의 가치 요인이 도출되었으며, 다음 표에서 [M]은 메타분석, [o]는 개방형 설문, [F]는 FGI을 통해 도출된 요인들을 의미한다.

〈표 4〉 도출된 공공 데이터 가치 요인

가용성	[M] [O] [F]	사용자 지원	[M] [O] [F]	저장성	[M] [F]
개인식별 가능성	[M] [O] [F]	생산성	[M] [O] [F]	적시성	[M] [O] [F]
객관성	[M] [O] [F]	예측성	[M] [O] [F]	적절성	[M] [O] [F]
경쟁력	[M] [O]	수요 주도성	[O] [F]	적합성	[M] [O] [F]
데이터 원천	[M] [O]	순환성	[O] [F]	문제 해결성	[M] [F]
기밀성	[M] [O]	시간 효율성	[M] [O] [F]	정확성	[M] [O] [F]
대응력	[O] [F]	시의성	[M] [O] [F]	지속성	[O] [F]
데이터 압축	[M] [O]	시장 반응성	[O] [F]	참신함	[M] [O] [F]
데이터 관리	[M] [O] [F]	무결성	[M] [O] [F]	참여성	[M] [F]
데이터 모델링	[M] [O] [F]	신속성	[O] [F]	창조력	[M] [O] [F]
데이터 오류	[M] [O] [F]	책임성	[M] [F]	책임성	[M] [F]
데이터 오염	[O] [F]	최신성	[M] [O] [F]	추적성	[M] [O] [F]
맞춤형	[O] [F]	업무 수용성	[O] [F]	통찰력	[M] [F]
메타 데이터	[M] [O] [F]	융합성	[M] [O] [F]	투명성	[M] [O] [F]
반응성	[M] [O] [F]	이식성	[M] [O] [F]	표준화	[O] [F]
범위	[M] [O] [F]	이해 가능성	[M] [O] [F]	해석 가능성	[M] [O] [F]
복구 가능성	[M] [O] [F]	일관성	[M] [O] [F]	혁신성	[M] [F]
분석성	[M] [O] [F]	일치성	[M] [O] [F]	현재성	[M] [O] [F]
비교성	[M] [O] [F]	재활(사)용성	[M] [O] [F]	형식	[M] [O] [F]
사용빈도/양	[M] [O] [F]				

5. 2차 공공데이터 가치요인 도출

5.1 설문을 통한 공공데이터 가치요인 도출

도출된 공공데이터 가치 요인에 대해 ‘매우 중요하다(5점)’에서 ‘전혀 그렇지 않다(1점)’로 5점 척도를 통해 설문을 진행했다. 본 연구의 검증을 위한 분석단위(Unit of Analysis)는 개인 단위로 응답자는 10대 ~50대 이상의 데이터 담당자들을 대상으로 했다. 총 202부의 설문 데이터 중 불성실 응답자 19명을 제외하고 총 183명의 설문 데이터를 수집하였다. 수집된 설문을 기반으로 요인에 대해 요인이 가진 특성을 바탕으로 그룹화하기 위해 SPSS를 통해 군집분석을 시행하였다. 본 연구에서는 대상 간 거

리에 의해 군집을 형성하는 계층적 군집 분석을 시행했고, 기술 통계를 통해 Data 표준화하여 ward의 방법으로 구성 가능한 군집들 모두에 대해 그 군집을 구성하는 대상들의 측정치 분산을 기준으로 분석을 시행하였다. 분석 결과 덴드로그램을 통해 본 연구에서는 5개의 군집으로 군집이 설정되었고, 군집 분석 결과는 다음과 같다.

도출된 군집 간의 다중 공산성 검증을 위해 상관 분석을 실시하였고, 실시 결과 상향적 가치와 인지적 가치의 상관관계수가 0.181로 상관관계가 발견되었으나, 0.2* 미만이어서 상관관계가 거의 없다고 해석할 수 있다(Rea and Parker, 2005). 각 군집에 대한 신뢰도 분석 결과 군집 5개 모두 Cronbach의 알파 값이 .7 이상으로 높은 신뢰 수준을 유지했다.

지각된 가치 이론을 기반으로 5개의 군집을 정의했고, [군집 1]은 사회적 가치(사회성/개방성), [군집 2]는 금전적 가치(효율성), [군집 3] 인지적 가치(생산성), [군집 4]는 기능적 가치(기능성), [군집 5]는 상황적 가치(편의성)로 정의하였다.

[군집 1]은 사회적 가치로써 정부, 기업, 시민(개발자, 분석가)이 데이터를 통해 도시의 생활 가능성 향상, 정부의 투명성 향상, 지속 가능성 향상 등을 통해 시민(사용자)에게 혜택을 제공하는 것을 의미한다. 즉, 사용자가 데이터를 이용할 때 느낄 때 얻어지는 가치다. 사회적 가치에는 경쟁력, 창조력, 맞춤형, 최신성, 참여성, 문제 해결성, 투명성, 사용량/사용 빈도, 현재성, 시장 반응성, 지속성, 참신함, 수요 주도성이 포함되어 있다.

[군집 2] 금전적 가치란 데이터를 사용의 전반적인 효율성을 의미하며 사용자가 데이터를 사용하면 효율적으로 의사결정을 진행하거나, 사용자들이 서비스를 이용하는데 운영 및 소요 시간, 비용 또는 노력에 대한 단축을 통해 사용성이 높아지는 것을 의미한다(Sheth, et al, 1991). 금전적 가치에는 적합성, 시간 효율성, 적시성, 데이터관리, 저장성, 분석성, 재활(사)용성, 가용성, 데이터 압축, 메타데이터가 포함되어 있다.

[군집 3]의 인지적 가치란 데이터를 통해 신규 제

〈표 5〉 군집분석 결과

군집 1	군집 2	군집 3	군집 4	군집 5
경쟁력	적시성	데이터 원천	데이터 오류	데이터 모델링
창조력	적합성	융합성	무결성	사용자 지원
맞춤형	시간 효율성	시의성	객관성	일관성
최신성	데이터 관리	대응력	반응성	업무 수용성
참여성	저장성	신속성	복구 가능성	큐레이션
문제 해결성	분석성	생산성	비교성	표준화
투명성	재활용성	예측성	일치성	이해 가능성
사용량	가용성	순환성	정확성	데이터 오염
현재성	데이터 압축	적절성	형식	이식성
시장 반응성	메타 데이터	통찰력		개인식별 가능성
지속성		범위		기밀성
참신함				추적성
수요 주도성				해석 가능성
				책임성

<표 6> 분석결과

Spearman의 rho		군집 1	군집 2	군집 3	군집 4	군집 5	
상관분석	군집 1	상관계수	1	-0.049	-0.015	0.005	0.007
		유의 확률	.	0.513	0.84	0.942	0.928
	군집 2	상관계수		1	-0.049	0.134	.181*
		유의 확률		.	0.514	0.073	0.015
	군집 3	상관계수			1	-0.011	-0.024
		유의 확률			.	0.881	0.752
	군집 4	상관계수				1	0.063
		유의 확률				.	0.396
	군집 5	상관계수					1
		유의 확률					.
Cronbach의 알파		군집1	군집 2	군집 3	군집 4	군집 5	
		.957	.934	.934	.97	.914	

*0.0~0.1: 거의 없음/0.1~0.2: 약한 상관관계/0.2~0.4 보통의 상관관계/0.4~0.7약한 상관관계/0.8~1.0 매우 강한 상관관계.

품/서비스, 새로운 알고리즘 등의 생산활동에 사용하면서 고려하는 가치로 사용자가 데이터를 사용함으로써 발생하는 편익이나 비용에 대한 가치를 의미한다. 즉, 총체적으로 측정하거나 호기심의 각성, 새로움의 제공, 지식에 대한 욕구를 충족시키는 능력의 결과로서 얻어지는 효용을 의미한다(Zeithaml, 1988). 인지적 가치에는 융합성, 시의성, 신속성, 생산성, 예측성, 순환성, 적절성, 통찰력, 데이터 원천, 데이터 범위가 포함되어 있다.

[군집 4] 기능적 가치란 기능, 품질, 서비스와 같은 물리적 기능이나 실용성에 관련된 개념으로, 사용자가 데이터를 사용할 때, 데이터의 성능, 구성, 신뢰성 및 가격과 같은 물리적 속성에 가치로 합리적인 경제에 의해 추구되는 가치라고 정의할 수 있다(Sweeney, 2001). 기능적 가치에는 객관성, 무결성, 데이터 오류, 반응성, 복구 가능성, 비교성, 일치성, 정확성, 형식이 포함되어 있다.

[군집 5] 상황적 가치란 선택을 내리는 사용자가 데이터를 사용하면서 상황에 따라 발생하는 정서적·감정적 상태를 의미하고, 사용자의 경험으로 인해 의도치 않게 발생하여 사용자의 의사 결정 과정에 있어서 매우 중요한 역할을 수행한다. 즉, 데이터

<표 7> 도출된 가치

사회적가치 (사회성)	금전적가치 (효율성)	인지적가치 (생산성)	기능적가치 (기능성)	상황적가치 (확장성)
경쟁력	적시성	데이터원천	데이터오류	데이터 모델링
창조력	적합성	융합성	무결성	사용자 지원
맞춤형	시간효율성	시의성	객관성	일관성
최신성	데이터관리	대응력	반응성	업무수용성
참여성	저장성	신속성	복구가능성	큐레이션
문제 해결성	분석성	생산성	비교성	표준화
투명성	재활 (사)용성	예측성	일치성	이해 가능성
사용빈도/ 양	가용성	순환성	정확성	데이터 오염
현재성	데이터압축	적절성	형식	이식성
시장반응성	메타데이터	통찰력		개인식별 가능성
지속성		범위		기밀성
참신함				추적성
수요주도성				해석 가능성
				책임성

를 이용한 후 개인이 느끼게 되는 심리적 감정이나 느낌이라고 정의할 수도 있으며, 데이터를 사용하면서 상황에 따라 추구되는 가치라고 정의할 수 있다. 상황적 가치에는 데이터 모델링, 사용자 지원, 일관성, 업무 수용성, 큐레이션, 표준화, 이해 가능성, 데이터 오염, 이식성, 개인식별 가능성, 기밀성, 추적성, 해석 가능성이 포함되어 있다.

5.2 AHP를 통한 공공 데이터 가치 요인 분석

AHP(Analytic Hierarchy Process) 기법은 1972년 Satty(1998)에 의해 개발된 다기준 의사결정 기법(Multi-Criteria Decision Making: MCDM)의 하나로 ‘분석적 계층화 과정’ 또는 ‘계층적 분석 과정/방법’이라고 해석할 수 있다(조근태, 2002). 즉 의사결정의 전 과정을 단계별로 분석하여 합리적인 의사결정을 지원해 주는 방법으로 현존하는 의사결정 방법 중 과학적이고 강력한 의사결정 방법이라 할 수 있다. 본 연구에서는 클러스터 분석 결과를 기반으로 데이터 사용자들이 각 속성 또는 차원에 대한 선호를 직접적으로 하는 방법인 AHP 방법을 통해 지표 간의 가중치를 분석하고자 한다. 이를 통해 응답자들의 주관적 판단을 논리적으로 정량화, 계량화하고, 의사결정의 일관성 유무에 대한 통계적 검증과 각 결정 요인들 간의 상호작용 정도를 측정하고자 한다.

본 연구에서는 요인분석 결과 도출된 5가지 가치를 기반으로 가치에 해당하는 요인을 포함하여 계층화 모델을 설정하였고, 설정된 모델을 업무 상황별 가중치를 통해 주요 요인을 도출하고자 하였다. 이를 도식화하면 다음 그림과 같다.

AHP 분석을 위해 필요한 조사 대상은 2년 이상 데이터 관련 업무를 수행하는 전문가를 대상으로. 설문 조사 방법은 온라인 설문 방식으로 6월 3일 ~ 6월 8일까지 총 16명을 대상으로 진행하였다. 이에 대한 정보는 다음 표와 같다. AHP 분석은 공개 AHP 분석 프로그램인 DRESS V.1.7(CHOISH Software)를 사용하여 수행하였다.

<표 8> AHP 전문가 정보

분류	분류	수
업무별	개발	8
	데이터분석	8
근무기간	2년 이상 4년 미만	4
	4년 이상 6년 미만	11
근무형태	6년 이상	1
	대기업	6
	스타트업	7
	공공기관(연구소, 학교, 정부)	4
	기타	2

AHP 분석 결과 전체적으로 사회적 가치(사회성) > 기능적 가치(기능성) > 인지적 가치(생산성) > 금전적 가치(효율성) > 감정적(편의성) > 상황적(확장성) 가치 순으로 도출되었다. 그러나 개발 업무의 경우 기능적 가치 > 사회적 가치 > 인지적 가치 순으로 도출되었고, 분석 업무의 경우 사회적 가치 > 금전적 가치(효율성) > 인지적 가치 순으로 도출되었다. 사회적 가치의 경우 개발과 분석 두 분야에서 전체적으로 높은 순위로 나타났다. 개발 업무의 경우 데이터 자체적인 기능적인 가치와 사회적



[그림 2] 계층화 모델

가치가 높게 도출됐지만, 금전적 가치, 즉 효율적 측면이 낮게 도출되었고, 분석 업무의 경우 사회적 가치와 금전적 가치인 효율적인 측면에 높았지만, 기능적 가치가 낮게 도출되었다. 반면 상황적 가치 측면에서는 둘 다 낮게 도출됨에 따라 데이터분석 또는 개발과정에서 보안에 대한 중요성은 높지만, 직접 사용하는 사용자 입장에서는 상황적 요인보다는 직접적인 요인인 기능 또는 효율성을 추구한다는 것을 알 수 있었다. 결과에 대한 상세한 내용은 다음 표와 같다. 사회적 가치의 경우 전체적으로 높은 순위를 나타내었고, 반대로 상황적 가치는 전체적으로 낮은 순위를 나타내었다.

〈표 9〉 연구결과

	가치	개발 업무		분석 업무		종합	
1	사회적 가치	0.18	2	0.49	1	0.48	1
2	금전적 가치	0.11	4	0.24	2	0.09	4
3	인지적 가치	0.15	3	0.13	3	0.21	2
4	기능적 가치	0.48	1	0.02	5	0.14	3
5	상황적 가치	0.06	5	0.06	4	0.02	5

6. 결 론

6.1 연구 요약 및 시사점

본 연구는 사용자 관점에서 공공데이터의 가치 요인은 무엇인가에 대한 의문에서 출발하였다. 본 연구에서는 기존 데이터의 가치를 측정해 대한 기준이 미비하여 개방형 설문, 메타분석, FGI를 통해 58개의 요인을 도출하고, SSPS를 통해 계층적 군집분석을 시행하여 사회적 가치, 금전적 가치, 인지적 가치, 기능적 가치, 상황적 가치로 총 5가지 군집으로 분류하였다. 사회적 가치는 사회적, 개방성을 중심으로 총 13개의 요인이 해당되었고, 금전적 가치는 효율성을 중심으로 10개의 요인이 해당되었다. 인지적 가치는 생산성을 중심으로 11개의 요인이 해당되었으며, 기능적 가치는 기능성을 중심으로 9개의 요인이 해당되었다. 마지막으로 상황적 가치에서는 데이

터 사용에 대한 확장성을 중심으로 14개의 요인이 포함되었다. 도출된 5가지의 군집은 신뢰도 및 타당도를 확보하였으며, 군집간의 다중공선성도 확보하였다. 그 결과 5가지의 군집을 데이터 사용 목적에 따라 AHP 분석을 통해 사용 목적에 따른 가중치를 측정하였다. 종합적으로 사회적 가치 > 인지적 가치 > 기능적 가치 순으로 도출되었으며, 개발 업무에서는 기능적 가치 > 사회적 가치 > 인지적 가치 순으로 도출되었다. 반면 분석 업무에서는 사회적 가치 > 금전적 가치 > 인지적 가치 순으로 중요도가 도출되었다. 이를 바탕으로 개발 업무에서는 데이터 자체적인 가치를 높게 평가하고 있는 반면, 분석 업무에서는 데이터의 개방성, 사회적 등 데이터를 통한 파급효과에 대해 높게 평가하고 있었다. 그러나 아직 데이터를 사용할 때 상황에 따른 확장적인 측면에 대한 가치는 미비하게 생각하고 있어 앞으로 데이터의 사용성 확장을 위한 연구가 필요할 것으로 생각된다.

본 연구의 학문적인 의의는 다음과 같다. 첫째, 사용자 관점에서 공공데이터를 측정할 수 있는 요인을 도출하였다. 실질적인 데이터 활용을 높이기 위해 데이터를 사용하는 사용자 관점에서 가치를 측정할 수 있는 요인을 개발함으로써 향후 데이터 활용도를 높일 수 있을 것으로 예상된다. 둘째, 데이터의 가치에 영향을 미치는 요인들을 대상으로 상위 차원을 도출하였다. 각 도출된 요인을 바탕으로 데이터 활용에 관한 연구를 확장할 수 있는 발판을 마련하였다. 마지막으로 데이터 가치에 관한 연구를 사용자 관점으로 확장했다는 점에 의의가 있다. 본 연구에서는 데이터를 실질적으로 활용하는 사용자의 관점에서 진행했다는 점에서 데이터 가치에 대해 실증 연구에 대한 근간을 제공했다고 볼 수 있다.

본 연구의 실무적 시사점 및 활용방안은 다음과 같다. 첫째, 본 연구는 데이터를 개방 또는 거래하거나 판매하는 공공기관이나 기업관점에서 데이터 활용도를 높이기 위한 근간이 되었고, 사용자의 기대 수준을 측정하기 위한 초석이 되었다. 구체적으로 본 연구에서 개발한 가치 요인을 적용하여 데이터

사용자들의 전반적인 기대하는 요인에 대해 파악할 수 있다. 둘째 본 연구 결과를 통해 사용자 관점에서 데이터 가치에 대한 평가를 통해 데이터의 활용을 높일 수 있다. 이를 통해 현재 데이터 포털에서 개방되고 있는 데이터들이 활용도 낮은 이유에 대해서도 파악할 수 있다. 셋째 데이터 거래 시 데이터에 대한 가격정보를 합리적으로 제안할 수 있다. 현재 데이터 가격 측정에 대한 가이드라인 부재로 인해 데이터 정보가 높다는 점이 데이터 거래 활성화의 장애로 도출되고 있다. 이러한 점을 반영하여 데이터 품질 측면뿐만 아니라 데이터 가치적인 측면을 포함하여 데이터의 가격을 합리적으로 제안할 수 있다.

6.2 연구의 한계 및 향후 연구

본 연구가 가지는 학문적, 실무적 의의에도 불구하고 다음과 같은 한계점을 지니고 있다. 본 연구에서는 대표적으로 3가지의 연구 한계를 갖고 있다. 첫 번째로 연구의 범위를 공공데이터와 데이터를 활용하는 사용자로 한정하였다. 그러나 분야별 사용자에 따라 데이터에 대한 요구 조건이 다르거나, 민간 데이터의 경우 공공데이터의 특성과는 다르기 때문에 보다 넓은 포괄적 데이터의 범위를 설정이 필요하다. 뿐만 아니라 데이터를 활용하는 측면에서 데이터를 관리하는 관리자 또는 데이터를 활용한 결과물을 사용하는 END USER 역시 데이터를 활용함으로써 보다 다양한 사용자의 관점이 포함되어 진행된다면 본 연구의 외적 타당도를 높일 수 있을 것이다. 둘째 본 연구에서 사용한 방법론에 대해 한계가 존재한다. 본 연구에서는 유사 요인 간의 군집화를 위해 군집분석을 시행하여 의미 있는 결과를 도출했지만, 실험적인 방법론을 통해 58개에 대한 새로운 군집화를 진행하여 보다 더 사용자 관점에서 군집을 진행하는 것이 필요해 보인다. 대표적으로 렌즈모델이나 Q 방법론 등 실험적인 UX 방법을 통해 사용자 관점에서 새로운 가치를 도출할 수 있을 것으로 기대한다. 그뿐만 아니라 AHP 분석에서 도출된 상위개념만을 계층화 모델로 설정하여 도출된 하위

요인들에 대한 AHP 분석이 필요하다. 세 번째로 군집화되어 도출된 가치와 요인들에 대해 검증이 필요하다. 일차적으로 도출된 58개의 요인이 데이터 가치에 영향을 미치는지에 대한 검증과 이차적으로 군집분석으로 도출된 가치에 포함된 요인들이 실제 도출된 가치에 영향을 미치는지에 도출된 가치에 대한 객관성이 높아질 것으로 예상된다.

본 연구를 시작으로 발전될 수 있는 향후 연구에 관해 기존 선행연구에서 사용되었던 데이터의 가치 측정 항목들과 본 연구에서 개발한 가치 측정 항목들을 직접적으로 비교해볼 수 있다. 두 번째로 사용자 중심의 방법론을 통해 새롭게 요인이 도출될 수 있다. 대표적으로 본 연구에서 진행한 FGI 대신 델파이 분석을 시행하거나 Q 방법론 및 실험방법을 통해 새로운 관점에서 도출된 요인들에 대해 군집화를 진행할 수 있을 것이다. 마지막으로 연구 범위를 포괄적인 범위에서 데이터 가치체계를 기반으로 다양한 사용자의 관점에서 데이터의 가치에 관한 연구를 추가로 진행하고자 한다.

참고문헌

- 강만모, 김상락, 박상무 “빅 데이터의 분석과 활용”, *정보과학회지*, 제30권, 제6호, 2012, 25-32.
- 관계 부처 합동, “혁신성장 전략투자 데이터·AI경제 활성화 계획(‘19~’23년)”, 2010.01.
- 김선호, 이창수, “마스터 데이터 품질관리 프레임워크”, *Entrue Journal of Information Technology*, 제9권, 제2호, 2010, 109-121.
- 김은지, 이태욱 “컴퓨팅 사고력 평가 도구로써 자기보고식 설문과 비버 챌린지 간의 상관관계”, *한국컴퓨터정보학회 학술발표논문집*, 제26권, 제1호, 2018, 113-116.
- 문동규, 김영희, “청소년의 자살생각과 관련된 유발변인의 메타회귀분석”, *상담학연구*, 제12권, 제3호, 2011, 945-964.
- 박주현, 류한영, “사물인터넷 서비스의 사용자 가치 요인”, *한국 HCI 학회 논문지*, 제11권, 제2호,

- 2016, 23-30.
- 성육준, “4 차산업혁명 시대 지역발전의 기회와 도전”, *한국지역정보화학회 학술발표대회 논문집*, 2017, 65-80.
- 성육준, “지능정보시대의 공공부문 빅데이터 활성화 방안”, *한국정책학회 춘계학술발표논문집*, 2017, 1-22.
- 송인국, “인터넷기반 공공데이터 활용방안 연구: 혜택, 기회, 비용, 그리고 위험요소 관점에서”, *인터넷 정보학회논문지*, 제16권, 제4호, 2015, 131-139.
- 송효진, 황성수, “정부 3.0 추진에 따른 공공데이터 개방과 지방정부의 방향성 모색”, *한국지역정보 화학회지*, 제17권, 제2호, 2014, 1-28.
- 송효진, 이태휘, 김형주, “모바일 통합 검색에서 온톨로지를 활용한 컬렉션 순서의 정확도 향상”, *정보과학회논문지: 컴퓨팅의 실제 및 레터*, 제18권, 제3호, 2012, 187-196.
- 오성삼, “선행 연구 결과 의 통합 과 분석 을 위한 메타 분석의 이론과 실제”, *Kōnguk Taehakkyo Chūlpánbu*, 2002.
- 이건학, 김감영, “공간통계 기법을 이용한 현주인구 추정 모델링”, *한국지도학회지*, 제16권, 제2호, 2016, 71-93.
- 이경서, 문혜정, “의미망 분석을 통한 공공데이터 가치 변화 연구”, *학술대회 발표논문집*, 2018, 5-22.
- 이은영, “고객기대수준의 측정척도 개발”, *마케팅연구*, 제34권, 제1호, 2019, 49-73.
- 장석인, “제4차 산업혁명 시대의 산업구조 변화 방향과 정책과제”, *국토*, 제424호, 2017, 22-30.
- 장은순, “행정정보의 통계적 활용을 위한 품질요건에 관한 연구”, *Journal of Digital Convergence*, 제12권, 제6호, 2014, 43-53.
- 전치홍, 심재웅, 조대곤, “공공데이터의 경제적 가치 측정 : 소비자 효용 관점”, *Information Systems Review*, 제20권, 제2호, 2018, 1-19.
- 정보화진흥원, “데이터 경제의 부상과 사회경제적 영향”, 2017.
- 정보화진흥원, “데이터 경제를 위한 우리의 준비”, 2018.
- 정인숙, 전성숙, 황선경, 김동희, 하주영, *체계적 문헌 고찰과 메타분석*, 서울: 수문사, 2011.
- 조근태, “기술대안의 전략적 평가를 위한 AHP 적용에 있어서 평가자 신뢰성을 고려한 가중치 통합”, *경영과학*, 제19권, 제2호, 2002, 139-154.
- 차영일, 최성규, 한경석, “민간의 공공데이터 활용을 위한 이용의도에 미치는 영향에 관한 실증적 연구”, *디지털융복합연구*, 제15권, 제6호, 2017, 9-17.
- 최광렬, 김창재, “빅데이터 성숙도 모델에 관한 연구”, *한국정보기술학회논문지*, 제12권, 제12호, 2014, 149-162.
- 한상규, 김은진, 김희웅, “공공데이터의 표준 비표준 비교를 통한 표준화 방안 연구”, *한국경영정보학회 학술대회*, 2019, 63-63.
- 허필선, 박광만, 박원주, 조기성, 류원, “공공정보 민간 활용 시장 및 파급효과”, *전자통신동향분석*, 통권 142호, 2013, 118-131.
- Ahn, D., J.S. Park, C.S. Kim, J. Kim, Y. Qian, and T. Itoh, “A design of the low-pass filter using the novel microstrip defected ground structure”, *IEEE Transactions on Microwave Theory and Techniques*, Vol.49, No.1, 2001, 86-93.
- Babin, B.J., W.R. Darden, and M. Griffin, “Work and/or fun: Measuring hedonic and utilitarian shopping value”, *Journal of consumer research*, Vol.20, No.4, 1994, 644-656.
- Bowker, G. C., *Memory practices in the sciences* (Vol. 205). Cambridge, MA: Mit Press, 2005.
- Costanza, R., R. de Groot, P. Sutton, S. Van der Ploeg, S.J. Anderson, I. Kubiszewski, ... and R.K. Turner, “Changes in the global value of ecosystem services”, *Global Environmental Change*, Vol.26, 2014, 152-158.
- DeLone, W.H. and E.R. McLean, “Information systems success: The quest for the depend-

- ent variable”, *Information Systems Research*, Vol.3, No.1, 1992, 60–95.
- Dodds, W.B. and K.B. Monroe, “The effect of brand and price information on subjective product evaluations”, *ACR North American Advances*, 1985.
- Doll, W.J. and G. Torkzadeh, “The measurement of end-user computing satisfaction”, *MIS Quarterly*, 1988, 259–274.
- Fornell, C., M.D. Johnson, E.W. Anderson, J. Cha, and B.E. Bryant, “The American customer satisfaction index: nature, purpose, and findings”, *Journal of Marketing*, Vol.60, No.4, 1996, 7–18.
- Gartner, Inc., “Big Data Fades to the Algorithm Economy”, *Forbes*, 2015.8.14. <https://www.forbes.com/sites/gartnergroup/2015/08/14/big-data-fades-to-the-algorithm-economy/?sh=3df8d72651a3>.
- Glass, G.V., “Primary, secondary, and meta-analysis of research”, *Educational Researcher*, Vol.5, No.10, 1976, 3–8.
- Goodhue, D.L. and R.L. Thompson, “Task-technology fit and individual performance”, *MIS Quarterly*, 1995, 213–236.
- Gutman, J. and M.K. Mills, “Fashion life-style, self-concept, shopping orientation, and store patronage—an integrative analysis”, *Journal of Retailing*, Vol.58, No.2, 1982, 64–86.
- Hu, Z., B.J. Deibert, and J. Li, “Luminescent metal-organic frameworks for chemical sensing and explosive detection”, *Chemical Society Reviews*, Vol.43, No.16, 2014, 5815–5840.
- IBM, “Analytics the widening divide original”, *MIT Sloan Management Review*, 2011
- Jason Lewis, “Planning Guide for Data and Analytics”, *Gartner Research*, 2018.
- Jensen, M.C., “Value maximization, stakeholder theory, and the corporate objective function”, *Journal of Applied Corporate Finance*, Vol.14, No.3, 2001, 8–21.
- Jetzek, T., M. Avital, and N. Bjorn-Andersen, “Data-driven innovation through open government data”, *Journal of Theoretical and Applied Electronic Commerce Research*, Vol.9, No.2, 2014, 100–120.
- Kavadias, S., K. Ladas, and C. Loch, “The transformative business mode”, *Harvard Business Review*, Vol.94, No.10, 2016, 91–98.
- Kim, E. and J. Lee, “An exploratory contingency model of user participation and MIS use”, *Information & Management*, Vol.11, No.2, 1986, 87–97.
- Kotler, P., G. Armstrong, M. Harker, and R. Brennan, “Marketing: an introduction (Vol. 1). Englewood Cliffs, NJ: Prentice-Hall, 1990.
- Lee, Y.W., D.M. Strong, B.K. Kahn, and R.Y. Wang, “AIMQ: a methodology for information quality assessment”, *Information & Management*, Vol.40, No.2, 2002, 133–146.
- Liljander, V., “Modeling perceived service quality using different comparison standards”, *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, Vol.7, 1994, 126–142.
- McAfee, A., E. Brynjolfsson, T.H. Davenport, D.J. Patil, and D. Barton, “Big data: The management revolution”, *Harvard Business Review*, Vol.90, No.10, 2012, 60–68.
- Miller, D., *The comfort of things*, Polity, 2008.
- Newbery, D.M., L. Bently, and R. Pollock, “Models of public sector information provision via trading funds, 2008.
- Park, J. and S.H. Han, “Defining user value: A case study of a smartphone”, *International*

- Journal of Industrial Ergonomics*, Vol.43, No.4, 2013, 274-282.
- Park, J., S.H. Han, H.K. Kim, S. Oh, and H. Moon, "Modeling user experience: A case study on a mobile device", *International Journal of Industrial Ergonomics*, Vol.43, No.2, 2013, 187-196.
- Parsons, T., "Illness and the role of the physician: A sociological perspective", *American Journal of Orthopsychiatry*, Vol.21, No.3, 1951, 452.
- Pipino, L.L., Y.W. Lee, and R.Y. Wang, "Data quality assessment", *Communications of the ACM*, Vol.45, No.4, 2002, 211-218.
- Rokeach, M., "The role of values in public opinion research", *Public Opinion Quarterly*, Vol.32, No.4, 1968, 547-559.
- Rowley, J., "The wisdom hierarchy: representations of the DIKW hierarchy", *Journal of Information Science*, Vol.33, No.2, 2007, 163-180.
- Sheth, J.N., B.I. Newman, and B.L. Gross, "Why we buy what we buy: A theory of consumption values", *Journal of Business Research*, Vol.22, No.2, 1991, 159-170.
- Simonson, I., "The BDT effect and future: A reply to John lynch and Norbert Schwarz", *Journal of Marketing Behavior*, Vol.1, No.1, 2015, 59-73.
- Snavely, H.J., "Accounting information criteria", *The Accounting Review*, Vol.42, No.2, 1967, 223.
- Strong, D.M., Y.W. Lee, and R.Y. Wang, "Data quality in context", *Communications of the ACM*, Vol.40, No.5, 1997, 103-110.
- Sweeney, J.C. and G.N. Soutar, "Consumer perceived value: The development of a multiple item scale", *Journal of Retailing*, Vol.77, No.2, 2001, 203-220.
- Wang, G., A. Gunasekaran, E.W. Ngai, and T. Papadopoulos, "Big data analytics in logistics and supply chain management: Certain investigations for research and applications", *International Journal of Production Economics*, Vol.176, 2016, 98-110.
- Wang, R.Y. and D.M. Strong, "Beyond accuracy: What data quality means to data consumers", *Journal of Management Information Systems*, Vol.12, No.4, 1996, 5-33.
- Wang, R.Y., "A product perspective on total data quality management", *Communications of the ACM*, Vol.41, No.2, 1998, 58-66.
- Wang, R.Y., Y.W. Lee, L.L. Pipino, and D.M. Strong, "Manage your information as a product", *MIT Sloan Management Review*, Vol.39, No.4, 1998, 95.
- Zuiderwijk, A. and M. Janssen, "Open data policies, their implementation and impact: A framework for comparison", *Government Information Quarterly*, Vol.31, No.1, 2014, 17-29.

◆ About the Authors ◆



이 상 은 (go20002002@naver.com)

연세대학교 정보대학원(석사)을 졸업하고, 현재 연세대학교 기술경영학 박사과정에 재학하고 있다. 주요 관심사는 공공데이터, 사용자 UX/UI, 스마트시티 등이다.



이 정 훈 (jhoonlee@yonsei.ac.kr)

영국 Manchester 대학교 전기 전자공학(학사), 영국 Manchester 대학교 정보시스템공학(석사) 및 영국 London School of Economics 경영정보학(석사), University of Cambridge, Institute of Manufacturing 산업공학경영(박사), 현재 연세대학교 정보대학원 정교수로 재직 중이다. 스마트시티, IT 성과관리 및 거버넌스 관련 다수 프로젝트에 참여하였으며 주요 연구 연구분야는 스마트시티, IT 성과관리 및 거버넌스, IoT, 빅데이터 서비스, 시민참여 등을 포함한다.



최 현 진 (zxsk56@naver.com)

홍익대학교 경영학과(학사), 현재 연세대학교 정보대학원 비즈니스 빅데이터 분석 트랙 정보시스템학(석사) 과정 중에 있다. 주요 관심분야는 Information Systems, Data analysis, IoT Business Model, Smart City 등이다.