

생태계 관점에서의 빅데이터 활성화를 위한 구조 연구*

조지연** · 김예진*** · 박건철** · 이봉규****

An Analysis of Big Data Structure Based on the Ecological Perspective*

Jiyeon Cho** · Taisiya Kim*** · Keon Chul Park** · Bong Gyou Lee****

■ Abstract ■

The purpose of this research is to analyze big data structure and various objects in big data industry based on ecological perspective. Big data is rapidly emerging as a highly valuable resource to secure competitiveness of enterprise and government. Accordingly, the main issues in big data are to find ways of creating economic value and solving various problems. However big data is not systematically organized, and hard to utilize as it constantly expands to related industry such as telecommunications, finance and manufacturing. Under this circumstance, it is crucial to understand range of big data industry and to which stakeholders are related. The ecological approach is useful to understand comprehensive industry structure. Therefore this study aims at confirming big data structure and finding issues from interaction among objects. Results of this study show main framework of big data ecosystem including relationship among object elements composing of the ecosystem. This study has significance as an initial study on big data ecosystem. The results of the study can be useful guidelines to the government for making systemized big data ecosystem and the entrepreneur who is considering launching big data business.

Keyword : Big Data, Ecological Perspective, Big Data Structure, Stakeholder, Ecosystem

논문투고일 : 2012년 10월 25일 논문수정완료일 : 2012년 12월 01일 논문게재확정일 : 2012년 12월 06일

* This treatise was supported by the project of Global Ph.D. Fellowship which National Research Foundation of Korea conducts from 2012.

** 연세대학교 정보대학원 박사과정

*** 연세대학교 일반대학원 기술경영학협동과정 박사과정

**** 연세대학교 정보대학원 교수

1. 서론

최근 스마트 기술의 급속한 발전과 함께 등장한 빅데이터(Big data)는 정보통신분야의 선두에 있는 한국정부와 기업의 핵심 경쟁력이 되어줄 주요 자원으로 인식되고 있다[4]. 역동적인 산업 패러다임 변화와 미래사회의 불확실성에 대응하고, 새로운 가치를 창출할 핵심 자원으로 빅데이터가 주목 받는 것이다. 그러나 가트너의 보고서에 따르면 Fortune 500대 기업 중 85% 이상이 빅데이터의 성공적인 활용에 실패할 것으로 예측하였다[23]. 즉 대부분 기업들이 빠르게 증가하는 복잡한 구조의 대용량 데이터를 수집하고 분석하여 기업의 의사결정에 적시에 활용할 수 있는 능력을 갖추지 못하였다는 것이다. 하지만 빅데이터 활용 실패는 데이터에 대한 기업의 관리 능력 부족 뿐 아니라 빅데이터를 둘러싼 복잡한 이해관계와 주변 환경 등도 관련이 있다.

빅데이터는 개인이 데이터를 생산하고 소비하는 역할을 동시에 수행하고, 제 3자가 빅데이터를 가공하여 서비스를 제공하는 등 복잡한 이해관계 구조를 형성하고 있다[28]. 또한 빅데이터는 데이터 자체의 생산과 사용에 있어 즉시성, 복잡성, 연결성 등의 특징으로 개별 기업 수준에서 비즈니스 전략을 수립하기 어렵다. 이러한 복잡한 빅데이터 구조 분석에 생태계 연구가 유용한데, 그 이유는 경계를 넘어선 여러 산업 간, 비즈니스 간 유기적으로 연결된 구조와 개별 개체들의 역할에 대한 종합적인 해석이 가능하기 때문이다. 이기식 외[20]의 소프트웨어산업 생태계 연구에서도 기술 발전에 따른 소프트웨어 산업의 컨버전스 현상, 즉 산업 간 경계가 와해되고 산업, 비즈니스가 유기적으로 연결되는 네트워크형 산업구조를 분석하기에 생태계 관점이 적절하다고 언급한다.

즉 빅데이터의 성공적인 활용과 장기적 관점의 생태계 선순환 구조 형성을 위하여, 빅데이터를 둘러싼 생태계에 관한 학문적 연구가 필요한 시점이다. 그러나 기존 빅데이터 관련 연구들은 빅데

이터의 의미와 실체, 분석 기술 및 사례 연구 등에 초점이 맞추어져 왔다[1, 6, 9, 28]. 최근 성공적인 빅데이터 활용을 위한 메커니즘 수립, 빅데이터 생태계 구성 방안 등에 관심이 증가하고 있지만, 빅데이터 생태계에 대한 학술적 논의는 충분히 이루어지지 못하였다[22, 24, 27, 31].

따라서 본 연구는 빅데이터에 대한 생태계학적 관점의 접근이 가능한가? 빅데이터 생태계의 구성 개체와 역할은 무엇인가? 등에 대한 문제의식 하에 빅데이터 생태계를 분석하는 것을 목표로 한다. 이를 위하여 기존 생태계 연구들을 검토하고 빅데이터의 특성을 반영한 생태계 분류체계를 제시하고자 한다. 본 연구는 향후 빅데이터 생태계에 대한 복잡하고 심화된 연구를 위한 기초자료로써, 빅데이터 생태계 형성 및 진화 방안에 대하여 유용한 시사점을 제공할 것으로 기대한다.

2. 이론적 배경

2.1 빅데이터 개념

빅데이터에 대한 초기 개념은 빅데이터의 양, 생성 속도, 다양한 형태 3V(Volume, Velocity, Variety)에 초점이 맞추어져 왔다. 하지만 최근에는 빅데이터의 가치를 찾기 위한 모든 연관 활동을 의미하는 것으로 범위가 점차 확대되고 있다. 데이터 크기에 초점을 맞춘 빅데이터의 정의는 기존 DB, SW로 저장, 관리, 분석하기 어려운 정도로 큰 규모의 자료를 뜻한다[36]. IDC에서는 다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치를 추출하고 초고속 수집, 발굴, 분석을 지원하도록 고안된 차세대 기술 및 아키텍처로 빅데이터의 범위를 데이터를 통한 업무 수행에 초점을 맞추어 정의내리고 있다[35]. 빅데이터는 그 의미뿐 아니라 관련 산업 분야도 확대되고 있다. 초기 빅데이터는 주로 인터넷, 포털, 통신 분야에서 트위터, 페이스북 북 등 개인 사용자에 의해 발생하는 소셜데이터를 지칭했다. 그러나 최근 빅데이터는 스

<표 1> 빅데이터 개념의 확대

구분	내용
빅데이터 정의	<ul style="list-style-type: none"> • 협의의 빅데이터 정의 : 기존 DB, SW로 저장, 관리, 분석하기 어려운 규모의 자료로 기존 데이터와 비교하여 크기가 매우 크고 급속도로 증가하는 특징을 가진 데이터를 의미, 문자, 이미지와 같은 비정형 데이터 등을 포함 • 광의의 빅데이터 : 대용량 데이터를 활용하고 분석하여 가치 있는 정보를 추출하고, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 전 과정을 의미
빅데이터 산업	<ul style="list-style-type: none"> • 빅데이터의 개념이 확장됨에 따라 빅데이터는 관련 사회·경제활동을 지칭하는 포괄적인 의미로 기업 활동, 정부 정책, 기술 개발 등의 활동 모두를 포함하는 산업의 범위의 개념으로 확대 됨 • 빅데이터 산업의 연관 범위는 기술 발전에 따른 데이터 생산량 폭증과 함께 다양한 분야로 계속해서 확대될 것으로 예상됨 • 최근에는 IT 분야 외에 제조, 의료, 금융 등 전 산업 분야에서 빅데이터가 생성되고 있으며, 빅데이터 활용 방안에 대한 관심이 증가하고 있음
빅데이터의 가치 확대	<ul style="list-style-type: none"> • 빅데이터는 사회 환경 변화를 신속하게 감지하고 대응하는 역량의 요소로, 내재되어 있는 가치 창출을 통하여 개인, 기업, 정부의 경쟁력의 원천으로 각광 받으며 빅데이터에 대한 기대가 증가하는 추세임

마트 디바이스의 보급 확대와 SNS 이용 활성화, RFID, M2M, 센서 기술 적용 등의 IT 발전에 따라 다양한 산업 분야에서 생성되고 있다[31]. 즉 인터넷을 기반으로 한 디지털 경제 확산의 결과로 빅데이터 관련 산업은 점점 확산될 것으로 전망되고 있다[24]. 이러한 빅데이터의 개념과 분야는 조직의 업무 특성과 데이터를 바라보는 입장에 따라 다양한 시각과 이해가 존재하지만 점차 그 범위는 확대되는 추세이다[28].

최근 빅데이터에 관한 시각이 확대됨에 따라 빅데이터를 넓은 의미에서 산업과 비즈니스의 범위까지 정의내리는 연구들도 있다. 양창준[13]의 연구에서는 빅데이터 산업이란 빅데이터 관련 사회·경제활동을 지칭하는 포괄적인 의미로 기업 활동, 정부 정책, 기술 개발 등의 활동 모두를 포함하며 이와 연관된 기술, 서비스 산업을 의미한다고 하였다. 이와 유사한 맥락에서 김상락 외[6]의 연구에서는 데이터 속에서 가치 있는 정보를 찾고 인포그래픽으로 이해하기 쉽게 표현하여 알기 쉽게 전달하고 정보를 원하는 사람과 기관에 판매하는 비즈니스 과정을 전부 포괄한다고 하였다. 이처럼 최근 많은 연구에서는 빅데이터를 데이터 자체를 넘어서 가치 있는 정보를 추출하여 활용하기까지의 일련의 과정과 관련된 기술, 도구 및 인력, 조

직, 산업으로 의미를 점차 확대하고 있다.

그렇다면 이러한 전 산업 분야에 걸쳐 빅데이터가 주목 받는 이유는 무엇일까? 빅데이터는 기존 데이터와 비교하여 저장, 검색, 관리, 공유의 단계에 분석과 추론의 과정이 추가된 기존 정보화 시대의 데이터와는 다른 역할을 수행하게 되었다[32] (<표 1> 참조). 빅데이터는 사회 환경의 빠른 변화를 신속하게 감지하고 대응하는 역량 원천 요소로서 역할을 수행할 수 있다는 점에서 주목받고 있다. 더불어 데이터의 양적인 팽창 뿐 아니라, 대용량, 비정형 데이터를 분석, 처리할 수 있는 기술도 병행하여 발전함에 따라 데이터 자체가 가진 자원으로서의 가치가 점점 증가하는 추세이다[9]. 이로 인하여 빅데이터 확보와 성공적인 활용을 위한 개인, 기업, 정부의 협력과 데이터를 체계적으로 수집 하고, 지속적인 정보와 지식의 창출과 축적, 이에 기반을 둔 핵심역량 강화에 대한 필요성이 강조되고 있다[32].

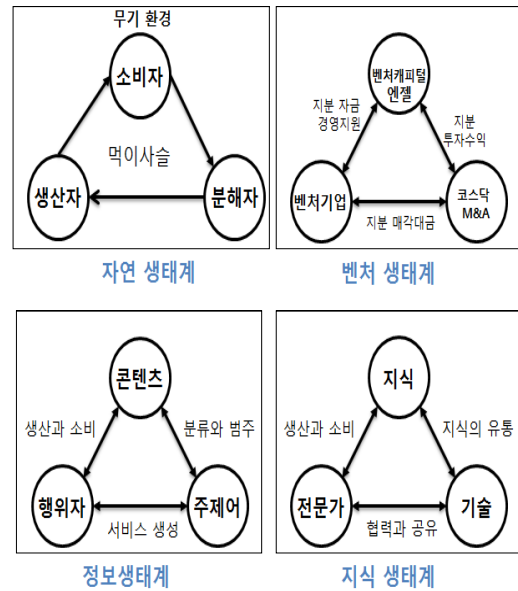
빅데이터는 앞서 논의한바와 같이 데이터 자체의 특성과 데이터 소유권을 둘러싼 다양한 이해관계가 연관되어 있다. 따라서 관련 하드웨어, 소프트웨어 등 기술적 요소와 더불어 자본, 인적 자원, 비즈니스 요소 등의 균형 잡힌 확보가 무엇보다 중요하다. 이를 위한 해법으로 빅데이터 생태계

구축이 거론되고 있는 것이다. 따라서 건전한 생태계 구축을 위하여 장기적인 관점의 데이터 확보와 분야를 넘어선 데이터 공유를 위한 기술, 시장, 산업에 대한 종합적인 연구가 필요한 시점이다.

2.2 생태계 연구

정보 사회과학 분야에 생태계 개념의 적용은 산업 전체의 종합적인 이해를 위한 목적으로 다양한 분야에 적용되어 왔다[10]. 생태계 이론에서 생태계란 생물 공동체와 이를 둘러싼 환경이 상호작용하는 공간을 의미한다. 1935년 영국의 식물생태학자인 텐슬리가 특정 환경에서 생존, 번식, 진화하는 생물의 관계를 연구한 결과로 생태계의 개념을 제시하였다[2, 21]. 생태계는 1차 생산자, 2차 소비자, 3차 분해자로 이어지는 먹이사슬이 형성되어 있고, 개체 간 상호작용에 의하여 생태계의 균형이 유지되고 외부환경에 적응하며 진화한다[2]. 생태계에서 외부 환경이란 개체에 영향을 미칠 가능성이 있는 외적인 모든 요소를 총칭하는 것으로 개체들이 어떻게 상호작용하며 환경의 변화에 적응하는지 연구한다. 종합하면 생태계란 유기체, 환경, 상호 의존성 등 자연 질서를 기초로 하는 생태학의 원리를 기본으로 한다[7].

생태계 모델에 기본이 되는 생태학 원리는 다음과 같다[15, 16, 26]. 첫째, 자기조직화 원리로 이는 생태계 외부의 개입, 통제 없이도 생태계 내에 개체가 상호작용을 통해 자생적으로 질서와 규칙을 만들고 성장하는 메커니즘의 원리이다. 둘째는, 상호의존성의 원리로 생태계 내의 개체들이 서로 의존적인 관계로 존재한다는 것이다. 셋째, 다양성의 원리는 생태계 내의 개체 간 특징과 역할의 차이와 그에 따른 먹이사슬의 관계를 나타낸다. 넷째는, 물질의 선순환 원리로 끊임없이 에너지가 유입되고 순환되며, 먹이사슬을 통해 에너지가 흐르는 순환성을 의미한다. 다섯째는 공진화의 원리로 상호 의존하는 개체 간 긍정적인 영향을 통해 진화한다는 원리이다. 마지막은 개체의 변종원리인



자료 : 고정민 외[2], 유계미 외[16], 이수상[19] 참조.
 [그림 1] 분야별 생태계 개념 비교

데 환경이 급변하는 경우, 적응하지 못한 생태계의 개체가 파괴되거나, 새로운 종류의 잡종이 나타날 수 있다는 것을 의미한다. 이러한 생태계의 기본 원리는 본 연구에서 생태계 모델에 적용되어 산업의 내 개체 간의 상호작용을 설명하는데 반영된다. 이러한 생태학 원리를 적용한 산업생태계 분석은 산업 내 이해관계자의 위치를 식별하고, 산업의 지속적인 발전과 진화를 위하여 어떠한 역할을 수행해야 하는지에 대한 시사점을 제공한다.

기존 생태계 연구들은 생태계 이론을 기본으로 분야별 생태계 모델을 제시하였는데, 사회과학 분야의 경우 벤처생태계, 미디어생태계, 정보생태계, 지식생태계 등의 분야에 적용되며 산업의 발전을 위한 종합적인 분석을 위한 이론으로 적용되어 왔다[7, 8, 10, 11, 19, 26]. [그림 1]은 자연 생태계의 기본 개념과 생태학 원리를 적용한 관련 연구들의 특징을 보여준다.

벤처생태계에 관한 고정민 외[2]의 연구에서는 벤처 기업과 그 주위 환경이 상호작용하는 공간으로 벤처생태계를 정의하였다. 벤처 기업, 벤처 캐

피털, 엔젤, 코스닥, M&A 등이 생태계 내의 주요 개체로 자본과 투자 수익 등을 토대로 상호작용하고 있다. 연구에서는 벤처 산업 내 개체들의 기능과 구조를 정의 내리고 그에 따른 국가별, 지역별 벤처 생태계를 진단하였다. 연구 결과 벤처 생태계의 중요 요인은 생태계 내의 개체와 주변 요인, 인프라 등이 상호작용 가능한 시스템 구축과 순환 시스템 구축을 위한 정부의 역할임을 강조하였다.

미디어 생태계는 인류, 미디어, 기술 간의 상호작용과 커뮤니케이션으로 구성된 생태계를 의미한다. 김유석[8]의 연구에서는 미디어 서비스의 특성을 반영하여 개체군을 구분하고 생태계 내의 콘텐츠 순환 모형을 도출하였다. 콘텐츠의 생산 및 전달자, 단순 수용자, 혁신적 소비자, 미디어 서비스 등이 포함되고, 최근의 미디어 생태계는 일방향의 가치사슬이 아닌 콘텐츠의 생산자와 소비자가 상호작용하는 관계를 가진다고 분석하였다. 이러한 과정에서 미디어 생태계 순환을 위한 분해자로서의 소비자의 역할을 강조하였다. 이와 유사한 성격의 문화콘텐츠 산업을 연구한 류준호 외[10]의 연구에서는 생태계의 생산, 소비, 재생산 가치구조를 기준으로 내부 구조를 분석하고 외부 요인을 분석하였다. 두 연구는 공통적으로 콘텐츠의 재생산을 위한 2차 생산자, 혁신적 소비자의 역할의 중요성을 강조하였다.

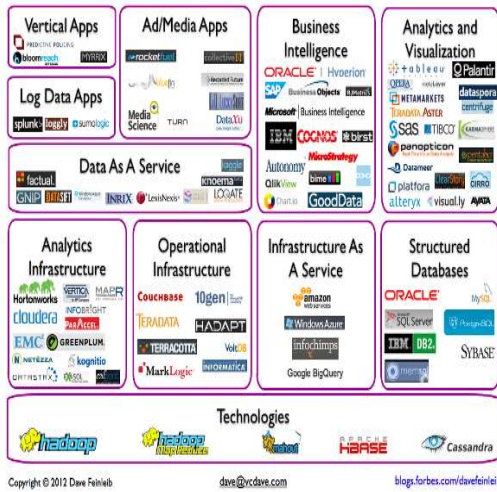
비교적 논의가 활발하게 진행되어온 정보생태계와 지식생태계의 경우 공통적으로 지식과 정보의 자연적 순환 과정에 초점을 맞추고 있다. 정보생태계는 새로운 정보 환경에서 각종 콘텐츠와 주제가 생산, 축적, 개방, 소비, 재생산과 같은 순환 과정에 기초하여 다양한 개체들이 참여하는 정보 환경을 의미 한다. 이러한 개념은 도서관 정보환경 관련 연구에 많이 사용되어 왔다[11, 19]. 정보생태계의 특징은 주요 개체들이 독자적으로 존재하거나 개체 간 다양한 연결 관계를 가지는 네트워크 구조라는 것이다. 지식생태계는 지식의 창조 활용을 위해 적합한 사회적, 조직적, 행동적, 기술적 환경을 제공해주는 방법과 환경을 의미하며, 지식

관리 및 교류, 학습, 지식 경영의 관점 등 다양한 분야에서 연구가 이루어져왔다[15, 16, 17, 26]. 관련 연구들은 지식 창출, 활용, 유통 및 확산과정을 이해하기 위한 개인의 사회적 연결망과 지식 교류 활동, 지식을 가진 자들의 협력적 경쟁관계를 분석한다. 정보생태계와 지식생태계에 관한 더욱 자세한 논의는 빅데이터 생태계 모델에서 더 자세히 언급한다.

이 외에도 생태계 관점의 산업 분석은 IT 발전에 따른 변화의 가속화로 인한 기업의 불확실성에 대한 대응 방안으로 환경 변화와 산업의 종합적인 이해를 위한 이론으로 각광받고 있다. 또한 IT 융합에 따라 복잡한 산업 구조를 가지는 통신, 소프트웨어, 클라우드 컴퓨팅 산업 등의 디지털 산업 분야의 생태계 연구로 확장되고 있다[3, 5, 20]. 생태계 관련 기존 문헌 연구를 통해 생태계 구조에 대한 올바른 이해가 이루어져야 균형 잡힌 생태계가 형성되고 생태계 내 개체들이 상호작용하며 공진화 할 수 있다는 것을 확인하였다. 따라서 선순환 구조의 빅데이터 생태계 정립을 위해서 생태계의 기본 개념을 이해하고 개별 개체들의 역할을 이해 할 필요가 있다는 것을 확인할 수 있다.

2.3 빅데이터 생태계 연구

빅데이터 생태계 관련 연구는 일부 논문이나 보도에서 빅데이터 생태계의 필요성을 제기하거나, 연구기관의 보고서에서 빅데이터 생태계의 주요 참여자를 제시한 연구가 대부분이다. 생태학 이론을 토대로 빅데이터를 분석하고 생태계 개체들의 역할과 상호작용 관계에 초점을 맞춘 학술 연구는 충분하지 않은 실정이다. 해외의 경우에도 기술 플랫폼 중심의 기업 생태계와 빅데이터를 보유한 기업과 하드웨어, 소프트웨어 기업의 파트너십을 보여주는 수준이다. [그림 2]는 해외의 빅데이터 관련 IT 인프라, 분석 기술, 애플리케이션 기업 등의 기술과 파트너 관계 등 빅데이터 관계자의 범위를 보여주고 있는 예이다. 국내의 경우 빅데이터 생



자료 : <http://www.forbes.com/sites/davefeinleib/2012/06/19/the-big-data-landscape/>.

[그림 2] 빅데이터 랜드스케이프

태계에 초점을 맞춘 정보통신 분야의 주요 연구 기관에서 발간한 연구들이 있다(<표 2> 참조).

데이터 생태계의 재정립에 초점을 둔 정보통신 정책연구원의 연구보고서에서는 ‘빅데이터 환경’을 민간 기업의 경영활동, 정부를 포함한 공공 부문의 혁신을 수반하는 패러다임의 변화라고 하였다

[24]. 이 과정에서 제기되는 정책적 이슈를 미디어 분야의 빅데이터에 초점을 맞추어 분석하고 있다. 연구 결과 빅데이터 패러다임에 대응하기 위하여 데이터 생태계 재정립이 필요하며, ‘데이터 공유 환경 조성’에 초점을 맞추고 데이터 유통 체계 혁신의 필요성을 강조한다. 특히 데이터 공유 환경 조성을 위한 데이터 지도와 데이터에 관한 정보가 집약된 데이터웨어 구축을 위한 개념도를 제시한다. 보고서에서는 데이터의 효율적인 생산체계와 생산된 데이터를 통제하고 빅데이터 환경에 대응할 수 있는 정부의 전담 조직의 필요성을 강조한다. 즉 빅데이터 생태계 내 촉매자로서의 정책기관의 역할을 강조하고 있다. 그러나 정책적으로 이슈가 되고 있는 데이터의 공유와 유통에 초점을 둔 생태계 연구라는 점에서 전체 빅데이터 산업을 이해하는 데에는 한계점을 가진다.

빅데이터와 관련한 한국정보화진흥원의 연구 보고서에서는 ‘빅데이터 서비스’를 중심으로 시장을 분석하였다[27]. 보고서에서는 빅데이터를 구성하고 있는 하드웨어, 소프트웨어를 포괄하는 모든 프로세스를 의미하는 거대 플랫폼으로 정의 내리고, 빅데이터를 재가공하여 목적에 맞게 활용 가능한

<표 2> 빅데이터 생태계 관련 연구

구분	빅데이터의 개념과 범위	주요 내용	생태계 분석의 시사점
정보통신정책연구원[24]	빅데이터 환경은 민간 기업의 경영활동, 정부를 포함한 공공 부문의 혁신을 수반하는 패러다임의 변화	<ul style="list-style-type: none"> 부가 가치 창출이 가능한 빅데이터 유통 체계의 혁신이 중요함 개인, 기업, 공공 기관으로 개체를 구분, 데이터 공유 개념도를 제시 	<ul style="list-style-type: none"> 빅데이터 생태계 환경의 선순환(유통)을 돕는 촉매자이자 중재자로서의 정책기관의 역할 강조
한국정보화진흥원a[27]	빅데이터를 구성하고 있는 하드웨어, 소프트웨어를 포괄하는 모든 프로세스를 의미하는 거대 플랫폼	<ul style="list-style-type: none"> 생태계 개체를 서비스를 기준으로 서비스 공급자, 빅데이터 사용자, 어플리케이션 공급자로 구분 빅데이터 서비스 프레임워크 기반 산업 분석 	<ul style="list-style-type: none"> 주체 간 상호작용 관점에서 시장 분석 빅데이터 생태계의 이해를 위해 서비스 공급자를 자세히 분석
한국과학기술기획평가원 [22]	빅데이터 개념은 데이터의 가치 활용과 효과의 의미가 확대되어 경제사회적 목적으로 대규모 데이터 기술과 도구를 다루는 것으로 확장	<ul style="list-style-type: none"> 스마트 시대의 스마트 네트워크 환경에 초점을 맞추어 이슈를 분석 데이터 생산자, 데이터 유통자, 데이터 소비자, 플랫폼으로 구분 	<ul style="list-style-type: none"> 개체 간 연결과 주변 환경을 포함한 전체적인 관점의 분석 균형적 생태계 조성을 위한 민·관·연 협력 체계 마련의 시급성 지적

동적자원으로써의 가치에 주목하였다. 특히 빅데이터 생태계에서 가장 중요한 구성요소라고 판단한 서비스 공급자에 초점을 맞추어 분석을 수행하였다. 연구 결과 별개의 영역에 있던 빅데이터 서비스 공급자들이 급격하게 경쟁관계에 놓이게 된다는 점을 발견하였다. 이 연구는 생태계 연구로서 시사점을 가지는 것은 빅데이터 서비스로 생태계의 범위를 확대하여 서비스를 주고받으며 연결하는 주체들 간 상호작용에 초점을 맞추었다는 것이다. 그러나 서비스 공급자에 초점을 맞춘 분석으로 빅데이터 생산자와 이용자로서의 개인 수준을 포함한 산업 전체를 이해하기엔 한계가 있다.

한국과학기술기평가원의 보고서에서는 데이터 생산자, 데이터 유통자, 데이터 소비자 및 솔루션 제공자로 구분하여 빅데이터의 균형적 생태계를 위한 시사점을 제시한다[22]. 보고서에서는 빅데이터가 갖는 속성, 특성을 기반으로 관련 기술 자체의 이해와 고유한 생명주기를 반영한 생태계의 이해가 필요하다고 언급하고 있다. 또한 데이터 생산자, 유통자, 소비자 및 플랫폼(솔루션 제공자)을 토대로 분석을 수행하였다. 결론으로 경제사회적 부작용과 불균형, 안보위협, 해킹 등의 문제를 제기하며 균형적 생태계 조성을 위한 법제도의 중요성을 강조하고 있다. 무엇보다 개인 수준의 데이터 생산자로서의 역할과 정부와 주변 환경을 포함한 전체적인 관점에서 이슈를 분석하였다는 것에 생태계 연구로서 시사점을 가진다. 하지만 빅데이터 생태계 개체 간의 복잡한 상호작용에 대한 분석에 한계점을 가진다.

이 외에도 빅데이터 관련 연구들에서는 빅데이터 활용을 위한 주요 3대 요소인 기술, 자원, 인력의 네트워크 구성과 데이터 공유 및 협력의 필요성과 시급성에 대하여 지적하고 있다[1, 4, 9, 13, 14]. 기존 연구들을 종합해보면 빅데이터 생태계 연구를 위해서는 빅데이터 자체의 흐름과 기술에 대한 이해, 빅데이터 기반 서비스를 통한 상호작용 관계, 균형적 생태계로의 성장을 위한 법/제도 등의 외부환경의 중요성을 확인할 수 있다.

3. 빅데이터 생태계 모델

3.1 생태계 개념의 적용

빅데이터 산업은 데이터의 생성과 분석, 재가공의 과정에서 데이터를 둘러싼 많은 개체가 활동을 하고, 빅데이터 서비스 제공을 위하여 여러 산업과 기업이 상호작용하는 복잡한 구조라는 것을 문헌 연구를 통해 확인하였다. 이는 앞서 살펴본 생태계이론 중, 정보생태계와 지식생태계 모델이 콘텐츠(정보, 지식)와 사람, 기술 등 개체 간 복잡한 네트워크 구조를 가진다는 면에서 유사한 성격을 가진다. 정보생태계와 지식생태계는 정보와 지식의 자연스러운 순환과 그에 따른 생성, 재생산을 위해 개체 간 상호작용, 협력과 연결성, 네트워크 구조적 관계의 중요성을 강조하고 있다. 다시 말해 정보나 지식을 원하는 사람이 정보를 원활하게 탐색하지 못하고 유통과정에 문제가 생기면 생태계가 소멸될 수 있기 때문에, 끊임없이 유용한 정보와 지식서비스를 만들어내야 한다. 이는 기존연구에서 알 수 있듯이 데이터의 원활한 유통과 가치 있는 빅데이터 서비스의 제공이 생태계 활성화에 필수적이라는 점에서 유사함을 확인할 수 있다. 즉 빅데이터 생태계의 활성화를 위하여 기존 생태학적 기본 원리와 정보생태계와 지식생태계의 구조적 특성을 빅데이터 생태계에 확장, 적용할 수 있다는 것을 시사한다.

정보생태계의 핵심은 정보들이 순환되고 새로운 가치를 만들어내는 것으로, 이수상[19]은 연구에서 행위자(개체)들이 생태계의 주체이며 정보 서비스가 주체들이 활동하도록 하는 주 에너지로, 서비스 제공을 위하여 생태계에 동적인 활력이 생기며 정보생태계가 성장하고 진화하게 된다고 하였다. 또한 정보생태계는 전체적으로 복잡한 형태를 나타내지만 구조적으로 개별 개체들이 상호 연결되어 있는 네트워크 형태를 가지는 것이 주요 특징이다. 이는 데이터의 유통과 공유를 통한 빅데이터 서비스 활성화를 추구하는 빅데이터 환경과 유

사하다고 볼 수 있다. 정보생태계에서는 상호 연결의 네트워크 특징, 즉 개체의 연결성에 주목한다. 각 개체들이 만들어내는 연결의 속성을 행위자와 콘텐츠가 생산과 소비로 연결되어 있고 정보 자체를 다른 정보와 관계를 가지는 유기체로 간주한다[11].

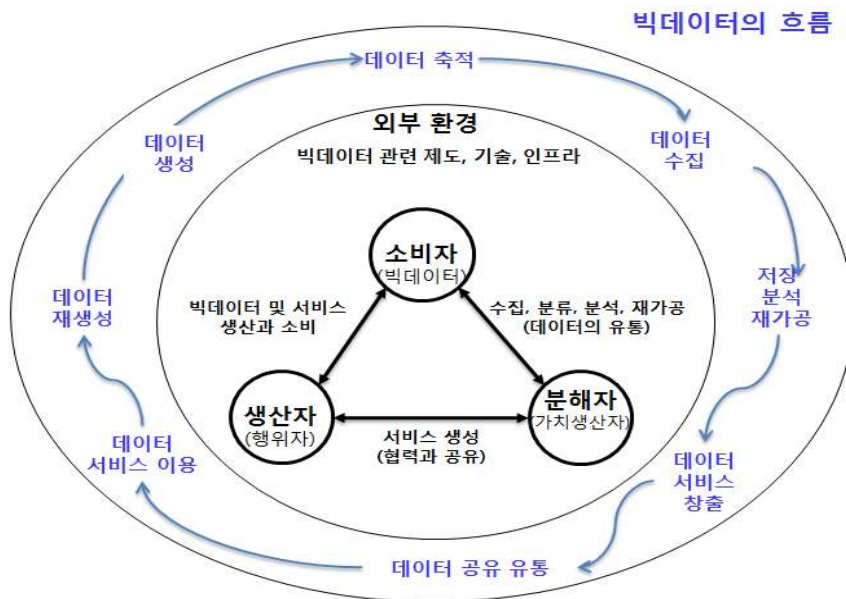
정보생태계의 주제어, 콘텐츠 개체는 분류와 범주 등의 의미에 따라 재분류되는 관계, 행위자와 주제어는 주제적 관심에 따라 새로운 정보서비스를 만들어 내는 것으로 설명한다[19]. 각 개체의 다음과 같은 역할과 관계를 형성한다. 우선 행위자 개체는 콘텐츠와 주제어, 즉 정보 자체를 생산, 관리, 소비하는 역할을 한다. 콘텐츠 개체는 정보와 정보 간의 관계를, 행위자 개체는 콘텐츠를 생산하고 소비하는 주체로서의 속성, 주제어의 생산 활동을 통해 다양한 관계를 만들어내는 관계, 마지막 주제어 개체는 행위자에 의하여 관심 있고 의미 있는 콘텐츠로 새롭게 분류되고 만들어내는 재생산의 속성을 가진다.

지식생태계도 정보생태계와 유사하게 지식창출

과 교류가 잘 이루어지는 네트워크의 구축을 중요하게 여기며, 지식생태계 내의 지식, 전문가, 기술 네트워크의 형성과 네트워크 간 연계의 중요성을 강조한다. 윤정선 외[17]의 연구에서는 지식생태계의 구성 개체를 인적 네트워크, 지식 네트워크, 기술 네트워크로 구분하였으며, 인적 네트워크는 조직 구성원, 이용자, 지식 네트워크는 전문가 등이 만들어내는 생산적 아이디어의 네트워크, 기술 네트워크는 지식의 생성, 공유, 이용을 위한 모든 기술을 포함하는 개념으로 설명하고 있다. 특징적인 것은 기술네트워크 개체를 포함하여 지식의 생성, 공유, 이용을 위한 커뮤니케이션을 위한 모든 기술을 포함한다는 것이다[17]. 이러한 정보생태계와 지식생태계의 구조적 특징과 개체의 분류 특징을 반영하여 빅데이터 산업에 다음과 같이 적용할 수 있다. [그림 3]은 본 연구에서 제안하는 자연생태계 기반의 빅데이터 생태계 구조의 개념도이다.

우선 빅데이터 개체는 정보생태계의 콘텐츠와 유사한 개념으로 빅데이터 자체를 의미한다.

이는 개인, 기업, 공공의 서로 다른 분야에서 생



[그림 3] 생태계 관점의 빅데이터 생태계 개념적 구조

산되고 있는 데이터를 의미함과 동시에 데이터와 데이터 간의 관계를 포함하는 개념이다. 행위자의 경우 빅데이터를 생산하고 소비하는 주체를 포괄하는 개체로 이해할 수 있다. 가치 생산자의 경우, 정보생태계의 주체어 개체와 유사한 개념으로 목적과 상황에 맞는 서비스 제공을 위하여 데이터를 수집, 분석, 하는 일련의 과정에 관련된 개체와 ‘빅데이터 서비스’의 제공, 유통, 중간 단계의 허브와 자연생태계의 분해자 역할을 하는 개체군으로 정의한다. 외부환경 요인의 경우, 정보생태계에서는 디지털 정보 환경, 정보 서비스 이용자 관리자로 구성된 정보 유통 환경 및 인프라로, 지식생태계는 불확실성 증가 환경, 지식 교류를 위한 IT, 디지털 환경, 기술 인프라를 중요한 환경 요인으로 여기고 있다. 빅데이터 경우 중요한 외부 환경 요인으로는 빅데이터 활용과 관련된 정부의 정책, 규제 환경 및 개인의 빅데이터에 대한 인식 등이 있을 수 있으며, 생태계 개체를 둘러싼 다양한 산업, 문화, 기술적 환경을 포함하는 개념이다.

3.2 빅데이터 생태계 구성 개체

본 연구에서는 생태학 원리를 토대로 정보생태계와 지식생태계의 개념을 참조하여 빅데이터 특성을 반영한 생태계 구성 개체를 <표 3>과 같이 분류하였다.

빅데이터 개체군은 각종 분야에서 생산되는 빅데이터 자체를 의미하며, 행위자 개체 안에는 빅데이터의 생산 소비의 주체인 빅데이터 생산자와 빅데이터 서비스 이용자로 구분한다. 가치생산자는 정보생태계의 주체어 개체와 유사한 개념으로 지식 생태계의 기술 네트워크를 포함한 개념으로 정의한다. 따라서 가치 생산자 개체군은 빅데이터 제공자, 빅데이터 서비스 제공자와 빅데이터 유통자로 분류하였다. 빅데이터는 저장, 검색, 관리 중심의 기존 데이터 관리에서 공유, 분석, 추론의 과정으로 확대되었으며, 가치 창출의 원천 요소로서 축적, 활용, 확산 및 공유, 가치 창출의 일련의 과정을 가지고 있다[32]. 따라서 본 연구에서는 이러

<표 3> 빅데이터 생태계 구성 개체와 개념

구조	개체	개념	개체 범위 및 역할
빅데이터	빅데이터	빅데이터 개체군은 공공, 민간, 각종 산업 분야에서 생성되는 빅데이터 자체와 목적에 맞게 재가공된 빅데이터와 데이터 간의 관계를 포괄함	
행위자	빅데이터 생산자	서비스 이용, 기업 활동을 통한 원시 데이터를 생산하는 개인과 기업	<ul style="list-style-type: none"> 개인 : 일상생활 활동 기업 : 단순 비즈니스 활동 정부 : 공공 빅데이터 생산
	빅데이터 서비스 이용자	빅데이터 서비스 이용자, 서비스 이용을 통한 원시 데이터 재생산	<ul style="list-style-type: none"> 개인 : 빅데이터 기반 서비스 이용자 기업 : 빅데이터 서비스 이용자 역할 정부 : 빅데이터 서비스 이용자 역할
가치 생산자	빅데이터 제공자	빅데이터 확보를 위한 수집, 저장, 가공 활동 관련 개체, 자연 생태계의 2차 생산자 역할	<ul style="list-style-type: none"> 기업 : 빅데이터 확보 및 활용을 위한 일련의 프로세스 과정 연관 기업 정부 : 빅데이터 제공자 역할
	빅데이터 서비스 제공자	빅데이터를 분석, 추론 및 빅데이터 서비스 기획, 제공, 빅데이터 유통 활동 등의 과정에 관련된 개체	<ul style="list-style-type: none"> 개인 : 빅데이터 분석 전문가 기업 : 빅데이터 분석, 추론을 통한 서비스 제공 과정의 연관 기업 정부 : 빅데이터 기반 서비스 제공의 역할
	빅데이터 유통자	새로운 가치 창출을 위한 빅데이터 공유, 재가공을 통한 빅데이터와 서비스 재생산 등 자연생태계의 분해자 역할을 수행	<ul style="list-style-type: none"> 개인 : 빅데이터 산업 활성화를 위한 전문가 활동 기업 : 데이터 공유를 통한 빅데이터 재가공, 2차 데이터 서비스 제공자 정부 : 생산자와 소비자의 상호작용을 촉진하는 연결고리의 역할
외부 환경	외부 환경	빅데이터 생태계 개체를 둘러싼 다양한 산업, 문화, 기술적 환경	

한 빅데이터의 순환 과정을 고려하여 다음과 같이 구성 개체의 범위와 상호작용 관계를 정의한다.

우선 빅데이터 개체는 정보생태계의 콘텐츠 개체와 유사한 의미를 가지며, 각종 분야에서 생성되는 빅데이터를 의미하고 행위자와 가치생산자의 목적에 맞게 축적, 개방, 소비, 공유되는 데이터간의 관계를 의미한다.

빅데이터 생태계의 행위자 개체군은 서비스 이용, 기업 활동 등을 통해 대용량의 원시 데이터를 생산해내는 개체들을 의미한다. 핵심 주체의 경우 개인과 기업으로 분류할 수 있는데 개인의 경우 일상생활에서 빅데이터를 생성하는 개체로 예를 들어, 스마트폰을 통해 SNS를 이용하거나 신용카드를 이용하는 등의 활동을 통해 빅데이터를 생산하는 역할을 한다. 기업과 정부의 경우 비즈니스 활동과 공공의 업무 활동을 통해 원시데이터를 생산하는 역할을 하게 된다. 본 연구에서 빅데이터 생산자 개체군의 기업과 정부는 별도의 부가 가치 창출을 위하여 빅데이터를 수집 가공하여 서비스를 제공하는 역할을 하지 않는 것으로 역할을 정의한다. 이러한 기업이 원시 데이터를 분류하고 가공하여 2차 서비스를 제공하는 역할을 하는 경우, 생태계의 가치생산자로 진화할 수 있다.

빅데이터 서비스 이용자는 빅데이터 서비스 이용을 통한 빅데이터 재생산의 역할을 수행한다. 개인과 기업들에 의해 빅데이터 서비스가 이용되면 이를 통해 소비자에 의해 자주 사용되는 데이터가 무엇인지 파악하고, 더욱 가치 있는 빅데이터를 찾아낼 수 있도록 해준다.

빅데이터 가치생산자는 빅데이터 활용을 위한 데이터 수집, 저장, 분석 및 가공의 역할을 수행하는 과정에 연관된 개체를 의미한다. 빅데이터 제공자 개체군의 기업과 정부는 비즈니스 활동 중에 생성되는 대용량의 원시데이터를 목적에 맞게 분류하고 저장하고 가공하며 기업의 경쟁력 확보를 위하여 활용한다. 예를 들어 유통 기업의 상품의 유통과정에서 발생한 빅데이터를 분석하여 유통과 프로세스를 개선하거나, 포털업체가 소셜데이터 분

석을 통해 광고 및 마케팅 활동에 활용할 수 있다. 정부 및 공공기관 또한 공공 업무를 수행하는 과정에서 발생한 데이터를 가공하여 필요로 하는 곳에 공공의 빅데이터와 서비스를 제공하는 경우 이 개체군에 포함된다.

빅데이터 서비스 제공자 개체는 빅데이터 소비자 역할과 새로운 빅데이터 서비스를 제공하는 역할을 동시에 수행하며, 빅데이터를 목적에 맞게 가공, 분석하여 빅데이터 서비스를 제공한다. 이는 정보생태계에서 정보서비스 제공을 위하여 끊임없이 진화하는 것과 유사하게 빅데이터 생태계의 순환 과정을 촉진시키는 개체군으로 해석할 수 있다.

빅데이터 서비스 제공자 개체군의 경우 기업이 자체적으로 원시데이터를 보유하지 않더라도 외부 기업이나 산업의 빅데이터를 가공하여 빅데이터 서비스를 제공하는 기업도 포함한다. 빅데이터 서비스 제공자 개체로서의 개인은 단순한 통계적인 분석을 수행하는 것이 아닌 빅데이터 서비스의 기획 및 분석, 추록 등의 역할을 수행할 수 있는 전문가를 의미한다.

빅데이터 유통자는 데이터 공유, 개방, 순환의 과정을 통해 빅데이터 생태계가 균형을 갖추고 발전할 수 있도록 역할을 수행하는 개체로 자연 생태계에서의 분해자로서의 역할을 나타낸다.

해당 개체군은 개인, 기업, 정부로 구분할 수 있으며 개인의 경우 빅데이터 산업 활성화를 위해 노력하는 전문가, 기업은 빅데이터를 공유하고 새로운 가치 창출을 위하여 생산자, 소비자 영역의 경계를 넘어 협력하고, 새로운 빅데이터와 서비스를 재생산하며 진화할 수 있도록 하는 역할을 한다. 정부는 빅데이터 순환 촉진자로서 생태계 내의 연결고리를 하는 정부기관을 의미한다. 정부의 역할은 빅데이터 생태계의 기반을 형성하고 빅데이터가 공유, 가치 창출을 할 수 있도록 선순환 될 수 있는 인프라를 구축하는 것이다.

본 연구에서 지식생태계의 기술네트워크를 의미하는 개체는 빅데이터 생태계에서 가치 생산자에 포함된다. 빅데이터 활용을 위한 데이터 저장, 처

리, 분석 연관 기업과 기술은 각 기술의 수준과 제공하는 기술 환경에 따라 빅데이터 제공자, 빅데이터 서비스 제공자, 빅데이터 유통자 개체군에 포함된다. 예를 들어 빅데이터 저장 하드웨어 기술과 관련 공급자는 빅데이터 제공자에, 빅데이터 분석을 위한 전문 통계 분석 기술 및 관련 업체는 빅데이터 서비스 제공자 개체군에 포함된다.

빅데이터 생태계의 외부 환경은 빅데이터 생태계 개체를 둘러싼 다양한 산업, 문화, 기술적 환경을 포괄하는 의미로 정보통신 기술 구축 환경, 개인의 라이프스타일, 빅데이터 관련 법/규제 등이 포함된다. 다음 절에서는 빅데이터 생태계의 구성 개체군의 핵심 주체와 이들 간의 상호작용관계와 이슈를 분석하고 시사점을 제시한다.

4. 빅데이터 생태계 시사점

4.1 국내 빅데이터 생태계 분석

본 절에서는 국내 빅데이터 생태계의 구성개체와 관계를 분석한다. 빅데이터 산업은 최근에 본격적으로 이슈화되고 있는 분야로 학술연구 자료가 충분하지 않아 정부 기관의 연구보고서와 최근 매체 기사를 위주로 분석하였다. 이를 바탕으로 빅데이터 생태계 내의 구성개체들을 둘러싼 이슈를 분석하고 시사점을 제시한다.

빅데이터 개체군은 현재 일반 개인, 다양한 산업 분야와 정부 및 공공 기관이 다양한 유형의 대용량 원시데이터가 포함되어 있다. 빅데이터가 발생하는 연관 산업과 주체에 따라 데이터의 형식과 특성은 상이하다. 예를 들어 생성 주체에 따라 컴퓨터 생산 데이터, 사람 생산 데이터, 관계 데이터로 데이터의 유형에 따라 정형, 반정형, 비정형 데이터로 분류된다[28]. 따라서 데이터 자체 간의 관계도 매우 복잡한 구조를 가진다. 예를 들어 국내 의료 분야의 빅데이터는 같은 산업 내의 데이터라도 공공기관, 민간 병원, 개인에 의해 발생하는 데이터는 데이터의 기술적인 특징 뿐 아니라 공유,

재생산 과정에 개인정보 보안 등의 문제로 데이터의 소유권 및 민감성 등과 관련하여 복잡한 관계를 갖는다.

국내 빅데이터 생태계의 행위자 개체는 자연생태계의 1차 생산자 수준에서 형성이 잘 되어 있다. 즉 빅데이터 생산자는 매우 잘 형성이 되어 있는데 수준 높은 정보통신 인프라를 기반으로 개인의 소셜미디어 이용 활동도 매우 활발하며, 기업의 각종 비즈니스 활동 과정에도 디지털화 되어 있는 대용량의 업무 데이터와 소비자 관련 데이터가 발생한다. 실제로 국내 디지털 데이터양은 2006년 2.9억사 바이트에서 연평균 약 57%의 높은 증가 추세를 보이고 있다[33]. 정부 또한 교통, 각종 센서 데이터 등 다양한 대용량 빅데이터를 보유하고 있다. 하지만 아직까지 원활하게 데이터가 공유되지 못하는 한계점을 가지고 있다. 기업 행위자 개체도 데이터 중심의 마인드가 확립되지 않아 일부 산업은 추후 빅데이터 서비스 제공자로 진화하기 위한 발판을 마련하지 못하고 있다. 빅데이터 서비스 이용자 수준의 개체는 아직까지 빅데이터 서비스 산업이 활성화되지 않아 상권분석 서비스, 소셜데이터 분석 기반 추천서비스 이용 수준이다.

국내 빅데이터 제공자 개체는 빅데이터를 자체적으로 보유하고 있는 통신사, 포털을 중심으로 형성이 되어 있다. 즉 국내 빅데이터 생태계는 자연 생태계에 비추어 1차 생산자와 2차 생산자가 균형 잡힌 구조를 형성하지 못하고 있다. 자연 생태계의 소비자(2차 생산자)인 빅데이터 제공자의 핵심주체인 빅데이터 생산보유 기업의 참여는 점점 증가하고 있지만 그 외 빅데이터 제공 과정의 경쟁력을 확보하고 있는 국내 기업이 거의 없는 실정이다. 현재 하드웨어, 소프트웨어, 플랫폼 기술 등이 포함되는 국내 빅데이터 제공자 개체군은 주요 글로벌 IT기업이 주도권을 가지고 있다. 대부분의 분석 솔루션의 외국 기업의 기술이며, IBM, 오라클, SAP, HP, 구글 등 주요 글로벌 기업들이 빅데이터 시장 전반에 경쟁력을 확보하기 위하여 기업을 인수하는 등 다각도로 준비하고 있

다. 이에 대응하기 위하여 국내 기업의 육성과 경쟁력 확보를 위해 노력하고 있으나 아직 많이 부족한 실정이다. 종합하면 빅데이터를 자체 생산하여 활용하는 기업은 다양한 반면에 빅데이터 제공 환경 구축 활동 즉 빅데이터 기술과 연관된 주체들은 전무한 실정이다.

빅데이터 서비스 제공자 개체군에 최근 국내 주요 대기업이 적극적으로 참여하고 있으며, 특히 통신 사업자를 중심으로 움직임이 활발하다. LG유플러스는 자사 가입자들이 사용하는 소셜네트워크 서비스 ‘와글’을 통해 정보를 수집하고 분석하여 고객 맞춤형 서비스를 제공하는 것을 목표로 하고 있다. KT도 고객센터에 접수된 정보와 요금청구 정보, 통화, 무선데이터, 애플리케이션 이용 기록을 분석하여 사업 영역을 확장하겠다는 계획을 가

지고 있다[12]. 이들 통신 기업은 타 분야 빅데이터 기업과 협력을 통하여 새로운 서비스제공을 위한 전략을 펼치고 있다.

대표적인 통신 사업자 KT는 신용정보 전문회사인 코리아크레딧뷰로와 제휴를 통해 빅데이터를 결합해 서비스를 출시하기 위한 움직임을 보이고 있다. KT의 지역별 유동인구 정보, 통화량 데이터와 코리아크레딧뷰로의 금융통계정보 결합을 통하여 새로운 서비스의 출시와 가치 창출이 가능한 것이다. 이처럼 최근 빅데이터 서비스 제공자 개체군에 속한 기업들은 서로 다른 분야에 있던 기업 간 제휴와 협력 관계를 통해 경쟁력을 확보하기 위하여 노력하고 있다[25]. SKT의 경우, 가입자들 위치로 유동인구나 상권 분석을 한 뒤 다른 영역의 사업자들이 서비스를 이용 할 수 있는 정

〈표 4〉 빅데이터 생태계 핵심 주체별 상호작용

세부 개체	핵심 주체 예시	데이터 처리 과정	플레이어 및 역할 예시
빅데이터 생산자	일반 개인	개인 빅데이터 생산	소셜미디어 이용자
	비즈니스 활동 기업	산업의 빅데이터 생산	제조 전문 업체, 유통 기업 등
	정부, 공공 기관, 지자체	공공분야의 데이터 생산	공공 업무 발생 데이터
빅데이터 서비스 이용자	서비스를 이용하는 개인	빅데이터 서비스 이용	소셜 분석 추천 서비스 이용
	서비스를 이용하는 기업	빅데이터 재생산	상권분석 서비스 이용 기업
빅데이터 제공자	빅데이터 생산 기업	빅데이터 보유(축적, 활용), 빅데이터분석 활용	통신사, 은행, 포털
	H/W, DB 장비 및 솔루션 기업	데이터 저장	EMC, IBM, 오라클, HP, SAP
	플랫폼 제공기업	데이터 수집, 저장, 처리	구글, 아파치, MS
	클라우드 서비스 기업	가공환경 제공	KT 클라우드, 델, 애플
	빅데이터 분석 과정 관련기업	데이터 분석 환경 제공	R, 넥스알, BI, DW 관련 기업
서비스 제공자	정부, 공공 기관, 지자체	공공 분야의 빅데이터 제공	가공된 공공 빅데이터 공개
	빅데이터 분석 전문가	빅데이터 서비스 기획, 분석	데이터 사이언티스트
	빅데이터 서비스 제공 기업	빅데이터 가공, 공유를 통한 새로운 서비스 및 가치창출	빅데이터 서비스 전문기업
	빅데이터 분석서비스 전문업체	빅데이터 분석, 추론	소셜 분석 전문기업 (그루터, 트윗믹스 등)
데이터 유통자	정부, 공공기관, 지자체	빅데이터 기반 공공 서비스	실시간 버스 도착 서비스
	빅데이터 관련 전문가(개인)	빅데이터 공유, 순환 빅데이터 서비스 재생산	전문가 포럼, 커뮤니티 활동
	2차 빅데이터 제공 기업		빅데이터 컨소시엄 구성
	정부 : 빅데이터 순환 촉진자 공공기관 및 지자체		빅데이터 전문 연구 기관 빅데이터 전담 부서

보를 판매한다.

아직까지 국내 빅데이터 서비스 시장이 성숙되지 않아 '빅데이터 서비스 제공'을 전문적으로 하는 기업은 많지 않다. 현재 서비스를 제공하는 기업은 빅데이터 생태계의 생산자 역할을 동시에 하는 대기업들로 자체적인 원시데이터를 보유하고 있다. 즉 아직까지 빅데이터(원시 데이터)를 보유한 기업 위주로 빅데이터 서비스 제공자로 진화하고 있으며, 추후 기업 간 데이터 불균형에 대한 염려가 제기된다. 빅데이터 가치생산자 개체군의 국내 빅데이터 분석 전문가, 정부, 공공기관 등의 핵심 주체는 향후 더욱 육성이 되어야할 것으로 분석되었다. 빅데이터 서비스 제공자 개체에 해당하는 분석 전문 인력은 데이터 서비스를 기획하고 데이터 성격과 산업을 이해하는 능력을 갖춘 빅데이터 사이언티스트를 의미한다. 국내의 경우 빅데이터 시장을 주도할 만큼 양질의 전문 인력이나 연구 역량을 확보하지 못하고 있다[1]. 정부 또한 데이터 통합의 문제, 전문 인력의 부족 등 다양한 이유로 빅데이터 기반의 전문적인 대국민 서비스는 제공하지 못하고 있다.

마지막 선순환 생태계 구조의 핵심적인 역할은 생태계내의 분해자가 수행하게 된다. 제안된 빅데이터 생태계 개체 중, 빅데이터 유통자에 해당한다. 특히 해당 개체군 내에서는 분해자로서의 정부 역할이 중요하다는 것은 기존 연구에도 언급되었다[24]. 빅데이터를 통한 새로운 가치 발굴을 위해서 폐쇄적인 산업 영역 간, 기업 간의 경계를 낮추어 데이터의 자유로운 공유와 융합이 필요한데 이러한 문제를 해결하기 위한 정부의 역할이 강조된다[18]. 이를 위해 우선 정부가 생산하는 데이터를 개방하고 공유하기 위한 방안을 수립하고 앞장서야 한다. 또한 개인 이용자와 기업 이용자는 빅데이터 서비스를 적극 이용하고 의견을 제시함으로써 빅데이터 서비스의 진화와 생태계 구축에 도움이 될 수 있다.

이러한 생태계 분해자의 성공적인 정착은 데이터의 유통과 공유, 협력에 기반을 둔 서비스 제공

을 가능하게 하며 이는 결국 수준 높은 빅데이터의 생산, 즉 데이터 품질의 제고와 다시 연결되어 선순환 구조를 형성할 수 있게 된다.

4.2 국내 빅데이터 생태계 시사점

빅데이터 생태계는 빅데이터가 1차 생산되고 재생산되고 재분류 되는 과정에서 새로운 의미가 부여되며 순환하는 과정을 거쳐 발전할 수 있다. 그러나 이러한 일련의 상호작용 과정에서 다음과 같은 이슈가 제기되고 있다.

우선 개인의 빅데이터 생산과 산업의 빅데이터 생산 과정에서 제기되는 이슈이다. 정보 활용에 대한 국민 인식에 의한 데이터 활용에 제한이 있을 수 있다. 현재 빅데이터와 빅데이터 생산자 수준의 국내 빅데이터 생태계가 잘 형성되어 있지만 최근 빅데이터가 '빅 브라더'를 연상하게 한다는 문제의식이 제기 되고 있다. 즉 데이터가 생성되고 순환되는 과정에 개인 정보화 사생활이 유출될 수 있다는 것이다. 그에 따라 정부의 데이터 공개를 최소화 하고 개인 정보를 비즈니스에서 재활용한다는 것에 대한 거부감이 증가되는 문제가 제기되고 있다. 이러한 부정적인 인식은 빅데이터 생태계의 기반이 되는 빅데이터 생산자의 형성을 저해할 수 있다.

두 번째는 기업과 정부 수준의 주요 이슈이다. 우선 기업의 데이터 활용 목표가 확립되지 않아 데이터 생성 과정에서 빅데이터의 가치를 창출하기 위한 체계적인 저장기 이루어지지 못하고 있다. 빅데이터 생태계의 생산자 수준에서는 다양한 데이터가 생성되고 있지만 빅데이터 제공자로의 진화를 위하여 이미 생산되어 있는 다양한 분야의 빅데이터들의 통합과 분석, 공유가 쉽지 않다. 예를 들어 공공 데이터는 원시데이터 자체가 부처별, 기관 별로 생산 주기와 생산 형태가 상이하여 데이터 통합의 어려움과 이슈가 제기 된다. 기업과 정부는 장기적인 관점에서 빅데이터 생태계 내의 빅데이터 제공자에서 서비스 제공자로 진화하

기 위해서 데이터 활용 목표를 분명히 해야 한다.

세 번째 이슈는 데이터 저장, 수집, 가공의 상호 작용 과정 중에 제기되는 문제이다. 빅데이터가 가지는 가장 큰 특징은 원시데이터 저장 형태가 상이하다는 점이다. 이 외에도 데이터 구조, 수집 주기 등이 다르기 때문에 데이터를 저장, 수집, 가공하는 과정에 많은 문제가 야기되고 있다. 이러한 문제는 빅데이터 제공자 개체의 빅데이터 서비스 제공자로의 진화에 장애요인이 될 수 있다. 전문적인 서비스 제공과 가치 창출을 위해서는 전혀 다른 분야의 상이한 데이터들이 자유롭게 융합될 수 있어야 한다. 현재 국내 빅데이터 생태계에서 가장 먼저 협력해서 해결해야 할 문제이다.

다음은 데이터 분석 환경 제공 및 분석 추론의 단계에 제기되는 이슈이다. 빅데이터 관련 자체 기술과 전문 인력 부족 문제이다. 앞서 언급한 바와 같이 국내 생태계 내에 가장 부족한 개체군이다. 빅데이터 제공자 개체군의 가장 기초가 되는 핵심 주체로 전문 인력 양성을 위하여 학계, 정부, 기업이 함께 노력해야 할 것이다.

넷째, 빅데이터를 공유하고 활용하는 관계에서 나타나는 데이터 품질에 대한 신뢰와 데이터 소유권 이슈이다. 우선 빅데이터가 유통되는 과정에서 불확실한 정보가 포함되는 경우 상호 간 신뢰가 확보되지 못하고 데이터 유통자 개체군의 성장이 저해될 수 있다. 향후 빅데이터 생태계의 가치생산자 개체군이 성장하고, 특히 빅데이터 서비스 제공자가 활성화되기 위해서는 빅데이터 품질 확보를 위해서 상호간 노력해야 한다. 이와 유사하게 빅데이터 기반 서비스 제공, 재생산, 데이터 순환 등의 전 과정에서 제기될 이슈로는 데이터 소유권 대한 이슈가 존재한다. 아직 국내에서 본격적으로 활성화되지 않은 서비스 시장이지만, 향후 빅데이터 서비스가 활성화 되고 2차, 3차로 가공된 데이터가 유통되는 산업이 형성된다면 데이터 소유권 분쟁이 발생할 수 있다[13]. 이러한 문제점을 미연에 방지하기 위하여 개인, 기업의 빅데이터 생산자 개체와의 협의가 먼저 이루어져야 한다.

결론적으로 국내 빅데이터 생태계의 체계적인 구축을 위해서는 현재 빅데이터 관련 개인, 기업, 정부가 각각 어느 개체에 속하는지 수행해야 할 역할을 인지하고, 상호 간 협력하고 해결해야 할 이슈가 무엇인지에 대하여 논의되어야 한다.

5. 결 론

본 연구는 최근 관심이 증가하고 있는 빅데이터 생태계 분석을 통하여 향후 빅데이터의 성공적인 활용을 위한 시사점을 제시하였다. 이를 위해 사회과학 분야의 생태계 이론을 검토하고 이를 기반으로 한 빅데이터 생태계 구조를 분석하였다. 특히 빅데이터 환경과 유사한 구조적 특징을 가지는 정보생태계와 지식생태계의 관점을 반영하였다. 이를 통해 빅데이터 산업 내 개별 개체의 역할과 데이터가 어떻게 상호작용하는지를 논의하였으며, 연구의 결과 및 시사점은 다음과 같다.

첫째, 빅데이터 생태계 활성화를 위해서는 빅데이터를 통한 가치 창출의 목표를 명확히 하고 현재 생태계내의 역할을 인지하고 전략을 수립해야 한다. 또한 전략 수립 과정에서 일시적인 데이터의 활용이 아닌 장기적인 관점에서 가치 창출을 위해서는 공유, 협력을 통한 빅데이터 생태계 인프라와 기반이 선결되어야 한다는 점을 인식해야 한다. 이를 위해 기업 간, 기업과 정부, 정부 부처 간 데이터 공유가 이루어져야 한다. 이는 현재 빅데이터 개체군에 속한 데이터 간의 관계 향상과 가치생산자 개체군의 형성에도 긍정적인 영향을 줄 수 있다. 강만모 외[1]의 연구에서는 빅데이터가 활성화 되는 경우 기업 간 양극화 현상에 대한 우려를 나타내며, 빅데이터를 보유한 기업이 독점적으로 활용하게 될 가능성이 높다고 언급하였다. 즉 데이터의 생산과 공유 과정에 각각의 여러 가지 이해관계와 이유로 데이터가 원활히 공유되지 못하면 생태계 형성에 부정적인 영향을 끼칠 것이다. 따라서 장기적인 관점에서는 공동의 가치창출을 목표로 하고 데이터 공유와 순환을 위한 노력

이 필요하다는 것을 인식하고 데이터 표준화를 위하여 노력한다면 생태계 내의 균형 잡힌 성장이 가능할 것이다.

둘째, 빠르게 증가하고 복잡한 구조적 특성을 가지는 빅데이터를 분석하고 활용할 수 있는 환경 구축을 위한 자연생태계에서의 소비자, 분해자로서의 능력을 강화해야 한다. 현재 외국에 비하여 국내 분석, 저장 자체 기술 및 솔루션은 부족하다. 생태계 분석 결과 빅데이터 제공자 개체 중 빅데이터 제공 환경 관련 핵심주체는 주요 글로벌 기업 위주로 형성되어 있으며 국내 빅데이터 생태계의 가장 취약한 부분으로 나타났다. 또한 가치생산자 개체군, 그 중에서도 빅데이터 서비스 제공자와 빅데이터 유통자가 가장 취약한 것으로 나타났다. 빅데이터의 가치 창출에 가장 핵심적인 역할을 할 개체인 만큼 빅데이터 생태계 활성화를 위하여 가장 육성되어야 할 개체군이다. 아직까지 한국이 원천 기술 부분에서의 경쟁력이 뒤처지고 있지만 뛰어난 디지털 인프라와 정부 주도의 적극적인 육성 전략을 통해 향후 글로벌경쟁력을 확보할 수 있기를 기대한다.

마지막 분해자로서의 정부, 외부 환경으로서의 정부 역할을 균형 있게 해야 한다는 점이다. 정부는 현재 빅데이터 생태계 내에서 매우 복합적인 역할을 해야 한다. 기존 사회과학분야에서의 정부는 외부환경으로서의 역할만 관계되었지만 빅데이터 생태계에서는 자연생태계의 생산자, 소비자, 분해자, 외부 환경으로서의 개체에 모두 포함되고 있다. 공공 정보의 개방을 통한 데이터 소비자로서의 역할을 수행해야 하며 동시에 데이터 생산자와 소비자가 잘 연결될 수 있는 분해자의 역할을 해야 한다. 무엇보다 빅데이터 생태계가 활성화될 수 있도록 기술, 인프라, 정책 제도 등의 외부 환경을 잘 조성하기 위한 역할을 수행해야 한다. 정보통신정책연구원[24]의 보고서에서도 정책 기관의 빅데이터 생태계의 선순환을 돕는 '촉매자'로서의 역할의 중요성을 지적하였다.

현재 국내 빅데이터 산업은 상생을 위해 체계적인 빅데이터 생태계를 구축하기 위한 초기 단계이다. 세계는 지금 데이터를 통한 가치창출과 경쟁력 확보를 위한 빅데이터 경쟁이 한층 더 치열해지고 있다. 따라서 본 연구에서 제안한 생태계 관점의 빅데이터 구조 분석은 성공적인 빅데이터 활용을 위한 산업, 개인, 기업, 정부 수준에서의 종합적인 이해를 도와 건전한 빅데이터 생태계의 형성에 시사점을 제공할 것으로 기대된다. 본 연구 과정의 한계점과 향후 연구 방향은 다음과 같다.

첫째, 본 연구에서 제안한 빅데이터 생태계 구성은 문헌 연구를 토대로 하였다. 이는 국내 빅데이터 생태계만의 특징을 엄밀하게 반영하지 못한 한계가 있다. 후속 연구에서는 국내 빅데이터 생태계 개체군 범위와 역할에 대한 연구가 필요하다. 또한 상이한 산업에 속한 빅데이터 기업 사례 분석을 통하여 개체군의 역할과 범위에 대하여 재분류하고 빅데이터 생태계에 포함되어야 할 추가 요인에 대하여 연구할 필요가 있을 것이다.

둘째, 본 연구에서 제안한 빅데이터 생태계 개체 분류를 토대로 개별 개체별 후속 연구가 필요하다. 본 연구는 전체 빅데이터 생태계 구조와 개체를 살펴보고자 하였기에 각 개체별로 구체적인 역할의 도출, 상호작용 과정에서의 이해관계 분석이 부족하다. 향후 연구에서는 설문 및 데이터의 분석 등을 통하여 빅데이터 흐름과 처리 과정에서의 개체의 역할 및 이해관계에 관한 연구가 의미 있다고 생각한다.

또한 본 연구에서는 빅데이터 개체 규명과 생태계 모델을 제시하는 것에 초점을 맞추어 외부 환경에 대한 분석이 충분히 이루어지지 못하였다. 빅데이터 생태계 내 개체 간 상호작용에 영향을 미치는 외부 환경에 대한 별도의 분석 연구가 필요할 것이다. 특히 본 연구 결과 정부의 역할이 매우 중요하게 나타남에 따라 빅데이터 생태계 내의 정부 역할에 대한 체계적인 연구를 진행해 볼 필요가 있다.

참 고 문 헌

- [1] 강만모, 김상락, 박상무, “빅데이터의 분석과 활용”, 『정보과학회지』, 제30권 제6호(2012), pp.25-32.
- [2] 고정민, 김정호, “벤처 생태계의 형성과 진화”, 삼성경제연구소, 『CEO 인포메이션』, 제240권(2000).
- [3] 박정호, 조지연, 이용석, 이봉규, “새로운 통신시장 활성화를 위한 모바일 생태계 통신정책”, 한국인터넷정보학회, 『한국인터넷정보학회 논문지』, 제12권 제4호(2011), pp.93-106.
- [4] 국가정보화전략위원회, “빅데이터를 활용한 스마트 정부 구현(안)”, 2011.
- [5] 김병일, 신현문, “클라우드 컴퓨팅 생태계 및 정책 방향”, 『전자통신동향분석』, 제27권, 제2호(2012), pp.137-148.
- [6] 김상락, 강만모, 박상무, “빅데이터가 여는 미래의 세상”, 『정보과학회지』, 제30권, 제6호(2012), pp.18-24.
- [7] 김성벽, “웹 미디어 경험과 커뮤니케이션 능력 (Communication Competence)에 대한연구 : 미디어생태학의 연구 관점을 중심으로”, 『중앙대학교 신문방송학과 박사학위 논문』, 2002.
- [8] 김유석, “미디어생태계의 ‘평형상태’를 위하여 -‘분해자’로서의 소비자에 대한 소고”, 『한국방송학회 봄철정기학술대회』, (2002), pp.3-18.
- [9] 김현곤, “빅데이터를 활용한 새로운 정보화”, 『Telecommunication Review』, 제22권 제3호(2012), pp.337-352.
- [10] 류준호, 윤승금, “생태계 관점에서의 문화콘텐츠 산업 구성 및 구조”, 『한국콘텐츠학회 논문지』, 제10권 제4호(2010), pp.327-339.
- [11] 손미, “정보생태계의 변화와 정보리더십 재조명 : 자원 기반 학습 환경 설계에의 함의”, 『교육정보미디어연구』, 제15권, 제4호(2009), pp.231-249.
- [12] 아시아 경제, “통신 3사, ‘빅데이터 사업’ 본격 시동”, available@ <http://www.asiae.co.kr/news/print.htm?idxno=2012081609530766914&udt=1>, 2012.
- [13] 양창준, “미래의 창, 빅데이터”, 『TTA Journal』, 제140권(2012), pp.16-23.
- [14] 손민선, 문병순, “빅 데이터 시대의 한국, 갈라파고스 되지 않으려면”, 동향/연구보고서, LG경제연구원, 2012.
- [15] 유영만, “지식생태학 : 지식기반사회를 위한 포스트 지식경영”, 『삼성경제연구소』, 2006.
- [16] 유재미, 오철호, “지식생태계 연구 경향, 한계 그리고 제언”, 『정보화정책』, 제18권, 제4호(2011), pp.3-21.
- [17] 윤정선, 한선화, 공수연, 이영훈, 백명수, 정재환, “인터넷 기반 지식생태계의 블루오션 전략”, 『한국기술혁신학회 학술대회』, (2005), pp.541-546.
- [18] 이봉규, “빅데이터로 하나되는 세상 만들기”, SBS 서울디지털포럼, 발표자료, 2012.
- [19] 이수상, “정보생태계 관점에서 본 도서관 2.0 서비스의 연구”, 『한국문헌정보학회지』, 제42권, 제2호(2008), pp.29-49.
- [20] 이기식, 오철호, “소프트웨어산업의 생태계에 관한 탐색적 연구 : 위기관인도출을 중심으로”, 『한국정책과 학회회보』, 제16권, 제1호(2012), pp.97-125.
- [21] 이지식, 권중욱, “기업생태계 관점에서의 한중 FDI 정책비교 연구”, 한국관세학회, 『관세학회지』, 제11권 제2호(2010), pp.259-277.
- [22] 전승수, “초연결 사회의 빅데이터 생태계 분석과 시사점”, 『한국과학기술기획평가원 이슈페이지』, 2012-10(2012), pp.1-36.
- [23] 전자신문, “내년 비즈니스 IT변화 핵심은 ... 모바일-클라우드 대세, 아웃소싱 타격, 생산기지 미주지역 이전 가속, 소셜 투자거품 꺼지고 빅데이터 활용실패 속출”, available@ http://www.etnews.com/news/computing/informatization/2526333_1475.html, 2011.

- [24] 정용찬, “빅데이터 혁명과 미디어 정책이슈”, 『KISDI Premium Report』, 2012-02(2012).
- [25] 중앙일보, “KT-코리아크레딧뷰로, 빅데이터 활용 확대 나선다”, available@ http://article.joinsmsn.com/news/article/article.asp?total_id=9437453, 2012.
- [26] 채승병, 양재석, 김선빈, “정책지식 생태계 활성화 전략의 행위자기반 접근”, 『제1회 복잡계 컨퍼런스』, 2006.
- [27] 한국정보화진흥원a, “빅데이터 시대 : 에코시스템을 둘러싼 시장경쟁과 전략분석”, 『빅데이터 전략연구센터 IT&Future Strategy』, 제4권(2012).
- [28] 한국정보화진흥원b, “빅데이터 시대의 데이터 자원 확보와 품질 관리 방안”, 『빅데이터 전략연구센터 IT&Future Strategy』, 제5호(2012).
- [29] 한국정보화진흥원c, “빅데이터 시대 : 공공부문 빅데이터 추진방향”, 『빅데이터 전략연구센터 IT&Future Strategy』, 제6호(2012).
- [30] 한국정보화진흥원d, “스마트 시대의 미래 변화 전망과 IT 대응 전략”, 2012.
- [31] 한국정보화진흥원a, “빅데이터의 개념 및 활용 가치”, 『월간 IT동향』, 제1권(2011).
- [32] 한국정보화진흥원b, “新 가치창출 엔진, 빅데이터 새로운 가능성과 대응 전략”, 『빅데이터 전략연구센터 IT&Future Strategy』, 제18권(2011).
- [33] 현대경제연구원, “현안과 과제 : 빅데이터의 생성과 새로운 사업 기회 창출”, 2012.
- [34] Brown, B., M. Chui, and J. Manyika, “Are you ready for the era of ‘big data?’”, *McKinsey Quarterly*, 2011.
- [35] Gantz, J. and D. Reinsel, “Extracting Value from Chaos”, *IDC IVIEW*, 2011.
- [36] McKinsey Global Institute, *Big data : The next frontier for innovation, competition, and productivity*, McKinsey and Company, 2011.

◆ 저 자 소 개 ◆

**조 지 연** (jy.cho@yonsei.ac.kr)

현재 연세대학교 정보대학원 박사과정에 재학 중이며, 정보통신·미디어 산업/정책을 전공하고 있다. 단국대학교에서 경영학, 경영정보학 학사, 연세대학교 정보대학원에서 정보시스템 석사 학위를 취득하였고, 주요 관심분야는 빅데이터, 컨버전스 전략, 방송통신융합정책 등이다.

**김 예 진** (lucky8619@yonsei.ac.kr)

연세대학교 정보대학원에서 정보시스템 석사 학위를 취득하였으며, 현재 연세대학교 일반대학원 기술경영학협동과정에서 박사과정에 있다. 학사는 중앙대학교 신문방송학과에서 취득하였다. 주요 관심분야는 IT경영, 방송통신융합정책 등이다.

**박 건 철** (parkkc@yonsei.ac.kr)

연세대학교 정보대학원에서 정보시스템 석사 학위를 취득하였으며, 현재 동대학원 박사과정에 있다. 또한 연세대학교 방송통신정책연구소에서 연구원으로 재직 중이며, 주요 관심분야는 ICT 정책, 교육정보화정책 등이다.

**이 봉 규** (bglee@yonsei.ac.kr)

현재 연세대학교 정보대학원 교수, 방송통신정책연구센터 소장으로 재직 중이며, 연세대학교에서 경제학사학, Cornell University에서 CRP 석사, 박사학위를 취득하였다. 주요 경력으로는 한성대학교 정보전산학부로 교수로 재직하였으며, 정보통신부 통신위원회 위원, 방송통신위원회 전과정책임의 위원 등을 역임하였다. 주요 관심분야는 방송통신융합과 ICT 정책 등이다.