

공공 영역에서의 빅데이터 활용 방안 연구 - 델파이 분석을 중심으로

김은진* · 박성욱** · 이용호*** · 정도범****

I. 서론

빅데이터가 2012년 세계경제포럼에서 ‘떠오르는 10대 기술’ 중 첫 번째로 선정된 이후, 2013년에는 정치·경제·사회 등 모든 분야에서 빅데이터 열풍이 전 세계적으로 거세게 확산되었다. 동시에 가트너 그룹의 기대 곡선(Hype cycle)에 따르면 2013년 초에 기술기대의 정점에 다다랐으며, 2014년 말에는 과장과 환상이 소멸되어 가는 환멸 단계에 진입한 것으로 평가되고 있다(Henschen, 2013). 환멸 단계의 기술이 성장 단계로 빠르게 진입하기 위해서는 해당 기술이 다른 영역과의 융합을 통해 동반 성장할 수 있도록 동기를 부여하는 것이 중요하다. 특히 빅데이터 활용을 통한 새로운 부가 가치 창출은 이종 영역 간의 데이터 공유를 통해 극대화될 수 있기 때문에, 데이터 생산자 및 소유자, 활용자 간의 상호작용과 선순환적인 발전이 이루어질 수 있는 생태계 조성이 필요하다.

빅데이터 생태계가 스스로 강화될 수 있는 자생력을 확보하기까지는 해당 기술 및 시장의 성숙을 촉진시킬 수 있도록 정부를 포함한 공공 영역의 의지와 지원 정책이 필요하다. 공공 영역에서의 빅데이터 활용 성공 사례가 도출되고 이와 관련되어 개발된 기술, 인프라, 정책 등이 민간 영역에 확산 및 적용됨으로서 빅데이터 생태계가 자생적으로 발전할 수 있는 기반이 마련될 수 있기 때문이다.

공공 영역에서의 빅데이터 활용은 민간에서의 비즈니스 수익 모델 관점이 아니라 글로벌 현안 및 사회 문제 해결에 초점을 맞추므로써 그 파급효과를 극대화 할 수 있다. 본 연구에서는 빅데이터 시장의 성장 단계로의 빠른 진입과 자생적 발전을 위한 공공 영역의 역할을 규명하기 위해 사회적 수요를 기반으로 주요 이슈 및 추진 과제를 도출하는 것을 목적으로 한다. 구체적으로 정책, 기술, 과학기술 빅데이터 활용, 사회 현안 해결 부문에서의 주요 이슈를 선정하고, 델파이 분석을 통해 전문가의 객관적 시각에 기반하여 이슈들의 우선 순위를 분석하였다.

본 연구는 정책자가 아닌 산·학·연 전문가의 시각을 반영함으로써 기존에 정부가 추진하고 있는 빅데이터 활용 강화 정책 또는 국정과제가 실질적인 사회 수요를 반영하고 있는지를 검증하고, 향후 정책 수립 및 신규 프로젝트 발굴의 기초자료로 활용할 수 있을 것으로 기대된다.

II. 델파이 분석

공공 영역의 역할에 집중하여 빅데이터 활성화 방안을 마련함에 있어 객관적 관점에서의 수요 발굴을 위해

* 김은진, 한국과학기술정보연구원, kimej@kisti.re.kr

** 박성욱, 한국과학기술정보연구원, supark@kisti.re.kr

*** 이용호, 한국과학기술정보연구원, stylee@kisti.re.kr

**** 정도범, 한국과학기술정보연구원, dbchung@kisti.re.kr

전문가 대상의 델파이 분석을 3회 수행하였다.

빅데이터와 관련된 보고서, 논문, 세미나 발표 및 관련 전문가들의 추천을 통해 전문성과 역량이 공식적으로 확인된 전문가 풀을 구성하였으며, 최종적으로 델파이 분석에 참여할 의사를 밝힌 20명을 분석 대상으로 선정하였다.

문헌 연구를 통해 도출된 부분별(정책, 기술, 과학기술 빅데이터 활용, 사회 현안 해결) 주요 이슈에 대한 중요성 및 시급성을 조사하고, 이를 기반으로 해결과제에 대한 우선순위를 도출하였다.

정책 측면에서는, 민간에서의 빅데이터 활용 기회 확대와 분석 역량 강화를 위해 해결되어야 할 문제들이 주요 과제로 제시되었다. 구체적으로 ① 공공데이터의 개방·활용, ② 개인정보 보호·관리, ③ 교육 및 전문가 양성이 핵심 이슈로 도출되었다.

[표 1] 정책 측면 - 빅데이터 주요 이슈의 중요성 및 시급성

관련 이슈		중요성 평균	시급성 평균
1	공공기관 보유 빅데이터의 개방·활용을 위한 정책 수립	4.63	4.53
2	빅데이터 저장·공동 활용 인프라 구축 및 제공 방안 수립	3.47	3.37
3	데이터 표준화 관련 정책 수립	3.79	3.68
4	데이터 품질 인증 관련 정책 수립	3.21	2.95
5	빅데이터 활용 사회 현안 해결 방안 수립	3.84	3.84
6	국가 차원의 전담 기구 설치 및 운영	3.63	3.47
7	빅데이터 산업 육성 정책 수립	4.00	3.79
8	지식재산권 보호와 관련된 정책 수립 (데이터의 재산 가치 보호)	3.47	3.11
9	빅데이터 활용에 따른 개인정보 보호 및 관리	4.32	4.16
10	교육 및 전문가 양성	4.11	4.26
11	데이터 분석 역량 지원 및 공유	3.79	3.79
12	중소기업의 빅데이터 활용 지원 정책 수립	3.58	3.21
13	공공데이터와 민간 데이터의 융합을 위한 정책 수립	4.16	3.68

기술 측면에서는, 수집·저장 부문에서의 실시간 데이터 수집, 처리·분석 부문에서의 데이터 분석, 서비스 부문에서의 가시화, 보안 부문에서의 개인정보 처리 및 보안기술이 핵심 이슈로 도출되었다.

[표 2] 기술 측면 - 빅데이터 주요 이슈의 중요성 및 시급성

기술 분야		중요성 평균	시급성 평균
1	실시간 데이터 수집	4.11	3.95
2	대용량 데이터 저장	3.74	3.63
3	비정형 데이터 처리	4.11	3.68
4	분산 데이터 처리	3.68	3.42
5	데이터 분석	4.84	4.79
6	데이터 전송	3.11	3.00
7	가시화	4.26	4.05
8	데이터 공유 환경	3.58	3.47
9	개인정보 처리 및 보안기술	4.21	4.21
10	저전력 고성능 컴퓨팅	3.28	3.06
11	빅데이터 클라우드 기술 및 플랫폼 개발	3.84	3.79
12	데이터 품질관리	4.21	3.58
13	데이터 표준화	4.00	3.74

과학기술 빅데이터 활용 측면에서는, 의료와 재난대응 부문에 대한 사회적 수요가 가장 높은 것으로 분석되었다. 1차 분석에서 제시된 총 18개의 활용 분야 중에서 13개 분야가 주요한 것으로 분석되었으며, 이중 2차 분석을 통해 8개 분야가 정제되었다. 최종적으로 3차 분석에서 분야별 주요 이슈에 대한 중요성과 시급성을 분석하였다.

[표 3] 활용 측면 - 주요 분야 및 핵심 이슈

주요 활용 분야	핵심 이슈	중요성 평균	시급성 평균
의료 (1순위)	1 국가 주요 질병 발병 및 발병 원인, 전파 경로 등 예측	4.32	4.11
	2 개인별 맞춤형 질병 예측	4.21	3.74
	3 의료 개인 정보 보호	4.05	3.68
	4 의료법 규제 완화	3.61	3.61
	5 의료데이터와 빅데이터 분석시스템과의 연계	4.32	4.21
치안/안보 (3순위)	1 치안 관련 데이터 실시간 수집 및 분석	4.32	4.16
	2 범죄 발생 지역 및 요인 예측	4.16	4.05
	3 비신뢰 데이터 선별	3.53	3.21
	4 정보 공유 제약사항 해결(개인 사생활 및 프라이버시 보호 등)	3.95	3.79
	5 범죄 관련 데이터와 공권력과의 연계	3.67	3.61
재난대응 (2순위)	1 환경 재난 및 생활 재난 관련 데이터 상시 모니터링/사전 대응	4.37	4.32
	2 과거 데이터 분석을 통한 재난 상황 예측	4.37	4.11
	3 재난 관련 데이터의 통합/연계/공동활용	4.42	4.11
주거교통 (4순위)	1 미래 공간 활용 가능성 예측	3.37	3.21
	2 도로 증설, 도로 교통신호 체계 수립 등 대중교통 계획에 활용	3.89	3.68
	3 교통난 해소, 매연과 소음 감소, 최적 노선 정비	4.32	3.89
	4 교통 관련 기관들의 데이터 통합·분석	3.89	3.79
교육/취업 (5순위)	1 데이터 기반 맞춤형 취업 및 고용 서비스 제공	3.63	3.42
	2 미래 수요가 높은 일자리 예측 및 인력 공급 계획 수립	3.63	3.42
	3 교육 및 고용 관련 기관간 데이터 공유를 통한 정책 수립	3.79	3.37
	4 고등교육에 필요한 다양한 데이터 제공	3.42	3.11

사회현안 해결 측면에서는, 재난대응·예방, 범죄예방, 학교폭력 예방, 질병증가/신종 질병 및 노인건강 등의 분야에서 빅데이터의 활용 가치가 높은 것으로 평가되었다.

[표 4] 사회 현안 해결 측면 - 빅데이터 주요 이슈의 중요성 및 시급성

사회 문제		중요성 평균	시급성 평균
1	노인건강관리	4.05	3.89
2	의료서비스 불균형	3.95	3.37
3	질병증가/신종질병	4.32	3.95
4	학교폭력	4.16	4.26
5	범죄예방	4.26	4.26
6	교통안전	3.89	3.79
7	식재료 안전	3.79	3.63
8	재난대응·예방	4.68	4.42
9	핵 안전	3.63	3.11
10	테러	3.53	2.95
11	에너지 절감	3.84	3.58
12	신재생 에너지	3.32	2.89
13	전력 수급	3.84	3.58
14	대기오염	3.89	3.58
15	수질 오염	3.79	3.58
16	생활·산업 폐기물 처리	3.37	3.26

III. 정책적 시사점

1. 공공 영역의 역할 재정립

정부와 민간, 또는 시장에서의 역할이 명확히 구분될 필요가 있는데, 정부 및 공공기관은 보유한 공공데이터의 적극적인 개방과 민간 활용을 위한 인프라 구성에 집중하고, 빅데이터를 활용한 서비스 구현 및 제공 주체는 민간 영역에서 담당하는 것이 바람직하다.

또한, 빅데이터를 통한 사회·경제적 가치를 확보하기 위해서는 규모의 경제와 같이 데이터나 분석 능력의 경제가 독점되지 않고 순환될 수 있는 환경 조성이 필요하다. 즉 데이터의 개방 및 공유뿐만 아니라 민간 영역으로의 분석 역량의 공유와 지원이 공공 영역의 주요 역할로 인식되어야 한다.

‘데이터의 공유’ 측면에서는, 데이터 자체의 개방 차원을 넘어서 공공데이터와 민간 데이터와의 융합이 가능한 기술 및 체계를 확립해야 해야 한다. 빅데이터를 통해 새로운 비즈니스 및 부가가치를 창출할 수 있을 것으로 기대되고 있으나 이는 데이터 생산자 또는 소유자가 해당 데이터를 공개한다는 전제를 바탕으로 한다. 현실적으로 개인이나 기업이 보유한 데이터의 개방을 강제할 수 없는 상황에서 빅데이터 활용 가치는 공공 분야에서 극대화 될 수 있다. 공공데이터의 단순 제공 방식을 통해서만 민간이 원하는 정보를 식별, 추출하여 분석하기가 매우 어렵다. 따라서 정부는 공공데이터의 개방을 단순한 제공의 행위를 넘어서 ‘서비스’로서 인식하여야 한다. 즉 민간이 필요로 하는 정보에 대한 수요 분석을 통해 공공 기관이 관련 데이터를 분석하여 민간 영역에서 필요로 하는 최종 결과 정보만을 제공하는 서비스 방식으로 발전할 필요가 있다.

‘분석 역량의 공유’ 측면에서는 분석 기술 및 전문가 확보가 어려운 중소기업과 창업자와 같은 개인이 빅데이터 분석을 통해 기존 사업 영역을 확장하거나 신규 서비스 창출이 가능하도록 분석 기술 및 서비스를 지원해야 한다. 이를 위해 서비스 구현에 필요한 데이터의 추출 및 분석을 직접적으로 제공하거나 분석 수요자와 전문가를 매핑시켜주는 플랫폼 구축이 필요하다. 또한 장기적 관점에서 국가 차원에서 데이터 과학자 양성을 위한 교육센터 및 프로그램을 운영하여 빅데이터를 통한 국가·사회 경쟁력을 확보할 수 있는 기반을 마련하여야 한다.

2. 가시화 및 시뮬레이션 기술 개발

사회 현안 문제 해결을 위해서는 센서데이터와 같은 방대한 양의 데이터 분석과 복잡하고 다양성이 높은 이종 데이터 간의 융합분석이 필수적이다. 대용량 데이터의 가시화 및 시뮬레이션을 위해서는 데이터의 분할 및 분배, 가시화 알고리즘의 병렬화, 실시간 렌더링, 이미지 병합, 원격 가시화 등의 요소 기술이 필요하다. 그러나 이와 같은 고도의 IT 역량을 현장 실무자가 확보하는 것은 매우 어렵다. 아울러, 주요 분석 소프트웨어 및 기술 신뢰도에 있어 대외의존도가 매우 높아 관련 산업에서의 글로벌 경쟁력 및 시장 확보에 한계가 존재한다. 이러한 문제점들을 극복하기 위해서는 빅데이터 차세대 핵심 기술 선점을 위한 국가R&D 투자가 촉진되어야 하며, 이를 통해 개발된 기술들은 민간으로 적극적으로 확산되어야 한다.

특히, 정책결정자들에게 데이터 분석 결과를 가시화하여 시각적으로 전달함으로써 빅데이터 활용 목적과 효과에 대한 전달력과 설득력을 제고시킬 수 있다. 텔파이 분석을 통해 데이터 활용 필요성에 대한 의사결정자의 인식 수준이 매우 낮은 것으로 분석되었는데, 이는 빅데이터 도입 및 프로젝트 실패의 주요 요인으로 작용할 수 있다. 가시화를 통해 최고 정책결정자들이 빅데이터 활용가능성과 중요성을 더욱 확실하게 인지하게 함으로써, 공공 및 민간 부문에서 빅데이터 활용 확산의 촉발적 계기가 마련될 수 있을 것이다.

3. 의료 및 재난 대응을 위한 빅데이터 활용

과학기술 공공데이터 활용 분야 및 사회 현안 해결을 위한 빅데이터 활용 분야로서 의료와 재난대응 부문에 대한 높은 수요가 존재하는 것을 확인할 수 있었다. 이는 국민의 행복과 안전을 보장하기 위해 의료와 재난대응 분야에서 빅데이터 활용을 위한 국가 차원의 노력이 필요하다는 것을 의미한다.

해당 분야에서의 빅데이터 활용 시나리오는 다양하게 도출될 수 있는데, 특히 시스템 기반의 의사결정을 통해 대국민 서비스로 연계되었을 때 활용 가치가 증가할 수 있다. 예를 들어 의료 분야에서는 기존에 연구자 및 의료진의 고유 영역이었던 질병 예측 및 진단을 빅데이터 기반의 시스템으로 지원하는 것이 가능하다. 재난대응 분야에서는 기존에 통계 데이터나 데이터의 단순 통합에서 벗어나 다양한 데이터의 연계 및 실시간 분석, 미래 상황에 대한 시뮬레이션 등을 통해 자연재해 및 인적재해의 예측과 사전 대응이 가능할 것이다.

참고문헌

Henschen, D. (2013), Big Data Meets through of Disillusionment: Garter, Information Week.

MIT Slon Management Review (2011), Big Data, Analytics and the Path from Insight to Value.

송위진 (2014), 사회문제 해결형 연구개발사업의 의의와 과제, 과학기술정책, 제24권 제2호.

오정수, 강대신, 하성도 (2014), 사회문제 해결형 연구개발을 위한 한국과학기술연구원(KIST)의 사업 기획 현황, 과학기술정책, 제24권 제2호.

미래창조과학부 (2013.11), 과학기술 빅데이터 활용 종합계획.

이상일 (2013.10), 창조경제의 신성장 동력 ‘빅데이터’ 어떻게 준비하고 만들어가야 하는가, 국정감사 정책자료집.

미래창조과학부, 방송통신위원회 (2013), 융합형 공공데이터 개방 서비스 방안 연구.

IT & Future Strategy (2012), 성공적인 빅데이터 활용을 위한 3대 요소-자원, 기술, 인력, 제3호.