

공공 빅데이터 플랫폼 이용 활성화 전략

김지용 (정부통합전산센터), 김대엽 (펜타시스템 테크놀로지)

목 차	1. 서 론
	2. 공공 빅데이터 플랫폼 운영현황
	3. 공공 빅데이터 분석현황 및 한계
	4. 공공 빅데이터 플랫폼 이용 활성화 전략
	5. 결 론

1. 서 론

정부의 운영패러다임이 ‘정부 1.0(정부 일방향) → 정부 2.0(국민· 정부 양방향) → 정부3.0(국민· 정부 개인맞춤형)’으로 전환되었다. 정부3.0은 정부가 신뢰받는 정부, 국민행복이란 비전을 실현하기 위해 국민 개개인에게 맞춤형 서비스를 제공하는 노력에 최대 역점을 두는 새로운 정부의 운영 패러다임을 말한다. 아울러 ‘데이터에 기반한 과학적 정책수립’ 지원을 위해 정부3.0 추진위원회, 정부부처 및 지자체 등에서 빅데이터 관련 업무를 수행하고 있다. 공공분야에서는 행정자치부가 빅데이터 관련 법· 제도 정비 및 빅데이터 사업을 발굴하고 있으며 행정자치부 소속기관인 정부통합전산센터에서는 범정부 기반의 빅데이터 플랫폼을 구축하여 정부 또는 지자체에서 수행하는 빅데이터 분

석 업무를 지원하고 있다. 민간분야에서는 미래창조과학부가 민간 빅데이터 활성화를 위해 빅데이터센터를 운영하고 있으며, 이를 통해 창업자, 중소기업, 대학 등에 Shared Service를 제공함으로써 산업 활성화를 촉진하고 있다. 더불어 일부 지자체에서도 빅데이터 전문조직을 구성하고 있으며 빅데이터 활용 관련 조례를 제정하고 있다.

정부3.0 추진위원회와 행자부가 조사한 공공데이터 개방 국민인식 조사결과(‘15년 9월~10월) 공공데이터 개방정책에 대해서는 약 64%가 인지하는 것으로 나타났으나, 개방되는 공공데이터의 양이 전반적으로 부족하다는 답변이 약 51%였다. 한편 공공데이터 개방 시 고려되어야 할 사항으로 ‘가치 있는 공공데이터 제공(37%)’과 공공데이터의 정확성, 최신성 및 표준화 등을 뽑았으며, 공공데이터 이용을 위한 정부의 지원책으로 50%이상이 교육제공과 IT 인프라 제공을 뽑았다.

이에 공공분야에서는 정부통합전산센터가 빅데

※ 본 내용은 필자의 주관적인 의견이며, 정부통합전산센터의 공식적인 입장이 아님을 밝힙니다.

이더 관련 IT 인프라 및 플랫폼 교육을 제공하고 있으며, 민간분야에서는 미래창조과학부가 그 역할을 하고 있다. 본 논문에서는 정부통합전산센터에서 운영 중인 빅데이터 공통기반 플랫폼을 중심으로 공공 빅데이터 플랫폼 이용활성화 전략을 논하고자 하며 논문의 구성은 다음과 같다. 2절에서는 공공 빅데이터 플랫폼의 전반적인 운영현황을 소개하고, 3절에서는 공공 빅데이터의 분석현황 및 한계점을 논한다. 마지막 절에서 공공 빅데이터 플랫폼 이용 활성화 전략을 제안한다.

2. 공공 빅데이터 플랫폼 운영현황

정부통합전산센터(이하 ‘센터’)에서는 부처·지자체 공무원이 공동으로 활용할 수 있는 ‘범정부 공통기반 플랫폼’(이하 ‘혜안’, www.insight.go.kr)을 행정 업무망을 통해 운영하고 있으며, 플랫폼 구성은(그림 1)과 같다. 혜안에서는 외부 공공기관정보와 민간정보를 Open API 및 크롤링(Crawling) 등의 방법으로 수집·연계하여 센터 내에 저장하고, 빅데이터 처리·분석·시각화 도구를 활용하여 분석한 결과물을 관리 및 제공하며, 제공된 결과물은 중앙부처·지자체의 정책수립에 활용될 수

있다. 혜안의 주요 서비스로는 누구나 즉시 서비스를 이용할 수 있는 언론스크랩과 실시간으로 변화하는 화제의 키워드 관련 연관 검색어의 추이 분석 결과를 제공하는 실시간 키워드 분석서비스 등 일반분석 서비스가 있으며, 전문 분석가가 대용량 데이터를 처리·분석·시각화 할 수 있도록 필요한 도구를 제공하는 전문분석 서비스가 있다.

아울러 센터에서는 빅데이터 분석능력이 부족한 공무원도 혜안을 활용하여 직접 분석할 수 있도록 ‘빅데이터 공통기반 활용모델’을 제작하여 빅데이터 분석 이용활성화에 앞장서고 있다. 주요 모델로는 위치기반 분석 모델, 소셜 분석 모델, 홈페이지 게시글 분석 모델이 있다. 위치기반 분석 모델은 국민생활과 밀접하게 연관된 지역·장소 데이터를 기존의 분석방식(통계표, 차트 등)에서 벗어나 GIS를 통한 위치 기반의 데이터분석 결과로 도출하는데 사용된다. 소셜 분석 모델은 센터가 수집하고 있는 소셜 데이터와 결합하여 여론동향을 파악하거나 정책방향을 설정하는데 적극 활용되고 있으며 홈페이지 게시글 분석 모델은 행정기관이 보유한 홈페이지 게시글의 비정형 데이터를 워드 클라우드, 랭킹, 소셜네트워크(SNA) 및 열지도(Heatmap) 등으로 결과를 표출하여 그 결과를 민원분석에 활용하고 있다.



(그림 1) 빅데이터 공통기반 플랫폼 구성도

3. 공공 빅데이터 분석현황 및 한계

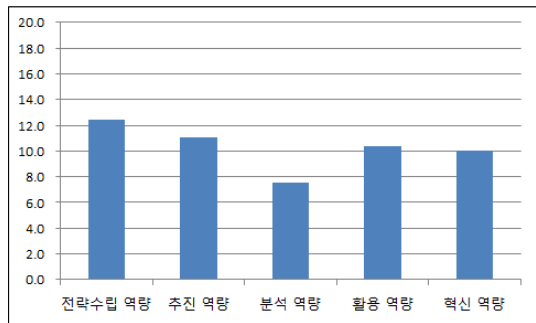
빅데이터 분석을 성공적으로 수행하기 위해서는 우선 분석의 목적과 목표를 명확하게 정의해야 하며, 보유하고 있는 데이터에 대해 충분히 이해하고 있어야 한다. 또한 데이터를 어떤 형태로 가공하고, 가공된 데이터를 어떠한 기법을 활용하여 분석할 것인지에 대한 전략도 미리 수립되어야 하며 분석 결과를 어떠한 형태로 시각화하고 활용할 것인지에 대한 계획도 구상해야 한

다. 즉, 빅데이터를 분석하고 활용하기 위해서는 데이터에 대한 이해와 더불어 데이터 가공 및 분석 기법, 분석결과 활용 전략 등과 관련된 전반적인 지식과 노하우가 요구된다.

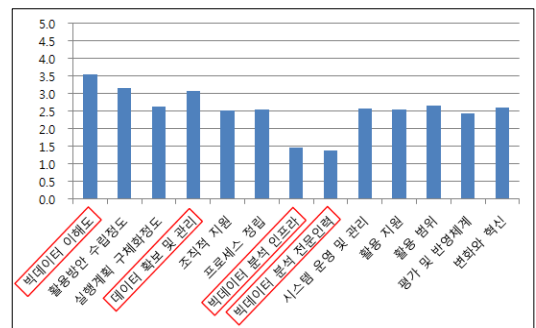
하지만 대부분의 정부부처나 지자체 공무원은 빅데이터 분석 방법 및 활용에 대한 지식과 노하우가 매우 부족한 것이 현실이다. 센터에서는 정부부처·지자체 공무원의 빅데이터 분석 역량을 향상시키기 위해 올해 두 차례의 빅데이터 분석 전문교육을 실시하였으며, 각 기관의 빅데이터 분석 역량을 파악하기 위한 설문조사를 실시한 바 있다. 설문조사 내용은 미래창조과학부가 주관하고 한국정보화진흥원 등이 공동으로 개발한 빅데이터 분석역량 진단모델을 기반으로 구성하였다. 설문 결과에 따라 기관의 역량 수준은 초기(initial), 인식기(repeatable), 정립기(defined), 관리기(managed), 혁신기(innovated)의 5단계 중 하나로 결정된다. 분석 역량 수준의 결정은 설문을 구성하는 다섯 가지의 역량(전략수립 역량, 추진 역량, 분석 역량, 활용 역량, 혁신 역량) 점수에 따라 결정되며, 이들 각각의 역량은 다시 그것을 구성하는 중핵목 점수에 따라 결정된다.

설문조사에 참여한 중앙부처·지자체 공무원의 응답 결과를 종합해 본 결과 약 90% 정도의 기관이 인식기(repeatable) 이하의 역량 수준을 보이고 있으며, 상위에 속한 기관 역시 핵심 역량(빅데이터 분석 인프라, 빅데이터 분석 전문 인력 등) 부분에서는 다소 미흡한 수준을 보이고 있다. (그림 2)와 (그림 3)은 설문에 참여한 기관 공무원에 대해서 다섯 가지 역량과 그것을 구성하는 세부 항목들의 평균 측정치를 보여주고 있다.

대부분의 기관이 빅데이터에 대한 이해도나 데이터의 확보 및 관리 차원에서는 어느 정도 준수 수준의 역량을 보유하고 있지만, 빅데이터 분석 인프라 구축 및 빅데이터 분석 전문 인력



(그림 2) 빅데이터 핵심 역량에 대한 평균 점수



(그림 3) 세부 역량에 대한 평균 점수

확보 등의 핵심 역량에서는 기대에 못 미치는 수준을 보이고 있다. 다시 말하면 빅데이터 분석의 필요성을 인식하고는 있지만 실질적인 분석 역량의 확보에서는 아직 미진한 부분이 많다는 것을 알 수 있다.

정부부처·지자체 공무원의 빅데이터 분석 역량이 데이터 기반의 과학적 행정을 구현하는데 있어 아직 많은 한계를 보이고 있는 가운데 센터에서는 빅데이터 공통기반 시스템 구축, 분석 서비스 제공, 분석과제 지원 및 전문교육 실시 등의 역할을 수행하고 있지만, 빅데이터 활성화 측면에서 여전히 많은 문제와 제약사항이 존재한다. 현존하는 문제점을 정리해보면 다음과 같다.

첫째, 대부분의 부처 및 지자체는 각 기관에서 보유한 데이터에 대해서 자체적으로 분석을 수행하기보다 센터와 같은 빅데이터 전문기관에

의존하는 경우가 대부분이며, 경우에 따라 별도의 분석사업을 발주하여 외부 업체에 위탁하기도 한다. 이러한 방법은 시간이나 비용 측면에서 효율성을 저하시키는 요인이 될 수 있다. 빅데이터 활용의 활성화를 위한 근본적인 방법은 각 기관이 빅데이터 전문 인력을 적극 양성하고 필요 시 언제든지 빅데이터 분석을 수행하고 정책 마련에 활용할 수 있도록 하는 것이다.

둘째, 부처나 지자체에서 센터에 빅데이터 분석을 의뢰하는 경우를 보면, 해당 기관장의 관심 사항으로서 일회성 분석으로 진행되는 것을 종종 볼 수 있다. 이와 같은 경우는 대부분 분석 요구사항이 명확하지 않거나 막연하게 분석을 요청하는 경향이 빈번하게 나타난다. 앞서 언급한 바와 같이 빅데이터 분석의 활용 효과를 높이기 위해서는 분석의 목적과 목표가 명확하게 정의되어야 한다. 만약 그렇지 않으면 분석의 활용도는 저하되고 단순히 이벤트성 과제로 끝날 가능성이 높아질 것이다.

셋째, 낯선 빅데이터 분석환경 또는 기법(R 프로그래밍, SNA분석 등)으로 인해 직접 빅데이터 분석을 시도하는데 많은 부담을 느끼는 경향이 있다. 빅데이터 분석은 통계 분야에 대한 지식과 고수준의 데이터 처리 기술 등을 요구한다. 해당 분야에 경험이 부족한 공무원은 빅데이터 분석에 어려움을 겪을 수밖에 없다. 이러한 한계를 해결하기 위해서는 빅데이터 분석에 대한 꾸준한 교육과 다양한 사례 공유 등이 필수적으로 요구된다.

마지막으로 빅데이터 분석의 경우, 다양한 출처의 데이터를 융·복합함으로써 그 효과를 극대화시킬 수 있다. 즉, 데이터 기반의 과학적 행정을 실현하기 위해서는 다양한 기관의 데이터들을 서로 연계하고, 그로부터 분석을 통한 새로운 인사이트(Insight)를 확보해야 한다. 하지만

각 기관의 데이터를 확보 및 연계할 수 있는 법적 근거가 부족한 관계로 분석에 필요한 데이터가 부족한 것이 가장 큰 걸림돌로 작용한다. 업무 협의를 통해 어렵게 확보된 데이터 역시 정확성이 떨어지고 표준화되지 않은 형태로 존재하기 때문에 효과적인 분석에 어려움으로 작용한다.

지금까지 살펴본 바와 같이, 국내 공공기관의 빅데이터 분석 환경은 매우 열악하며, 이를 개선하기 위한 보다 근본적인 시도와 노력이 필요하다. 비단 한 기관만의 노력이 아닌 범정부적인 문제인식과 해결 노력이 수반되어야 한다.

4. 공공 빅데이터 플랫폼 이용 활성화 전략

올해 각 중앙부처·지자체에서는 기관의 고유 업무를 기반으로 다양한 분야에서 빅데이터 관련 추진계획이 마련되고 있다. 이에 따라 본 고에서는 기 구축된 해안시스템을 이용해 데이터 기반의 과학적·선제적 정책수립 지원 및 분석 전문인력 양성을 위해 센터가 선도적 역할을 할 수 있는 부문에 한정해 아래와 같이 활성화 제고 방안을 마련코자 한다.

첫째, 행정기관 등의 장은 전자정부법에 따라 다른 행정기관 등과의 상호연계 또는 공동이용과 관련한 전자정부사업 및 지역정보화사업을 추진할 때 중복투자 방지 등을 위하여 중앙사무관장기관의 장과 사전에 협의하여야 한다고 명시되어 있으나, 올해 상반기에 추진되는 정부부처·지자체별 빅데이터 사업 대부분이 플랫폼 구축사업이거나 여론 동향분석을 위한 서비스로 이는 센터에서 운영중인 플랫폼 및 빅데이터 서비스와 상당히 유사하다. 따라서 센터에서 기 구축한 플랫폼을 중앙부처·지자체에서 활용하도록 유도하기 위해 사전협의를 통한 빅데이터 사

업의 중복성 검토 과정을 강화하거나 중앙부처·지자체에서 추진하는 빅데이터 관련 사업을 센터와 먼저 협의할 수 있는 절차 등을 마련해야 한다.

둘째, 정부에서는 정부업무평가기본법에 따라 매년 행정관리역량평가 및 지자체 합동평가를 실시하며, 평가지표를 통해 국정의 효율적인 업무수행을 평가한다. 센터에서도 중앙부처·지자체에서 해안시스템을 활용한 빅데이터 분석업무를 적극 추진할 수 있도록 평가지표를 마련하여 데이터 기반의 과학적 정책을 수립할 수 있도록 지원하여야 한다.

셋째, 빅데이터 분석의 이용활성화를 위해 센터에서는 기 구축한 3종의 공통기반 활용모델 이외에 중앙부처·지자체의 업무 수요를 파악해서 상시적으로 활용할 수 있는 공통기반 활용모델을 적극 발굴하여야 한다.

넷째, 최근 공공분야에서는 기상·기후 빅데이터 플랫폼(기상청)을 통해 기상·기후 데이터와 타분야 데이터를 융합·분석할 수 있는 환경을 제공하고 있으며, 빅데이터 캠퍼스(서울시) 등을 통해 창업 및 정책 등의 아이디어 발굴을 지원하고 있다. 민간분야 역시 누구나 빅데이터 플랫폼을 활용하여 국민의 라이프스타일 트렌드를 분석하고 그 결과를 다양한 측면에서 활용할 수 있는 플랫폼 개방이 진행되고 있다. 따라서 해안시스템의 이용자 확대 및 다양한 분석지원을 위해 행정 업무망에서만 이용되는 시스템을 정부·지자체 산하의 연구소 직원이 활용할 수 있도록 개방하는 추진계획을 마련해야 한다.

다섯째, 센터에서는 빅데이터 분석역량 강화를 위해 매년 상하반기 전문교육을 실시하고 있으나 전문 분석도구 및 공통기반 활용모델의 실습시간이 제한된 관계로 단시간 내에 해안시스템을 이해시키기 어렵다. 따라서 온라인으로 언

제, 어디서나 상시적으로 서비스를 제공받을 수 있도록 지원하여야 하며, 이를 공무원 교육 전문기관과 연계해서 온라인 해안시스템 강좌를 적극 홍보해야 한다.

마지막으로 정부부처·지자체에서 추진하려는 빅데이터 분석사업 및 플랫폼 구축사업은 일반적인 정보시스템 구축사업과 상이하며 공공기관의 분석사례 또한 부족한 상황이다. 따라서 센터에서는 분석전문 인력을 활용하여 수요기관에 맞춤형 컨설팅을 제공할 수 있어야 하며, 기 구축된 분석결과(서비스)를 해안시스템을 통해 공유하여 다양한 아이디어가 도출될 수 있도록 적극 지원하여야 한다.

5. 결 론

지금까지 해안시스템의 운영현황과 중앙부처·지자체 공무원이 빅데이터 분석 시 당면하는 애로사항 및 이에 대한 이용 활성화 제고방안을 알아보았다. 올 상반기 정부에서는 빅데이터 관련 30여종의 보도자료를 배포하였고 주요 내용은 빅데이터를 활용한 대민서비스 지원이다. 공공 빅데이터 서비스를 추진하기 앞서, 정부부처·지자체에서는 다음 사항을 유의하여야 한다.

첫째, 데이터에 기반한 공공서비스를 일정한 수준이상으로 국민에게 제공하기 위해 데이터의 품질을 정의하고 객관적인 측정기준을 마련해야 한다. 특히 국민 중심의 맞춤형 서비스 구현을 위해서는 서비스 품질에 대해 상시 측정하고 그 품질을 향상시키려는 노력이 절실하다.

둘째, 중앙부처·지자체에서는 데이터 사이언티스트 인력확보가 중요하다. 이는 빅데이터를 도입하는 기관이 추진전략 및 수행계획을 수립하는데 있어 유의미한 분석 결과가 필요하기 때문에 데이터 사이언티스트의 역할이 매우 중요

하다고 볼 수 있다.

구글의 알파고(AlphaGo)와 이세돌의 바둑 대국 이후, 'AI(Artificial Intelligence)'에 대한 관심이 급증하고 있다. 정부에서도 빅데이터와 인공지능 기술의 결합을 통한 미래예측 시스템 구축 및 신규 지식 서비스 제공 방안을 마련하고 있다. 인공지능에서 우수한 판단을 하기 위해서는 데이터가 필수이다. 빅데이터의 궁극적인 목표는 인공지능의 완성일 수 있다. 빅데이터 기술이 주목받는 이유는 정보처리에 있지 않으며 우수한 정보처리를 바탕으로 의미있는 결과를 도출할 수 있는 점이 빅데이터가 주목받는 이유이다. 이에 따라 센터에서는 전문교육 과정을 통해 분석전문가를 양성하고 분석역량 강화를 위해 가이드 북을 제작 및 배포하였다. 또한 다양한 분석사례를 통해 기관에서 필요로 하는 분석 아이디어를 도출할 수 있도록 지원체계를 마련하였다. 향후 더욱 강화된 지원체계를 통해 중앙부처·지자체 공무원이 해안시스템을 이용할 수 있도록 지원할 예정이다.

한 제언, 2014년 5월

[9] 빅데이터 분석활용센터, 빅데이터 역량 진단모델, 2014년 2월

저 자 약 력



김 지 용

이메일 : coolguy@korea.kr

- 2008년 고려대학교 소프트웨어공학과 (석사)
- 2004년~서울시, 국립과학수사연구원, 행정자치부, 정부통합전산센터
- 관심분야: 정보시스템 표준화, EA



김 대 엽

이메일 : dykim@penta.co.kr

- 2015년 충남대학교 컴퓨터공학과 (박사)
- 2005년 충남대학교 컴퓨터공학과 (학사)
- 2016년~펜타시스템테크놀로지 책임컨설턴트
- 관심분야: 객체지향 분석 및 설계, 컴포넌트기반 개발 방법론, 소프트웨어 품질관리, 영향분석, 기능점수, 빅데이터

참 고 문 헌

- [1] 행정자치부, www.moi.go.kr
- [2] 정부3.0, www.gov30.go.kr
- [3] 빅데이터 공통기반 플랫폼(혜안), www.insight.go.kr
- [4] 국가법령정보센터, www.law.go.kr
- [5] 정부3.0 추진위원회, 정부3.0-데이터 기반의 과학적 행정, 2016년 4월
- [6] 한국정보기술학회지, 빅데이터가 인공지능에 미친 영향, 2016년 제14권1호
- [7] 행정자치부, 공공서비스 수준진단 및 품질제고 방안 연구용역 결과보고서, 2016년 8월
- [8] 정보통신산업진흥원, 빅데이터 추진전략을 위