

스마트 물관리를 위한 빅데이터 거버넌스 모델*

Big Data Governance Model for Smart Water Management

최영환[†] · 조완섭¹ · 이경희²

충북대학교 경영정보학과¹, 빅데이터학과²

요 약

스마트 물관리 분야에서도 빅데이터 분석을 통해 경쟁력을 강화하려는 요구가 급증하면서 빅데이터에 대한 체계적인 관리(거버넌스)가 중요한 이슈로 부각되고 있다. 빅데이터 거버넌스는 데이터의 품질보장, 프라이버시 보호, 데이터 수명관리, 데이터 전담조직을 통한 데이터 소유 및 관리권의 명확화 등의 데이터 관리를 평가하고(Evaluation), 지시하며(Direction), 모니터링(Monitoring) 하는 체계적인 관리활동을 의미한다. 빅데이터 거버넌스가 확립되지 못하면 중요한 의사결정에 품질이 낮은 데이터를 사용함으로써 심각한 문제를 야기할 수 있으며, 개인 프라이버시 관련 데이터로 인해 빅브라더의 우려가 현실화될 수 있고, 폭증하는 데이터의 수명관리 소홀로 인해 IT 비용이 급증하기도 한다. 이러한 기술적인 문제가 완비되더라도 데이터 관련 문제를 전담하고 책임지는 조직과 인력이 없다면 빅데이터 효과는 지속되지 못할 것이다. 본 연구에서는 빅데이터 기반의 스마트 물관리를 위한 데이터 거버넌스 구축모델을 제시하고, 실제 물관리 업무에 적용한 사례를 소개한다.

■ 중심어 : 데이터 통합, 데이터 거버넌스, 스마트 물관리, 데이터 관리체계

Abstract

In the field of smart water management, there is an increasing demand for strengthening competitiveness through big data analysis. As a result, systematic management (Governance) of big data is becoming an important issue. Big data governance is a systematic approach to evaluating, directing and monitoring data management, such as data quality assurance, privacy protection, data lifetime management, data ownership and clarification of management rights. Failure to establish big data governance can lead to serious problems by using low quality data for critical decisions. In addition, personal privacy data can make Big Brother worry come true, and IT costs can skyrocket due to the neglect of data age management. Even if these technical problems are fixed, the big data effects will not be sustained unless there are organizations and personnel who are dedicated and responsible for data-related issues. In this paper, we propose a method of building data governance for smart water data management based on big data.

■ Keyword : Data Integration, Data Governance, Smart Water Management, Data Management System

2018년 12월 10일 접수; 2018년 12월 10일 수정본 접수; 2018년 12월 31일 게재 확정.

* 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (2017-0-00253, 국제표준 기반 오픈 데이터 유통 플랫폼 확장 기술 개발)

† 교신저자 wscho@chungbuk.ac.kr

I. 서 론

2010년 초부터 사회전반에서 데이터 기반 혁신을 위루기 위한 다양한 정책이 국가 주도로 이루어지고 있다. 미국은 데이터 주도의 사회, 유럽연합은 데이터 주도의 경제, 경제협력개발 기구는 데이터 주도의 혁신을 강조하고 있다. 2016년 다보스포럼에서는 수많은 데이터를 연결하고 분석, 활용하는 4차 산업혁명을 중요하게 다루었고, 선진국 뿐 아니라 중국 등에서도 사회의 모든 분야에서 데이터를 활용한 혁신을 이루어 나가고 있다.

물관리 분야에서도 빅데이터를 활용하여 스마트 물관리를 실현하는 연구가 최근에는 중요한 이슈로 등장하고 있다[7]. 미국의 지질조사국에서 운영 중인 국가수자원정보시스템은 물관련 데이터를 저장, 검색하기 위한 통합시스템이다. 이 시스템은 다양한 유형의 실시간 물관리 데이터를 15~60분 간격으로 수집하고, 상호 연계하여 제공하고 있다. 영국 런던의 스마트워터 그리드는 기존 수자원 관리시스템의 한계를 극복하기 위해 첨단 정보통신기술을 이용하는 고효율의 차세대 물관리 시스템이다. 이 시스템은 물 수요와 공급량을 실시간으로 정확하게 관리해 수자원의 지역적, 시간적 불균형을 해소하는 지능형 물관리시스템으로 자리 잡고 있다. 한국에서는 수자원공사와 환경부를 중심으로 스마트 물관리 실현을 위한 빅데이터 수집과 활용에 중점을 두고 다양한 사업을 추진하고 있다[10].

국가적인 차원의 스마트 물관리를 위해 각 기관에서는 다양한 종류의 데이터를 활용하여 분석을 수행해 왔으나, 지속적으로 빅데이터 활용 효과를 개선하고, 예측정확도를 높이는 방법은 한계에 이르고 있다. 스마트 물관리에 있어서 기초적인 관측 데이터 뿐 아니라 이를 기존의 경영 및 기술적 데이터와 통합하여 분석하고, 자연현상의 복잡함을 고려한 빅데이터 기반의

분석이 지속적으로 개선되어 나가는 선순환 체계를 구축해 나가는 것이 강조되어야 할 것이다.

빅데이터 거버넌스(bigdata governance)는 다양한 소스로부터 유입되는 물관리 빅데이터를 체계적으로 관리하고 지속적으로 최적화시켜 지속가능하고 점진적으로 최적화되어 가는 빅데이터 실현에 중요한 역할을 수행한다. 빅데이터 거버넌스란 빅데이터의 수집에서 활용에 이르는 전체 과정을 평가(evaluation), 지시(direction), 모니터링(monitoring)하는 관리체계를 의미한다 [ISO]. 전체 분석시간의 80% 정도를 차지하는 빅데이터의 수집 및 정제 단계의 효율성 증진, 여러 소스로부터 유입되는 데이터의 표준화 및 시간-공간적 일관성 유지, 데이터 품질 및 신뢰도 증진 등이 뒷받침될 때 빅데이터 분석 결과를 활용한 정확한 의사결정 지원이 가능하게 된다. 이러한 문제를 해결하려면 데이터 품질이나 표준화 문제를 지속적으로 평가하고 문제점 도출과 개선을 지시하며, 일상업무에서 모니터링 하는 체계가 갖추어져야 하고, 이를 위한 조직과 제도가 빅데이터 거버넌스의 일부로 구축되어야 한다.

본 연구는 스마트 물관리를 위한 빅데이터 거버넌스 구축방안을 제시하고자 한다. 빅데이터 활용에서 필요한 빅데이터 거버넌스 프레임워크를 개발하고 구축방안을 제시한다. 이를 위해 일반적인 빅데이터 거버넌스 프레임워크로부터 물관리 분야에 적용 가능한 물관리 빅데이터 거버넌스 프레임워크를 도출하고, 물정보 통합, 데이터 품질관리, 비즈니스 창출의 단계적인 과정에서 빅데이터 거버넌스의 구축과 활용을 논의한다. 물기업은 물관리 빅데이터를 활용하여 생산원가를 절감하고 물산업을 육성하며 재해 재난을 예방할 수 있다. 스마트 물관리를 위해 빅데이터를 도입하고자 하는 물기업에게 빅데이터 거버넌스의 필요성과 구축모델을 종합적으로 제시함으로써 지속적인 빅데이터 활용과 최

적화를 이루는데 기여할 것이다.

본 논문의 구성은 다음과 같다. 제 2장에서는 스마트 물관리에서의 빅데이터 활용과 이를 지원하는 빅데이터 거버넌스에 관해 소개한다. 제 3장에서는 스마트 물관리 분야에서 빅데이터 활용에 필요한 빅데이터 거버넌스 프레임워크와 구축방안을 제시한다. 제 4장에서는 실제 스마트 물관리 현장을 대상으로 빅데이터 거버넌스 구축 방안을 제시한다. 제 5장에서는 전체 논문을 요약하고, 향후 연구방향을 소개한다.

II. 관련 연구

본 장에서는 물관리 기관들이 생성 및 활용 중인 데이터 현황을 살펴보고, 빅데이터 거버넌스에 관한 관련연구들을 소개한다.

2.1 물관리 빅데이터 현황

물에 대한 관심과 물산업의 중요성이 증가함에 따라 다양한 계층의 요구에 맞는 신뢰할 수 있는 물정보의 통합적 제공이 중요하며, 이를 위해 물기관은 데이터 통합, 공유 기반을 마련하고 대국민 서비스를 강화하여야 한다.

기존의 물정보는 부서나 목적별로 분산 관리 중이고, 정보를 제공하는 시스템도 기관들이 별도로 구축 운영하여 단편적인 정보만 제공하고 있다. 국민들 대상의 유용한 물정보 서비스를 위해서는 물순환 전반에 대한 정보가 필요하지만 기관별 시스템 구축으로 서비스에 한계가 있다. 대국민 정보서비스를 목적으로 물기관이 각종 시스템을 개별운영하고 있으나 기초자료나 공급자 관점에서의 정보제공으로 국민의 정보이용은 미흡한 실정이다. 국민들이 이해하기 쉽고 실생활에 필요한 물정보를 제공하는 수요자 관점에서의 정보제공이 필요하다.

국토교통부는 표 1과 같이 홍수통제소, K-wa-

ter, 유량조사사업단을 통해 물관리 데이터를 수집, 저장, 활용하고 있다. 홍수통제소는 국가하천의 강우, 수위, 유량 등 실시간 수문자료를 이용하여 홍수를 예·경보하고, 하천수의 취수와 인허가를 총괄한다. 유량조사사업단은 국가하천의 수위와 유량조사, 치수시설조사 등 14개 항목을 관리하고 있다.

〈표 1〉 국토교통부 주요 데이터 현황

정부부처	유관기관	주요 데이터
국토교통부	홍수통제소	수문 : 국가하천 강우 수위, 유량 자료관리 이수 : 하천 취수 및 인허가 사항 치수 : 홍수 예·경보, 홍수지도 본류구간
	유량조사사업단	수문 : 국가하천 수위, 유량조사 유역 : 치수시설조사 등 14개 항목

환경부는 표 2와 같이 국립환경과학원이 수질자료 관리와 평가를 시행하고, 물환경연구소가 오염총량지점의 유량과 수질관리를 담당한다. 유역환경청이 유역 내 하천, 호수, 산업단지의 수질을 관리하고, 한국환경공단은 전국에 수질자동측정망을 설치관리하고 있으며, K-water는 강수량, 수위, 유량, 유사량 등 실시간 수문자료와 유역 이수조사 146개 항목, 호수 및 하천의 수질조사 데이터, 국가 지하수 관측소의 운영데이터, 댐 보의 운영데이터, 상하수도의 운영데이터를 관리하고 있다.

농림수산식품부는 표 3과 같이 한국농어촌공사에서 강수, 수위, 유량, 토양수분량 등 수문정보와 농업용수의 수질, 지하수의 수위와 수질, 댐 보의 저수위와 방류량 등 운영데이터를 공유하고 있다. 산업통상자원부는 한국수력원자력에서 강수, 수위, 발전 댐의 유량을 관리하고 있다. 지자체도 산하기관을 통해 지방하천의 갈수량, 수위, 유량, 수질과 지하수의 개발이용, 상하

〈표 2〉 환경부 주요 데이터 현황

정부 부처	유관 기관	주요 데이터
환경부	환경과학원	수질 : 자료 관리 및 평가
	물환경연구소	수문 : 유량(오염총량지점) 수질 : 일반, 총량 및 퇴적물측정망
	유역환경청	수질 : 일반측정망 (하천, 호소 및 산단)
	한국환경공단	수질자동측정망 설치·관리
	한국수자원공사	수문 : 강수량, 수위, 유량, 유사량 등 유역 : 이수조사 등 146개 항목 수질 : 일반측정망(호소 및 하천) 지하수 : 국가 지하수 관측소 수위, 수질 등 댐·보 : 저수위, 방류량 등 운영실적 상수도 : 광역/지방 공급량, 수질 등 운영

수도의 운영데이터를 공유하고 있다.

이와 같이 다양한 부처의 기관들에 파편화되어 저장관리되는 물관련 빅데이터는 물의 흐름이나 서비스에 따라서 통합되어 관리되어야 만 빅데이터 분석을 통한 고부가 가치 창출이 가능하게 될 것이다. 즉, 빅데이터의 연계 및 통합이 중요한 과제로 남아 있다.

〈표 3〉 주요 물기관 주요 데이터 현황

정부 부처	유관 기관	주요 데이터
농수산부	농어촌공사	수문 : 강수, 수위, 유량, 토양수분량 수질 : 일반측정망(농업용수) 지하수 : 농촌지역 지하수 정보 댐·보 : 저수위, 방류량 등
산업부	한수원	수문 : 강수, 수위, 유량(발전용댐 상류)
지자체	산하기관	수문 : 강수량, 수위, 유량(지방하천) 수질 : 일반측정망(하천, 호소, 도시관류) 지하수 : 지하수 개발, 이용허가 상수도운영 : 상수도통계

2.2 빅데이터 거버넌스

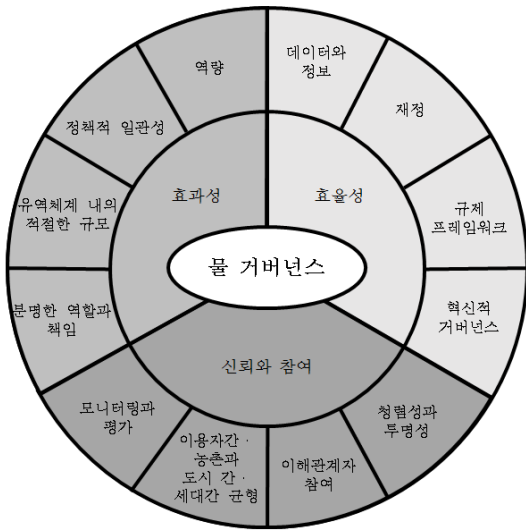
빅데이터 거버넌스(Big Data Governance)란 전사적인 차원에서 보유하고 있는 데이터에 대해 관리 정책, 지침, 표준, 전략 및 방향을 수립하고 데이터를 관리할 수 있는 조직 및 서비스를 구축하는 데이터 관점에서의 IT 관리 체계(IT 거버넌스)를 의미한다[1,2]. 빅데이터 거버넌스의 궁극적인 목적은 고품질 데이터의 확보와 관리를 통해 의사결정자가 필요로 하는 데이터를 적시에 공급하는 체제를 확립함으로써 조직의 다양한 가치 창출에 지속적으로 기여하는 것이다. 마치 인체의 혈액이 전신으로 공급되어 생명체가 건강하게 유지되듯이 빅데이터가 항상 수집, 통합되고, 분석되어 필요한 사람에게 적시에 전달되도록 하는 것이다.

Ⅲ. 스마트 물관리를 위한 빅데이터 거버넌스

본 장에서는 스마트 물관리를 위한 빅데이터 거버넌스 프레임워크를 제안하고, 이를 기반으로 실제 국내지역을 대상으로 시범적용한 결과를 소개한다.

경제협력개발기구는 2015년에 물 거버넌스 원칙을 그림 1과 같이 발표하였다. 물관리는 대부분 지역적인 문제로 시작되지만 동시에 국가적인 문제로 확대되며, 중요한 의사결정과 물관리 정책수립과정에서 수많은 이해관계자들이 개입하고 있다. 이러한 물관리 거버넌스의 문제를 해결하고자 물 거버넌스 원칙은 효과성, 효율성, 신뢰와 참여라는 세 가지 측면에서 각각 네 가지의 세부원칙을 제시하고 있다.

효과성은 분명하고 지속가능한 물정책의 목적과 목표를 규정하고, 국가는 각 부서의 역할과 책임을 명확히 부여하며, 이들 간의 협력방안을 조성해야 함을 의미한다. **효율성**은 물관리



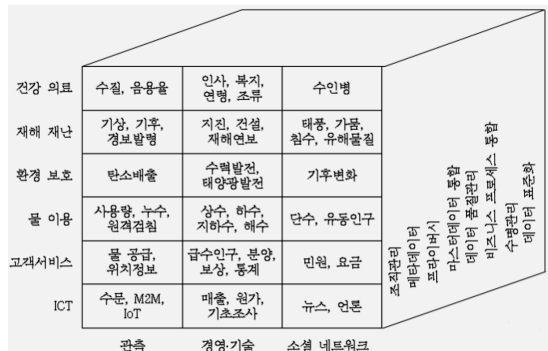
〈그림 1〉 경제협력개발기구의 물거버넌스 원칙

에 있어서 최소한의 사회적인 비용으로 편익과 복지를 극대화하는 것이다. 물 관련 데이터와 가공된 정보는 적절한 시기에 일관성 있게 제공되어 활용할 수 있어야 하고, 물 정책을 개선하고 평가하는데 활용할 수 있어야 한다. 물 거버넌스에 대한 **신뢰와 참여**는 국민의 신뢰를 형성하고 규제 프레임워크의 청렴성과 투명성을 확보하여 이해관계자들의 참여를 확대하는 것이다. 물관리 정책의 수립과 실행에 있어서 물 사용자 간, 도시와 농촌 간, 세대 간 균형 있는 참여를 장려하고, 물정책과 거버넌스에 대하여 정기적으로 모니터링하고 평가하며, 그 결과를 국민과 공유하여 개선방향을 지속적으로 발전시켜야 한다. 특히 관측 데이터 등 빅데이터 수집과 통합 분석을 통해 이해당사자간의 신뢰와 참여를 높여야 한다.

이러한 경제개발협력기구의 물거버넌스 원칙의 많은 부분이 물관리 빅데이터의 수집과 연계 및 통합분석을 통해 실현될 수 있을 것이다. 빅데이터 분석을 통해 효율성을 높일 수 있으며,

이해 당사자 간의 신뢰 증진과 참여도를 높일 수 있기 때문이다. 즉, 물거버넌스의 효율성과 신뢰 및 참여의 증진에 기여하는 빅데이터 활용은 빅데이터 거버넌스 구축을 통해 이루어질 수 있다.

본 연구에서는 스마트 물관리를 위한 빅데이터 거버넌스의 프레임워크를 그림 2와 같이 3차원 큐브(cube) 모델로 제시한다.¹⁾ X축은 물관에 필요한 데이터의 유형으로 관측데이터, 기존의 경영 및 기술 데이터, 소셜 네트워크 데이터로 구분하였다. Y축은 물관리를 통해 실현되는 물산업을 나타내며, 건강의료부터 ICT산업까지 6가지로 구분하였다. Z축은 데이터 거버넌스 원칙들로서 거버넌스 조직부터 데이터 표준화까지 8가지로 구분하였다.



〈그림 2〉 물관리 빅데이터 거버넌스 프레임워크

3.1 물관리 빅데이터 유형 (X축)

빅데이터 유형은 효율적인 물관리를 위해 기업에서 자체적으로 생산 관리하는 데이터와 외부에서 수집되는 각종 데이터를 유형별로 구분한 것이며, 현장의 센서에서 수집되는 정형 데이터, 업무를 처리하면서 자동으로 생산하는 정형 및 비정형 데이터, 소셜 네트워크에서 소비

1) 일반적인 빅데이터 거버넌스 프레임워크는 Sunil 이 제시하였으며[Sunil], 본 연구에서는 이를 물관리 영역에 적합한 형태로 제시한 것임.

자가 생산하는 소셜 네트워크 데이터들이 있다.

표 4는 는 물관리 빅데이터의 유형을 세가지로 구분하고, 각 데이터에 대한 종류와 특성을 요약한 것이다.

〈표 4〉 물관리 빅데이터 유형

	관측	경영.기술	소셜 네트워크
활용 분야	댐, 수도, 지하수, 수질실험	행정, 기술, 설계, 지방상수도	재해재난, 민원, 언론
데이터 종류	유량, 압력, 수위, 수질, 강우량, 누수 등	인사, 재무, 감사, 법무, 고객, 위험 등	태풍, 가뭄, 단수, 요금, 보상, 분양 등
데이터 형태	정형	비정형	비정형
수집주기	실시간 수집, 1분/30분/1시간 저장	분야별 업무처리시 생산 및 저장	수시
품질관리 시스템	HDIMS, DQMS, GIMS, LIMS	없음	없음

〈표 5〉 물관리 빅데이터 유형별 물산업 경영성과

구분	관측	경영.기술	SNS
생산 원가 절감	물 사용량 절감, 원격검침, 유수율 향상, 음용율 향상	매출액 증가, 급수인구 확대, 요금 현실화, 점검정비 향상	유동인구 예측, 무단수 공급
환경 보호	탄소배출 최소화, 신재생에너지 확대	분양 확대, 적기 보상, 골재 판매, 자원 순환	물복지 향상
재해 재난/건강 의료	홍수예측, 가뭄예방, 도시침수 예방, 기상 예측, 강우량 예측 수질정보	재해연보 활용, 기후변화 대응, 지진 대응	공정 뉴스, 민원 예방, 조류 예방, 유해물질 대응

3.2 물 관련 산업 (Y축)

물기반 산업(비즈니스)를 생산원가절감(물이용, 고객서비스, 물이용)과 환경보호 및 재해재난/건강의료로 구분하여 데이터 유형별로 표 5와 같이 분석하였다.

3.3 빅데이터 거버넌스 원칙

거버넌스 원칙은 빅데이터를 통합관리하기 위한 체계적인 메커니즘으로 데이터 조직관리, 메타데이터 관리, 프라이버시, 마스터데이터 통합, 데이터 품질관리, 비즈니스와 프로세스 통합, 데이터 수명관리, 데이터 표준화를 말한다.

조직관리는 데이터과학자를 빅데이터 거버넌스 위원회에 포함하여 제도와 규정을 보완하는 것이 중요하고, 메타데이터는 물관리 데이터의 다양성과 복잡성이 증가하고 있어서 데이터 종류와 출처 등을 체계적으로 관리해야 할 필요성이 증가하는 것이다.

빅데이터의 프라이버시 문제는 빅브라더 방식을 위해 중요한 이슈이다. 특히, 관측데이터의 경우 각 고객 가정의 프라이버시가 노출될 수 있으므로 비식별화는 물론 해킹 등에 대비해서 관리해야 한다.

데이터 품질관리는 데이터가 많아지고, 빅데이터 분석결과를 중요한 의사결정에 활용하는 사례가 증가함에 따라 더욱 중요성이 강조되고 있다.

기존의 마스터 데이터 관리에서도 일반적인 고객 데이터나 기타 마스터 데이터에 소셜 네트워크 정보나 관측정보를 추가로 활용함으로써 마스터데이터의 가치를 증진시켜야 한다.

빅데이터 분석결과는 비즈니스를 통해 이익을 창출하고 규제를 준수하게 되며, 관리비용을 절감하는데 활용되어야 하고, 그 효과를 측정할 수 있어야 하며, 지속적인 평가와 모니터링을 통해 최적화되어 나가야 한다.

폭증하는 데이터의 보관과 폐기정책을 통해 데이터 관리비용을 절감하도록 데이터 수명을 관리하고, 빅데이터의 수집, 저장, 공유를 위해 데이터 표준화를 체계적으로 실행해야 한다.

VI. 관측정보를 재해재난에 활용한 거버넌스 모델

본 절에서는 임실군 지역을 대상으로 스마트 물관리 빅데이터 거버넌스 프레임워크를 적용한 결과를 소개한다. 거버넌스 프레임워크의 모든 영역을 다루는 것은 너무 방대하므로 (X축 3개 * Y축 6개 * Z축 8개 = 144 항목이 존재함) 여기서는 (관측, 재해재난, 품질관리/표준화/메타데이터) 중심으로 연구하였다.

4.1 관측 데이터의 유형

임실군 지역에는 국가하천 1개소, 지방하천 38개소, 소하천 317개소이며, 주요하천으로는 섬진강, 오수천, 임실천, 둔남천, 유천 등이 있다. 임실군에 위치하고 있는 용담댐과 섬진강댐에 대한 설치된 우량관측소는 총 36개소이며, 수위 관측소는 16개소, 경보시설은 4개소이다. 여러 관측소에서 우량, 수위는 실시간으로 수집하여 저장하고, 경보시설은 경보 기준에 맞게 원격 방송하여 국민이 대피하도록 운영하고 있다.

4.2 재해 재난 업무

섬진강 지역을 대상으로 홍수 모니터링이라는 재해재난을 예측하는 문제를 실시간 관측정보와 과거 축적된 데이터 및 하천유역 정보를 결합하여 수행하였다.

홍수 모니터링 분석을 위해 활용되는 관측정보는 섬진강 지역에서 자체 운영하는 관측소를 포함하여 국토교통부 홍수통제소, 해당 지방자치단체, 기상청 자동기상관측소 등 유관기관과

우량, 수위, 영상, 기상 정보를 통합하여 재해재난 예측에 활용하고 있으며, 통합된 빅데이터는 데이터의 신뢰성을 증가시킨다.

홍수모니터링 기준은 강우강도와 수위상승 정도를 조합한 기준을 적용하였다. 강우강도는 강우의 불확실성을 감안하여 과거 관측 자료를 분석하여 예비정보 기준으로 활용하고, 수위상승 기준은 현장에서 계측되는 수위 값을 기준으로 직접 알람 기준으로 활용하였다.

4.3 데이터 거버넌스 원칙

4.3.1 메타 데이터

임실군 지역에서 홍수 모니터링에 관련된 데이터의 메타 데이터를 수집하고, 이를 저장 관리하는 것이 중요한 과제이다. 관측장비의 경우 대부분 IoT 센서들이므로 이들에 대한 메타 데이터를 구축해야 한다. 센서들이 메이커나 규격 및 정확도 등에 관한 정보를 제조사 혹은 장비 매뉴얼 등으로부터 확보하여 메타 데이터로 축적해야 한다. 또한, 각 하천별로 메타 데이터를 풍부하게 수집하여 관리함으로써 홍수 모니터링 정보와 결합하여 예측의 정확도를 높이는데 활용되어야 할 것이다.

4.3.2 데이터 품질관리

홍수 모니터링과 예측에 활용되는 데이터는 분석결과가 사람의 안전과 관련되므로 실시간 성과 정확도가 매우 중요하다. 대규모로 유입되는 실시간 정보의 품질을 높이는 작업은 상당한 노력과 비용이 들어가지만 재해재난 방지를 위해 필요한 것이다.

빅데이터에 대한 품질관리를 사람의 수작업에만 맡길 수는 없다. 각 장비의 특성을 고려하고, 해당분야 전문가들과 협의하여 관측장비가 생성하는 데이터의 품질 기준을 마련하고, 데이터 수집단계에서 자동으로 품질을 검증하고 보

정하는 작업이 이루어지는 것이 바람직하다. 이를 위해서는 품질기준이나 기기의 특성 및 관련 분야의 지식 등을 메타 데이터로 저장하는 것이 중요하다. 최근의 빅데이터 수집기나 DBMS들은 대부분 품질관리를 자동적으로 수행하는 기능이 내재되어 있으므로 이를 적극적으로 활용함으로써 실시간 대용량 데이터의 품질을 자동으로 (혹은 반자동으로) 개선할 수 있다.

4.3.3 표준화

수년간에 걸쳐 센서 등을 강 유역에 부착하여 관측데이터를 수집할 때 다양한 (버전이나 제조사 혹은 가격이나 품질 등의 측면에서) 센서가 현장에 부착될 수 있다. 이러한 환경에서 수집되는 데이터는 품질이 상이할 수 있으며, 그 기능도 상이할 수 있고, 기상 악조건에서는 작동하지 않을 수도 있다. 또한, 홍수 모니터링에 사용되는 데이터가 여러 기관에서 수집되는 경우에도 각종 코드 정보나 품질 등이 상이한 경우가 발생한다.

데이터 거버넌스 위원회에서는 이러한 점을 감안하여 각종 기기정보의 표준화는 물론 관련 기관 정보의 표준화를 통해 실시간 홍수 모니터링과 예측을 통한 재해재난 예방이라는 공동의 목표를 달성하도록 평가, 지시, 모니터링 기능을

강화해야 한다.

관측 데이터 수집과 연계 통합 후 분석된 결과는 최종적으로 경보 기준과 비교하여 경보를 발령하는데 활용되어야 한다. 경보 기준은 하천 유역의 지형적인 특성과 인구통계학적 특성을 감안하여 표 6과 같이 세밀하게 정의되어야 하며, 과거 축적된 데이터와 전문가의 노하우가 결합되어 만들어진다. 이러한 데이터도 표준화 후 메타 데이터로 축적하여 지속 개선되어 나가야 할 것이다.

V. 결론

빅데이터의 활용이 강조될수록 빅데이터에 대한 체계적인 관리(거버넌스)가 중요한 문제이다. 실시간으로 쏟아지는 빅데이터의 분석 결과로부터 미래에 관한 통찰력을 얻고, 중요한 의사결정에 활용하기 위해서는 데이터의 품질보장, 프라이버시 보호, 데이터 수명관리, 데이터 소유 및 관리권의 명확화, 조직의 역할 재정립 등이 함께 이루어져야 한다.

본 연구는 물관리 빅데이터 거버넌스 구축모델을 제시하고 특정지역에 대하여 실증적 연구를 수행했다. 스마트 물관리 분야에서 빅데이터의 활용사례가 급증하고 있으며, 이에 따라 지속적인 빅데이터 활용과 고도화를 위해 빅데이터 거버넌스의 구축이 중요한 과제이다. 본 연구에서는 스마트 물관리에서 빅데이터 거버넌스 구축을 위한 프레임워크를 3차원 큐브 형태로 제시하고, 물관리 부문의 데이터 유형과 물관리 데이터 산업 및 데이터 거버넌스 원칙을 각각 X, Y, Z 축으로 모델링하였다. 실증적 연구를 위해 임실군 지역을 대상으로 재해재난 업무를 빅데이터 활용으로 해결하는 업무에서 빅데이터 거버넌스 구축 방안을 제시하였다. 향후 과제로는 3차원 프레임워크의 각 요소(cell)을 채워나가는 방식으로 각 데이터 유형별로

〈표 6〉 홍수재해 모니터링 경보 기준

경보 기준	세부 내용	관촌교 경보 기준
관심	평시 수위보다 1.0m 상승 또는 고수부지 수위	1.70m
주의	제방 설계홍수량에 대한 50% 도달 또는 보행자 통행제한 수위	3.55m
경계	제방 설계홍수량에 대한 70% 도달 또는 차량 통행제한 수위	4.09m
심각	제방 설계홍수량 유하 시 계획홍수위 (여유고 0.5~1.0m)	4.76m
제방 월류	하천수위가 해당지점 제방을 월류하는 수위	7.26m

산업별로 빅데이터 거버넌스를 완성해 나가야 할 것이다.

참 고 문 헌

- [1] 강만모, 김상락, 박상무(2012), “빅데이터 분석과 활용”, 정보과학회, 30(6), 25-32.
- [2] 김승현, 박재홍, 김이환, 박주석(2015), “국내기업을 위한 빅데이터 적용방안 도출 및 기대효과 분석”, 정보기술아키텍처연구회, 12(1), 159-170.
- [3] 김신곤, 조재희(2013), “지방자치단체의 빅데이터 도입을 위한 제안”, 한국지역정보학회지, 13-41.
- [4] 김옥기(2014), 데이터 과학, 어떻게 기업을 바꾸었나?, 이지스퍼블리싱.
- [5] 김용연, 이경희, 조완섭(2017), “TTS 빅데이터를 활용한 도시 교통네트워크 구조분석”, 한국빅데이터학회, 2(2), 1-7.
- [6] 김진섭(2006), “데이터표준화 사례를 통한 데이터 품질 향상에 대한 연구”, 한국정보과학회, 33(2c), 210-213.
- [7] 김진영(2015), “물관리에서 빅데이터 활용 가능성 탐구”, 물정책과 경제, 24, 59-75.
- [8] 배덕효, 이병주(2011), “대우역 홍수예측을 위한 연속형 강우유출 모형 개발”, 한국수자원학회, 51-64.
- [9] 신철균, 조완희(2017), “임실군 홍수모니터링 기준 수립”, 한국수자원공사
- [10] 이상호(2015), “글로벌 경쟁, 스마트 물 서비스”, 한국수자원공사, 5-14.
- [11] 정동원 외 4인(2016), “빅데이터 센싱 객체 메타데이터 관리”, 전자정보연구정보센터, 23, 804-807.
- [12] 정승호, 정덕훈(2013), “공공기관의 데이터 품질에 영향을 미치는 요인에 관한 연구”, 정보처리학회, 251-266.
- [13] 최영환, 김영렬(2017), “IoT 기술을 활용한 스마트 물관리 필요성에 관한 연구”, 한국산업정보학회, 11-18.
- [14] 최영환, 조완섭(2018), “열화상이미지와 딥러닝을 활용한 자동 누수탐사기술 연구”, 한국정보기술학회, 1-9.
- [15] 홍석민, 장압(2017), “ICT 기반의 물정보 통합관리시스템 개발 연구”, 대한환경공학회, 723-732.
- [16] K-water 통합물정보처(2017), K-water 전사 데이터 관리전략. K-water
- [17] Sunil Soares(조완섭, 우준식, 신정길, 조재용, 김상하, 주재은 옮김)(2014), 빅데이터 거버넌스, 홍릉과학출판사.

저 자 소 개



최 영 환(Young-Hwan Choi)

- 1993년 : 경상대학교 전자공학과 (학사)
- 2006년 : 충남대학교 전자정보통신공학과 (석사)
- 관심분야 : 빅데이터, 거버넌스, 스마트 물관리, 사물인터넷



조 완 섭(Wan-Sup Cho)

- 1997년 ~ 현재 : 충북대학교 경영정보학과 교수
- 2001~2002 : DB Center, Univ. of Florida Post-Doc. 연구원
- 1996년 2월 : KAIST 전산학 (박사)
- 관심분야 : 데이터베이스, 비즈니스 인텔리전스, 빅데이터, 블록체인



이 경 희(Kyung-Hee Lee)

- 2008년 ~ 현재 : 충북대학교
빅데이터학과 교수
- 2004년 : 충북대학교 컴퓨터
과학과 (박사)
- 관심분야 : 빅데이터, 알고리
즘, 데이터 마이닝