

국내 전력산업에서의 빅데이터 플랫폼 성과 평가 방법론

Methodology for Evaluating Big Data Platforms Performance in the Domestic Electronic Power Industry

조치선¹ · 이난규¹ · 함유근^{2*}

(주)엑셈¹, 건국대학교 경영대학 경영학과²

요약

국내 전력산업이 스마트 그리드화 되면서 이로 인해 발생하는 빅데이터를 활용하여 수요관리, 시설물관리, 대고객서비스 등을 위한 빅데이터 플랫폼들이 도입되고 있는 추세이다. 그러나 빅데이터 프로젝트의 속성상 실제로 빅데이터 플랫폼의 활용이 업무 프로세스 상에서 정착되기 위해서는 많은 시간과 업데이트가 필요하다. 따라서 기존에 알려져 있거나 이론적인 평가 방법으로는 초기 빅데이터 플랫폼의 성과를 평가하기는 적절하지 않다. 본 논문에서는 빅데이터의 규모, 다양성, 속도에 따른 정보의 완전성/충분성, 정보의 신뢰성/정확성, 정보의 적합성/관련성, 정보의 상세성/구체성, 정보의 비교가능성, 정보의 불편성, 정보의 적시성 등 특정 정보의 7 가지 품질 측면에서 전력산업에서 초기 빅데이터 플랫폼의 성과를 평가하는 방법론을 제시한다.

■ 중심어 : 빅데이터 플랫폼, 성과 평가 방법론, 전력산업, 정보 품질

Abstract

As the domestic electric power industry becomes a smart grid, big data platforms for demand management, facility management, and customer service have been deployed. However, due to the nature of the big data project, big data platforms take time to realize their value in the business processes. Therefore, it is not easy to evaluate the performance of the initial big data platforms using the known or theoretical evaluation methods. In this paper, we propose a methodology of big data platform performance evaluation based on specific information quality such as information completeness/sufficiency, information reliability, information relevancy, information comparability, information unbiasedness, timeliness of information, related to the volume, diversity, and velocity of big data.

■ Keyword : Big Data Platform, Performance Evaluation Methodology, Power Industry, Information Quality

I. 서론

전력산업은 관련 설비/기기들 간에 네트워크로 연결되어 하나의 지능화된 그리드(Grid)를 이루면서 발전 중이다. 해외에서는 이에 속한 자산들이 상호호환되면서도 업무 영역별로 단절된 데이터의 흐름을 극복하고 전력 그리드의 성능을 높이는 인텔리전스 제공을 위한 분산형 빅데이터 플랫폼들이 도입되는 추세이다[1]. 국내에서도 현재의 전력산업 환경에서는 흩어져 있는 자산은 물론, 데이터 및 통제 관리들을 여러 플랫폼을 통해 공유하여 상호호환성을 높이고 비용을 절감하여 문제에 신속히 대응하는 것이 중요한 목표가 되고 있다. 예를 들어 설비/기기들 간에 P2P(peer-to-peer) 커뮤니케이션과 지금까지 독립적으로 운영되던 시스템들 간의 데이터 통합으로 완성된 커뮤니케이션 노드와 표준 메시징 아키텍처로 분산형 플랫폼에서의 전압 관리 기능을 스마트화하는 것도 이러한 빅데이터 플랫폼 구축으로 가능하게 된다[2].

국내외 전력산업에서는 각 지역에 설치된 설비/기기들의 센서에서 발생하는 데이터들을 지역적으로 수집해 통합하거나 다양한 데이터 스트림들을 본사로 모두 보내서 이런 데이터를 한 곳에 모아 통합하는 시스템들이 구축되고 있다 [3].¹⁾ 이에 따라 본사에서 중앙집중화된 데이터 통합을 위해서 표준화된 인터페이스 개발의 필요성도 높아지고 있다. 데이터들은 중앙에서 취합하는 것 이외에도 지역적인 통합관리가 필요한데 그 이유는 지역에서 발생한 문제 해결은 지역적 관리가 보다 신속한 대응 방안이 되기 때문이다.

전력산업에서 빅데이터의 활용 가능성에 따라 관련 투자가 증가하면서 과연 빅데이터 플랫폼 구축이 적절한 성과를 내고 있는지에 대한 우려도 높아지고 있다. 하지만 지금까지 개발에

대한 논의에 비춰어 평가에 대한 논의는 매우 부족한 것이 현실이다. 전력산업에서 빅데이터를 보다 효율적이고 효과적으로 활용하기 위해서는 이제 빅데이터 플랫폼에 대한 적절한 평가 방법에 대해 고민해야 할 시점이다. 이에 본 논문에서는 이론적 바탕 하에 현실적으로 초기에 있는 국내 전력산업의 빅데이터 플랫폼 평가 방법을 제시하고자 한다.

II. 문헌 연구

2.1 전력산업 빅데이터 플랫폼

전력산업에서 사용되는 빅데이터 플랫폼은 기존 데이터 이외에 외부 데이터 및 비정형 데이터 (특히 관련 장비로(예를 들어 AMI(Advanced Metering Infrastructure))부터의 센서 데이터) 등을 통합해 분석 도구를 이용하여 각종 발전 및 배송전 관련 관리와 대고객 서비스에 사용되는 시스템을 말한다[4][5][6].

〈표 1〉 전력산업에서 빅데이터 분석 도구 활용 예상 시나리오[7]

- 고객 수준에서 수요 대응의 계산 및 검증
- 최종 소비자의 소비 분석을 통한 전력 도난/사기 탐지
- 미터기에서부터 현금 수납까지 분석
- 소비자 사용 패턴 분석
- 규제감독 당국에 서비스 신뢰성 및 성과 보고
- 변압기 과부하 탐지 및 서킷 분석
- 예측적이고 예방적인 유지보수로 자산 수명의 연장
- 고객군(주거용 및 상업용)별 부하(load) 프로파일 특징화 및 정의
- 지역별 전력 공급 체계 통합에 의한 업무 부담 경감 평가를 위한 계산, 검증, 및 감사
- 부하 감소 프로그램을 위한 공급 능력 한계치 예측 (수요 대응 및 예측가능한 부하 관리 프로그램)

이러한 플랫폼은 잠재적으로 전력수요공급관

1) <https://blog.kepco.co.kr/1502>

리, 수익관리, 시설관리, 대고객 서비스, 전력부정사용감시, 법규준수 등에 사용될 수 있다 [7]. 전력수요공급관리를 위해서는 수요 예측 및 공급 역량 관리가 필요하다. 정확한 예측은 적시적이며 경제적인 수요 및 공급 관리에 필수적이다. 전력회사는 전력 소비, 날씨, 및 공급 제한 요인들에 대한 데이터를 이용할 경우 보다 효과적으로 전체 운영을 관리할 수 있다.

수익관리 측면에서 전력회사는 에너지 사용 관련 빅데이터를 이용해 소비자들이 에너지를 효율적으로 사용하고 절약하도록 하는 교육 프로그램을 개발하고, 테스트하며, 모니터링 할 수 있다 [8]. 이런 활동들은 적시적인 분석과 데이터를 기반으로 한 인사이트가 필요하다. 모든 전력미터기들이 사용된 전력 규모를 정확히 반영하도록 하는 것은 수익 극대화를 위해 기본적인 사항이다. 즉, 전력회사들에게 능동적인 프로세스 모니터링은 청구되지 않은 전력 사용 및 덜 청구된 미터기를 파악하기 위한 핵심 적인 요소이다.

법규 준수도 중요한 빅데이터 플랫폼 활용 분야이다 [9]. 법규적인 준수사항 이행을 위해 운영 지표에 대한 일관적이고 완전하며 정확한 리포팅을 하는데 빅데이터 플랫폼이 사용될 수 있다. 이런 부분에 소홀할 경우 범칙금을 유발할 수 있다.

자산 유지보수 및 관리는 전력 그리드와 전력 공급 장비의 잠재적 문제를 발견하는 능력은 고가 자산의 수명을 연장시키고 예상하지 못한 서비스 중단을 방지할 수 있다 [9].

고객 케어 및 관리 측면에서 전력 사용자들은 앱을 통해 기본적인 청구 내역, 결제, 과거 사용 내역 등 이상을 전력공급사가 인터랙티브하게 서비스하기를 기대하고 있다 [10]. 전력공급사는 빅데이터를 이용해 고객에게 맞춤형 요금제를 개발하고 에너지를 효율적으로 사용한 고객에게 인센티브를 제공하는 방안을 마련할 수 있다.

사기 및 손실 방지도 빅데이터 플랫폼이 사용될 수 있는 매우 중요한 분야이다 [11]. 전력회사는 사기나 도난으로 손실되는 전력 탐지를 위해 전력 사용 현황에 관한 거의 실시간 정보가 필요하다. 이러한 능력의 확보는 재무적인 실적 개선에 상당한 기여를 한다.

2.2 빅데이터 성과 평가

성과 평가란 사전에 정의된 (혹은 기대되는) 목표의 달성 여부를 평가하는 과정으로 결과(사용자 및 경영자 시각)에 중점을 둘 수도 있고(즉, 효과성 중심), 과정에 중점(관리자 시각)을 둘 수도 있다(즉, 효율성 중심) [12]. 투입된 자원들이 원하는 결과물로 변환되는 효율성을 측정하거나(관리자 시각) 산출물의 품질(얼마나 고객들에게 잘 제공되었고 이들이 얼마나 만족하는지) 및 결과물이 의도했던 목적에 맞는 프로젝트의 활동 결과인지 여부(사용자 및 경영자 시각)를 평가할 수 있으며 목표와 상황에 따라 두 가지 방향 모두 혹은 하나만을 선택할 수도 있다.

2.3 빅데이터의 효과

2.3.1 기존적인 효과

기존에는 다차원 분석 도구, 대시 보드, 정형화 도구, 정형화된 리포팅 도구 등 많은 방법이 있지만 이들에 대해 사용자들은 데이터 검색, 데이터 확장성, 데이터 품질, 데이터 통합, 실시간 정보 제공 등에 불만이 있었다 [13]. 즉, 데이터의 품질 이외에 시차차를 가지는 데이터 분석과 통합되지 않은 데이터, 분석 범위의 협소 등 기존 데이터 분석의 한계는 기본적으로 빅데이터가 제공되어야 해결 될 수 있는 문제이다. 따라서 일반적인 데이터 분석 결과와 다르게 빅데이터 플랫폼의 성과는 빅데이터의 개념인 데이터의 규모, 데이터의 속도, 데이터의 종류 등 3V 과 관련이 깊다. 그 배경을 좀 더 살펴보자.

2.3.2 규모(Volume)로 인한 효과

데이터의 규모는 보통 원하는 결과를 얻기 위해 사용되고 분석되는 데이터의 양을 의미한다 [14]. 당연히 빅데이터에서 규모는 이러한 결과를 얻기 위해 적합한 데이터가 충분하다는 뜻이다. 기업들은 데이터 호수(Data Lake) 등을 구축하여 비용 대비 효과적으로 가능한 한 많은 데이터를 저장하고 다양한 데이터 주제 영역에 대해 더욱 폭넓은 분석을 할 수 있으며 또한 수년 동안의 과거 기록들을 토대로 더욱 깊이 있는 분석도 가능해진다[15].

빅데이터 플랫폼은 단지 현재의 데이터 분석 수요를 감당하는데 그치지 않고 더 많은 사용자, 더 많은 데이터 활용 및 더 많은 질의를 가능하게 한다. 에너지 산업의 경우, 예를 들어, 1년 동안의 정전 관련 데이터를 저장하신 대신에 과거 30년 동안의 데이터를 저장하면 시간 변수와 관련된 유지보수 문제들을 보다 쉽고 정확히 분석할 수 있다[4].

기본적으로 빅데이터로 표본을 대상으로 데이터 분석을 할 필요가 없으며 모집단 전체를 대상으로 분석을 행할 수 있다. 예를 들어 모든 전력 설비의 잠재적인 고장 원인들에 대한 분석이 가능하다. 또한 고객에 대한 전방위적인 시각을 확보할 수 있어 신뢰성 있는 고객 프로파일 작성이 가능하며 상세한 수준까지 고객 니즈나 선호도를 파악해 고객을 이해하고 고객이 제기한 불만/문제에 보다 효과적으로 대응이 가능하다[16]. 고객에 대해 더 많은 데이터를 확보할수록 이들에 대한 인사이트 더욱 많아진다. 그리고 이는 새로운 고객 관계를 만들고 기존 고객 관계를 유지하며 향상시키는 의사결정을 더욱 효과적으로 행할 수 있게 한다.

2.3.3 속도(Velocity)로 인한 효과

빅데이터 플랫폼으로부터, 그리고 일반적으로 가장 많은 기대 효과를 예상하는 부분은 속

도(즉 빠른 검색, 정보 제공, 분석, 해결책 등)와 비용 절감 부분이다[17]. 새로운 빅데이터 플랫폼이 기존 시스템보다 더 빠르고, 더 효과적으로, 그리고 더 적은 비용으로 과업을 수행하기를 기대한다. 예를 들어 어떤 정보를 원할 경우 이전에는 얻지도 못했으며 얻는다고 해도 수일이 걸렸던 것이 이제는 수분, 수초로 단축된다.

한 조사에 따르면 기업들이 빅데이터로 성과를 보는 것 중 가장 큰 비중을 차지하는 부분이 데이터를 실시간으로 활용할 수 있기 때문으로 나타났다[13]. 즉 빅데이터 플랫폼의 도입으로 데이터를 수집하여 활용하는데 소요되는 시간이 단축되면서 조직은 빅데이터의 효과를 (의사결정 혹은 문제 발견의 신속화 등) 여러 면에서 볼 수 있다. 실시간으로 데이터가 관리되고 관련 검색이 신속히 이루어지는 부분은 시간적으로 관련성이 높은 정보를 제공할 수 있어 정보의 적시성과 적합성을 향상시킨다고 할 수 있다.

더 신속하게 데이터를 수집하고 분석해 업무에 활용할수록 빅데이터 플랫폼은 더 많은 사용자의 질문을 처리해 원하는 답을 구할 수 있다. 신속하게 분석할 수 있는 능력이 있으면 기관은 적절한 의사결정은 적절한 시점에 할 수 있는데 최근 들어 이러한 능력은 대고객 서비스와 관련해서 중요성이 커지고 있다(예를 들어 급작스러운 수요 증가에 대응). 즉 문제 해결 시간의 경우 데이터 속도는 매우 중요한 문제 해결 요소로 이런 문제 유형에서 빅데이터의 가치는 점차 더 높아지고 있는 추세이다. 특히 효율적인 자원 관리를 위한 재고 파악 등은 정보의 신뢰성과 상세성, 적시성 등이 중요하다[18].

보통 기업에서 많은 데이터를 분석하는 경우 데이터웨어하우스를 구축하지만 데이터의 규모는 물론 비정형 데이터까지 수집하고 데이터 수집과 분석의 속도를 중시할 경우 한계에 직면한다[19]. 분석 결과를 얻는데 시간도 오래 걸리고 이에 따라 사용자들의 불만은 높아지며 결국은

데이터를 활용하는 것 자체가 이들에게 외면당할 수 있다. 현재 국내외에서도 많은 기관/기업들이 데이터웨어하우스를 구축하고 있지만 빅데이터를 활용하는 데는 성능과 확장성 면에서 한계를 보이고 있다.²⁾ 결국 빅데이터 플랫폼을 구축하는 목적도 이런 병목 현상을 극복하고 사용자가 필요할 때 데이터 분석 및 그 결과를 제공하는데 있다.

2.3.4 다양성(Variety)으로 인한 효과

더욱 다양한 형태의, 그리고 더욱 다양한 소스로부터의 데이터를 사용할 수 있는 빅데이터의 다양성은 정보의 완전성과 관련성, 그리고 불편성(unbiasedness)을 높이는데 매우 중요하다 [20][21]. 빅데이터의 다양함으로 인해 정보의 형태와 내용 면에서 과업에 보다 관련성이 높고 더 많은 변수들을 고려할 수 있는 정보(즉, 정보의 완전성)를 제공할 수 있다. 특히 문제의 파악(예를 들어 정전) 및 원인 탐색의 단계에서 데이터의 소스가 다양할수록 더 많은 가설들을 설정할 수 있어(예를 들어 변전소의 기계적 고장이나 기상 관련 등) 대안 창출에도 유리하다.

특히 기존에는 접근할 수 없었던 고객의 반응이나 행태에 대한 데이터는 고객 관련 의사결정의 질을 높일 수 있게 한다. 예를 들어 소셜 미디어 데이터는 고객이 무엇을 좋아하며, 더욱 중요하게 무엇을 싫어하는지 말해준다[22]. 그리고 소셜 미디어 데이터는 고객의 떠날 가능성을 탐지하는 조기 정보 시스템에 기여할 수 있다. 전력 산업도 고객의 요금에 대한 불만 등을 모니터링하는데 소셜 미디어 데이터의 활용이 가능하다.

물론 이러한 데이터 다양성의 장점을 살리기는 것은 데이터의 통합이 적절히 달성되었음을 전제로 한다. 데이터 소스와 형태가 다양해지는 것은 그 만큼 데이터를 상호연결하고 호환되게

하기 전에는 의미 있는 정보를 만들기 어렵다는 의미이기도 하다[23]. 이런 한계를 극복한 경우에만 실제로 실용 가능한 정보를 생산해 낼 수 있다.

III. 평가 방법 개발

3.1 전통적 IT 프로젝트 평가 방법

지금까지 IT 프로젝트의 성과를 측정하는 방법으로는 경제적 효과 및 사용자 만족도를 측정하는 방법들이 일반적이었다. 투자수익률, 내부수익률 등 재무지표를 이용한 방법이 가장 이상적이나 IT 프로젝트가 실제로 효과를 보이는 데는 시간의 경과가 필요하기 때문에 프로젝트의 완료 시점에서 경제적 효과를 평가하는 것은 쉽지 않다. 반면에 사용자 중심의 평가 방법은 이런 시간적 요인에 덜 영향을 받기에 다양한 IT 프로젝트에서 사용되고 있다. 그리고 이전 IT 프로젝트들이 IT 인프라와 관련된 전산부서 업무에 집중되었으나 조직 내 IT 활용이 확산되면서 현업부서 사용자를 위한 IT 프로젝트들이 증가하였고 그 결과 사용자 중심의 평가 방법들이 나타나기 시작했다. 사용자 만족도 측정도 기존 평가 방법에서 일반적으로 사용되는 방법이다.

정보시스템의 성공을 관련된 선행연구들은 정보시스템의 성공을 그 개념의 범위에 따라 각자 나름대로의 다양한 정의를 제시하고 있다. 일반적으로 정보시스템의 성공(IS Success)은 정보 시스템에서 사용자 만족에 영향을 미치는 품질로 시스템의 정확성과 효율성 등 기술적 수준을 의미한다[24]. 그 외에도 시스템의 온라인 처리시간 혹은 반응속도(응답성), 안정성, 신뢰성, 시스템 정확성, 반응시간, 반환시간, 데이터 정확성, 완전성 등 매우 다양한 측면을 평가하기도 한다[25].

2) http://www.dt.co.kr/contents.html?article_no=2013101002011060718002

최근 연구에서는 DeLone & McLean[26]의 업데이트된 연구 모델을 사용해 성과 평가를 하는 추세이다[27]. 이 모델에 따르면 빅데이터 플랫폼을 평가하는 방법을 세 가지로 정리할 수 있다.

첫 번째 방법은 “순수 효과(NET BENEFITS)”를 평가하는 것으로 이는 빅데이터 플랫폼의 비즈니스적인 가치를 평가하는 것이다. “순수 효과”는 가장 진정한 성과를 평가하는 것이지만 실제로 빅데이터 플랫폼이, 예를 들어 기관의 비용 절감에 얼마나 기여했는지는 쉽게 파악되지 않기에 이를 직접 평가하기는 어렵다. 특히 빅데이터 프로젝트들은 효과가 조직에 직접 나타나는 데 기존 프로젝트들 보다 오랜 시간이 소요되기 때문에 이 방법으로 성과를 평가하는 것은 실질적으로는 효과적이지 않다.

두 번째는 잠재적 플랫폼 사용자들이 플랫폼을 실제로 사용했는가 혹은 사용한 의도가 있는가(“INTENTION TO USE” 혹은 “USE”)를 평가하는 방법이다(혹은 사용자 만족도로 성과를 평가하는 방법). 예를 들어 빅데이터 플랫폼을 얼마나 자주 사용하는지를 측정해 평가한다. 이런 방법은 빅데이터 플랫폼의 활용 목적과 활용 가능성이 사용자들에서 잘 알려져 있는 상황에서 활용될 있는 방법으로, 그렇지 않다면 이 방법 역시 빅데이터 플랫폼이 설치되었다고 해도 이의 활용도를 잠재적 사용자들이 알지 못해 성과가 나타나기 어려운 면이 있다.

세 번째 방법은 위 두 가지 방법의 한계로 인해 플랫폼 자체를 평가하는 방법으로 DeLone & McLean[25]의 모델에서는 (1) 정보의 품질(Information Quality), (2) 플랫폼을 사용하기가 얼마나 쉬운가(인터페이스 측면의 Service Quality), (3) 플랫폼이 얼마나 업무에 유용한가(즉 기능적인 측면의 System Quality)를 평가한다. 본 연구에서 초기에 있는 전략산업 빅데이터 플랫폼들을 고려해 세 번째 방법을 중심으로 성과 평가

방법을 도출한다.

정보의 품질은 사용하기 적절함(fitness for use)을 의미한다[29]. 정보의 품질에는 다양한 요소들이 있으나 일반적으로 적시성, 정확성(혹은 신뢰성), 업무 관련성(혹은 적합성), 완전성(혹은 충분성) 등 네 가지 측면을 정보 품질의 주요 요소로 볼 수 있다. 적시성은 정보 검색이나 데이터 분석이 신속하게 진행되어 필요한 시점 또는 시간 내에 정보가 제공됨을 의미한다. 정확성은 말 그대로 오류나 누락이 없으며 또한 정보의 편향성도 낮음을 의미한다. 그리고 업무 관련성은 정보의 내용이 업무와 관련됨을 의미하는데, 업무에 적용할 수 없는 정보가 많을 경우 관련성이 떨어진다. 예를 들어 특정 지역에 대한 전력 수요 분석을 원했으나 다른 지역이든지 종합적인 전국에 대한 분석 결과는 업무 관련성이 적다. 완전성은 요구했던 내용의 주요 이슈/주제들을 모두 커버했는지에 관한 기준으로 예를 들어 기기 유지보수 문제 분석을 위해 기계적 결합과 관련된 부분을 분석할 경우 모든 주요 기기들에 대한 정보 분석이 반영되었는지 여부가 이에 해당한다. 여기에 더해 빅데이터 분석에서 시각화의 중요성으로 인해 이해가능성도 완전성의 일부로 볼 수 있다.

이러한 네 가지 품질 측면에서 더해 의사결정에서 정보의 사용이 증가함에 따라 문제 파악과 대인 창출 등의 측면에서 정보의 불편성, 비교 가능성, 상세성/구체성 등도 중요한 요소로 부각되고 있다(<표 1>).

플랫폼 사용의 용이성은 조작의 용이성, 학습의 용이성, 유지보수(업그레이드 용이성), 접근의 용이성, 보안관리(개인정보 보호)의 용이성 등을 의미한다. 이 부분은 플랫폼의 목표 사용층을 어떻게 정하느냐에 따라 달라지는데 예를 들어 전문가 수준의 데이터 분석을 위한 플랫폼이라면 일반 사용자들에게는 조작의 용이성 평가 결과가 낮게 나올 수 있다.

〈표 1〉 정보 품질 요소들 ([28]을 토대로 저자들이 수정 및 보완)

정보 품질	내용
· 적시성	정보가 의사결정자에게 의미를 갖도록 적시에 제공될 수 있어야 한다.
· 충분성/완전성	의사결정을 지원하기에 충분한 크기의 데이터여야 한다.
· 구체성/상세성	기업 내 여러 부문의 상세데이터를 적절히 통합해 요약한 데이터여야 한다.
· 이해가능성	데이터는 의사결정자가 이해하기 쉬운 방법으로 표현되어야 한다.
· 불편성	데이터베이스의 구성에 설계자의 어떠한 편견도 개입되어서는 안 된다.
· 적합성/관련성	의사결정에 적합하지 않은 데이터가 데이터베이스에 있어서는 안 된다.
· 비교가능성	적합성이 확보된 다른 정보와 비교할 수 있어야 한다.
· 신뢰성/정확성	데이터의 무결성을 확인해야 한다.
· 중복성	중복데이터가 경우에 따라 요구되기도 한다.
· 비용효율성	데이터의 획득에 소요되는 비용과 그 가치에 대해 평가가 필요하다.
· 계량가능성	수치로 표현할 수 있는 수준까지 계량화가 필요하다.
· 표현의 적정성	데이터는 적절한 형식으로 표현되어야 한다.

플랫폼 자체의 유용성은 정보의 유용성이 아닌 플랫폼의 응답시간, 다양한 문제에 적용 가능한 유연성 등 기능의 유용성이다. 이 부분도 플랫폼의 원래 목표가 특정한 분야에 대한 문제 해결이라면 유연성 등이 높게 평가되기는 어렵다. 반면에 기관 내 다양한 사용자들의 다양한 문제를 해결하기 위해 구축된 플랫폼이라면 이 요인에 대한 평가의 가중치는 높아질 수 있다.

3.2 초기 빅데이터 플랫폼 성과 평가 방법

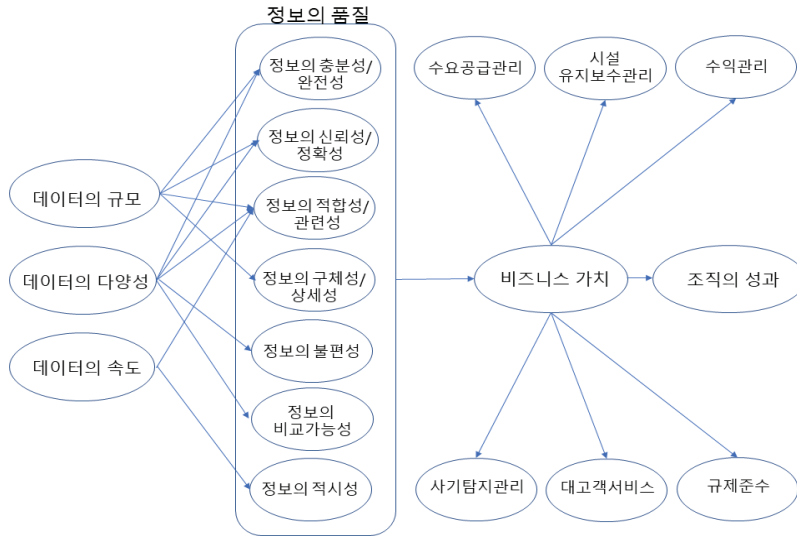
빅데이터 프로젝트들은 보통 조직 내 데이터의 소유 및 접근/관리 권한 등 복잡한 이해관계를 해결해 데이터를 통합하고 명확한 목표를 설정해야 하기 때문에 앞서 언급한 비즈니스 가치의 실현까지는 많은 어려움이 있다.³⁾ 이런 현실적인 문제로 최근 해외에서는 빅데이터 플랫폼의 평가 시 두 가지 부분에 중점을 두는 경향이 있다. 즉, 플랫폼 구축 전후로 바뀐 사용자들의 정보 접근/사용 행태의 변화, 혹은 정보에 대한 태도와 플랫폼 이용과 직접적으로 관련된 KPI(key performance indicator) 등이다[30]. 궁극

적으로 후자의 방법이 ‘평가’라는 목적에 부합하지만 실제로는 플랫폼의 구축 초기에는 시간적으로 이의 비즈니스 측면의 효과가 구현되기 어려워 이런 방법에는 어려움이 따른다. 최근에서는 빅데이터 플랫폼의 효과 실현 시간이 기존 시스템들보다 오래 걸린다는 지적도 있다.⁴⁾

이러한 두 가지 접근법을 앞서 살펴본 이론적 평가 방법과 비교해 보면 전자(정보 접근/사용 행태 변화 혹은 정보에 대한 태도)는 정보 품질에 중점을 둔 것으로 볼 수 있다. 물론 시스템 품질 측면도 관련되나 정보 품질에 비교가능성 등은 시스템 품질이 일부 반영된 요소라고 할 수 있다. 그리고 후자인 KPI를 사용하는 것은 이론적 방법론에서 ‘순수 효과’, 즉 비즈니스 가치를 평가하는 것이라고 하겠다. 예를 들어 ‘고객 만족도’는 KPI이자 고객에 대한 기여를 나타내는 비즈니스 가치이기도 하다. 그리고 앞서 언급한 빅데이터를 활용한 전략산업 관련 각종 업무와 관련된다. 반면에 현실에서 기업들은 이론적인 방법 중 ‘사용 의도’에 대한 관심은 낮은 것으로 보인다. 빅데이터 플랫폼의 활용이 전사적으로 정착되면 KPI를 통한 평가가 의미를 가질 수 있다. 즉, 예를 들어, 데이터가 고객 지원

3) <https://www.infoworld.com/article/3393467/4-reasons-big-data-projects-failand-4-ways-to-succeed.html>

4) <https://pubsonline.informs.org/doi/10.1287/orms.2019.06.08/full/>



<그림 1> 전력산업 빅데이터 플랫폼 성과 평가 모델

이나 판매와 같은 비즈니스 프로세스의 일부가 되면 조직은 고객 만족, 판매액 및 기타 사용 중인 지표를 측정할 수 있다. 이 부분은 <표 2>의 관련 업무의 향상을 평가하는 측정 도구로 사용될 수 있다. 빅데이터 플랫폼 활용 전과 후를 비교해 조직에 어떤 성과가 있는지를 평가할 수 있으며, 또한 데이터를 수집하고 활용해 어떤 기회를 가질 수 있을 지도 파악할 수 있다. 예측 분석 도구들은 적절한 시점에 고객의 요구/불만에 대응하고 판매 기회를 포착할 수 있도록 해 할인이나 관련 제품 제공으로 고객 만족이나 매출을 높일 수 있다.

본 연구에서는 국내 기업이나 기관에 실제 도입된 빅데이터 플랫폼들이 아직 초기에 있다는 전제 하에, 빅데이터의 활용과 보다 관련성이 직접적인 정보의 품질에 초점을 맞춘 방법론을 제시한다. 이런 방법론의 타당성은 앞서 빅데이터 플랫폼의 3V(데이터 규모, 속도, 다양성) 측면에서 효과를 이론적으로 논의한 부분과 업무 관련성이 높은 정보 품질 측면(7 가지)에 근거를 둔다. 다시 말해 <그림 1>의 방법론에 따르면 빅데이터 플랫폼의 도입 성과를 평가하기 위해서는 먼저 빅데이터의 규모와 다양성, 그리고 속

도의 적절한 도입을 전제로 해서, 그 효과를 정보의 품질 측면에서 평가하는 것이다.

빅데이터의 3V는 각기 다른 특성으로 인해 정보의 품질에 영향을 주는 부분이 다르다. 규모는 사용자가 해결을 원하는 문제에 대해 많은 부분을 커버하는 것은 물론 보다 상세하며 표본 크기의 확대로 신뢰성 있는 정보를 제공한다. 따라서 정보 품질 면에서 정보의 충분성, 신뢰성, 구체성, 적합성 등에 기여한다.

그리고 데이터의 다양성은 다양한 소스들로부터의 다양한 형태의 데이터로 인해 해결 문제에 대해, 특히 고객, 협력사/관계사 등 외부 소스들로부터의 데이터로, 다각도적인 시각을 제공하여 정보의 충분성, 적합성, 신뢰성, 그리고 불편성 및 비교가능성에 기여한다. 즉 내부에 존재하지 않은 다양한 외부 소스, 특히 최종 고객에 대한 데이터는 정보의 적합성에 크게 기여한다. 또한 외부로부터의 데이터는 내가 부족한 부분을 보완하기 때문에 정보의 완전성/충분성을 높인다. 더욱이 외부의 시각이 담긴 데이터는 문제 해결 과정에서 발생할 수 있는 있는 주관적 견해, 즉 편견을 줄이는데 기여한다.

<표 2> 전력산업 빅데이터와 관련된 정보 품질(Information Quality) 요소 및 관련 업무

3V	빅데이터의 차별성	관련 정보 품질 요소	주요 관련 업무
데이터 규모	<ul style="list-style-type: none"> · 관심 주제의 일부가 아닌 전체에 대한 데이터 · 오랜 기간 축적된 데이터 · 낮은 단위의 상세한 데이터 · 그 동안 저장하지 못했던 데이터 	<ul style="list-style-type: none"> · 정보의 충분성/완전성 · 정보의 신뢰성/정확성 · 정보의 구체성/상세성 · 정보의 적합성/관련성 	<ul style="list-style-type: none"> · 공급관리 · 수요관리 · 유지보수관리 · 수익관리 · 규제준수
데이터 다양성	<ul style="list-style-type: none"> · 다양한 내부 소스들로부터의 데이터 · 다양한 형태의 데이터(특히 시각적인 데이터) · 고객 데이터 · 협력사/관계사 데이터 · 공공 데이터 · 기타(기상, 위치, 결제 데이터 등) 	<ul style="list-style-type: none"> · 정보의 충분성/완전성 · 정보의 적합성/관련성 · 정보의 신뢰성/정확성 · 정보의 불편성 · 정보의 비교가능성 	<ul style="list-style-type: none"> · 공급관리 · 수요관리 · 유지보수관리 · 사기탐지관리 · 대고객서비스 · 수익관리 · 규제준수
데이터 속도	<ul style="list-style-type: none"> · 가장 최근의 데이터 · 짧아진 데이터 업데이트 주기 	<ul style="list-style-type: none"> · 정보의 적시성 · 정보의 적합성/관련성 	<ul style="list-style-type: none"> · 사기탐지관리 · 대고객서비스 · 유지보수관리

데이터의 속도는 실시간으로 수집되어 처리, 활용되는, 말 그대로 시기적절한 정보와 관련된다. 정보의 업데이트 주기를 단축하는 것도 데이터의 속도와 관련된다. 이는 보다 현 문제와 관련성이 높은 정보를 제공하여 정보의 적합성을 높일 수 있다. 그리고 데이터의 속도는 어떻게 데이터를 확보하느냐에 따라 결정되는 데이터의 규모나 다양성과 달리 기술적으로 해결되어야 하는 부분이다. 데이터의 규모와 다양성을 처리하기 위해서도 이를 위한 시스템이 필요하지만 특히 데이터 속도는 특히 기술적인 발전에 따라 크게 영향을 받는다. 사물인터넷 등 센서 데이터의 확산으로 데이터 속도에 의한 정보 품질 향상 기회가 증가하고 있다.

반면에 <표 1>의 정보 품질 요소들 중 이해가능성, 중복성, 비용효율성, 계량가능성 표현의 적절성은 수집한 데이터와 관련되기 보다는 시스템의 설계와 분석도구와 보다 더 관련되는 것으로 판단된다. 따라서 이들 요소는 빅데이터와 관련된 정보의 품질 향상을 평가하는 경우에서 제외하였다.

비즈니스 가치는 <표 2>의 관련 업무의 향상

부분으로 수요공급관리, 시설유지보수관리, 수익관리, 사기탐지관리, 대고객서비스, 규제준수 등의 향상을 뜻하며 조직의 KPI와 관련된다. 마지막으로 조직의 성과 ROI 등 재무적인 성과를 뜻한다.

무엇보다도 이러한 평가 방법이 의미를 갖는 것은 빅데이터를 실제로 유용한 정보로 변환하기까지의 과정이 대단히 복잡하고 어렵기 때문이다. 비정형 데이터를 실시간으로 처리해야 하는 부분은 물론 여러 데이터 원천으로부터 데이터를 통합하는 과정은 결코 단순하지 않다. 더욱이 각기 다른 데이터의 단위들을 표준화하여 비교 가능하게 정리하는 과정도 복잡한 작업이다. 따라서 빅데이터가 의미 있는 정보로 전환되었다고 한다면 그 자체만으로도 충분히 평가받아야 한다. 그런 의미에서 실질적으로 빅데이터를 활용하였으며 이를 사용자들이 필요로 하는 수준의 정보 품질로 구현하였는가는 빅데이터 플랫폼의 성공 여부를 결정한다. 본 평가 방법은 이와 같은 배경에 따라 제한된 상황이지만 현실적인 측면에서 의미 있게 사용될 수 있는 방법이다.

IV. 결 론

대부분의 국내에서 빅데이터 플랫폼을 구축한 조직의 경우 아직은 그 활용이 초기 단계에 있다. 전력산업도 마찬가지이다. 핀테크, 전자상거래, 게임 등 인터넷기업이나 일부 금융회사를 제외하고는 대부분의 기업이나 기관들은 빅데이터에 기반을 둔 다양한 예측 모델을 통해 예방적인 관리를 하는 차원이나 이를 이용한 새로운 고객 서비스를 제공하는 수준에는 못 미치고 있다.

빅데이터 투자의 타당성을 계속 확보하기 위해서는 이러한 현실적인 상황을 고려한 빅데이터 플랫폼의 평가 방법이 필요하다. 본 논문에서는 빅데이터의 규모, 다양성, 속도와 관련된 정보의 품질에 바탕을 둔, 즉 플랫폼에 사용되는 빅데이터 자체를 중심으로 한 평가 방법을 제안하였다. 다시 말해 본 논문에서는 빅데이터를 접근해 검색, 분석할 수 있는 적절한 플랫폼을 전제로 이를 평가하는 방법론을 제안하였다, 그리고 그 과정에서 몇 가지 특정 정보의 품질 측면이 반영되어야 함을 주장하였다.

이는 단순히 시스템의 기술적인 수준이나 AI 등 첨단 분석 도구에 초점을 두기 보다는 순수하게 빅데이터의 활용과 관련된 직접적인 단기성과를 중심으로 한 평가 방법이다. 따라서 조직의 실질적인 경영성과에 대한 빅데이터 플랫폼의 기여도를 평가하는 데는 한계가 있는 방법이다. 하지만 현실적인 상황에서 빅데이터 플랫폼에 대한 실질적인 평가가 필요한 경우 경영성과 위주의 빅데이터 플랫폼 기여도 평가 보다는 사용자의 문제해결이나 의사결정에 필요한 정보의 품질 수준을 평가하는 것이 보다 효과적일 것으로 판단된다. 한 연구에 따르면 빅데이터 플랫폼 구축 실패의 가장 큰 원인은 “충분하지 않은 조직 내부에서의 연계, 중간 관리자의 적응 및 이해 부족과 저항”으로 알려져 있다

[31]. 사용자들이 정보의 품질을 인정한다는 것은 이런 조직 내 장애물들을 극복하고 빅데이터 플랫폼이 성공적으로 도입되었다는 뜻이 된다. 물론 빅데이터 자체만으로 정보의 품질이 보장되었다고는 할 수 없고 이러한 평가가 내려지면 사용자들에 대한 적절하고 충분한 교육 및 훈련이 제공되어야 하며 경영자의 지원도 필요하다. 그럼에도 불구하고 ‘평가’의 문제는 결과(정보의 품질)로 판단할 수밖에 없고. 그렇기에 본 방법론은 실무적으로 의미가 있다고 하겠다.

참 고 문 헌

- [1] Accenture (2016) “Optimizing Grid Performance through Advanced Operations”.
- [2] Wilcox, T., Jin, N., Flach, P., & Thumim, J. (2019). A Big Data platform for smart meter data analytics. *Computers in Industry*, 105, 250-259.
- [3] Alahakoon, D., & Yu, X. (2015). Smart electricity meter data intelligence for future energy systems: A survey. *IEEE Transactions on Industrial Informatics*, 12(1), 425-436.
- [4] Zhou, K., Fu, C., & Yang, S. (2016). Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews*, 56, 215-225.
- [5] Ku, T. Y., Park, W. K., & Choi, H. (2018, July). Demand response operation method on energy big data platform. In 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN) (pp. 823-825). IEEE.
- [6] Guilan, W., Guoliang, Z., Hongshan, Z., & Hongyang, L. (2016). Real-time big data technologies of energy internet platform. In 2016 IEEE International Conference on Power System Technology (POWERCON) (pp. 1-6). IEEE.

- [7] Capgemini (2012) "Smart Analytics for the Utility Sector".
- [8] Bhattarai, B. P., Paudyal, S., Luo, Y., Mohanpurkar, M., Cheung, K., Tonkoski, R., ... & Manic, M. (2019). Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions. *IET Smart Grid*, 2(2), 141-154.
- [9] IS Group. (2012). Managing big data for smart grids and smart meters. IBM Corporation, whitepaper (May 2012).
- [10] Fotopoulou, E., Zafeiropoulos, A., Terroso-Sáenz, F., Şimşek, U., González-Vidal, A., Tsiolis, G., ... & Skarmeta, A. (2017). Providing personalized energy management and awareness services for energy efficiency in smart buildings. *Sensors*, 17(9), 2054.
- [11] Jiang, R., Lu, R., Wang, Y., Luo, J., Shen, C., & Shen, X. (2014). Energy-theft detection issues for advanced metering infrastructure in smart grid. *Tsinghua Science and Technology*, 19(2), 105-120.
- [12] Ostroff, C., & Schmitt, N. (1993). Configurations of organizational effectiveness and efficiency. *Academy of management Journal*, 36(6), 1345-1361.
- [13] IDC (2016) "Big Data: Turning Promise Into Reality".
- [14] Tole, A. A. (2013). Big data challenges. *Database systems journal*, 4(3), 31-40.
- [15] Stein, B., & Morrison, A. (2014). The enterprise data lake: Better integration and deeper analytics. *PwC Technology Forecast: Rethinking integration*, 1(1-9), 18.
- [16] Qi, Q., & Tao, F. (2018). Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *Ieee Access*, 6, 3585-3593.
- [17] Marinakis, V. (2020). Big Data for Energy Management and Energy-Efficient Buildings. *Energies*, 13(7), 1555.
- [18] Hofmann, E. (2017). Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. *International Journal of Production Research*, 55(17), 5108-5126.
- [19] Mittal, A. (2013). Trustworthiness of big data. *International Journal of Computer Applications*, 80(9).
- [20] Mao, R., Xu, H., Wu, W., Li, J., Li, Y., & Lu, M. (2015). Overcoming the challenge of variety: big data abstraction, the next evolution of data management for AAL communication systems. *IEEE Communications Magazine*, 53(1), 42-47.
- [21] Al-Salim, A. M., Lawey, A. Q., El-Gorashi, T. E., & Elmirghani, J. M. (2017). Energy efficient big data networks: Impact of volume and variety. *IEEE Transactions on Network and Service Management*, 15(1), 458-474.
- [22] Capgemini (2017) "The deciding factor: Big data & decision making," Capgemini Reports, 1-24.
- [23] Kadadi, A., Agrawal, R., Nyamful, C., & Atiq, R. (2014, October). Challenges of data integration and interoperability in big data. In 2014 IEEE international conference on big data (big data) (pp. 38-40). IEEE.
- [24] DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research*, 3(1), 60-95.
- [25] Li, E. Y. (1997). Perceived importance of information system success factors: A meta analysis of group differences. *Information & management*, 32(1), 15-28.
- [26] Delone, W. H., & McLean, E. R. (2003). The

DeLone and McLean model of information systems success: a ten-year update. *Journal of management information systems*, 19(4), 9-30.

- [27] Aldholay, A., Isaac, O., Abdullah, Z., Abdulsalam, R., & Al-Shibami, A. H. (2018). An extension of Delone and McLean IS success model with self-efficacy. *The International Journal of Information and Learning Technology*.
- [28] Kahn, B. K., Strong, D. M., & Wang, R. Y. (2002). Information quality benchmarks: product and service performance. *Communications of the ACM*, 45(4), 184-192.
- [29] Wang, R. Y., & Strong, D. M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12(4), pp 5-33.
- [30] Fanning, K. (2016). Big Data and KPIs: A Valuable Connection. *Journal of Corporate Accounting & Finance*, 27(3), 17-19.
- [31] McShea, C., Oakley, D., & Mazzei, C. (2016). The reason so many analytics efforts fall short. *Harvard Business Review*.

저자 소개



조치선(Chisun Cho)

- 2001년 2월 : 한국해양대 물류시스템공학과(공학사)
- 2003년 8월 : 건국대학교 일반대학원 경영학과 MIS · 경영정보 전공(경영학 석사)
- 2008년 8월 : 건국대 일반대학원 비즈니스 인텔리전스학과 박사 수료
- 관심분야 : 분석 모델링 설계와 Automated Machine Learning, 유통, 금융 · 보험 모델링



이난규(Nangyu Lee)

- 1987년 3월 : 한경대학교 컴퓨터공학과(학사)
- 1989년 5월~현재 : SE (System Engineer), Mobile, Bigdata Arch.
- 관심분야 : 실시간 대용량 빅데이터 활용, 에너지 빅데이터 설계&활용



함유근(Yukun Hahm)

- 1984년 2월 : 고려대학교 정경대학 통계학과 (경제학학사)
- 1994년 5월 : Boston University (경영학박사)
- 1998년 3월~현재 : 건국대학교 경영대학 경영학과 교수
- 관심분야 : 빅데이터 활용 전략, 빅데이터 비즈니스 모델