한국철도학회 철도저널 18권 5호, 2015년 10월, pp. 87–91

빅데이터 기법을 통한 철도산업 KPI개발[1)]

신 창 기
Parsons Brinckerhoff Korea Transportation Div.
부장
shinck@naver.com

1. 서론

"여의도에서 근무를 하는 성팀장은 지난 밤 과음으로 늦은 아침을 연다. 지난밤 회식자리에서 오랜만에 과음을 하였고, 어떻게 집으로 귀가하였는지 알 수가 없다. 아무 생각 없이 스마트폰을 꺼내서 지난 밤 타고 온 택시의 정보와 이동경로를 확인하고, 늦은 출근 준비를 한다. 얼마 전 구입한 스마트시계로 심박수, 혈압, 맥박 등을 체크하고, 전 날 소모한 칼로리를 포함한 스마트 폰에서 알려주는 건강지수를 체크한다. 집을 나오면서 출입카드로 문을 열고, 지능화 대중교통 관리시스템을 통하여 수많은 카메라와 정산시스템을 거쳐 여의도 사무실에 들어온다. 각종 SNS를 통하여 지인들의 소식을 접하고, 회사내 포탈을 통하여 지난달 개인역량 자료들을 확인한다. 1인당 매출, 수주기여율, 1인당 생산력 등을 확인하고, 현재 새롭게 시작한 프로젝트의 성능지표(KPI)를 개발하려고 자리에 앉는다".

이렇듯 우리의 생활전반에 걸쳐 정보의 바다 또는 홍수라는 표현으로 우리 생활에 모든 조그만 활동들이 기록되고 남겨지고 있다. 자기가 생성한 수많은 자료들은 다양한 방법으로 기록된다. 여기까지는 정보화시대의 당연한 모습들이다. 과거 이러한 자료들은 마케팅 기법에 단골로 활용되어 왔었다. 연령대별 소비자들이 많이 밀집하는 곳, 어떠한 상품이 어떤 계절에 잘 팔리는지, 어떠한 형태의 포장을 선호하는지 등, 1차원적인 자료의 취합과 분석을

통한 기법들을 활용하였다. 이러한 기법들의 형태는 과거 활용 가능한 자료들의 단순성과 제한적인 이유로 기초적인 소비특성만 나오게 되었으며, 이러한 자료들의 수준 및 단 시간에 방대한 량의 자료가 쏟아지는 현재에는 다양한 형태의 분석들이 생겨나고 있다.

빅데이터 기법이라는 것은 이렇게 방대한 데이터의 수집, 저장, 관리, 분석할 수 있는 역량을 넘어서는 대량의 정형 또는 비정형 데이터 집합 및 이러한 데이터로부터 가치를 추출하고 결과를 분석하는 기술을 의미한다. 하루에 트위터는 3억여건, 페이스북은 4억여건, 유투브는 1초에 한시간 분량이 업로드 되는 데이터의 홍수 시대에 이러한 방대한 자료를 분석하는 일은 우리 생활에 아주 깊숙이 들어와 있다.

또한(이러한 자료들의 분석은 절대적인 기준이 아니라, 상대적인 기준이여서) 기존의 정보 부족의 시대에서 수행

1) 키워드: 빅데이터, 데이터마이닝, 자산, 관리운영, 자산관리, KPI, 핵심성능지표

되었던 샘플링조사 및 가설을 세워 일반화하는 과정을 거치지만, 빅데이터 기법은 현실정보를 분석하는 과정이므로, 기존의 샘플링조사의 폐단을 많이 검증하고 있으며, 실질적인 자료들의 방향성을 예측하여 실현가능한 예측을 하게 해준다.

우선 빅데이터란 무엇인가? 라는 기초적인 질문에서 시작하여 생활속 관련 예를 들어 설명을 하고, 철도분야에서 어떠한 형태로 빅데이터 활용방안을 수행하고 있는지를 확인하고, 앞으로 어떠한 부분에 집중적으로 활용해야 할지를 소개하기로 한다.

2. 우리생활 속의 빅데이터 - 야구기록

빅데이터의 활용 형태를 야구에 비교하여 설명해 보기로 하자. 1982년 국내에서 프로야구가 출범하면서 수많은 기록들이 현재까지 쏟아져 나온다. 그 중 가장 쉽게 타자를 평가하는 데이터로 타율, 홈런, 타점을 기본으로 출루율, 장타율 같은 전통적인 지표들로 타자를 평가하였다. 이러한 지표들은 미국인 빌 제임스가 창시한 야구의 데이타를 통계학적/수학적 분석을 통해 평가하는 세이버메트릭스를 통하여 단순한 개인적 취미차원을 넘어선 분석적인 지표들로 활용하기 시작하였다. 이러한 지표들 중 WAR(Wins Above Replacement) 지표는 기본적으로 하나의 '통합지표'라 할 수 있다. 예전에는 선수의 가치를 평가하기 위해서 수비, 타격, 피칭 등 일일이 나누어 봐야 했지만, WAR지표가 고안된 이후에는 선수의 포지션과 관계없이 일괄적인 평가가 가능하게 되었다. 또한 OPS(On-base percentage Plus Slugging)는 출루율과 장타율의 합으로 정해지는 간단한 방법임에도 타자들의 능력과 득점기여도를 효과적으로 판단할 수 있는 대중적인 지표이다. 간략하게 소개하면 타자의 능력은 얼마나 많은 안타를 때리느냐와 얼마나 득점 확률이 높은 포지션에 진출하느냐를 종합적으로 놓고 판단하는 것이다.

그렇다면 왜 현대 야구의 수많은 지표들 중에서 WAR와 OPS를 예로 들었을까? 이유는 이 두 지표는 가장 손쉽게 선수의 능력을 판단하여 팬과 시장에서 선수의 가치를 판단하는 중요한 가치로 활용되는 대표 지표이지만, 두 지표가 생성되는 과정은 매우 다르다. WAR과 같은 지표는 세이버메트릭스 시스템에서 전문적인 통계기법을 통하여 만들었으며, 전문가가 아니면 그 많은 방대한 양의 데이터를 처리하여 손쉽게 만들 수 없는 지표이지만, OPS의 경우는 타율과 출루율만 합해서 누구나 손쉽게 만들어 낼 수 있는 지표이다.

OPS는 WAR에 비교해서 간단한 산출방법을 통해 만들어지지만, 훌륭한 결과를 도출할 수 있는 능력을 가진 지표임을 입증할 수 있는 근거는 쉽게 찾을 수 있다. 일례로 메이저리그 유명타자를 대상으로 "만일 내가 타자기록 중 하나를 골라 1위를 할 수 있다면, 어떤 기록에서 1위를 차지하고 싶은가?"라는 질문에 OPS지표가 당당히 1위를 하였다고 한다. 물론 OPS가 절대적인 평가기준이 될 수는 없다. 부분적으로 균형을 맞춰야 할 부분도 있으며, 반론의 여지가 될 수 있다는 지표임을 야구계에서도 수용하고 있는 지표이다. 그래서 이를 보완한 wOBA와 같은 지표들이 생성되기도 했다.

하지만, 이러한 야구지표 예를 통해서 필자가 말하고 싶은 것은, 시스템의 발전에 따라 엄청난 양의 데이터들이 현장에서 생성되어 소멸되고, 과거보다 현재에는 이러한 데이터들을 활용하여 현재를 평가하고 미래를 예측하는 지표들의 개발이 필요하다는 것이다.

이러한 지표들은 과거에 활용되던 지표보다 훨씬 세분화되고 구체화되어 시장을 움직이는 중요한 숫자로 활용되며, 이러한 지표들은 전문가들이 복잡한 수식을 통해 만들어지는 지표들에서부터 손쉽게 만들었지만 중요한 의미를 가진 지표들까지 그 활용도는 무궁무진 하다는 것이다.

3. 철도분야 운영성과의 기준 – KPI 지표

이를 바탕으로 본고의 주제인 철도시장을 반영하여 생각해 보자. 현재 도시철도, 아니 광역철도, 아니 우리나라

모든 철도시스템의 계획, 설계, 공사, 운영을 아우르는 전 단계에서 생성되는 수많은 데이터들은 어떻게 관리가 되고, 어떻게 기록되고, 어떻게 활용이 되는지에 대하여 필자의 경험을 통해 기고하여 본다.

철도시장을 포함한 현재 산업계에서 활용되고 있는 빅데이터를 이용한 작업 또는 그와 유사한 형태의 결과물을 살펴보면, 대부분이 수요예측, 통행분산, 소비자의 성향분석 등과 같은 목적으로 활용이 되고 있다. 당연히 교통과 여행, 고객서비스측면에서 활용가치가 높을 것으로 판단한다. 이는 빅데이터의 활용가치가 아직 산업전반에 생소하고 이를 철도와 접목시키기 위한 노력이 제한적임을 간접적으로 보여주고 있다.

최근 개통한 민자 철도분야에 전문운영사들이 진출하면서 사업자(발주처)와 체결한 운영계약, 흔히 관리운영계약이라고 일컫는 계약에는 운영사로서 운영 및 유지보수를 책임지면서 전문운영사를 관리하는 지표를 설정하여 그들의 수준을 평가하는 지표로 적용되고 있다.

이러한 지표들은 최초 민자사업영역에서 운영사업자로 체결한 관리운영계약서가 기본이 되어 유사하게 타 사업에서도 운영사를 관리하는 지표로 계약에 명시되었으나, 이러한 지표들이 현재 운영사의 운영 및 유지관리 상태의 적정성을 평가하기에는 아주 미흡한 수준이다. 이런 문제는 현장에서 현실적인 문제로 개선의 필요성이 대두되지만, 각 사업의 전문 운영사들은 계약적인 제약을 벗어나고자 희망하고, 사업자는 이를 검토할 능력이 없다는 문제로 이러한 지표들의 개발이나 개선은 되지 않고 있다.

〈그림 1〉 리포팅툴을 이용한 도식화

하루에도 수많은 자료들이 생성되고 소멸되는 운영, 유지보수, 안전, 고객자료들은 차량, 선로, 통신, AFC, CCTV등이 가지고 있는 데이터 및 정보와 이러한 정보들을 관리하는 정보관리시스템(MIS)으로 관리되며, 이러한 관리시스템은 운영 및 유지관리 형태를 포함한 검수 및 시설물관리시스템 형태로 개발되어 활용되고 있다.

하지만 관리시스템은 초기 버전과 비교하여 많은 기능들을 탑재하고 있으나, 대부분 기존의 관리시스템 플랫폼 위에 유저인터페이스를 개선하거나 자동적인 관리 기능을 개선시키는 형태로 발전해 오고 있다. 일부 기능개선이나 필요한 형태의 Reporting tool을 이용한 형태로 발전은 하고 있으나 기존 시스템상의 관리되는 기본 지표도 시스템 상으로 데이터를 받아 후처리하는 수준의 시스템으로 관리하는 작업장도 있다.

현장에서 수집되는 방대한 자료들이 가지는 의미가 무엇인지, 이 자료들을 어떻게 조합하여 의미 있는 지표를 만들지, 개발된 지표관리, 지표의 개선, 지표의 유지 등의 활동이 필요하다.

<표 1>은 거의 모든 민자철도(경전철)에서 사용되고 있는 성과지표에 대한 관리목표를 보여주고 있다. 관리 항목들은 대부분 유사한 형태이며, 성과목표의 수준은 사업자의 시스템 성능에 따라 달리 적용하고 있다.

이러한 지표는 우선적으로 운영사의 운영 적정성을 개량적으로 판단하기 위한 지표로 개발하여 사용하고 있지만, 시스템성능과 관련된 지표와 고객만족 및 청결상태 등의 대승객서비스를 기초로 한 서비스성능 지표로 크게 나눌 수 있다. 이들을 산정하는 수식, 기준, 용어의 정의, 산정예외조항 등 많은 숙제들이 현장에는 남아 있지만, 이 기고에서는 이러한 핵심성능지표(KPI)를 통해서 운영수준을 개량화 한다는 것을 소개하는 것에 의미를 둔다.

이러한 KPI는 조직의 전략에 부합하기 위한 일종의 BSC(Balanced Scorecard)의 기법으로부터 도출된 지표들이며, 이는 대부분의 조직구도에서 조직의 목표와 전략을 도출하기 위한 기초적인 활동에서 비롯된다. 이렇듯 전략

〈표 1〉 예 : 경전철 관리운영 계약내 성과지표]

분 야	측정항목		성과목표	비 고
안 전 운 행	사상자 발생		1.23건/백만km	자살/자해 제외
	열차가용성		99%	–
	열차정시성		98%	–
	열차운전장애		0.80건/백만km	–
고 객 만 족	외부고객만족도		90점	–
청 결 상 태	역사 청결상태		98%	–
	전동차 청결상태		98%	–
	전동차량		SA=99%	–
설비 RAM 목표값	PSD		A=98%	–
	AFC	승차권발매기	A=98%	–
		자동 개집표기	A=98%	–
		중앙전산기	MTBF=5,000시간	단위 : 운영시간 기준, 내구성 : 15년
	EL, EC, MW		A=99%	승강기 : 법규사항 준수

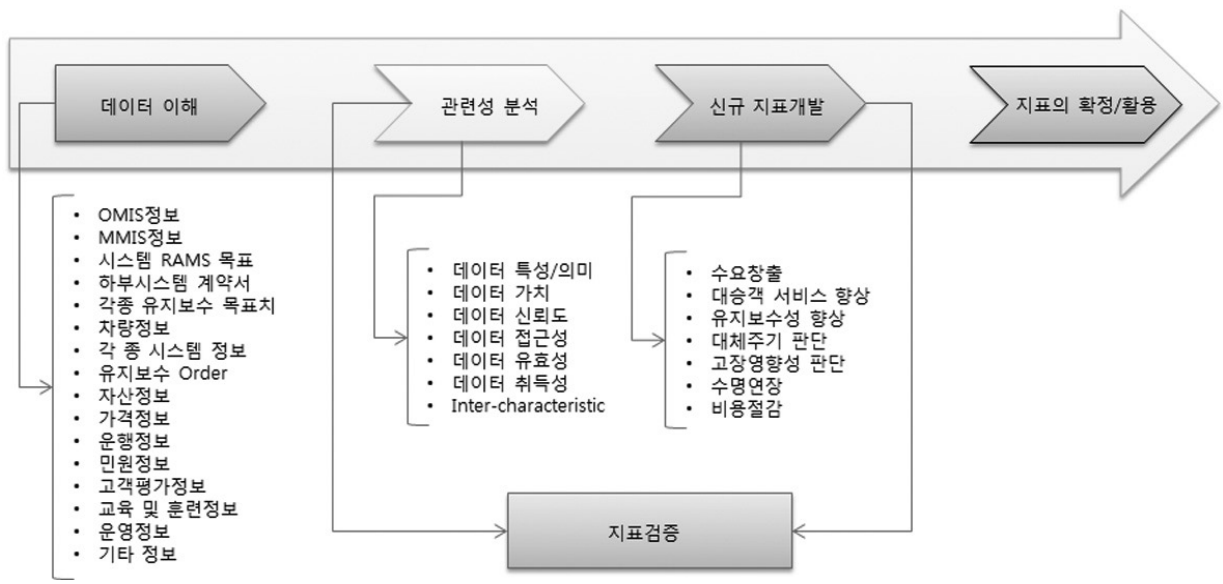
※ 성과목표의 설정은 임의로 조정하였음

적 개량지표의 필요성 및 작성지침과 관련한 깊은 이해는 "철도건설에서 BSC의 KPI도출과 BPM을 위한 프로세스 KPI도출_류시욱, 안병준, 류명욱_2006 철도학회 논문"을 통하여 쉽게 접근할 수 있을 것이다.

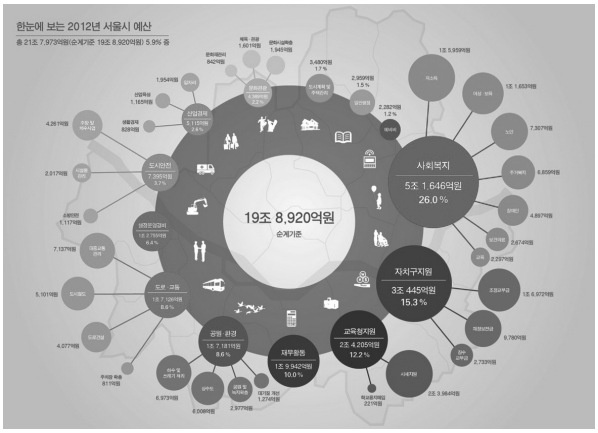
따라서 이러한 KPI는 조직의 전략과 비즈니스를 위한 기본적인 조직관리 도구이며, 우리 철도시장에 이미 깊숙이 자리 잡고는 있으나, 기존의 지표를 개량하지 못하거나, 효과적인 지표를 개발하지 못하거나, 경영, 재무적인 영역을 벗어나 시스템 유지보수분야로의 확장을 못하거나, 전문적인 인력을 보유하지 못하는 등, 세계 선두의 인프라를 확보한 하드웨어 기술력 뒷면에는 초라한 소프트웨어 분석기술을 가지고 있는 모습으로 필자의 눈에는 비춰지고 있다.

마치 이런 모습은 다양한 데이터를 가공하여 효과적인 선수평가와 작전이 뒷받침되는 메이저리그에, 타율, 타점, 홈런 등으로 선수를 평가하여 야구를 하는 격이라고 할 수 있을 것이다. 질 좋은 배트, 고급 글로브, 멋진 고글과 같은 하드웨어는 메이저리그의 모습을 흉내 낼 수는 있지만, 전략에 의한 효율적인 야구는 기대하기 힘들 것이다. 전체 전력을 짜내어 단기간 성적은 나오겠지만, 방대한 데이터를 효과적인 분석을 통해 지표를 기초로한 효율적인 야구는 장기간 시합을 최고의 전력으로 가능케 하는 지속성을 보장한다.

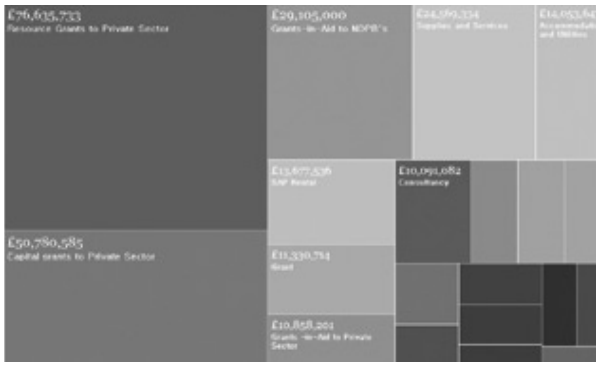
따라서, 우리 철도가 건설 및 운전운영을 통하여 효율적인 자산관리를 통해 안정적인 운영의 지속성을 부여하기 위해서는 보다 의미 있고, 세련된 현장지표의 개발에 집중



〈그림 2〉 지표개발 프로세스



〈그림 3〉 한눈에 보는 2012년 서울시 예산



〈그림 4〉 Tree Map을 이용한 예산정보 시각화

해야 하며, 이러한 지표의 개발의 첫 걸음은 데이터의 분석에서 출발한다.

우선 현장에서 생성되는 지표들의 정의와 의미를 정확히 파악을 해야 하며, 각 지표를 구성하는 데이터간 연관성을 이해해야 하며, 새로운 지표를 생성하여 분석을 통한 검증을 순화시켜 의미를 부여해야 한다. 이와 같은 작업을 통해 도출된 결과 값은 여러 가지 리포팅 툴을 이용하여 도식화되며, 다양한 시각화 작업을 통하여 사용자에게 손쉽게 전달될 수 있도록 해야 한다.

4. 결론 - 빅데이터를 활용한 철도KPI의 개발

우리 철도가 직면하고 있는 문제 중, 하드웨어적인 문제가 아닌 소프트웨어적인 문제를 생각해 보면, 현장에서 쏟아져 나오는 자료들에서 어떠한 유용한 정보를 캐내어 실시간적인 진단과 조치를 통해 운영의 안전성과 비용의 효율화를 도모해야 할지 생각할 수 있다. 데이터 운영관리상의 경쟁력을 확보하기 위해서는 이에 대한 정확한 이해와 이를 전담으로 관리하는 조직이 필요하며, 지속적인 활동을 통해 현장에 전달할 수 있는 프로세스가 기반이 되어야 한다.

결국 이러한 힘은 소프트웨어가 뒷받침되어야 하며, 이러한 소프트웨어 위에서 철도의 경쟁력이 생성되는 것이다. 우리는 우리철도가 직면하고 있는 안팎으로의 어려움을 잘 알고 있다. 이러한 문제는 비단 철도뿐 아니라 대부분의 국내시장이 직면한 어려움들이며, 이를 극복하기 위해서는 우리가 가지고 있는 수많은 자료들을 기반으로 이를 효율적이고 활용성 있게 관리할 수 있는 능력을 키워야 한다.

♣ 참고문헌

[1] 철도건설에서 BSC의 KPI도출과 BPM을 위한 프로세스 KPI도출_류시욱, 안병준, 류명욱_2006 철도학회 논문
[2] OPS 야구에서 타자를 평가하는 간단하면서도 효과적인 지표