

PART 02

빅데이터 분석 지형을 바꾸다 원스톱 플랫폼 ‘펜타호’가 남다른 까닭

SNS, 인터넷, 이메일 등 정형 데이터뿐만 아니라 비정형을 포함한 새로운 구조의 데이터를 처리하는 빅데이터 분석 플랫폼에 대한 요구 조건이 까다로워지고 있다. 단절되거나 산재되어 있는 여러 시스템에서 진행되는 작업으로는 빅데이터를 통해 효율적으로 비즈니스 가치를 창출하기 어렵다는 판단에서다.

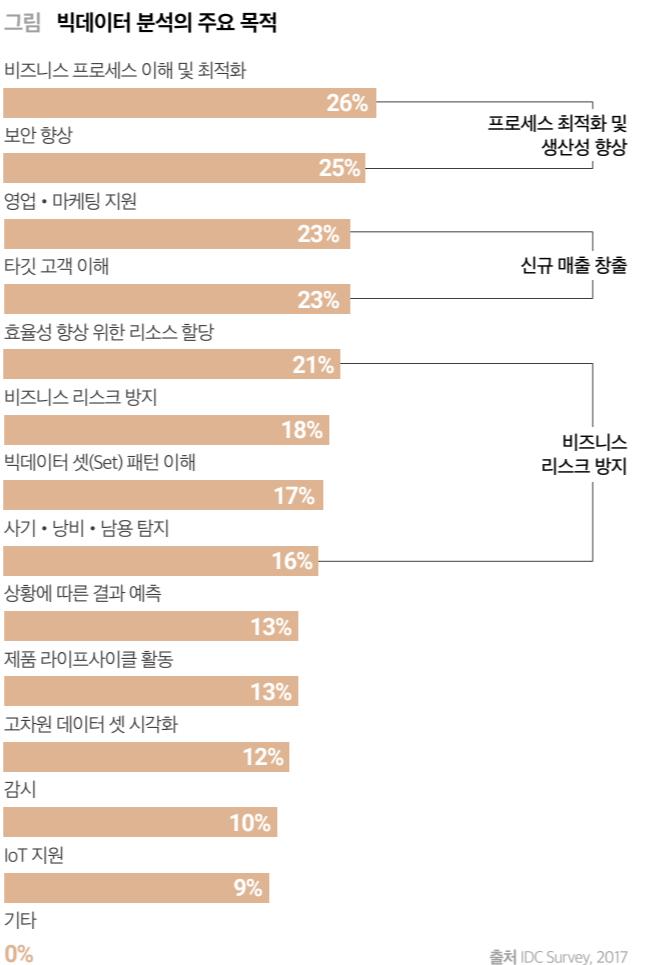
빅데이터에서 가치를 이끌어낼 수 있는, 제대로 된 빅데이터 분석 플랫폼 ‘펜타호(Pentaho)’를 알아보자.

시장조사업체 IDC에 따르면, 전 세계 데이터 규모는 2025년에 163泽字节(Zetabytes)에 이를 것이라고 한다. 2016년 데이터 발생량을 기준으로 10배에 달하는 예상치다. 또한 이 데이터 중 4분의 1 이상이 실시간으로 발생하고, 그 중 95% 이상이 IoT(사물인터넷) 데이터가 차지할 것으로 전망되고 있다.

이처럼 데이터는 끊임없이 증가한다. 관계형 데이터부터 비정형 텍스트, 센서 데이터, 머신 데이터까지 그 종류도 무척이나 다양하다. 기업은 더 나은 성과를 얻고 새로운 비즈니스 성과를 내기 위해 이러한 데이터를 충분히 이해하고 활용할 수 있어야 한다. 디지털 트랜스포메이션을 추진하는 이유도 이 때문이다.

디지털 트랜스포메이션의 핵심은 데이터를 통해 기업을 혁신시키고 비즈니스 성과를 높여 매출 확대를 이루는 일이다. 이에 데이터의 소스, 형태, 규모 등의 조건에 제약 받지 않고, 기업이 보유하고 있는 모든 데이터를 통합하여 분석하며, 이를 비즈니스 확장을 위한 인사이트로 활용할 수 있다면 좋을 것이다. 하지만, 데이터의 수집, 정제, 분석 및 활용까지 단계별로 각각의 솔루션을 도입하기엔 기업의 입장에서는 IT 투자비용에 대한 부담감이 클 수밖에 없다.

데이터의 수집, 정제, 분석 및 활용까지 한 번에 해결할 수 있는 펜타호라면 이들 요건을 충족시킬 수 있다.



펜타호는 전 세계 1,800여 고객을 확보한 빅데이터 분석 플랫폼으로, 정형 및 비정형 데이터에 대해 수집부터 변환, 적재, 분석, 시각화, 그리고 머신러닝까지 일련의 과정을 모두 제공하는 원스톱 플랫폼이다.

다양한 소스의 데이터에 쉽게 접근해 읽어오고, 이를 즉각적으로 가공할 수 있다. 더 나아가 이를 바탕으로 최신의 머신러닝 및 인공지능(AI) 알고리즘을 접목시킴으로써, 최선의 결과를 예측하고 시뮬레이션이 가능하다. 물론 이 모든 작업은 하나의 워크플로우로 통합되어, 조직의 일부가 아닌 전체가 효과를 얻고 추가적인 수익을 창출할 수 있다.

펜타호는 다양한 데이터 소스들을 추출, 변환, 적재 할 수 있는 Pentaho Data Integration(DI)과 고급 분석 및 시각화 기능을 제공하는 Pentaho Business Analytics(BA)로 구성되어 있다. 일반적인 빅데이터 솔루션과는 달리 오픈 소스를 기반으로 하여, 비용 효율적이며 우수한 커스터마이징을 제공한다. 펜타호는 아래와 같이 네 가지 특징이 있다.

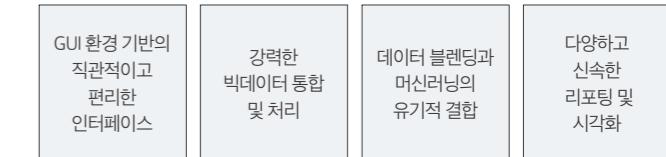


그림 펜타호 구현 원리



Q&A

시중에 펜타호와 비교되는 플랫폼이 있는가.

펜타호는 데이터 블렌딩²⁾부터 시각화까지 전 과정을 제공하는 업계 유일의 앤드-투-엔드 빅데이터 분석 플랫폼이다.

배치 분석이 아닌 실시간 분석에도 적용할 수 있다.

당연하다. 최근의 트렌드는 배치 데이터 처리 및 분석뿐 아니라 실시간으로 데이터를 수집하고 분석한 결과를 통해 인사이트를 확보하는 것이다. 펜타호는 실시간 데이터 분석을 위한 다양한 기능을 제공한다.

기준에 다양한 데이터 취합 및 저장 시스템이 구축되어 있는 상황에서 펜타호를 적용하고자 할 때 어려움은 없나.