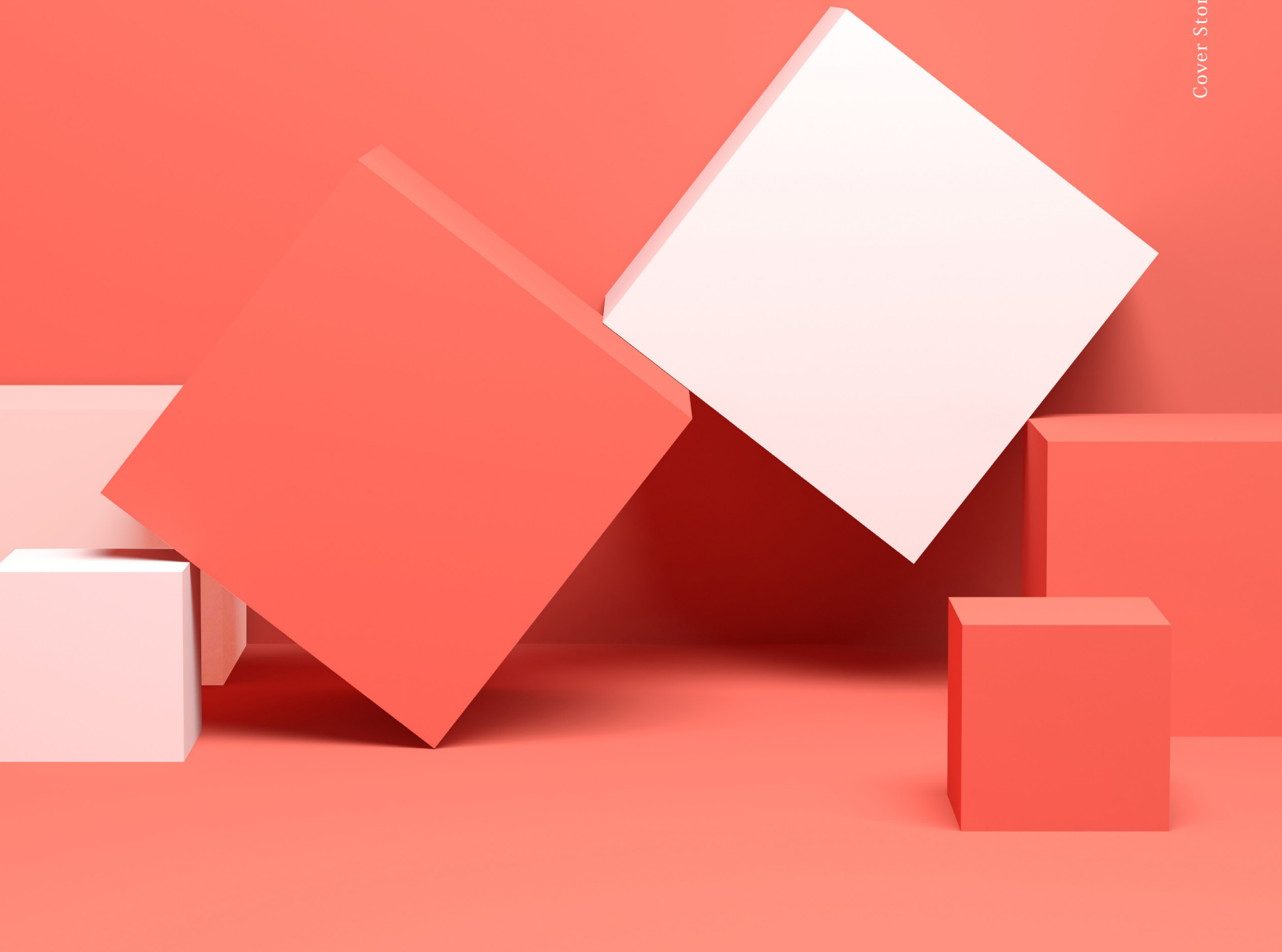




빅데이터 분석 시장에 펜타호(Pentaho) 바람이 분다

비즈니스
가치를 높이는
데이터 바로 알기

“

빅데이터가 없다는 건
고속도로에서 눈과 귀를
막고 운전하는 것이나
마찬가지다.

제프리 무어(Geoffrey Moore),
매니지먼트 컨설턴트 겸 이론가

“

데이터는
21세기의 휘발유,
분석은 이 휘발유를
연소시키는 엔진이다.

피터 손더가드(Peter Sondergaard),
가트너 리서치 수석 부사장

“

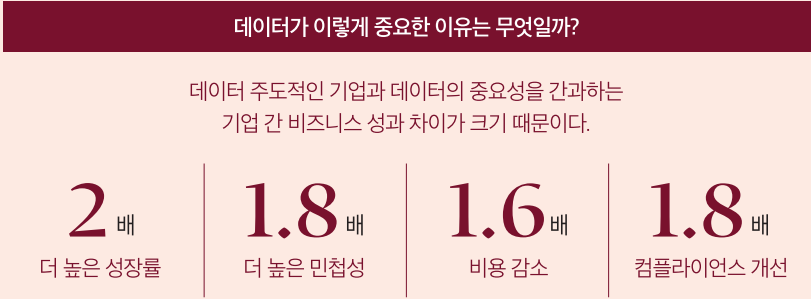
목표는 데이터를
정보로, 정보를 통찰력으로
전환하는 것이다.

칼리 피오리나(Carly Fiorina),
전 HP CEO

“

데이터를 충분히 괴롭혀라.
그러면 뭐라도
얻을 수 있다.

파올로 매그라시(Paolo Magrassi),
전 가트너 부사장, 리서치 부문 이사



데이터의 중요성은 이미 오랫동안 논의되어 왔다. 이제 빅데이터의 핵심은 데이터 수집이 아닌, 수집한 데이터를 어떻게 활용할 것인지에 달려있다. 특정 비즈니스의 과제를 해결할 답을 찾는 데 데이터를 어떻게 이용할 것인지에 대해 집중할 필요가 있다. 또한 새로운 기술과 인프라를 준비하는 것과 더불어, 이를 어디에 쓰고 어떤 결과물을 만들어 낼 지에 대한 고민이 중요하다.

데이터 자체만으로는 가치가 없다는 것을 이미 모두가 잘 알고 있다. 이제는 데이터를 제대로 활용하는 것에 가치가 있다. 인공지능과 머신러닝을 통한 데이터 분석을 통해 고객의 니즈를 만족시키는 것이 중요하다.

데이터 활용성
극대화하는 펜타호

Written by
이진석 수석
DATA사업팀 빅데이터사업 파트

현실적인 빅데이터 전략을 찾는 이들에게 ‘데이터옵스(DataOps)’라는 용어가 화두가 되고 있다. 데이터옵스는 원활한 협업 환경과 문화를 전제로 한 데이터 운영 전략 및 프로세스를 뜻한다.

펜타호는 데이터옵스를 추구하는 조직에 필요한 워크플로우를 더 잘 지원하도록 진화해 개별 프로젝트의 경험을 여러 담당자가 공유할 수 있게 됐다.

“데이터 수집과 저장, 분석, 시각화 과정을 워크플로우 단위로 처리하는 점은 펜타호의 가장 큰 특징이다. 따라서 빅데이터 활용에 초점을 둔 운영 시스템 개선 프로젝트에 최적이다.”

빅데이터 활용 시 데이터 수집, 저장, 보호를 넘어 ‘운영’이 중요해진 이유는?

과거 개별적인 업무가 주목적일 땐 빅데이터의 수집, 저장, 보호가 중심이었지만 결과물을 재활용할 수 없는 경우가 많았다. 프로젝트 결과물을 보고 다른 사람이 그 데이터를 재활용하려고 하면, 이미 했던 작업을 처음부터 다시 해야 한다. 왜 그런 결과가 나왔는가 보여주는 정보가 남아 있지 않기 때문이다. 프로젝트 결과물의 재활용에 관심이 커지면서 운영이 주요 이슈가 되었다.

데이터 운영, 어떻게 하면 잘 할 수 있나?

데이터 운영은 데이터를 잘 쓸 수 있도록 제공하는 툴과 프레임워크뿐 아니라 관련 문화도 포함한 개념으로 확장되어야 한다.

효성인포메이션시스템은 데브옵스(DevOps)와 마찬가지로 적합한 데이터를 적시에 올바른 장소로 가져올 수 있는 프로세스를 자동화한다. 데이터의 잠재력을 최대한 활용할 수 있는 전략 실행 방법으로 데이터옵스(DataOps)를 제시한다.

효성인포메이션시스템이 기업을 어떻게 도와줄 수 있나?

효성인포메이션시스템이 제공하는 펜타호는 빅데이터 분석 플랫폼으로, 세계 약 2,000여 고객사를 지원하는 오픈소스 기반 플랫폼이다. 데이터 소스

를 추출, 변환, 적재(ETL)하는 ‘펜타호 데이터 인티그레이션(PDI)’과 고급분석 및 시각화를 지원하는 ‘펜타호 비즈니스 애널리틱스(PBA)’로 구성되어 있다. 정형·비정형 데이터의 수집, 변환, 적재, 분석, 시각화, 머신러닝 과정까지 처리가 가능하다.

펜타호의 가장 큰 특징은 데이터 수집과 저장, 분석, 시각화 과정을 워크플로우 단위로 처리한다는 점이다.

워크플로우 기반의 협업이 제공하는 이점은 무엇인가?

과거 데이터 과학자와 데이터 엔지니어는 한쪽이 지시하면 다른 쪽이 만들어 주는 관계였다. 그러나 하나의 워크플로우에선 양쪽이 데이터가 어떻게 존재하는지 확인하면서 더 쉽게 작업할 수 있다.

데이터 분석 플랫폼으로서 펜타호의 차별성은?

오픈소스 커뮤니티가 활성화돼 있고 유용한 플러그인을 많이 갖추고 있다는 점이다. 커뮤니티에서 신기술과 제품에 대한 아이디어를 많이 얻고, 새로운 기술을 활발하게 테스트하고 보완한다. 기업들은 펜타호와 유저 커뮤니티 그룹이 개발한 플러그인을 ‘펜타호 마켓플레이스’에서 다운로드 받고 설치할 수 있다.

내년 2월에 출시 예정인 펜타호 9.0의 개선점은?

올 7월에 출시된 펜타호 8.3 버전은 멀티클라우드 지원 부분이 많이 개선됐고, 데이터 가시성도 향상됐다. 내년에 출시될 펜타호 9.0은 데이터 파이프라인 공유, 관리, 모니터링을 위한 새로운 데이터 플로우(flow) 서비스를 지원할 예정이다. 데이터 파이프라인 마찰을 줄이고 엣지에서 클라우드까지 고급 데이터 발견, 접근, 통합 기능을 제공할 예정이다.

효성인포메이션시스템의 펜타호를 포함한 빅데이터 사업 성과는?

최근 국내 대형 게임 회사에서 펜타호를 많이 도입하고 있다. 조직 특성상 다른 산업보다는 기술과 환경 변화에 대응이 빨라 신기술 도입이 용이한 편이며, 펜타호의 플러그인 지원 기능 등이 유용한 분야다. 반정형 데이터나 게임 로그 데이터의 대용량 처리를 위해서도 펜타호가 최적이다.

2020년 빅데이터 시장 전망과 향후 사업 추진 방향은?

과거 프로젝트는 빅데이터 저장에 무게 중심을 두었고 단발성 프로젝트가 많았다. 그러나 앞으로는 ‘활용’에 초점을 둔 운영 시스템 개선 프로젝트가 많아질 전망이다. 기업은 데이터 파이프라인을 구성하고 데이터오피스 채택을 원할 것이다.

중장기적으로는 사물인터넷(IoT)과 관련된 빅데이터 처리 인프라 구축 시장이 빠르게 확산될 전망이다. IoT 분야에서는 센서 데이터와 같은 반정형 데이터가 많이 발생하기 때문에, 제조 기업이 이들 데이터를 활용하려 할 것이다. 현재 진행 중인 프로젝트도 IoT, 실시간, 엣지, 반정형 데이터 분석과 활용에 초점을 맞추고 있다.

효성인포메이션시스템에는 ‘루마다’라는 IoT 플랫폼 솔루션이 있다. 이 솔루션에도 펜타호가 포함되어 있다. 펜타호의 분석 영역을 확장하는 데 유용한 솔루션이다. 효성인포메이션시스템은 향후 엣지단, 반정형 및 실시간 센서 데이터를 포함한 IoT 분야의 데이터 활용 영역에 집중할 계획이다.

↓ 펜타호의 빅데이터 플랫폼, 루마다

