



COVER STORY⁰⁴



빅 데이터
프로젝트를 하고싶은데
무엇을 어떻게
해야 하나요?

Big Data 접근 방안

주요 고려사항

빅 데이터 접근에 대한 가장 중요한 고려 사항으로는 새롭고 가치있는 서비스를 만들어낼 수 있는지의 여부와 이를 지원할 수 있는 인프라 및 인력을 확보할 수 있는가를 들 수 있습니다.

Key Point #1 VAS(Value-Added Service) 제공 여부

- 사내 정보의 통합과 외부 정보의 연결을 통해 기존에 없던 새로운 VAS를 만들어 낼 수 있는가?
- VAS의 제공 여부는 경쟁 관계에서 key differentiator가 될 수 있음

Key Point #2 VAS를 제공할 수 있는 분석 역량 (인프라, 인력)

- VAS를 구성하는 다양한 내,외부 데이터를 통합/분석할 수 있는 빅 데이터 관리 인프라의 구성
- 통합된 빅 데이터를 다각도로 분석할 수 있는 적정 분석 인력의 확보

빅 데이터 프로젝트를 효율적으로 수행하기 위해서는 먼저 문제를 명확히 정의해야 하고, 해당 문제의 분석에 활용될 수 있는 데이터를 정확히 선별할 수 있어야 하며, 이를 효율적으로 처리할 수 있는 아키텍처를 설계해야 합니다.

문제에 집중

- 기업 내의 어떤 문제점을 빅 데이터 기반으로 풀 것인가?
- 유사한 사례에는 어떤 것이 있는가?

데이터 범위와 분석 방법론 구성

- 해당 문제의 해결을 위해 어떤 데이터를 사용해야 하는가?
- 데이터의 이면에 숨겨져 있는 어떤 원리를 파악하여 문제를 해결할 것인가?
- B2C 빅 데이터는 대부분 고객의 행위 결과 데이터
→ 데이터 이면에 숨겨져 있는 행동 심리의 인문학적 통찰 필요
- B2B 빅 데이터는 기계적인 프로세스의 결과 데이터
→ 수리적 모델의 검증과 숨겨진 변수/데이터의 확인이 더욱 중요
- 이를 위해 어떤 분석 방법론을 사용할 것인가?

빅 데이터 처리 IT Architecture 설계

- 데이터 이동을 최소화 할 수 있는 설계 필요
- DB 중심의 비즈니스 데이터처리 + 비정형의 빅 데이터 처리
- Hybrid 구조의 IT Architecture 필요

빅 데이터 프로젝트를 새롭게 기획할 때에는 우리회사가 빅 데이터의 어떤 부분들을 필요로 하는지, 그리고 잘 수행할 수 있는 환경을 만들 수 있는지가 중요합니다.

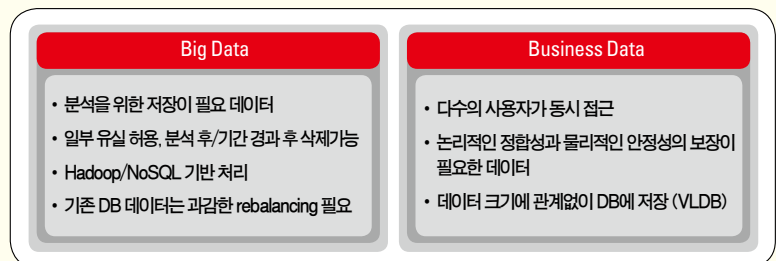


기존 DB 데이터와 새로운 빅 데이터를 어떻게 구분하여 저장/관리하나요?

Big Data 접근방안

주요 질문과 대응 방안 01

빅 데이터를 효율적으로 처리하기 위해서는 데이터의 특성에 맞추어 가장 효율적인 장소에 저장하고 처리하여야 합니다. 복잡한 연계 분석과 동시 처리가 필요한 데이터는 DB에서, 그렇지 않은 경우에는 Hadoop 혹은 NoSQL에서 처리합니다.



빅 데이터와 비즈니스 데이터의 특징을 잘 이해하고 데이터의 특성에 맞춘 효율적 장소로의 저장과 관리가 무엇보다 중요합니다.



빅 데이터 분석을 위해
기업은 무엇을
준비해야 할까요?

Big Data 접근방안

주요 질문과 대응 방안 02

빅 데이터 분석을 효율적으로 수행하기 위해서는 분석 인프라와 이를 활용할 수 있는 역량이 있는 인력이 반드시 필요합니다. 기술은 인프라 관리 기술과 분석 기술로 나누어 검토해야 합니다.

Big Data 관련 주요 기술

Big Data 인프라 구축 및 운영 기술	Big Data 분석 기술
과목별 충분한 문항을 레벨별 개발	Big Data Visualization
NoSQL 기반의 트랜잭션 처리	R 기반 통계 처리
관련 인프라의 안정적인 운영 방안	경영 과학 기법 이용 VAS 구현

주요 고려 사항

업무 분석의 리더십 확보	빅 데이터 처리 관련 업무에 대한 철저한 이해를 기반으로 빅 데이터 분석 전반에 걸친 리더십을 확보할 필요가 있음
VAS 설계와 분석 역량 확대에 집중	가치 있고 차별화된 부가 서비스의 설계에 최우선 목표를 둬, 그런 후 이를 위한 분석 역량의 확대에 우선 집중함
인프라 운영의 리스크 최소화	데이터의 급격한 증가에 따른 관리의 어려움을 예상하여 인프라 운영의 리스크를 최소화할 수 방안을 선행 검토

빅 데이터 분석을 위한 우리기업의 역량 확인이 선행되어야 하겠군요.



SNS 분석을 위해서는
무엇을 고려하여야
하나요?

Big Data 접근방안

주요 질문과 대응 방안 03

SNS 분석을 위해서는 시장/사회에 대한 이해가 선행되어야 합니다.

제품의 선호도 분석만으로는 시장의 변화를 선행할 수 없으며, 제품을 넘어 고객/시장/사회에 대한 분석이 우선되어야 합니다.

그런 후 SNS 분석을 수행할 수 있는 통합 인프라를 전사 또는 그룹 관점에서 구성해야 합니다. 빅 데이터 분석을 부서마다 개별적으로 수행하면 데이터의 중복 등으로 인해 효율성이 떨어집니다.

전사 관점에서 통합적으로 수행하는 것이 좋습니다.

SNS 분석이 필요 없는
제조 현장에서는
빅 데이터를 어떤
방법으로
활용해야 하는가요?

제조 현장에서는 빅 데이터 수집/저장과 시각화에 우선 집중해야 합니다.

이를 이용해 현업 담당자와 데이터 기반으로 문제 협의를 수행합니다.

그 후 특정 문제 해결을 위한 분석 프로세스를 구체화해야 합니다.

빅 데이터 분석은 크게 고객의 선호도 분석을 위한 SNS 분석과 제조 등의 업무 프로세스 분석으로 나누어 볼 수 있습니다.



Big Data 저장

- Hadoop을 활용한 설비 관련 환경 변수값의 저장
- 현업의 필요에 따라 저장 데이터의 종류와 기간 결정
- 저장의 의미: 전수 데이터 샘플이나 일부 데이터가 아님

Big Data Visualization & 현업의 판단

- 수집된 전수 데이터를 다양한 형태로 시각화
→ data pruning & outlier pin-pointing 가능
- 현장의 장인에 의한 빅 데이터 의미 판단과 프로세스 개선 가능

문제 발생시 Big Data Analytics 수행

- 문제(품질 오류, 오작동, 수율 하락 등) 발생시 다양한 분석을 통한 key factor analysis 가능
- 다양한 통계 및 경영 과학 기법의 활용