



빅데이터 기반 비즈니스 가치 창출 여정

Modern Data Architecture on AWS

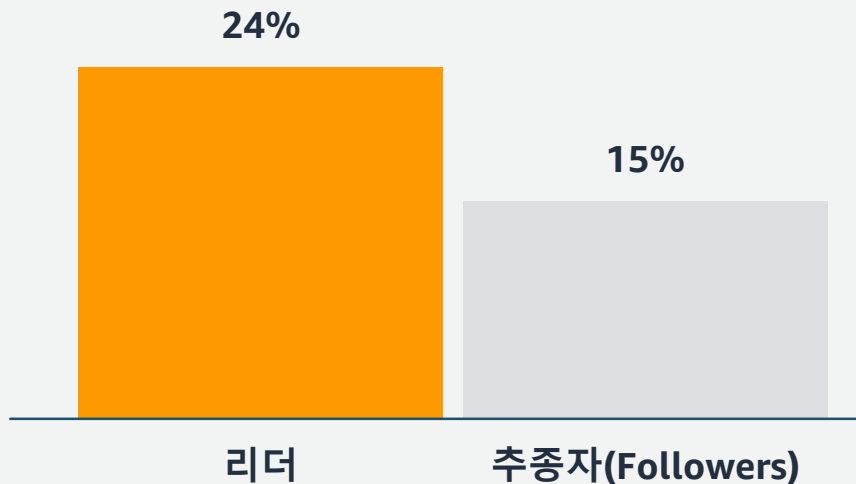
김효정 AWS 솔루션즈 아키텍트

June 28, 2018



리더(Leader)는 데이터를 핵심 차별화 전략으로 활용

자생적(Organic) 매출 성장



데이터에서 비즈니스 가치를 성공적으로 창출하는 조직은 경쟁업체보다 우수한 성과를 보입니다. 애버딘(Aberdeen)의 조사에 따르면 데이터 레이크 (Data Lake)를 구현한 기업은 다른 동종 기업 대비 자생적 매출 성장률이 9 % 증가하였습니다.*

*Aberdeen: Angling for Insight in Today's Data Lake, Michael Lock, SVP Analytics and Business Intelligence

데이터로 부터 비즈니스 가치를
창출하는 방법은?

A사의 비즈니스 확대 여정 – B2B에서 B2C로

2005

A사는 헤드셋을 제조. 2005년부터
유통 및 소매 채널로 판매하는
전통적인 B2B 사업 운영



2018

2018 년에 첫 번째 제품인 “Smart Buds”
출시. 무선 헤드셋은 음성 통화,
GPS 추적, 심장 박동 모니터 내장.
Bluetooth를 통해 사용자 휴대폰과
동기화. 모바일 앱은 장면 감지 지원



비즈니스 민첩성을 위해 필요한 다양한 질문

품질 문제가 생산 원가에 미치는 영향은 ?

시간 경과에 따른 공급업체 실적을 볼수 있을 까?

재고 보유 비용을 줄이려면?



재무팀

신규 출시 제품에 대한 실적이 좋은 지역은?

제품 카테고리별 수요 예측 결과는?

우리 제품에 대한 사회적 감성 분석 결과는 ?



제품팀

새로운 질문 대응을 위한 현대화된 데이터 아키텍처 디자인의 필요성

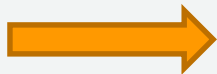
온프레미스

막대한 선결제 비용
피크치 중심 용량 설정
장기간의 구현 시간 필요



AWS 클라우드

사용한 만큼만 지불(Pay as you go)
데이터 파이프라인과 엔진의 분리
신속, 반복가능한 실험 플랫폼





비즈니스 목표 1 : 현대화 및 통합

특정 질문에서 시작, 필요로 하는 데이터 확보(Work backwards)

품질 문제가 생산 원가에 미치는 영향은 ?

시간 경과에 따른 공급업체 실적을 볼수 있을 까?

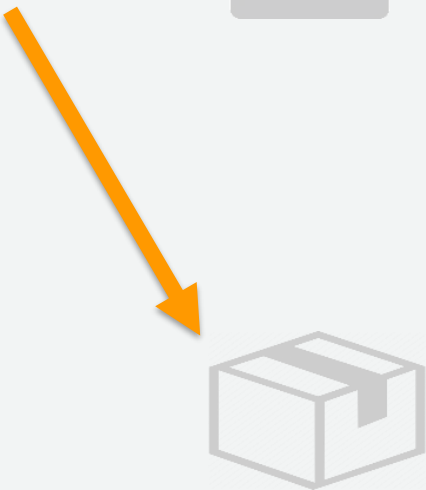
재고 보유 비용을 줄이려면?



재무팀



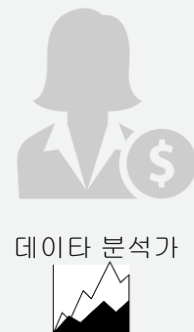
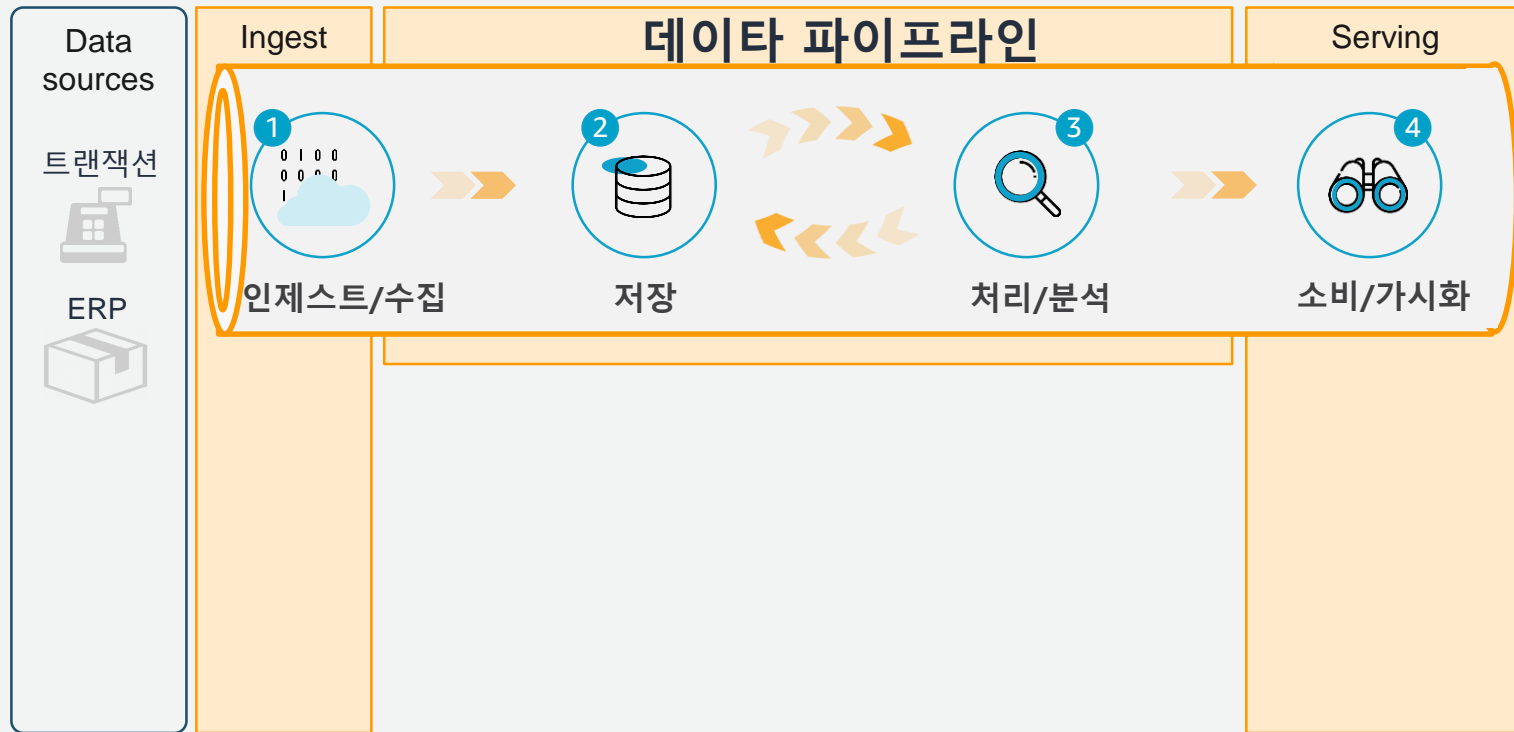
주문 히스토리 /
반품 (CRM)



재고 / 생산
(ERP)

현대화된 데이터 아키텍처

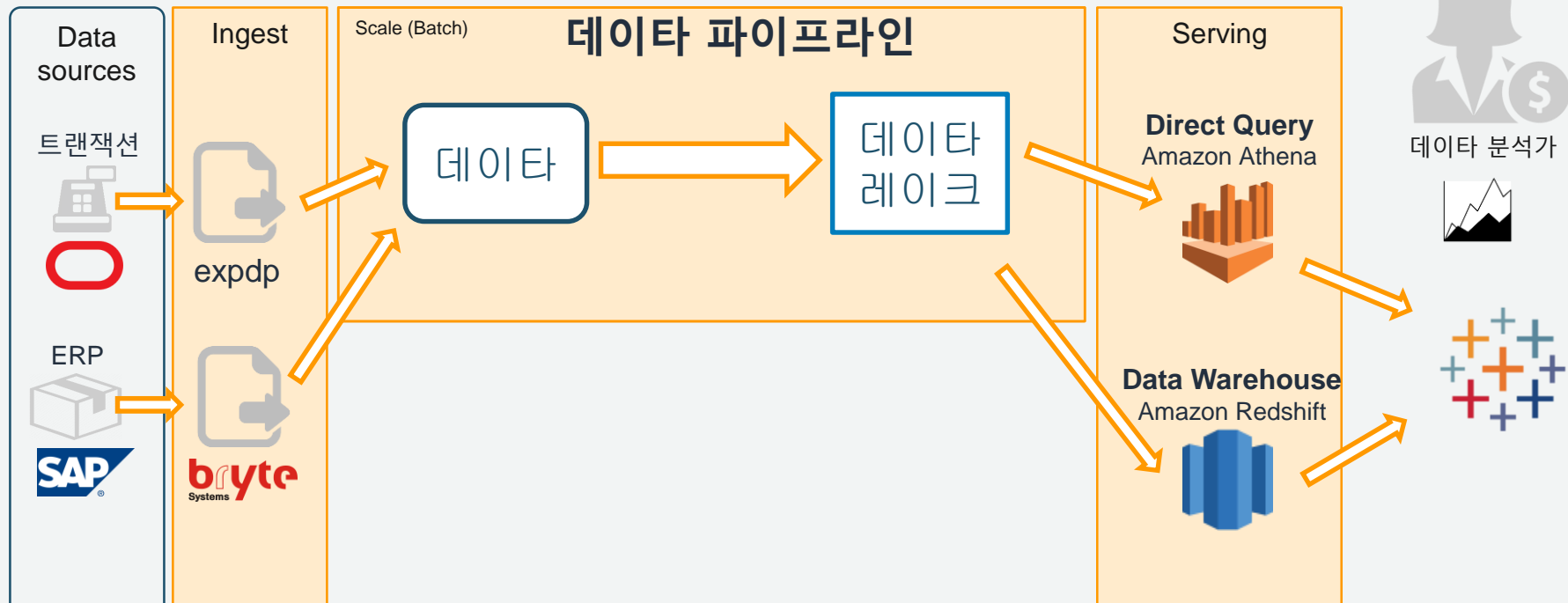
비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스



작게 시작해서 반복

현대화된 데이터 아키텍처

비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스



데이터 분석가는 분석을 위해 필요한 **작은 양**의 데이터 요청
데이터가 충분치 않으면, 또 하나의 **작은 슬라이스**의 데이터를 **데이터 레이크**로 업로드 요청

Amazon Redshift— 데이터 웨어하우스

빠르고, 강력하며, 간단한 완전 관리형 데이터 웨어하우스로 상용 DW 대비 10분의 1의 비용
기가바이트에서 엑사바이트에 이르는 규모의 대규모 병렬 처리

확장성 있고 빠름



컬럼 기반 스토리지 기술로
I/O 효율 개선 및
조회 성능 확장

오픈 파일 포맷 지원



Amazon S3의 모든 오픈
데이터 포맷 지원으로
변환 불필요
부착된 디스크(최신 SSD)
상에서 최적화된 데이터 포맷
기반 분석

보안



모든 것을 감사;
데이터 암호화;
많은 인증 및 컴플라이언스

저비용



테라바이트당 1년 \$1,000로
전통적인 상용 DW대비
10분의 1 비용;
시간당 \$0.25에서 작게 시작



좋은 데이터 레이크의 특징



모든 것을 수집



원하는
데이터 검색



유연한 접근



미래 대비

필요한 특정 질문에서 시작, 이를 위해 필요한 데이터를 확보

신규 출시 제품에 대한 실적이 좋은 지역은?

제품 카테고리별 수요 예측 결과는?

우리 제품에 대한 사회적 감성 분석 결과는 ?



제품팀



주문 히스토리 /
반품 (CRM)

데이터 레이크에 존재

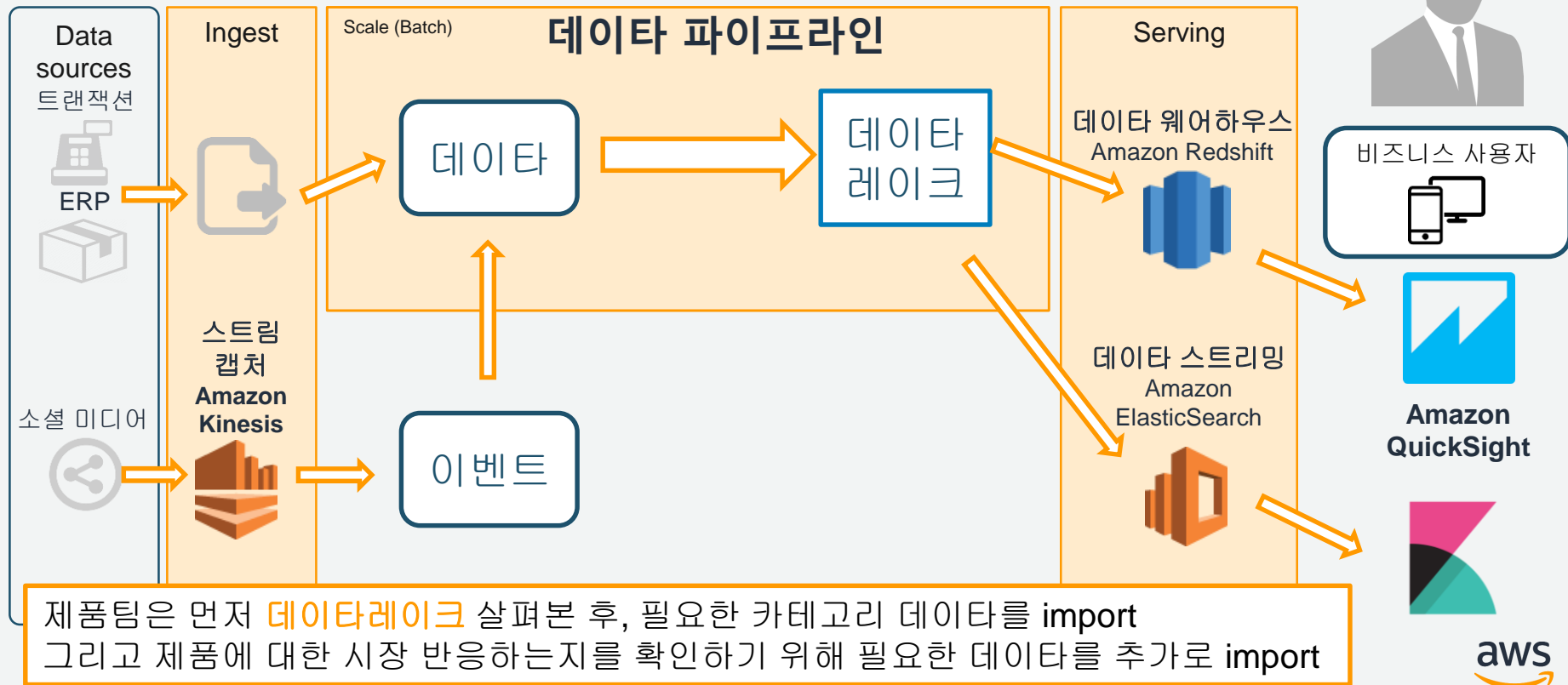


트렌드 정보 /
소셜 멘션(Mentions)

실험, 확인 그리고 확장

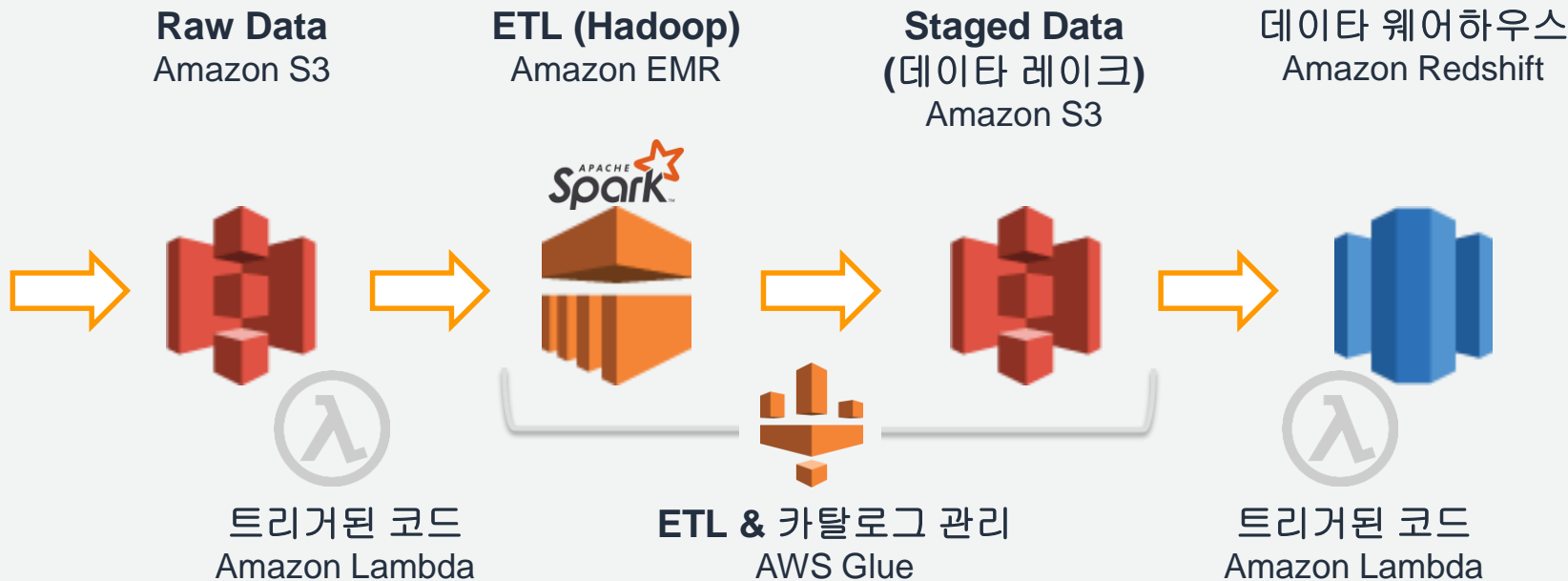
현대화된 데이터 아키텍처

비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스



일반적인 데이터 파이프라인 구성

분리된(Highly decoupled) 구성으로 향상된 확장성, 내결함성(fault tolerant) 및 비용 최적화 제공



AWS 제공 다양한 스토리지 빌딩 블록

데이터 이동



하이브리드 스토리지



데이터 스트리밍



파일 동기화



WAN 가속



프라이빗 네트워크



제3자 애플리케이션



물리적 어플라이언스

스토리지 타입



블록



파일



오브젝트 아카이브

데이터 보안 및 관리



데이터 발견 및 보호



데이터 시각화



서버리스 컴퓨팅



자동화



감사 추적



모니터링 및 메트릭



접근 제어



암호화

가장 광범위한 스토리지 서비스

데이터 이동



AWS Storage Gateway



Amazon Kinesis
Firehose



EFS File Sync



S3 Transfer Acceleration



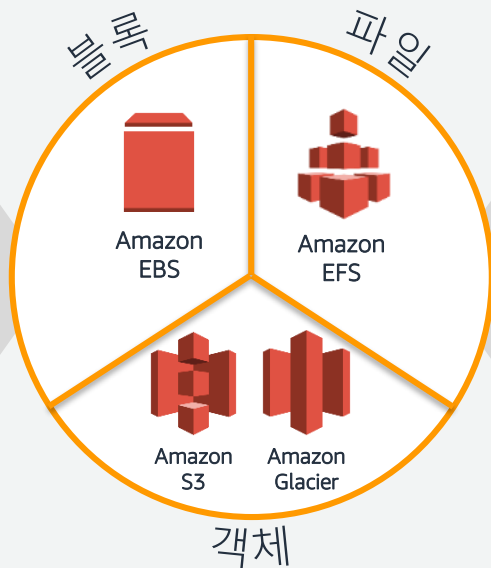
AWS Direct Connect



타사 애플리케이션



AWS Snow 제품군



데이터 보안 및 관리



Amazon Macie



AWS QuickSight



AWS Lambda



AWS CloudFormation



AWS CloudTrail



AWS CloudWatch



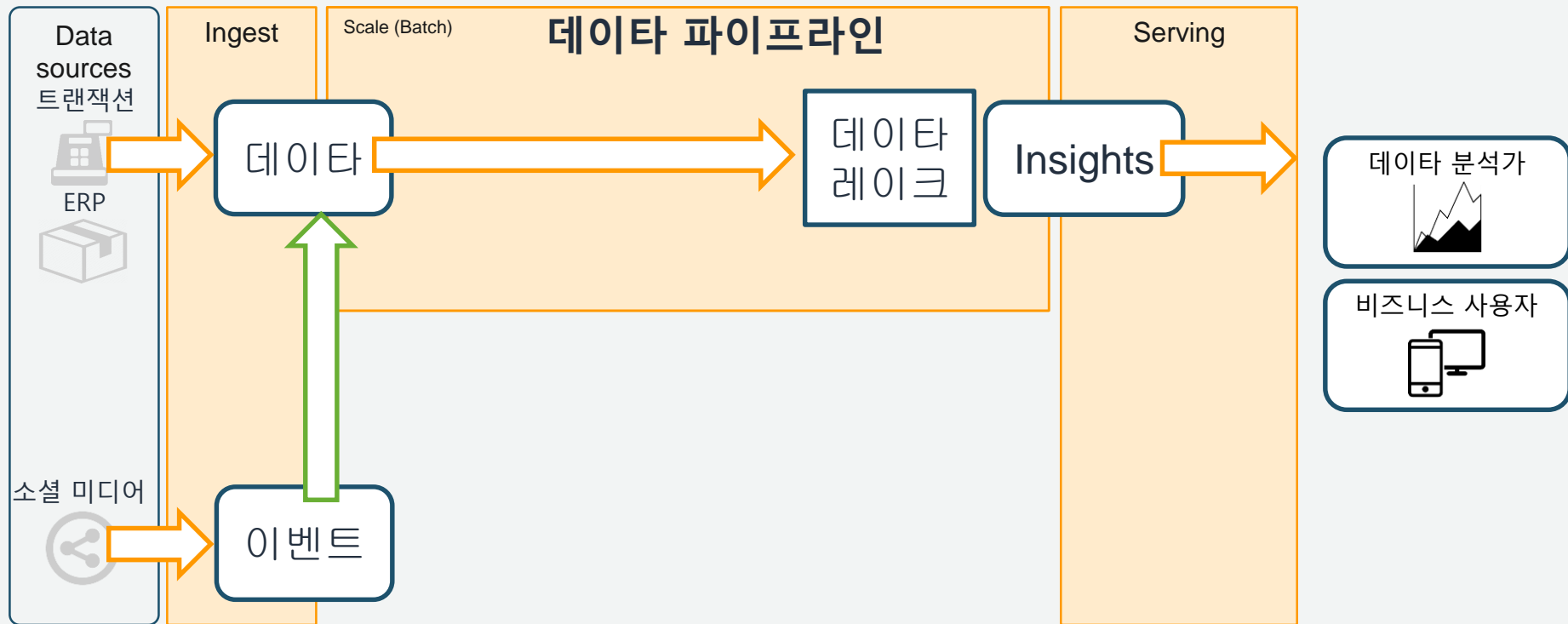
AWS IAM

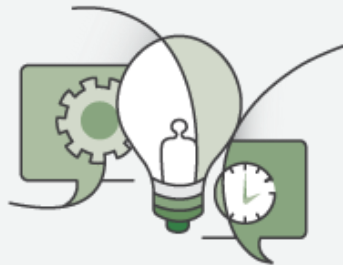


AWS KMS

현대화된 데이터 아키텍처

비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스





비즈니스 목표 2: 신규 매출 창출을 위한 혁신

A사는 소비자와 처음으로 직접 접촉하게 됨

소비자들이 우리 제품을 **사용하는 방식**은?

소비자들이 언제 **무엇에 귀기울이지**는 이해하려면?

어떤 타입의 소비자들이 사용중이고 사용이 늘어나는지?



마케팅

사용량 기반 **고객 세그먼트**는?

장소 및 HR(High Resolution)별 **음악 선호도**를 예측할 수 있는지?

음성 명령에 추가적인 **신호**이 존재하는지?

사진 내 장면으로 부터 **사용자 행동**을 유추 가능한지?



데이터 사이언티스트



필요한 특정 질문에서 시작, 이를 위해 필요한 데이터를 확보

소비자들이 우리 제품을 사용하는 방식은?

소비자들이 언제 무엇에 귀기울이지는 이해하려면?

어떤 타입의 소비자들이 사용중이고 사용이 늘어나는지?



마케팅



등록, 사용[시간/장소]
(모바일 앱)



미디어 소비
(파트너 API)



필요한 특정 질문에서 시작, 이를 위해 필요한 데이터를 확보

사용량 기반 **고객 세그먼트**는?

장소 및 HR(High Resolution)별 **음악 선호도를** 예측할 수 있는지?

음성 명령에 추가적인 **신호**가 존재하는지?

사진 내 장면으로 부터 **사용자 행동을 유추** 가능한지?



데이터 사이언티스트



등록, 사용[시간/장소]
(모바일 앱)

데이터 레이크로 적재



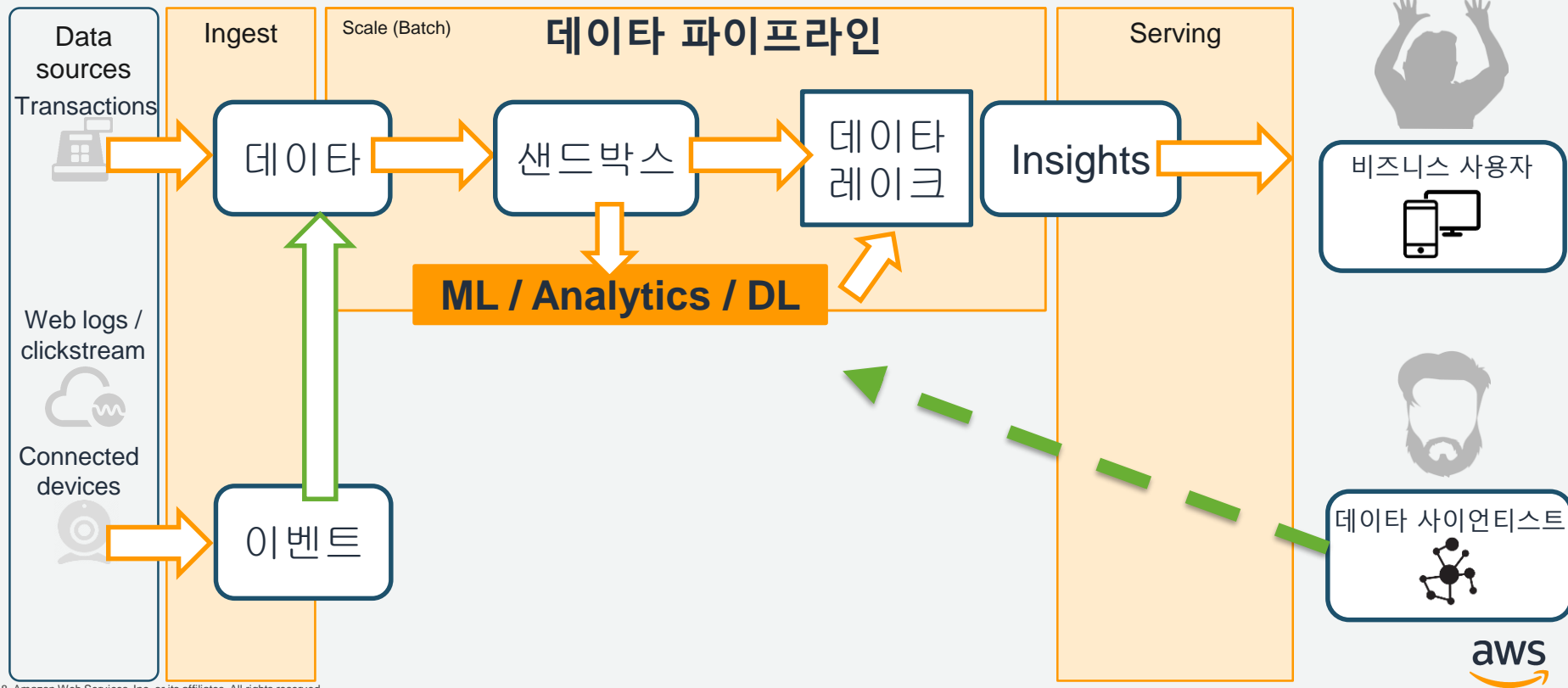
Hi-Res, 음성, GPS,
이미지(디바이스 데이터)

데이터 레이크로
이동 여부 판단 필요

샌드박스 – 신속, 저비용, 낮은 리스크

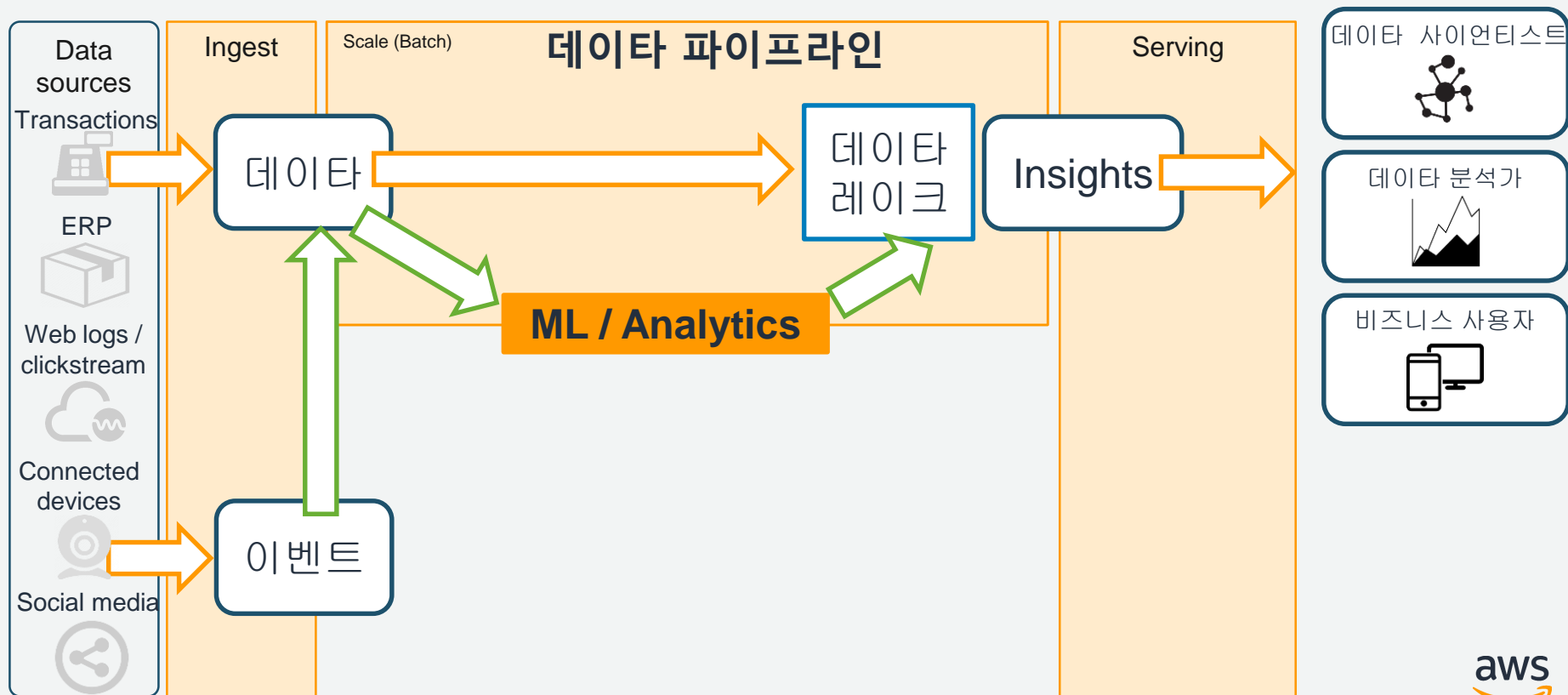
현대화된 데이터 아키텍처

비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스



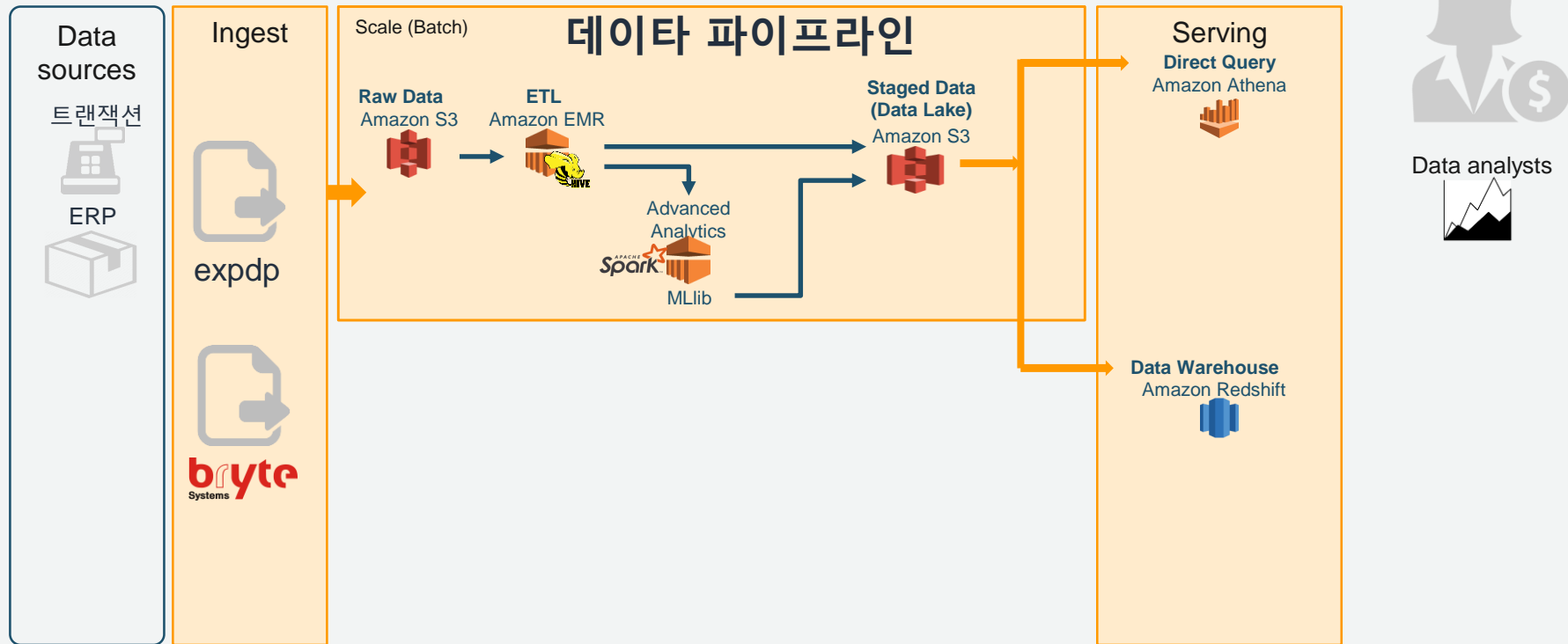
현대화된 데이터 아키텍처

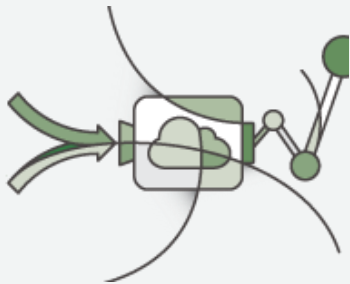
신규 매출 창출을 위한 혁신



현대화된 데이터 아키텍처

비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스





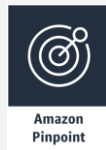
비즈니스 목표 3: Real-time engagement

A사는 개인화된 라이프 사운드트랙을 소비자에게 제시

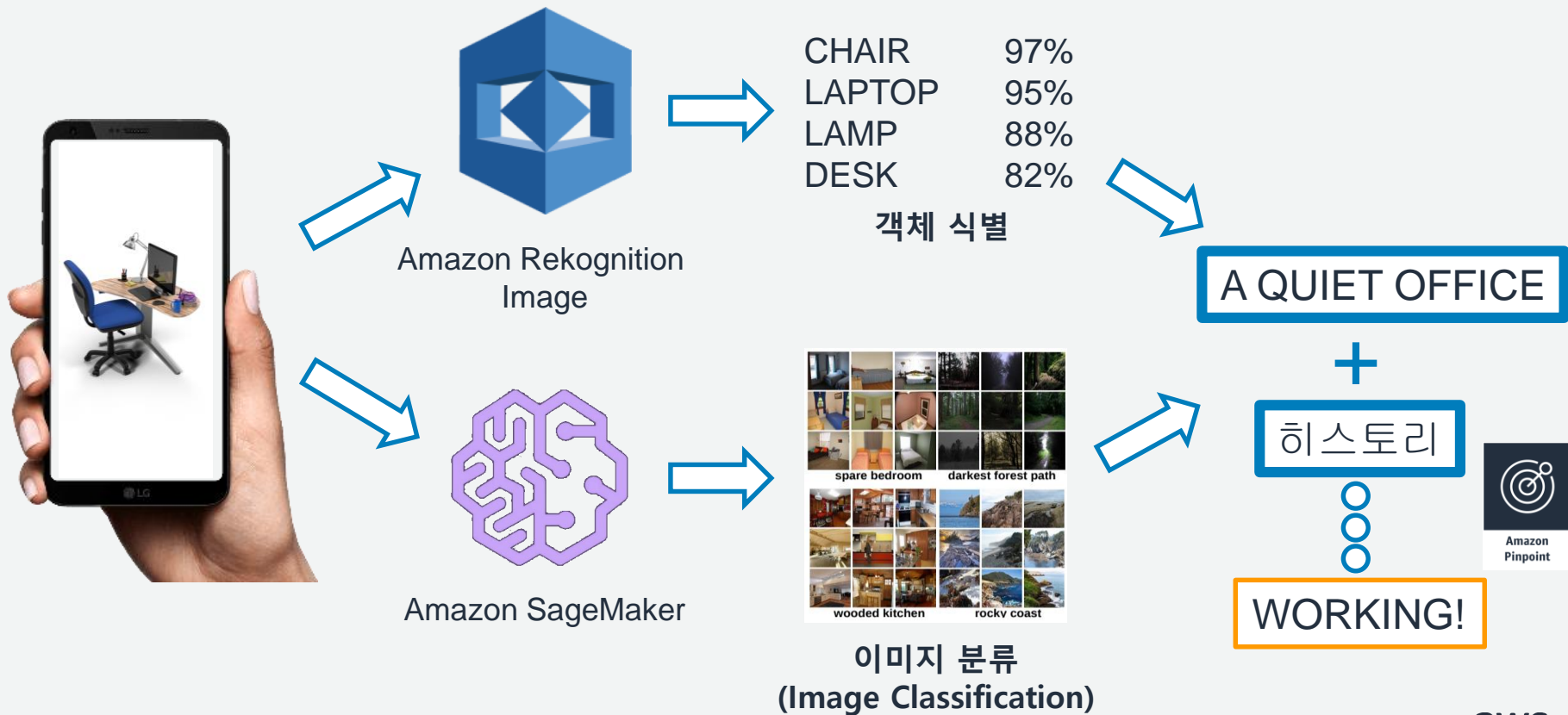
과거 선호도, 비슷한 행동 방식의
사용자 및 감지 환경을
기반으로 개인화



음성으로 뮤직 추천 요청

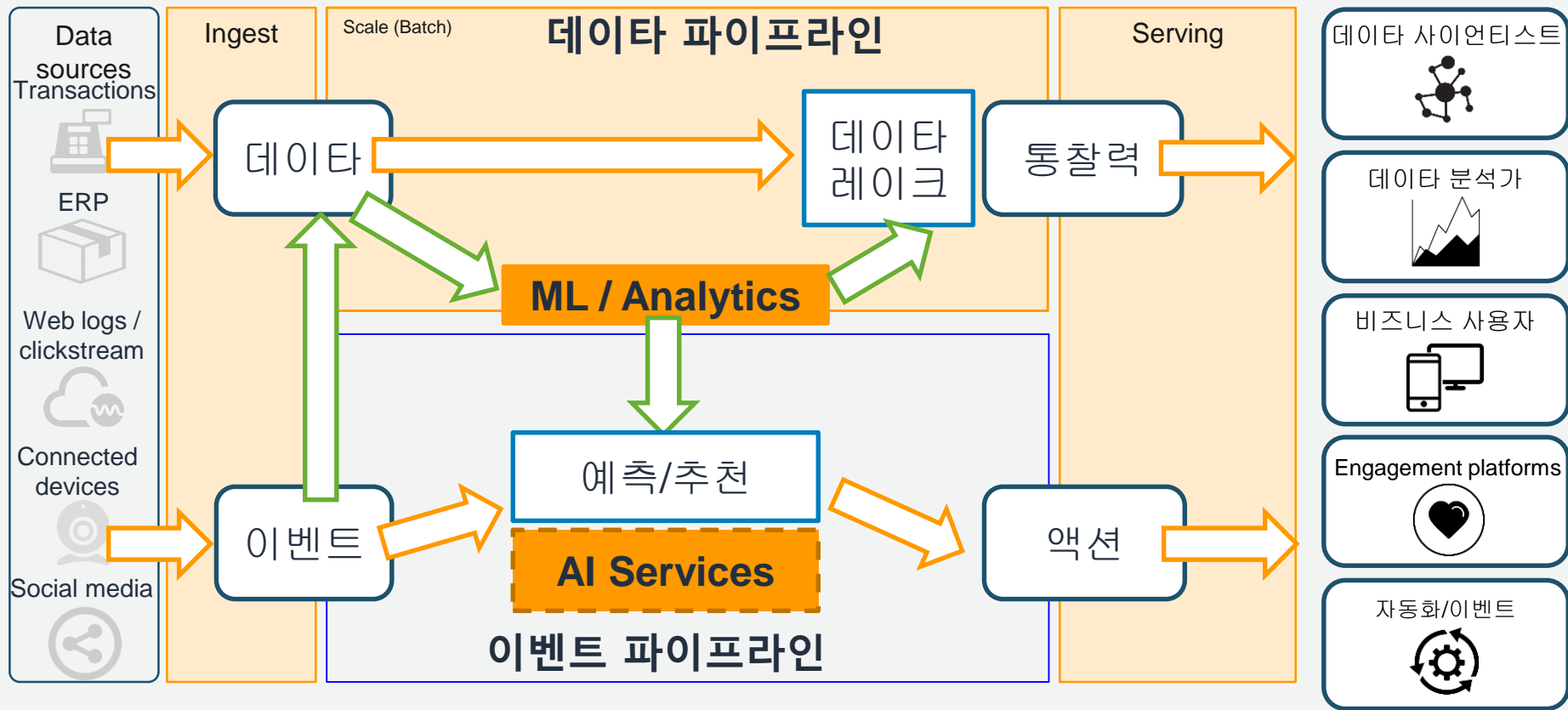


모바일 앱으로 소비자 활동 식별



현대화된 데이터 아키텍처

실시간 Engagement 및 Interactive 고객 경험



요약 - AWS는 데이터 레이크 접근을 위한 모든 서비스를 제공



수집



저장



분석 / 처리



시각화 / 활용



Kinesis
스트리밍 데이터



Direct Connect
데이터 센터와 연결



Snowball
벌크 데이터 로드



Database Migration Service
Oracle, Netezza 등의
데이터 импорт

더 많은 방법들..



Glue
데이터카탈로그와 ETL



Amazon S3
안전하고, 비용
효율적인 스토리지

어디서든 활용 가능한..



Redshift
데이터 웨어하우스



EMR
비정형 데이터 처리,
Apache Spark



Athena
ad-hoc 쿼리



QuickSight
시각화, BI



SageMaker
머신러닝 플랫폼

더 많은 방법들..

내부 사용자와 시스템

고객 대상 서비스

다양한 솔루션과 연동



데이터 중심 아키텍처 기반 비즈니스 목표 달성



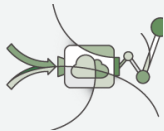
목표 1 : 현대화 및 통합

- 비즈니스 업무 개선을 위한 통찰력, 신규 디지털 서비스



목표 2 : 신규 매출 창출을 위한 혁신

- 개인화, 수요 예측, 리스크 분석



목표 3 : Real-time engagement

- 인터랙티브 고객 경험, 이벤트 중심 자동화, 사기 탐지

비즈니스 가치 창출을 위한 아이디어가 있나요?



고객 참여 및 채택에
직접적으로 영향을 주는
짧은 목록의 프로젝트를
작성

신규 아이디어를
테스트하고 데이터
레이크를 구축할 수 있는
간단한 데이터파이프
라인에서 시작

솔루션즈 아키텍트 및
Professional services 팀에게
지원 요청



Thank You!

