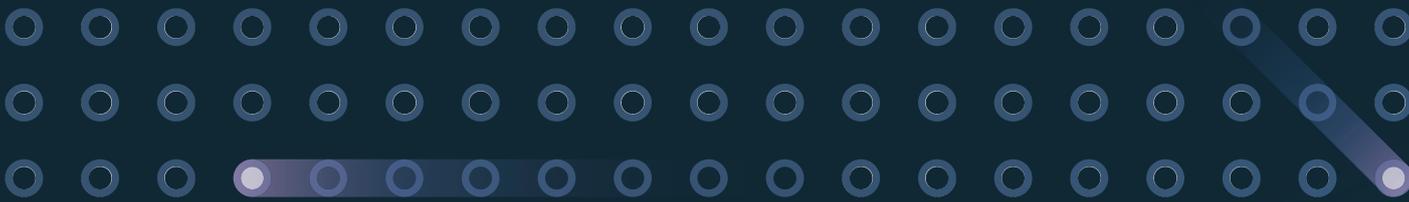
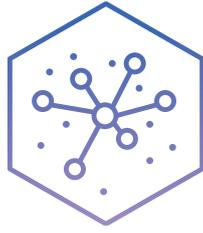


## Cloud Pak for Data

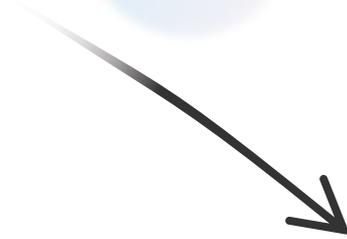
정보 아키텍처를 단순화하여 데이터 처리





IBM Cloud Pak for Data는 데이터의 수집, 거버넌스, 분석을 한 곳에 통합하고 간소화할 수 있는 단일 통합 플랫폼입니다. 기업은 통합된 클라우드 네이티브 아키텍처를 통해 데이터에서 비즈니스 인사이트를 얻어낼 수 있습니다. IBM Cloud Pak for Data는 IBM, 오픈소스, 써드파티 마이크로서비스 추가 기능을 통해 확장 가능하며 고객의 고유한 데이터 및 AI 환경에 쉽게 맞춤 적용할 수 있습니다.

골칫거리 데이터에서



신뢰받는 데이터로



먼저 현 상황을  
파악해 봅시다.



## 문제

우리는 우리가 생성하고있는 데이터를 활용하지 못하고 있습니다.

모든 기업은 오늘날의 환경에서 데이터 중심이어야 합니다.

데이터가 폭발적으로 생산되고 있는 오늘날, 매일 전 세계적으로 250경 바이트의 데이터가 생성됩니다.

하지만 기업에서 생성된 데이터 중, 평균적으로 약 70%가 전혀 사용되지 않습니다.

선두 기업 임원들의 79.4%가 데이터 중심 스타트업 또는 기업이 가져올 시장의 격변을 두려워하며, 73%의 기업만이 자사의 미래 데이터 전략을 확신하고 있습니다.

IA(Information Architecture)  
없이는 시도  
없습니다.

## 원인

대부분의 기업이 완전한 데이터 중심 기업이 되기 어려운 이유

1. 대부분의 기업은 다수의 다양한 레거시 시스템을 보유하고 있으므로 모든 데이터 소스를 연결하고 활용하기가 어렵습니다.
2. 산업계는 데이터를 보호하기 위해 보다 엄격한 규제를 마련하고 있습니다.
3. 디지털 전환으로 인해 기업은 과거와는 다르게 전략적으로 사고해야 합니다.
4. 데이터 팀은 비즈니스 인사이트를 찾아 내는 작업을 시작하기 전에 단순히 데이터 세트를 찾고, 준비하고, 검증 및 관리하는 데 50~80%의 시간을 사용하고 있습니다. 정보를 공유하지 않는 개별적 팀의 규모 증가와 업무 복잡성도 비효율성에 기여하는 요인입니다.
5. 97.2%의 임원들이 빅 데이터 및 AI 이니셔티브를 수립하거나 시작한다고 말하는 주된 이유는, 지난 10년 동안 포춘 1,000대 기업 중 거의 3/4이 바뀌었기 때문입니다. 많은 기업들이 데이터 중심 기업의 모범이라 할 수 있는 Facebook이나 Amazon과 같은 회사로 대체되고 있습니다. 이러한 데이터 중심 기업은 데이터와 AI를 사용하여 비용을 절감하거나 비즈니스를 성장시키는 한편 산업계 전반을 재구성하고 있습니다.

데이터 중심의 비즈니스를 실행하는 것이 날이 갈수록 어려워지고 복잡해지고 있습니다. 하지만 데이터 중심의 비즈니스 실행은 분명 필요합니다.

# IBM 접근법

뛰어난 AI 및 분석 성과를 달성하기 위해서는  
응집력있는 데이터 전략이 필요합니다.

IBM의 접근법은 데이터가 디지털 전환을 촉진하고 AI가 데이터의 가치를 드러내며 하이브리드 클라우드를 통해 누구나 데이터에 접근할 수 있는 3가지 핵심 요소를 바탕으로 합니다. 이러한 요소들이 데이터와 AI로부터 진정한 비즈니스 가치를 얻어내기 위한 전략을 이끌어냅니다.

## 성공적인 AI로 가는 여정



### Infuse(AI 확산)

신뢰성과 투명성 있게 AI를 운용



### Analyze(분석/머신러닝)

어디에서든 AI로 인사이트를 확장



### Organize(거버넌스/품질)

신뢰할 수 있는 분석 기반을 마련



### Collect(수집)

데이터를 단순화하고 접근성을 개선

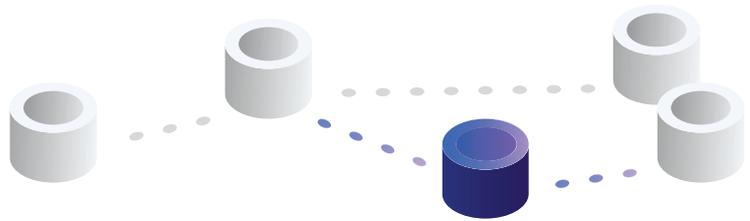
## AI의 미래는 유연합니다.

하이브리드 멀티 클라우드 접근법에서 모든 것이 시작됩니다.

클라우드를 넘어서는  
세상을 바라봅니다.



“정보 아키텍처를 단순화하여 올바른 데이터를 업무에 쉽고 빠르게 활용하세요.”



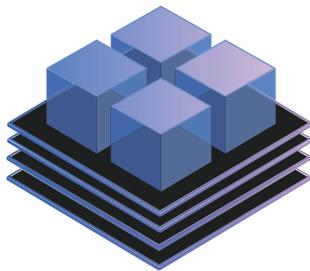
정보 아키텍처(IA, Information Architecture)를 간소화하는 방법



이동하거나 복제할 필요 없이 여러 소스에 걸친 데이터를 탐색하고, 연결하고, 통제 관리하고, 활용합니다.

정제, 매칭, 메타데이터 생성과 같은 반복적인 작업을 자동화하여 데이터 준비 시간을 80%까지 단축합니다.

모든 퍼블릭 클라우드, 하이브리드 클라우드, 프라이빗 클라우드와 Red Hat OpenShift 간의 배포 유연성을 극대화 합니다.



모든 데이터 사용자가 협력하고 여러 개의 분석 애플리케이션과 모델에 접근할 수 있게 하여 서로 고립화된 업무 관행을 타파합니다.

45개 이상의 서비스가 통합된 플랫폼 환경을 통해 팀의 수많은 작업 워크플로우와 운영 관리를 중앙화 합니다.

숙련된 데이터 팀이 혁신을 이끄는 비즈니스 가치를 찾기 위한 빅 데이터 기반 분석에 더 많은 시간을 투자할 수 있게 합니다.

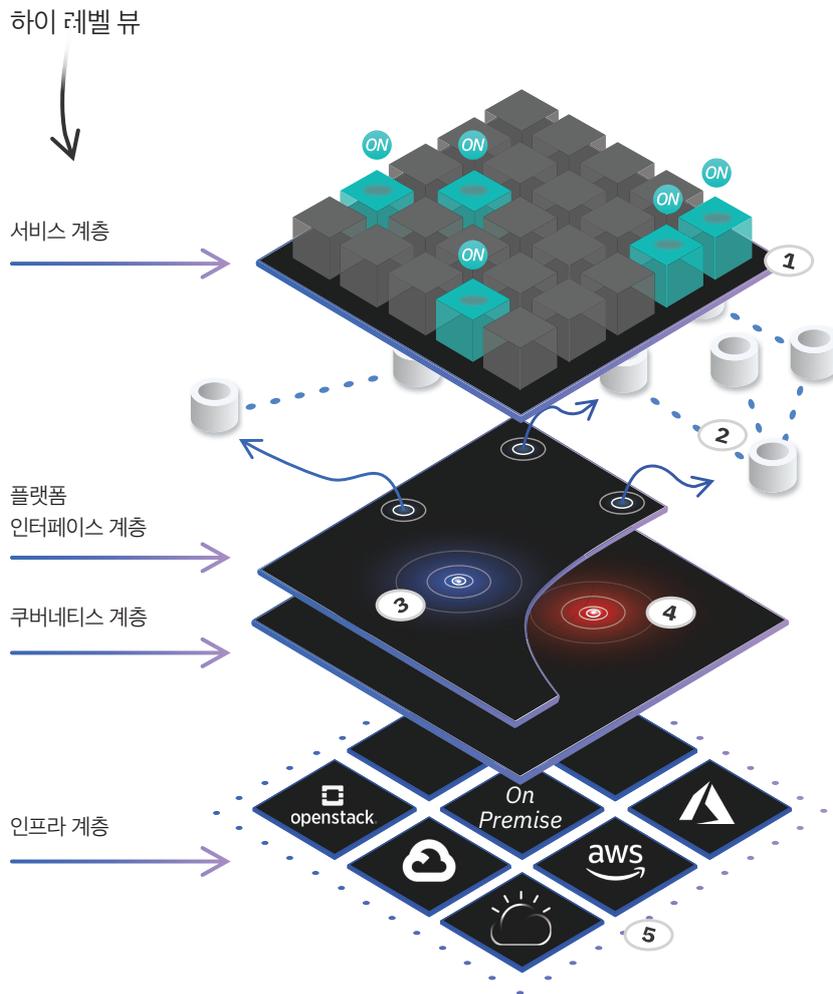
데이터를 완전히 활용하는 방법



# 플랫폼

## 완벽하지만 단순한 구조

Forrester Wave의 선도적인 전사적 인사이트 플랫폼(enterprise insights platform)이 어떻게 구축되었는지 살펴보세요.



### 1. 서비스 에코 시스템

한 번의 클릭으로 IBM 및 써드파티가 제작한 45개 이상의 분석 서비스 및 템플릿에 액세스하고 이를 배포합니다.

자세한 내용은 8페이지 참조

### 2. 데이터 가상화

데이터를 이동할 필요 없이 빠르고 쉽게 여러 데이터 소스를 쿼리합니다.

자세한 내용은 7페이지 참조

### 3. 플랫폼 인터페이스

데이터 관리, 데이터 거버넌스 및 분석 업무를 통합하여 효율성을 높이고 리소스 활용도를 개선하는 단일 플랫폼으로 가치 실현 시간을 단축하세요.

### 4. OPENSHIFT<sup>™</sup> by Red Hat

혁신적이고 신속한 배포 전략을 위한 선도적인 하이브리드 클라우드, 엔터프라이즈 컨테이너 플랫폼을 활용하세요.

### 5. 하나의 클라우드에 종속되는 현상에서 벗어나,

어떤 클라우드든 상관 없는 IBM의 'Any Cloud' 방식으로 모든 클라우드 인프라를 활용하세요.

위에 참조된 페이지에서 자세한 내용을 확인하세요.

# 데이터 사일로 문제 해결

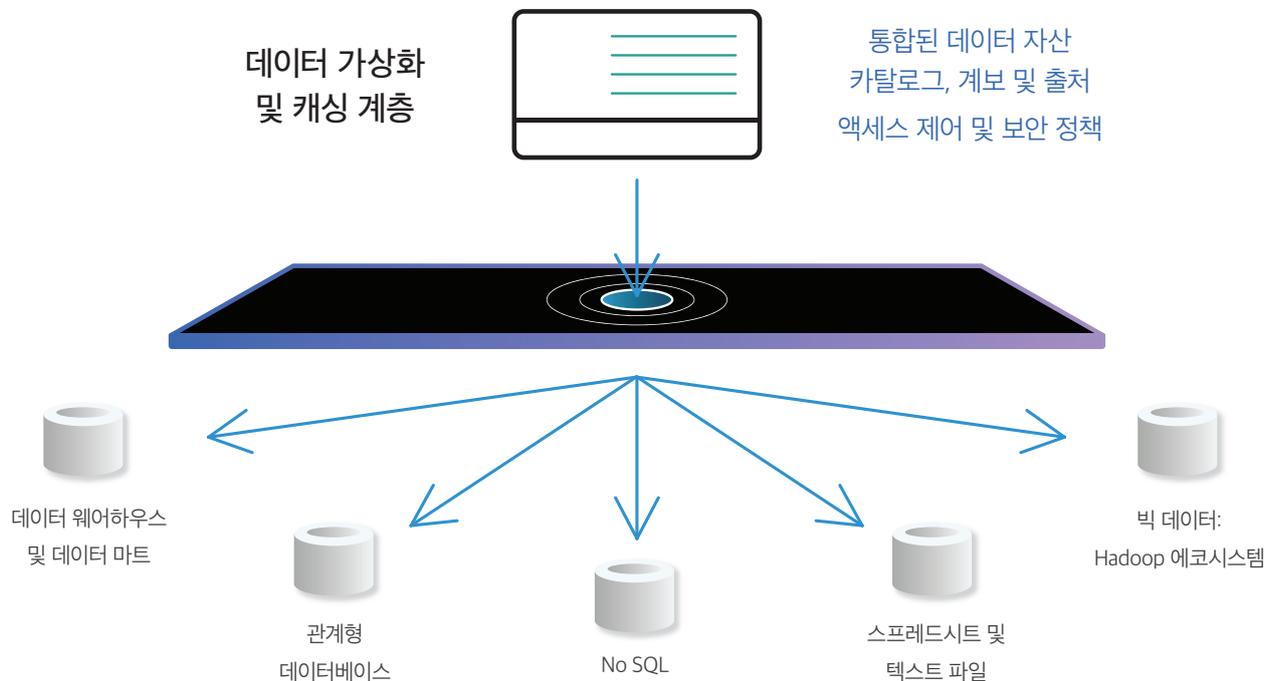
## 기존 인프라를 대폭 변경할 필요 없이 여러 데이터와 팀을 서로 연결하세요.

데이터 사일로는 데이터에서 얻을 수 있는 인사이트를 감추고 있을 가능성이 많습니다. 결과적으로 숙련된 데이터 팀이 이러한 데이터 리소스를 검색하고 이동하고 관리하는 지루한 일을 하게 됩니다. 이러한 잔업이 하나의 이니셔티브에 할애하는 시간 중 80%를 차지하는 경우도 많습니다.

데이터 가상화 기능은 이러한 데이터 사일로를 연결하여 마치 단일 데이터 세트인 것처럼 만들어 줍니다. 또한, 데이터가 위치한 서버에 대해 분석 쿼리를 수행한 다음 그 결과를 본래의 애플리케이션에 반환합니다.

데이터는 복제되지 않습니다. 데이터는 소스에만 존재합니다.

보고서 읽기



# 통합 서비스 메뉴

통합된 엔드 투 엔드(end-to-end) 분석 서비스를 통해 AI로 가는 과정에서 비즈니스를 성장시키고 확대하세요. 배포는 간단하며 프리미엄 추가 서비스는 유연한 라이선스 모델을 갖추고 있습니다. 바로 이용할 수 있는 여러 기능들을 알아보세요.

- Cloud Pak for Data 기본 기능
- + 프리미엄 추가 기능 액셀러레이터 또는 기존 라이선스 트레이드업

추가 라이선싱



## Collect (수집)

### 데이터 가상화 ●

위치와 관계없이 무엇이든 쿼리(여러 데이터 소스에 걸친 가상화된 데이터) / 내장된 거버넌스 기능을 통한 데이터 소스 & 메타데이터 자동 탐색 · 분산 병렬 처리

### Db2 Warehouse ●

인메모리 최적화된 열 기반 엔진 · SQL, Spatial, XML 및 JSON 지원 · 페타바이트 규모로 확장, 여러 DB와 호환 및 이식 가능

### Db2 AESE +

Db2 Advanced Enterprise Server Edition은 트랜잭션, 웨어하우스, 복합 워크로드에 적합

### Db2 Event Store ●

이벤트 중심 애플리케이션을 위한 스트리밍 데이터를 신속하게 처리하고 분석하도록 설계된 메모리 최적화 데이터베이스

### PostgreSQL ●

데이터 카탈로그

### Streams ●

IBM Streams 추가 기능을 통해 스트리밍 데이터를 처리하는 애플리케이션 개발 및 실행. IBM Streams를 사용하면 방대한 양의 이동 데이터를 지속적으로 빠르게 분석하여 비즈니스 인사이트 및 의사결정 속도를 향상시킬 수 있습니다.

### MongoDB +

교차 플랫폼 문서 지향 데이터베이스 프로그램 (cross-platform document-oriented database program)



## Organize (거버넌스/품질)

### 데이터 발견 ●

### 데이터 통합 ●

DFD & 데이터베이스에서 서비스 포함  
- 작업 생성, 업데이트, 삭제 · 연결 생성  
- 작업 컴파일 · 작업 로그

### Watson Knowledge Catalog Pro ●

기업의 데이터 거버넌스 플랫폼과 긴밀하게 통합되는 개발자를 위해 설계된 오픈소스 오브젝트 관계형 데이터베이스

### Infosphere DataStage Edition +

ETL 도구 및 IBM Information Platforms Solutions 제품군과 IBM InfoSphere의 일부

### Data Catalog ●

Information Governance Catalog & Watson Knowledge Catalog에서 서비스 포함  
- 비즈니스 용어 가져오기(UI) · 정책 & 규칙 작성  
- 정책 & 규칙 가져오기(UI) · 자산 탐색기 · 자산 검색  
- 그래프 탐색기 · 주석, 등급

### Infosphere Information

### Regulatory Accelerator ●

비용과 복잡성을 줄이도록 설계: 머신러닝을 통해 규정과 관련된 용어 체계에서 주요 용어, 사용 가능한 정의, 정책 및 규제를 추출하여 해당 프로세스와 관련된 수작업 감소



## Analyze (분석/머신러닝)

### Watson Studio ●

- 환경(Jupyter, RStudio, Zeppelin 등)
- 스크립트 & 작업 자동화 · 머신러닝 프레임워크 & Spark
- 이미지 & 패킷 관리
- 모델 관리& 개발 · Git 버전 제어

### Cognos Dashboard ●

보고, 모델링, 분석, 탐색, 대시보드, 스토리, 이벤트 관리를 통합하여 조직의 데이터를 이해하고 효과적인 비즈니스 의사 결정을 내릴 수 있습니다.

### Watson Studio Premium +

SPSS Modeler, Data Refinery · Decision Optimization · Model Builder(AutoML)  
SPSS/Notebook용 Hadoop 서비스, SPSS SQL Pushback · WML  
고급 학습(배치 학습, HPO, 분산 딥러닝) · 지속적 학습 서비스

### Watson OpenScale ●

시를 그 수명주기(Lifecycle)에 걸쳐 운영하고 자동화할 수 있는 오픈 플랫폼



## Infuse (AI확산)

### Watson API Kit +

- Watson Knowledge Studio · Natural Language Understanding
- Speech to Text · Text to Speech

### Watson Discovery +

- Watson Knowledge Studio · Watson Explorer
- Watson Discovery

### Watson Assistant +

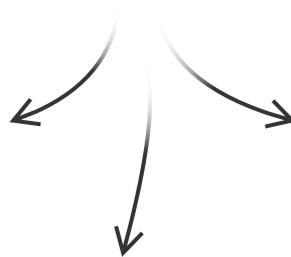
애플리케이션, 디바이스, 채널에 대화형 인터페이스 구축

### Cognos Analytics +

비즈니스에 대한 인사이트를 쉽게 시각화, 분석, 공유할 수 있는 비즈니스 인텔리전스 및 분석 솔루션

### 클라우드 네이티브 에디션

어떤 클라우드에서든 지원 가능



### 엔터프라이즈 에디션

어떤 클라우드에서든 지원 가능

### Cloud Pak for Data시스템

소프트웨어 + 하드웨어 최적화되고 검증된 하이퍼 컨버지드 시스템

