

*We Provide more than Solution*

Leading Enterprise Data Management and Analytics Company :  
Solution Provider for Data Integration and Quality Management

# 빅 데이터 시대의 전략적 기업 데이터 거버넌스



# Introduction



데이터 기반의 의사결정 **Think!!!**

**모든 데이터와 정보는  
정합성이 확보되어야 하며  
신뢰성이 동반되어야 한다**

# Big Data?

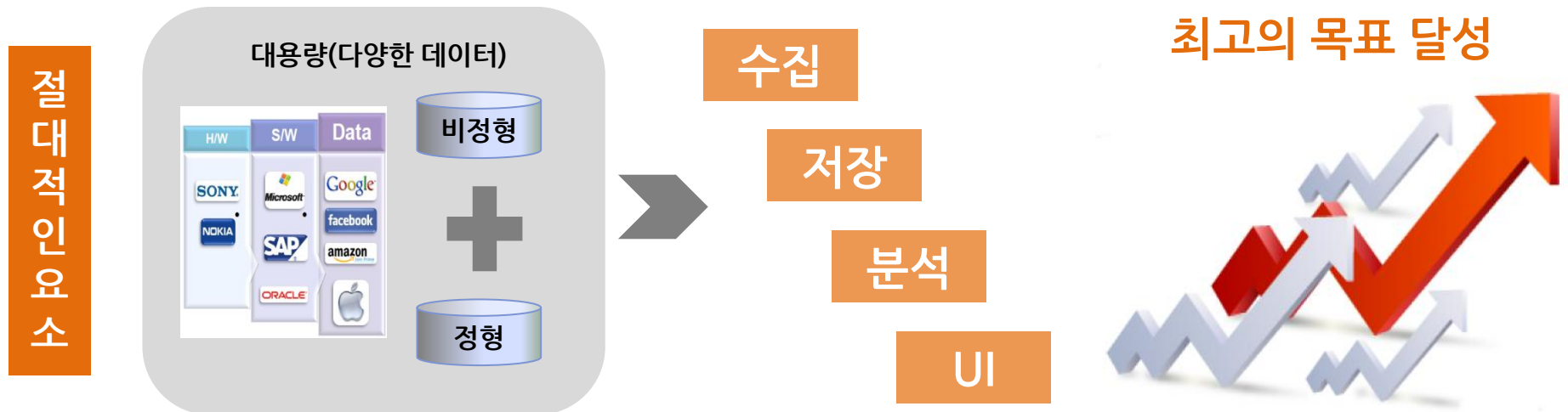


빅 데이터란 무엇인가?

데이터를 통해 공유하고 활용할 수 있는 정보의 집합체!

Data 활용 증가는 영업효율향상과 매출증대로 이어진다!!

데이터를 사용하고 공유할 수 있어 모든 사람이 행복해진다!!!



## 빅 데이터의 서비스

# 다양한 서비스층의 욕구를 분석하여 적시에 정보 공유와 활용

### 대용량(다양한 데이터)

- 그룹 차원의 실시간 로그 데이터(거래, 서비스, 마케팅)
- 업종의 경계를 넘어서는 다양한 데이터(고객)

### 정확도(더 많은 정보와 발견)

- 고객 접점 확대 및 연계 영역에 대한 분석과 예측
- 다양한 리스크 분석 및 그룹 차원의 브랜드 목표 관리

### 파워(분야의 확대)

- 고객 접점에서 맞춤형 비즈니스 모델 개발
- 비즈니스 변경 및 통합에 대한 유연성 확보

Real-time Information System

Analysis & Prediction

Mash-up Service

## 비정형 데이터와 정형 데이터를 접목시켜 비즈니스에 적용한 빅 데이터 활용 사례 (해외)

구분	Machine/Sensor 데이터 분석	소셜미디어/마케팅/동향/로그분석	사기/범죄/재해 예방 및 적발
제조	<ul style="list-style-type: none"> <li>GM/Volvo : Early Warning 시스템에서 초기 1천대 생산 후 결함 있는 Lots 파악에 활용</li> <li>Daimler : 텔레매틱스 데이터 분석에 활용</li> <li>제조설비 예방정비활동 시 사전 예측을 통한 교체 예상 부품 주문으로 다운타임 최소화</li> </ul>	<ul style="list-style-type: none"> <li>Barns &amp; Noble : 웹 행동분석, 개인화 추천 등</li> <li>네슬레 : 고객 선호도를 SNS 투표를 통해 분석</li> <li>아고다닷컴 : 해외호텔 최저가 예약 서비스</li> <li>Ebay : 통합 Collaboration, BI, DataHub 적용</li> <li>유통업 : 카테고리 단위로 연관 분석에 활용</li> </ul>	N/A
금융	<ul style="list-style-type: none"> <li>美 Progressive 보험사 : OBD 주행 데이터에 근거한 할인을 적용(자동차)</li> </ul>	<ul style="list-style-type: none"> <li>상황인식 서비스 기반의 1회성 보험 상품 판매                             <ul style="list-style-type: none"> <li>- 도쿄 해상 : 고객상황에 적절한 보험 상품 Offer</li> <li>- NTT 도코모 : 고객성향, 위치정보, 선호도 제공</li> </ul> </li> <li>Discover 사 : 금융 가맹점 간 연관성 분석에 활용</li> <li>Intuit 사 : 결제대행정보 기반 상권분석 정보 제공</li> </ul>	<ul style="list-style-type: none"> <li>보험/카드사 등 금융권 사기 적발 및 예방에 활용</li> </ul>
ICT	<ul style="list-style-type: none"> <li>통신사 : GPS(위치) 기반 Profiling Service 제공</li> </ul>	<ul style="list-style-type: none"> <li>AT&amp;T : 개인 맞춤형 서비스, SNS 분석</li> </ul>	N/A
유통/서비스	N/A	<ul style="list-style-type: none"> <li>Fedex : 소셜 커뮤니케이션을 통한 분석 시스템 구축</li> <li>Walmart : 소셜 데이터를 통한 투자 수익 증대</li> </ul>	N/A
에너지	<ul style="list-style-type: none"> <li>원유 등 에너지 자원 비축량 평가 및 수요 예측</li> <li>디지털 유전 분석(Upstream)</li> </ul>	N/A	N/A
공공	<ul style="list-style-type: none"> <li>美 남가주 Edison 사 : Smart Grid 분야 적용</li> </ul>	<ul style="list-style-type: none"> <li>美 대선 : 오바마 캠프에서 유권자 분석에 활용</li> <li>Blue Cross &amp; Blue Shield of Minnesota (비영리 건강보험사) : 고성능 분석솔루션 도입/활용</li> </ul>	<ul style="list-style-type: none"> <li>美 IRS(국세청) : 탈세(부정환급) 방지</li> <li>Terrorist Network Detection</li> <li>범죄 발생 예측 분석</li> </ul>

# Data Governance?





## 데이터 거버넌스란 무엇인가? 안정적인 정보서비스를 제공하기 위한 조직 및 역할 체계 구축

### Data Governance

#### Structure

- 경영전략과 연계된 IT전략 수행 및 자원 활용 극대화를 위한 조직 구조

#### Process

- IT서비스 수행 및 고객 대응, IT서비스 관리를 위해 요구되는 내부 업무 프로세스

#### People

- 전략 달성을 위한 조직 구조 하에서 프로세스 수행을 위해 요구되는 역할 및 책임

#### Alignment and Communication

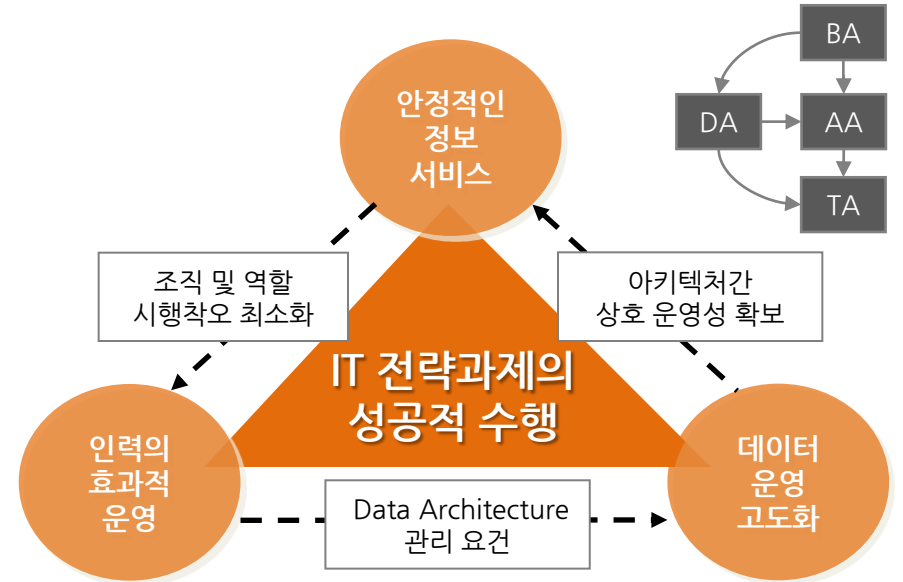
- 대 고객 및 IT부서원간 IT전략 달성을 위한 의사 결정 과정

#### Tools, Metrics and Investment Appraisal

- IT관리체계 수행 효과성 향상을 위한 자동화 도구, 이에 따른 성과 측정지표

Source: Gartner

### Data Governance Architecture



- Information-as-a-Service (IaaS) : 전사 관점의 정보 서비스의 균일성 보장
- 현업이 데이터 품질 관리의 Initiative를 갖도록 데이터 관리 전략 및 협업체계 수립
- 아키텍처 상호 간 표준 및 원칙 준수에 대한 통제 강화

## 데이터 거버넌스 기반으로 정보 활용 체계 정립

### 데이터 통합 관리 체계 구현

1

관리체계(표준, 구조, 흐름, 품질) 통합

- 자산으로써 데이터와 관리체계 확보
- 데이터 운영 정책 및 데이터 관리 원칙 마련

2

마스터 데이터 관리 통합

- 핵심정보 발굴, 식별, 통합, 관리 및 활용 체계
- 전사 마스터 데이터 통합 관리체계 구축

3

데이터 통합 방안 및 빅데이터 분석 모델

- 데이터 수집(Collect), 분석 모델(Mart), 검증규칙(Rule), 구현(view) 데이터 통합 분석 아키텍처

### 데이터 거버넌스 Single View

1

정보 활용도 제고로 정보 가치 창출

2

신속한 의사결정 지원 체계

3

업무의 생산성 및 효율성 제고

4

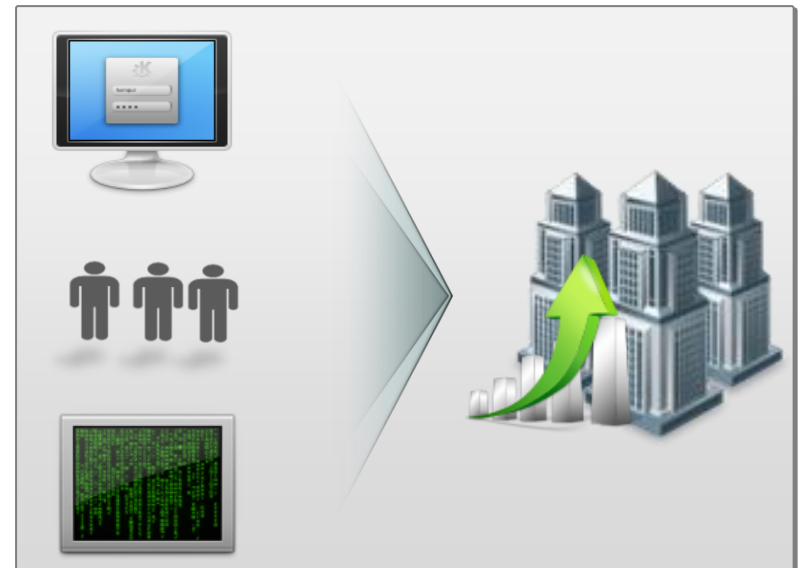
통합 데이터 관리 관제 센터

## 데이터 거버넌스의 목표

# 기업이 제공하는 정보 활용을 극대화하여 기업의 자산가치 창출

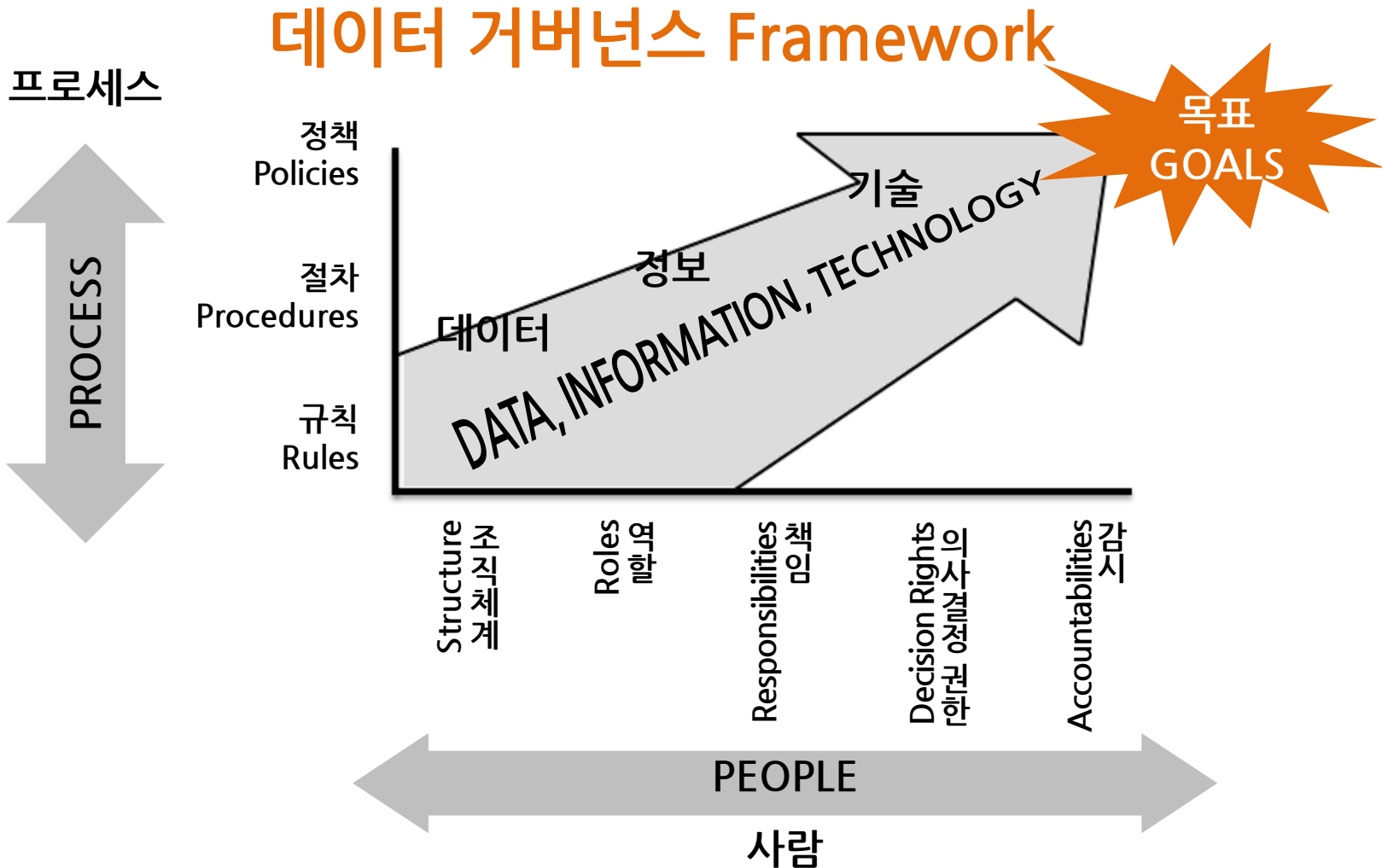
- 데이터 거버넌스(Data Governance)는 기업에서 사용하는 데이터의 가용성, 유용성, 통합성, 보안성을 관리하기 위한 정책과 프로세스를 다루며 관리규정
- 데이터 거버넌스(Data Governance)란 전사적인 차원에서 보유하고 있는 모든 데이터의 관리에 대한 정책, 지침, 표준, 전략 및 방향을 수립하고 데이터를 관리할 수 있는 조직 및 서비스를 구축하는 데이터 관점에서의 IT 관리 체계

**“조직, 프로세스, 기반기술을  
효과적으로 관리하여, 데이터를 기  
업의 비즈니스를 위한  
핵심 자산으로 최적화하고,  
활용을 극대화하기 위한 방안”**



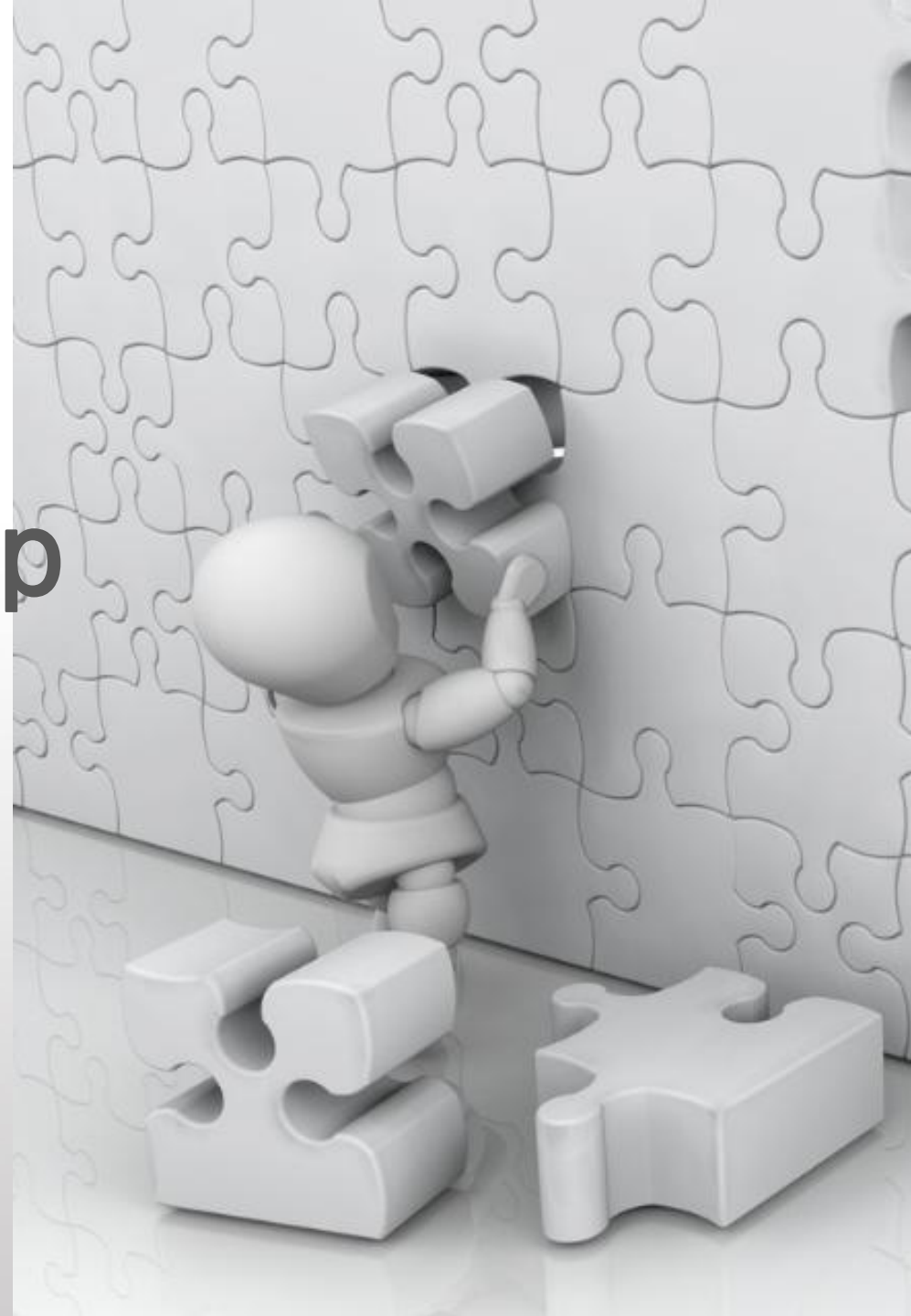
# Data Governance Framework





\* TDWI(The Data Warehousing Institute): DW(Data Warehousing)를 포함한 데이터 관리 분야의 전문기관으로서 매년 국제 컨퍼런스 개최, 세미나 등 사업을 수행  
[참고 문헌] Data Management International, "DMBOK(Data Management Body Of Knowledge)", (<http://www.dama.org>)

# How We Can Set Up Data Governance?



## 데이터 거버넌스 체계 수립 필수 요건

# “데이터 관리 체계/정책 수립과 프로세스의 일관성 있는 적용”

정보화 성숙에 따른 경쟁 환경 변화 : 정보 활용을 통해 경쟁 우위 확보 시도

데이터 품질 유지를 위해 조직, 권한, 역할, 프로세스 정립의 중요성 인식

데이터 서비스 관점에서의 통합 상호 운용성 필요

데이터를 가치 창출과 자산으로 인식하고 지속적인 유지관리 필요성 인식

통합 데이터 관리와 활용에 대한 성과지표와 통제 기반 조성

## 데이터 거버넌스 체계 수립을 위한 근본적인 숙제

### 근본적인 문제점

#### 전사 데이터 관리체계 미흡

- 데이터를 누가 만들고, 어떻게 사용하는지는 파악 어려움
- 특정 데이터의 변경이 다른 작업에 미치는 영향도 파악 힘들

#### 전사 데이터 표준 부재

- 각 시스템 별로 관리되고 있는 데이터에 대한 전사적인 표준화 기준이 없음

#### 데이터 품질관리 정책 부재

- 데이터 품질측정기준 및 모니터링 체계 부재
- 데이터의 Ownership 미 규정

#### 정보의 이해도/활용도 부족

- 원하는 데이터의 위치 및 특성 파악을 위해 많은 시간 소요
- 비즈니스 룰에 대한 정보 부재

### 시사점

- 전사관점의 표준/품질 및 데이터의 기획/실행/통제를 위한 데이터 관리체계 수립

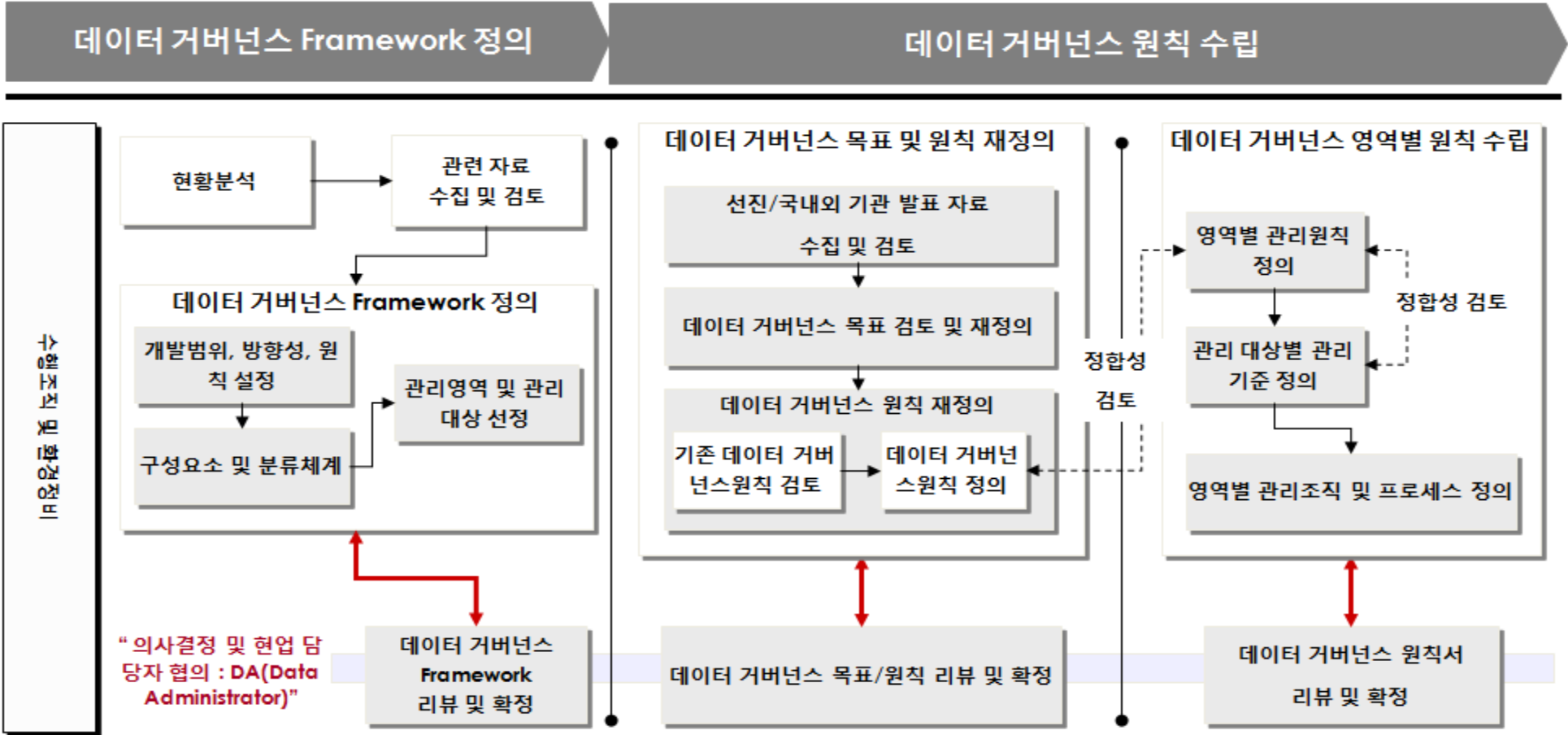
- 전사 데이터를 체계적으로 관리할 수 있는 조직 및 프로세스 정의

- 사용자 및 운영자의 정보활용 용이성을 위한 메타데이터 체계 구축
- 전사 공용의 참조데이터의 single view를 제공하기 위한 마스터 데이터 체계 구축

- 품질의 지속적 측정 및 향상이 가능한 데이터 품질관리 체계 구축



## 데이터 거버넌스 체계 수립 방안 (예시) 데이터 관리 영역과 대상을 선정하여 영역별 원칙서 작성



# Strategic Data Governance



## 전략적 데이터 거버넌스 데이터 관리 체계 융합과 관제 센터 구축

### 데이터 거버넌스 기반의 데이터 관리 체계 융합

데이터 거버넌스 정책, 통합 view, 워크플로우, 데이터 품질 성과 관리, 보안관리, DB 관리, 모니터링

법, 규정, 지침

통합 포털

통합  
워크플로우

데이터  
측정 / 평가

데이터  
관리 계획

데이터  
흐름 관리

마스터 데이터 관리

### 데이터 솔루션 / 툴

#### 표준/메타관리

- 표준 용어/단어
- 표준 도메인
- 표준 코드
- 마스터 데이터

#### 데이터 품질

- 품질진단/개선
- 품질관리계획
- 품질측정, 평가
- 모니터링
- 기준정보

#### 영향도 관리

- DB 영향도
- APP 영향도
- BI 영향도
- DATA CRUD

#### 형상 관리

- 데이터 표준
- 데이터 구조
- 물리 DB

#### 흐름 관리

- 데이터 추출
- 데이터 이행
- 데이터 배포
- 데이터 흐름

#### 툴/통합 관리

- 모델링 툴
- 데이터 플로우 툴
- 그룹웨어
- 레포지터리 통합툴

### 데이터 소프트웨어 융합

Single Sign On  
SW

Quality  
SW

MetaData  
SW

Impact  
SW

Extract/Transfer/  
Loading SW

RealTime  
Extract SW

MDM  
SW

Big Data

## 전략적 데이터 거버넌스 구축 방안 전사통제 조직, IT전담 조직, 업무 조직 간의 역할 구성원 정의

### 수행방안

#### 데이터 거버넌스 조직 설계 시 중점 고려사항

- 주요 시사점 및 개선 방향에 근거하여 고품질 데이터 확보를 위한 데이터 전담 조직과 구성원 정의
- 현황 분석을 통하여 데이터 거버넌스 프로세스, 조직, 인력 운영체계, 규정 및 지침 현황을 파악하여 주요 이슈 및 시사점을 도출함
- 고품질 데이터를 유지하기 위한 데이터 관련 원활한 협업체계 유지
- 데이터 분석, 데이터 표준 및 데이터 모델 통제, 데이터 상시 감시 및 사전/사후 데이터 정비 등의 활동을 주관하는 데이터 전담 조직 설계
- 데이터 오류에 대한 핵심 현안 발생 시 책임을 갖춘 의사결정기구 도입 및 역할과 임무 명시

### 역할 및 조직 구성원 정의

데이터 거버넌스 체계

1

통제조직, 전담조직, 업무조직 간 상호 협업

일관성 있는 데이터 표준 유지

2

데이터 분석 표준, 품질 및 정제기반 프로세스 지원

데이터 통합 모델 관리

3

통합, 품질 상시감시 및 모니터링 시스템 체계

모델-DB 간 정합성 확보

구조적 데이터 품질 관리

데이터 값, 활용 품질 관리

데이터 오류 재발 방지

데이터 오류 재발 방지

## 전략적 데이터 거버넌스 기대효과: 데이터 관리 측면 데이터 품질 향상과 지속적 유지관리

- ▶ 데이터 중복에 따른 비용을 줄이고 데이터 품질을 향상시키며 어플리케이션 개발 및 변경 시 데이터의 신속한 파악과 유연한 대응을 가능케 한다.

### 데이터 통제 강화

- 회사 자산으로서 전사적인 데이터 파악 능력 강화
- 데이터 수집과 배포 흐름 체계 강화

### 데이터 품질 향상

- 데이터 본질 및 특성에 따른 데이터 거버넌스
- 데이터 생성 및 갱신 노드의 명확화

### 데이터 중복 최소화

- ODOP(One Data One Placement) 구현
- 데이터 이동의 최소화

### • 서비스 향상을 위한 비즈니스 지원 가능

비즈니스 변화에 따라서 어플리케이션을 개발 및 변경을 하고자 할 때 기반이 되는 데이터의 신속한 파악과 유연한 대응이 가능함

### • 데이터 중복 비용 감소

데이터 중복은 스토리지 비용을 증가시킴  
중복 데이터 간의 데이터 이동을 위한 프로세스의 증가로 필요한 컴퓨팅 용량이 증가함

### • 성능 측정 능력 증진

명확한 데이터 위치 및 이동 방법은 정확한 성능 측정을 가능하게 함

### • 프로그램 개발 생산성 증진

효율적인 데이터의 분산 및 집중은 프로그램 개발을 용이하게 함

### • 데이터 품질 향상

데이터 중복 감소는 데이터의 정확성과 정합성을 향상시킴  
향상된 데이터 품질은 고객 서비스 및 경영 의사 결정 능력을 향상시킴

## 전략적 데이터 거버넌스 기대효과: 데이터 품질 측면 데이터 품질의 향상과 품질 이슈에 대한 체계적인 관리

### 원천 데이터의 품질 향상

- Source 에서 Target 간의 데이터 흐름 관리를 통해 상시 데이터 정합성 검증 기대
  - 원천 데이터 업무 규칙 도출 가능
  - 발견된 오류에 대해 원천 데이터를 시정조치 함으로써 데이터의 품질이 향상됨

### 데이터 품질 상시 모니터

- 대상 데이터의 주기적인 품질 모니터링을 수행함
- ETCL 작업 결과에 대한 Feedback 후 조치를 통해 ETCL 작업 결과에 대한 품질 관리 효과가 기대됨
  - ETCL 작업 결과 검증, SQL, ETL Script, XML 관리

### 사용자 측의 데이터 품질에 대한 확신

- 사용자 중심의 데이터에 대한 정합성 유지 관리 가능
  - 사용자 입장에서 품질에 문제가 있다고 판단되는 데이터에 대해 관리/개선을 요구할 수 있는 체계를 마련함

### 품질 이슈에 대한 체계적인 대응과 조치

- 품질 이슈 발생 시 정의된 조직과 프로세스에 따라 시스템을 통해 체계적으로 대응하고 조치할 수 있음
  - 조직에 새로운 품질 이슈가 제기되면 대상 식별부터 진단, 개선방안 수립, 사후관리 등의 일련의 프로세스를 통해 이슈에 대한 해결방안이 제시되고 실행

## 전략적 데이터 거버넌스 기대효과: 데이터 활용 측면 정보 자원의 가치 상승, 자본 및 운영 비용 감소, 신뢰성/업무 생산성/객관성 확보 가능

### 경영 성과 측면

데이터의 가치 제고와 비즈니스 생산성 개선으로 전략적 활용과 고객 만족을 고취하고 고객 자산 가치 증대와 비용 구조 감소로써 핵심 가치 공유

### 운영 효율 측면

통합 Single View, 실시간서비스, 연관성 공유로 업무 효율 및 데이터 활용성 향상

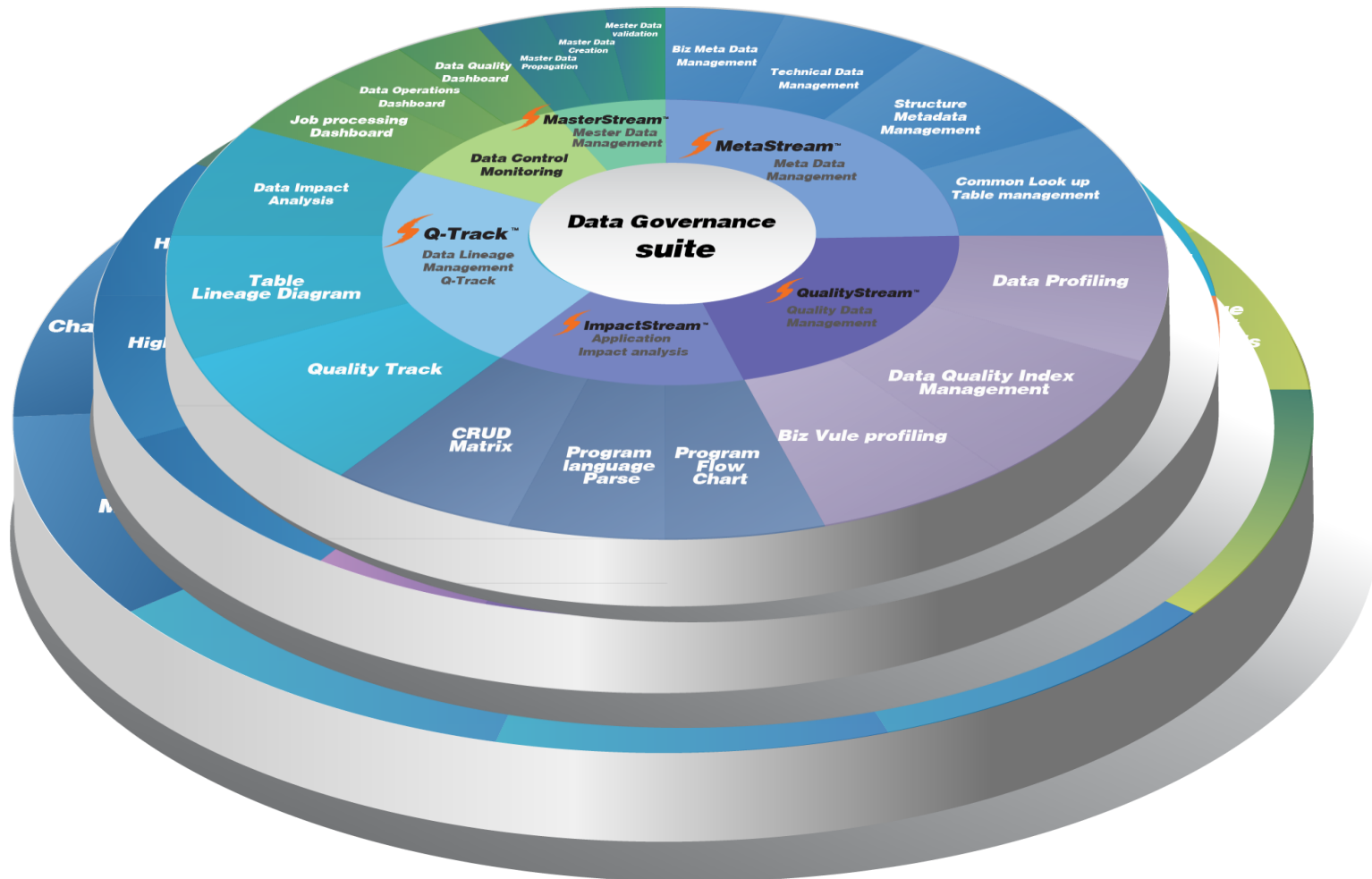
### 데이터 분석 측면

자동 매핑정의서 관리를 통한 데이터 흐름 및 추적, 데이터 간 연간 분석, 데이터 교차 분석, 데이터 분석 체계화 기반 마련

### 데이터 품질 측면

실시간/준 실시간 데이터 품질 관리 모니터링과 정제 기반 마련

## (주)데이터스트림즈 데이터 거버넌스 기반의 데이터 관리 Framework





# Q & A



DataStreams Corp. [www.datastreams.co.kr](http://www.datastreams.co.kr)

*Thank You!*



Leading Enterprise Data Management and Analytics Company :  
Solution Provider for Data Integration and Quality Management