

**IBM Information**  
On Demand

2009

>>> Comes To You

**IBM**

# 비즈니스 이노베이션을 위한 IBM 뉴 인텔리전스 인프라

## 스마트한 정보 통합 플랫폼



**INNOVATE.OPTIMIZE.**  
**PERFORM.**

Unlock the business value of your information.

윤석인 부장  
한국 IBM 소프트웨어그룹

# The Trusted Business Information Challenge

*Business Leaders Face a Growing Number of Challenges*

## Business Challenges

- Customer loyalty & churn
- Industry Consolidation
- Rapid Globalization
- Supply Chain

## Information Challenges

- Exploding volume & variety
- Data is highly fragmented

### *Business Impact*

Enterprises have 14 different customer databases on average<sup>2</sup>

The cost of a lost or incomplete customer record is \$182<sup>4</sup>

66% indicate profitability is negatively impacted by bad customer data<sup>1</sup>

30% of retail item information is incorrect, leading to 3.5% in lost sales<sup>3</sup>

One Telco Company reports 90% monthly churn

AMEX fined \$100M due untimely reporting

# Industry Leaders investing to Optimize Information Assets

*Trusted Information Enables Innovation*



## Insightful

Derive meaning from information changes



## In Context

Real-time delivery of relevant information when and where it's needed



## Complete

Related information reconciled into a single and holistic view



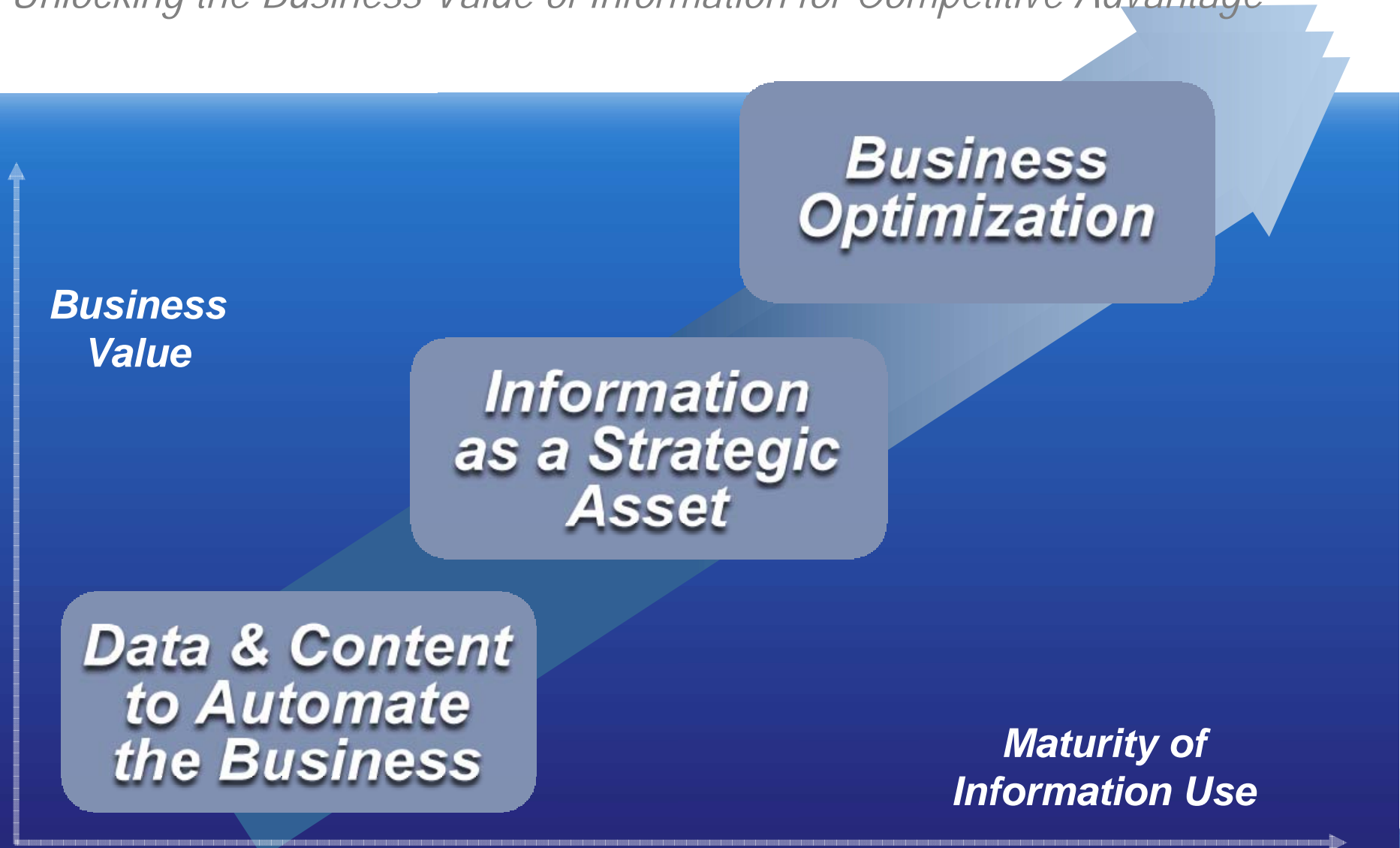
## Accurate

Complex and disparate data transformed, cleansed and delivered



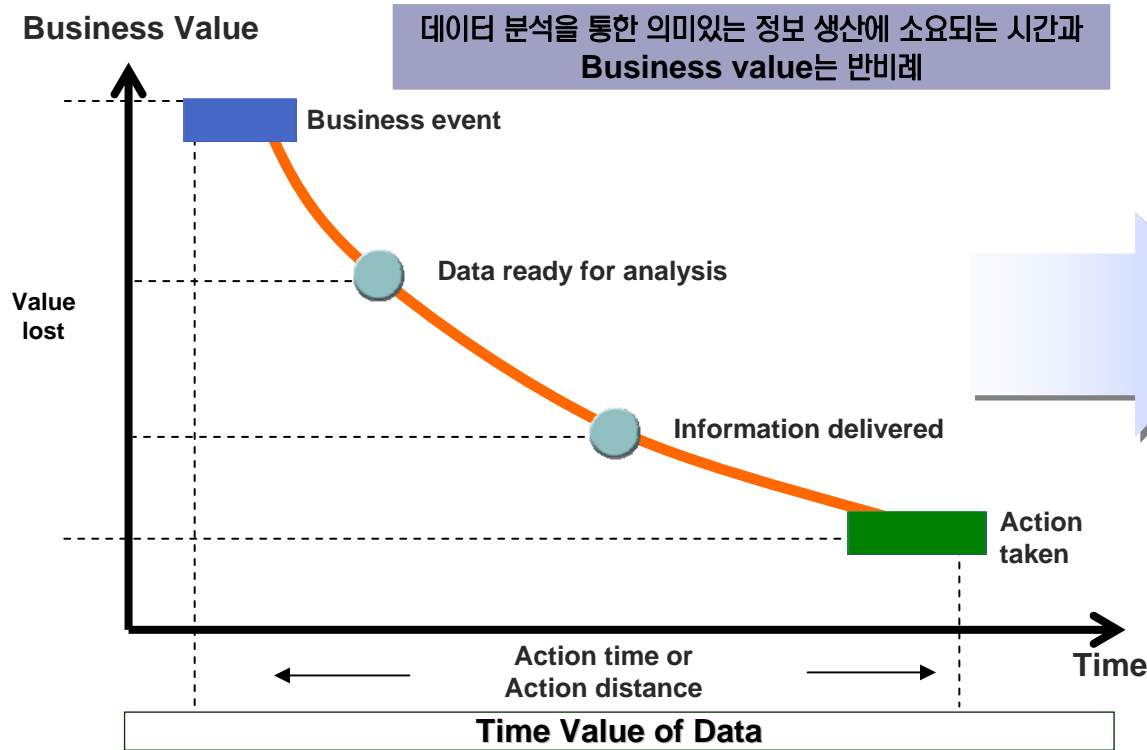
# Accelerate to the Next Level

*Unlocking the Business Value of Information for Competitive Advantage*

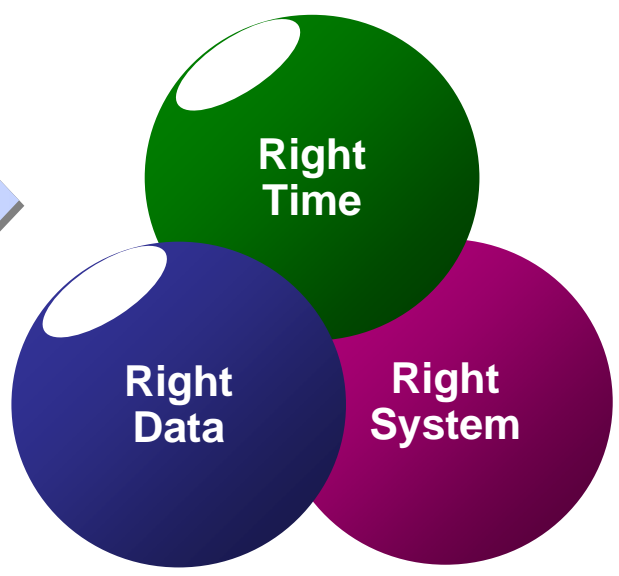


# Business Value 증대를 위한 데이터 통합 요건

급변하는 업무 환경과 고객 요구에 적절히 대응하기 위해 양질의 데이터를, 적절한 시점에, 필요한 시스템으로 전달 할 수 있어야 합니다.



이상적인 데이터 통합의 요건



Source: Richard D. Hackathorn, Bolder Technology, Inc.



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE. PERFORM



# InfoSphere Information Server



**Right Data**

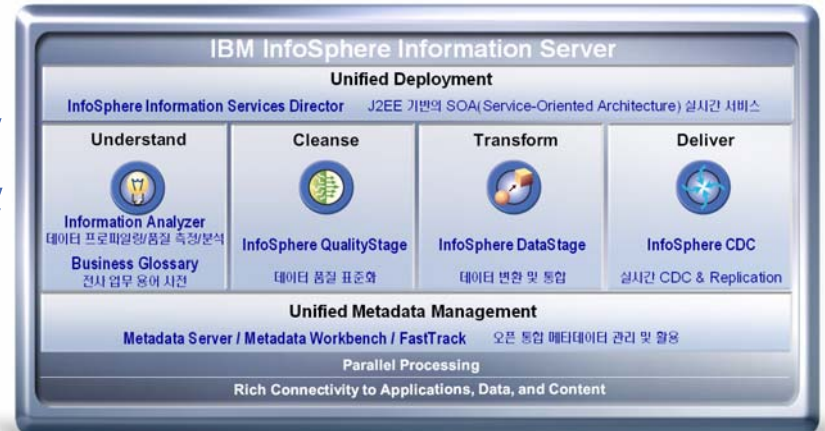
*이기종 시스템에 산재된 각종 정보를 이해하고, 정제 및 정비하여 통합하고 전달*

**Right System**

*통합 메타 데이터를 중심으로 구조화하여 정보 활용 극대화 및 비즈니스와 IT 연계*

**Right Time**

*대용량 데이터에 대한 배치 처리를 위한 무한 확장성 및 실시간 처리를 위한 CDC 기술*



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**

# 기업의 각 담당자가 느끼는 정보 통합 기술의 한계?



**Business Users**

비즈니스 용어와 개념을 시스템을 통해서 파악할 수 있는 방법은 없습니까?



**Subject Matter Experts**

우리가 사용하고 있는 정보의 소스를 문서화하여 쉽게 이해하고 활용할 수 있는 방법은 없는 겁니까?

제가 관리하고 있는 모든 데이터에 대해서 일어나는 모든 변화를 파악하고 보고할 수 있는 좀 더 쉬운 방법이 없을까요?



**Architects**

정보 통합 프로세스를 설계하는 좀 더 진보되고 가속화할 수 있는 편리한 방법은 더 기대할 수 없는 건가요?



**Data Analysts**



**Developers**

데이터베이스에 발생된 변경사항에 따른 영향도를 어떻게 파악할 수 있을까요?



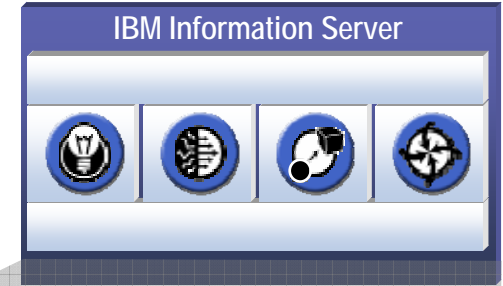
**DBAs**

도출된 개발 요건이 전통적인 배치처리 뿐만 아니라 실시간 처리 방식을 필요로 하는데, 어떤 솔루션을 적용해야 할까?



# Information Server : 전사적인 협업체계

## 통합된 메타데이터를 활용한 역할 기반 솔루션



**Business Users**



**Subject Matter Experts**



**Architects**



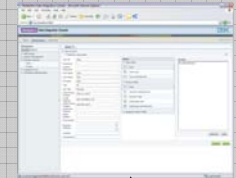
**Data Analysts**



**Developers**



**DBAs**



### Unified Metadata Management



**Design Operational**

- 통합 프로세스 단순화
- 정보에 대한 신뢰도 확보
- 변화 관리 및 재활용성 극대화
- 컴플라이언스 확보





# 데이터 품질 분석 : IBM InfoSphere Information Analyzer

- 애플리케이션, 데이터베이스, 파일기반 소스에 대한 데이터 및 프로세스 중심적 분석
- 컬럼, 컬럼 간, 테이블 간에 신뢰할 수 있는 상세한 프로파일링 정보 제공
- 프로파일링 결과로부터 메타데이터 생성
- 프로파일링 결과를 IBM Information Server 전체에 걸쳐 활용되도록 공유 가능
- 30여 가지의 다양한 Report Template 제공



Subject Matter Experts



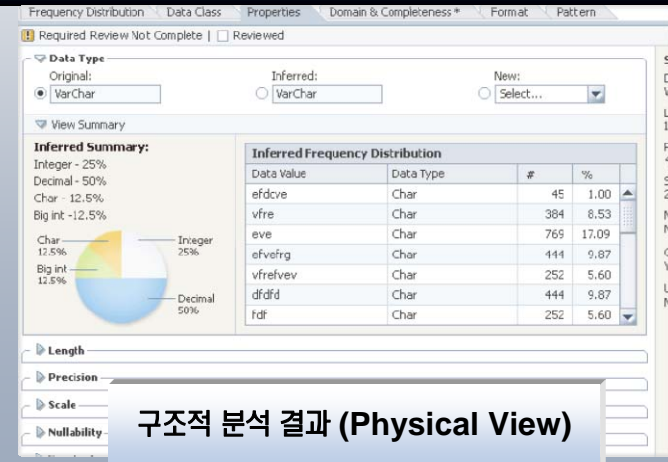
Data Analysts

Understand



IBM InfoSphere™ Information Analyzer®

소스 데이터의 구조를 분석하고, 데이터 통합과 품질 관리 규칙에 부합하는지 여부를 모니터링 함



구조적 분석 결과 (Physical View)



# 데이터 품질 분석 : 직관적인 데이터 프로파일링

**Project Domain Analysis Status**

- 45% - Reviewed/Analyzed
- 20% - Analyzed
- 35% - Not Yet Done

**Cross-Domain Candidate Pair**

	Base Column	Paired Column
Column	SS_NUM	SS_NUM
Table	BANK_ACCOUNTS	BANK_SAVINGS
Sources	JKLW_DB/JK_BANK1	JKLW_DB/JK_BANK2
PrimaryKey(D PK/S PK)	No	No
Foreign Key(D FK/S FK)	No	No
Data Class	Text	Quantity
Data Type	STRING	STRING
Length	12	12
Precision	12.00	12.00
Scale	0	0
Cardinality	199	998
Unique	No	No
Constant	No	No
Definition	No	No
Redundant Status	No	No

**Paired to Base:**

#:	4	%:	0.40 %	Common Domain:	Yes
----	---	----	--------	----------------	-----

**Base to Paired:**

#:	4	%:	2.01 %	Common Domain:	Yes
----	---	----	--------	----------------	-----

**Common Domain #:**

Legend: Base Column (Green), Paired Column (Blue)



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**

# 비즈니스 메타데이터 : IBM InfoSphere Business Glossary

- 웹 기반의 비즈니스 메타데이터에 대한 작성, 관리 및 공유
- IT 자상에 비즈니스 용어 제공
- 역할과 책임에 따른 관리

Database = DB2

Schema = NAACCT

Table = DLYTRANS

Column = ACCT\_NO

data type = char(11)



Technical



Business

GL Account Number

The ten digit account number. Sometimes referred to as the account ID. This value is of the form L-FIIIIVVVV.



Subject Matter Experts



Business Users

Understand



IBM InfoSphere™ Business Glossary®

비즈니스 어휘를 생성 및 관리함과 동시에 이를 물리적 소스와 연계함

eporting Help | Log U

Overview Manage Custom Attributes x Manage Business Terms x

Overview

Welcome to WebSphere Business Glossary

International Business Machines (NYSE:IBM) is the leader in enterprise data integration. Customers and partners worldwide use its information integration tools to confidently transform data into accurate, reliable and complete business information to improve operational performance and decision-making across every critical business dimension. Our comprehensive end-to-end solutions provide on demand data integration complemented by our professional services, industry expertise, and methodologies.

USA Finance Manufacturing Customer Relationship

Europe Finance Manufacturing Customer Relationship

Data Model Top-level category, which groups all the elements by type

Asia Finance Manufacturing Customer Relationship

See All Categories

**비즈니스 뷰 (Business View)**



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**

# 비즈니스 메타데이터 : 웹 기반의 업무 지원 솔루션

**JK LIFE & WEALTH**

**Total Account Balance by Gender & Marital Status**

**Expected Loss per Score Range / Line of Business**

**Credit Cards**

26.422	24.2972973
26.21612903	24.45769231
27.736	22.12
24.80606061	22.79387755
24.59074074	25.384375
27.01142857	22.22
26.83333333	25.175
25.62	23.51304348
26.81636364	23.85882353
23.5483871	27.42075472
24.56041667	24.83783784
28.66129032	23.82264151
24.666	22.08571429
25.60404048	24.28979592
26.42727273	24.446875
27.9325	25.31956522

**Home Equity**

**Profession Group**

- Customer Profession
- Craftsman
- Factory worker
- Farmer
- Mechanic
- student/retired/unempl
- Accountant
- Executive
- Sales Rep

**IBM Information Server Business Glossary Anywhere**

Line of Business

JK Wealth »

**Line of Business**

**Short Description:** JK Life & Wealth division

**Long Description:** BANK1 and insurance are called divisions of JK Life & Wealth. A division is responsible for the creation of the organization's products and/or services. It is not IT, HR and Accounting

**Steward:** Data Analyst (Lt. Commander) Data Soong

**Defined by:** Undefined

**Status:** Accepted

**General Information**

**Synonyms**

- Business Group

**Abbreviation**

- LOB

**Example**

- Banking, Insurance

**Related IT Assets**

- JK\_BANK1
- IBM-09E579C16AA
- JKLW\_DB

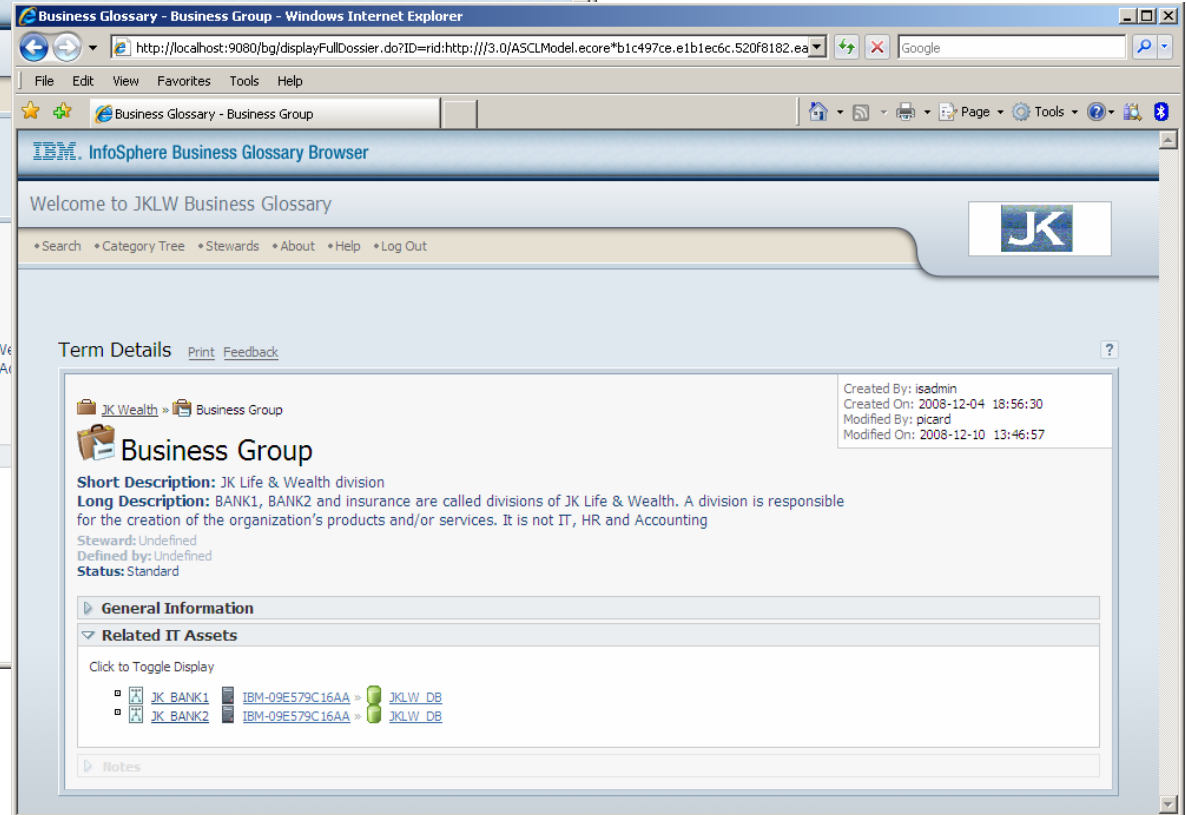
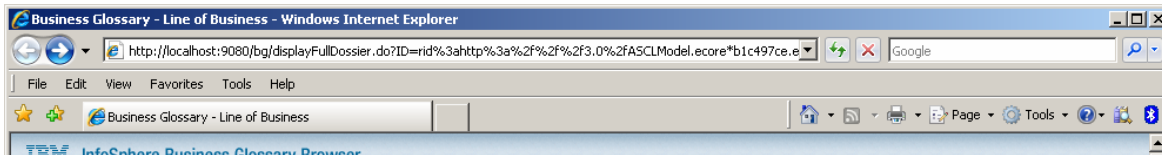
Nov 13, 2008 1 9:19:34 PM

- 폭넓은 협업체계를 지원하는 웹 기반의 업무지원 솔루션
- 업무 담당자가 업무 용어에 대한 정의, 동의어, 사용예, 분류체계를 정의하고 협업 가능
- 보고서 기능, 탐색 기능, 영향도 분석 기능, 주석 기능 등 풍부한 메타데이터 관리 및 표현 기능 제공
- BG Anywhere를 통해 SAP, COGNOS, MSTR, BO 보고서에서 비즈니스 용어 참조





# 비즈니스 메타데이터 : 메타데이터를 통한 협업 체계 구축



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**

# 데이터 표준화 : IBM InfoSphere QualityStage

- DataStage와 통합된 데이터 품질 정제 기능
- 표준화 및 복잡한 매칭과 최적화 로직 (survivorship)을 정의하는 GUI 기반 제공
- 정제 및 표준화되고, 중복이 제거된 정보 확보
- 신뢰할 수 있는 단일 버전 정보 제공
- 국내외 주소 및 개인/법인명 표준화 규칙 제공



Subject Matter Experts



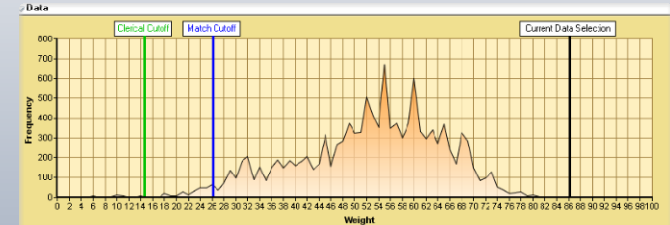
Data Analysts

Cleanse



IBM InfoSphere™ QualityStage®

소스 데이터 필드를 표준화 및 수정하고, 단일 뷰 생성을 위해 소스 간에 레코드를 매칭함



SetID	RecordType	PassNumber	Weight	GenderCode	FirstName	MiddleName
4111	XA	1	85.57	IM	OTIS	GARLAND
4111	DA	1	85.57	IM	OTIS	GARLAND
5311	XA	1	84.08	IM	NICHOLAS	T
5311	DA	1	84.08	IM	NICHOLAS	T
3420	XA	1	81.88	IM	CLAUD	LADALE
3420	DA	1	81.88	IM	CLAUD	LADALE
3420	XA	1	81.88	IM	CLAUD	LADALE
3420	DA	1	81.88	IM	CLAUD	LADALE
4025	XA	1	80.20	IM	CLIFF	H
4025	DA	1	80.20	IM	CLIFF	H
7308					CHIN	CHIN
7308					CHIN	CHIN
2262					GILLIS	GILLIS
2262					GILLIS	GILLIS
6979					NICHELLE	NICHELLE
6979					NICHELLE	NICHELLE

비주얼한 매칭을 디자인  
(Visual Match Rule Design)



# 데이터 표준화 : 텍스트 데이터 정제 및 표준화

데이터 리엔지니어링은 기업 내 데이터 활용을 극대화하기 위해 현행 데이터의 분석을 통해 데이터를 근본적으로 재설계하고 표준화



# 데이터 변환 및 전달 : IBM InfoSphere DataStage

- Top-Down 방식 설계 인터페이스와 다양한 변환 기능이 내장되어 있어 데이터 흐름을 코딩 없이 설계
- 객체의 재사용성을 높여 개발 생산성 및 운영 효율성 증가
- 설계의 변화 없이 병렬 처리를 활용
- IBM InfoSphere DataStage는 다양한 요건의 데이터 통합 프로세스에 적용 가능
  - 대용량 배치성 데이터 이동
  - 이벤트형 적시성(Right Time) 응답
  - 서비스 지향 아키텍처(SOA)



Developers



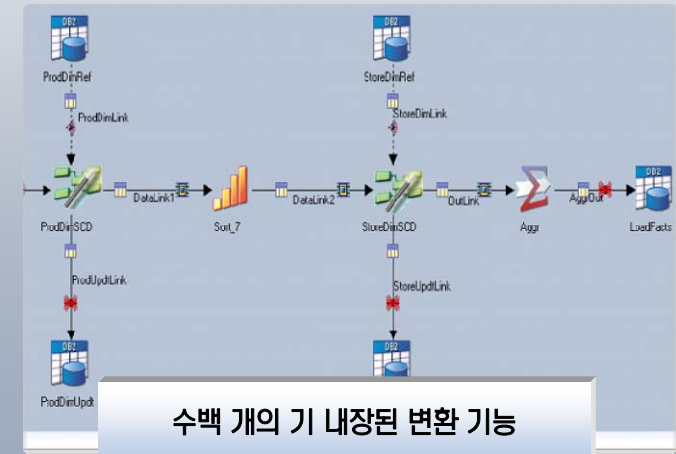
Architects

Transform



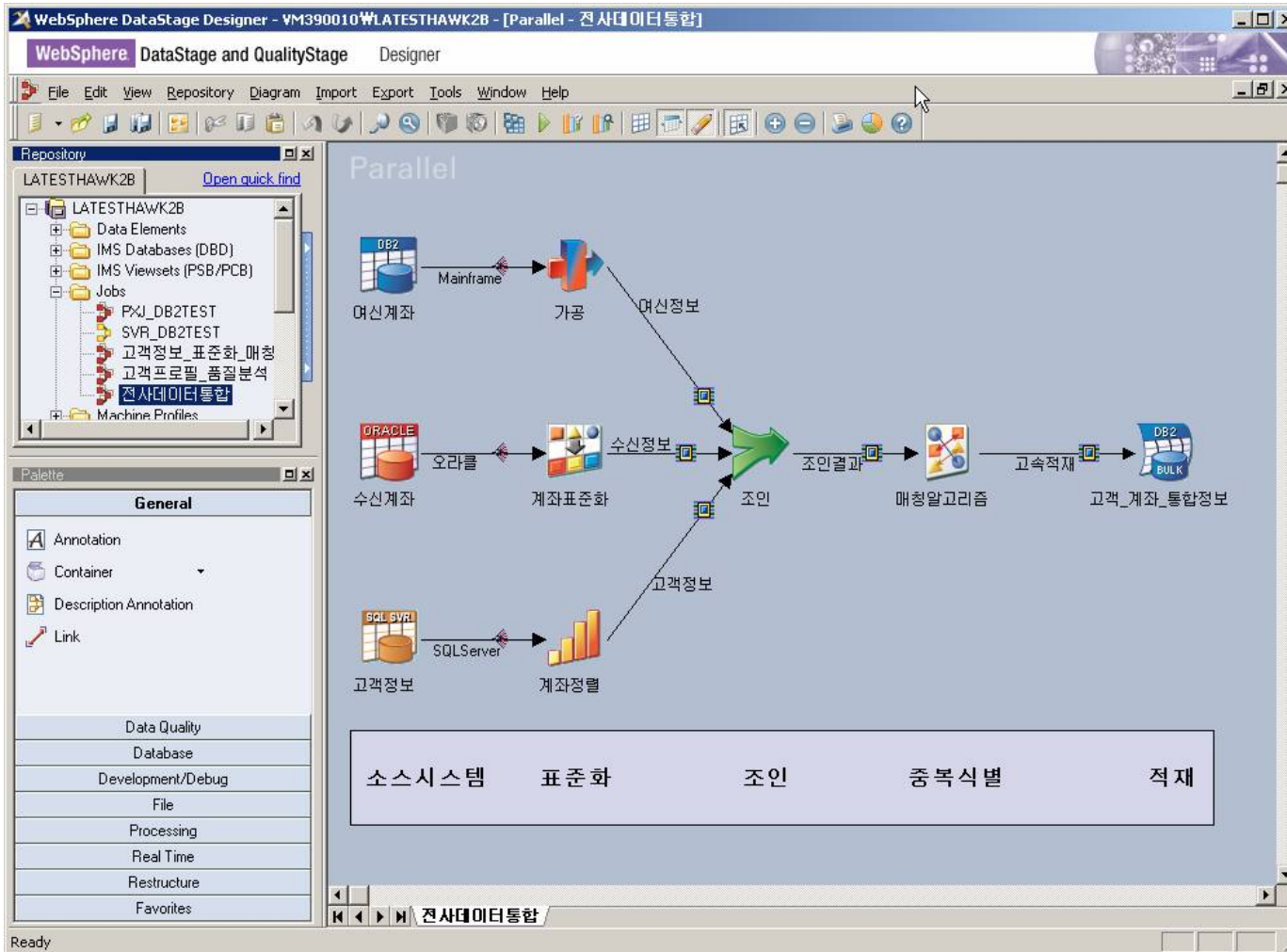
IBM InfoSphere™ DataStage®

어떠한 양의 정보도 도식화된 논리를 통해 배치 혹은 실시간 형태로 변환하고 aggregate 가능





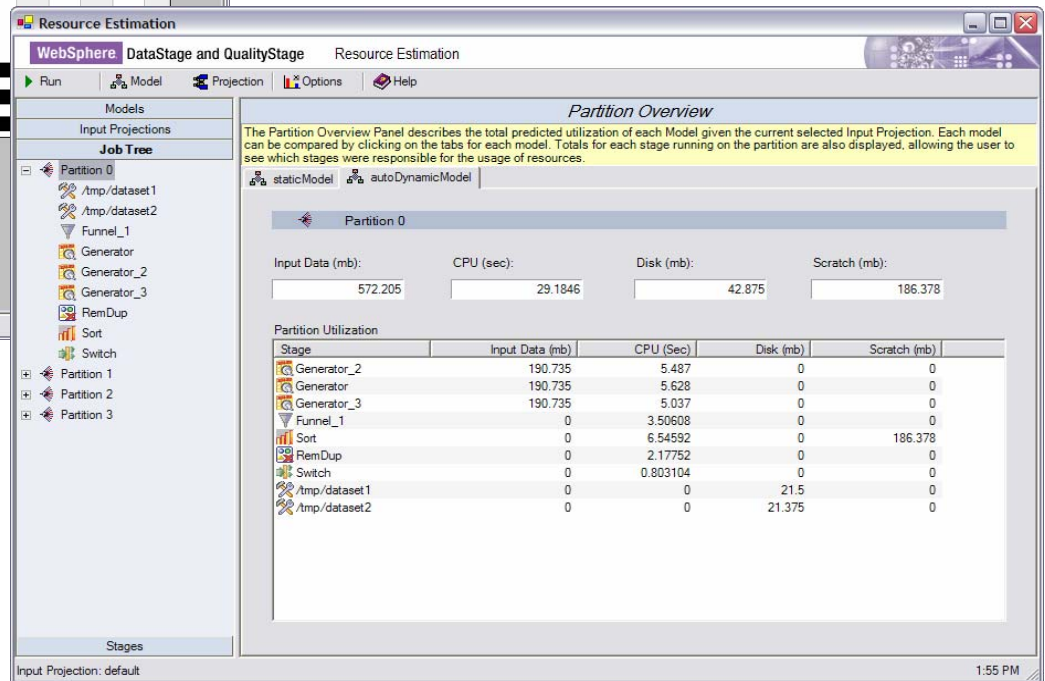
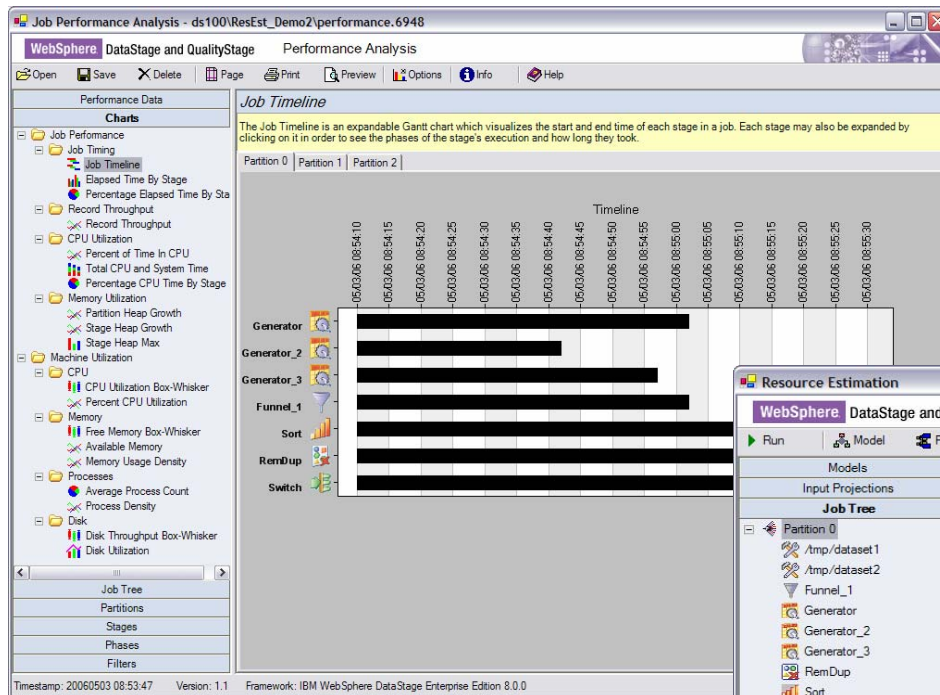
# 데이터 변환 및 전달 : 전사 통합 프로세스 설계



- 정보 통합 프로세스 설계
  - GUI 환경, 유연성
  - Top down Approach
  - 변환, 정제 및 통합
  - 다중 소스의 데이터 통합
  
- 통합 프로세스 적용
  - 실행 및 모니터링
  - 프로젝트 관리 및 제어
  - 생성된 객체의 재사용성
  - 설계된 Job의 테스트, 디버깅 및 적용
  
- 전사 통합 프로세스 관리
  - 메타데이터 참조, 편집 및 보고서 작성
  - 개발된 컴포넌트 공유
  - 감사 결과 레포트
  - 공유 메타데이터 활용



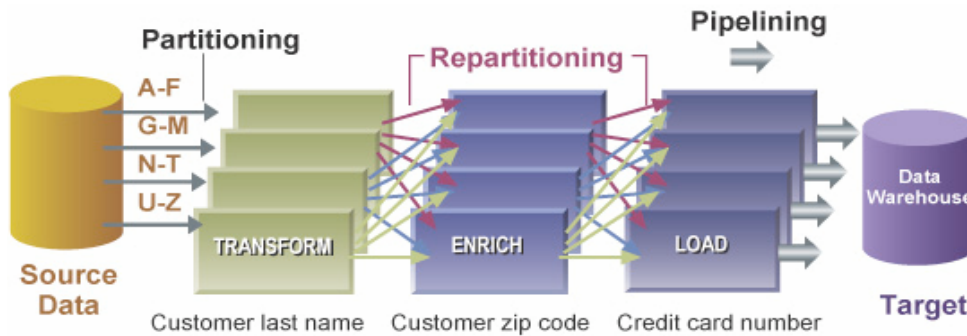
# 데이터 변환 및 전달 : 시스템 자원 사용률 및 성능 모니터링



# 데이터 변환 및 전달 : 모든 상용 병렬 서버 아키텍처 지원

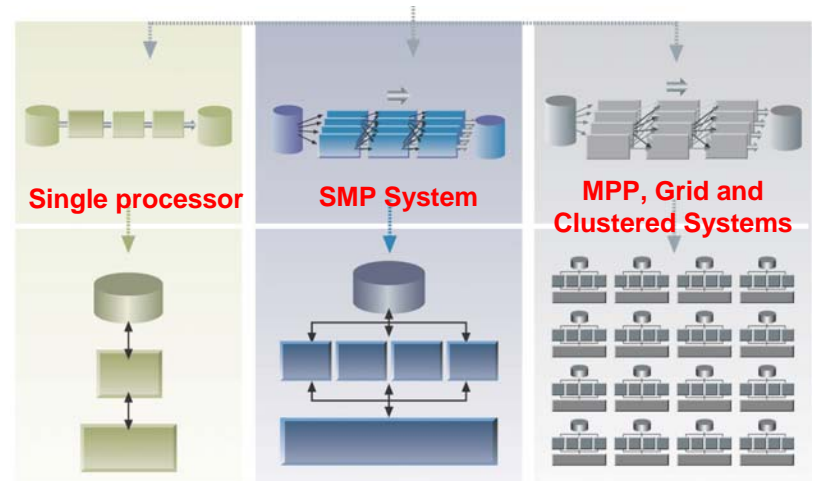


*업무로직에 충실하게 개발*



*실행은 병렬처리*

*기존 운영 플랫폼에서  
운영 중인  
단일 로직으로 설계된 Job은  
설계 변경 없이  
GRID 컴퓨팅 환경의  
처리능력을  
최대한 활용하여 운영*

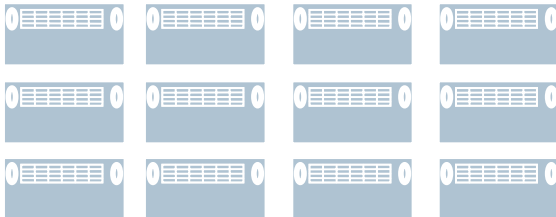


# 데이터 변환 및 전달 : Grid 환경 지원

*Simplify deployment, management & expansion*



- **Grid-기반의 전사 정보 통합 솔루션으로서 하드웨어, 소프트웨어가 통합된 구성**
  - 신속하고 쉽게 구성 : ROI 보장
- **확장성이 용이한 Grid 컴퓨팅 구현**
  - 전사적인 데이터를 단일한 환경으로 통합하기 위한 솔루션 적용을 단순화하여 빠르게 적용
  - 전통적인 SMP 형태로는 극복할 수 없는 대용량 정보 통합 요건을 해결할 수 있는 가장 비용 효과적인 솔루션
  - 에너지 효율성





# CDC : IBM InfoSphere CDC

- DB 로그에서 직접 변경데이터 추출
- 로그 기반의 데이터 복제 기술을 적용하여 운영시스템의 성능에 대한 영향을 최소화
- 변경 데이터 추출을 위해서 기존 운영 중인 어플리케이션에 로직을 변경할 필요 없음
- IBM Information Server 내에서 DataStage와 완벽하게 연동



Developers



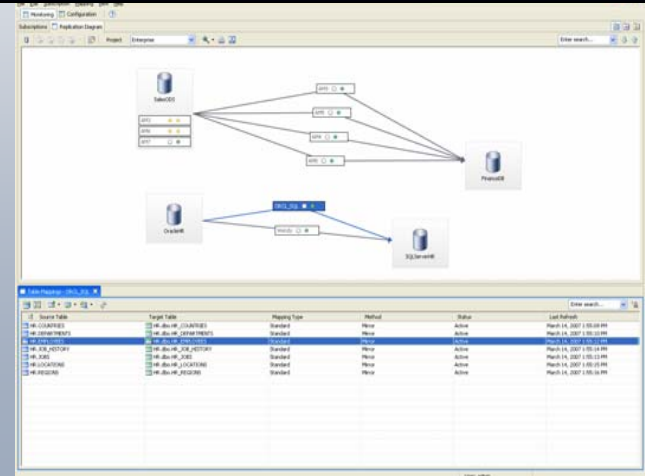
DBAs

Deliver

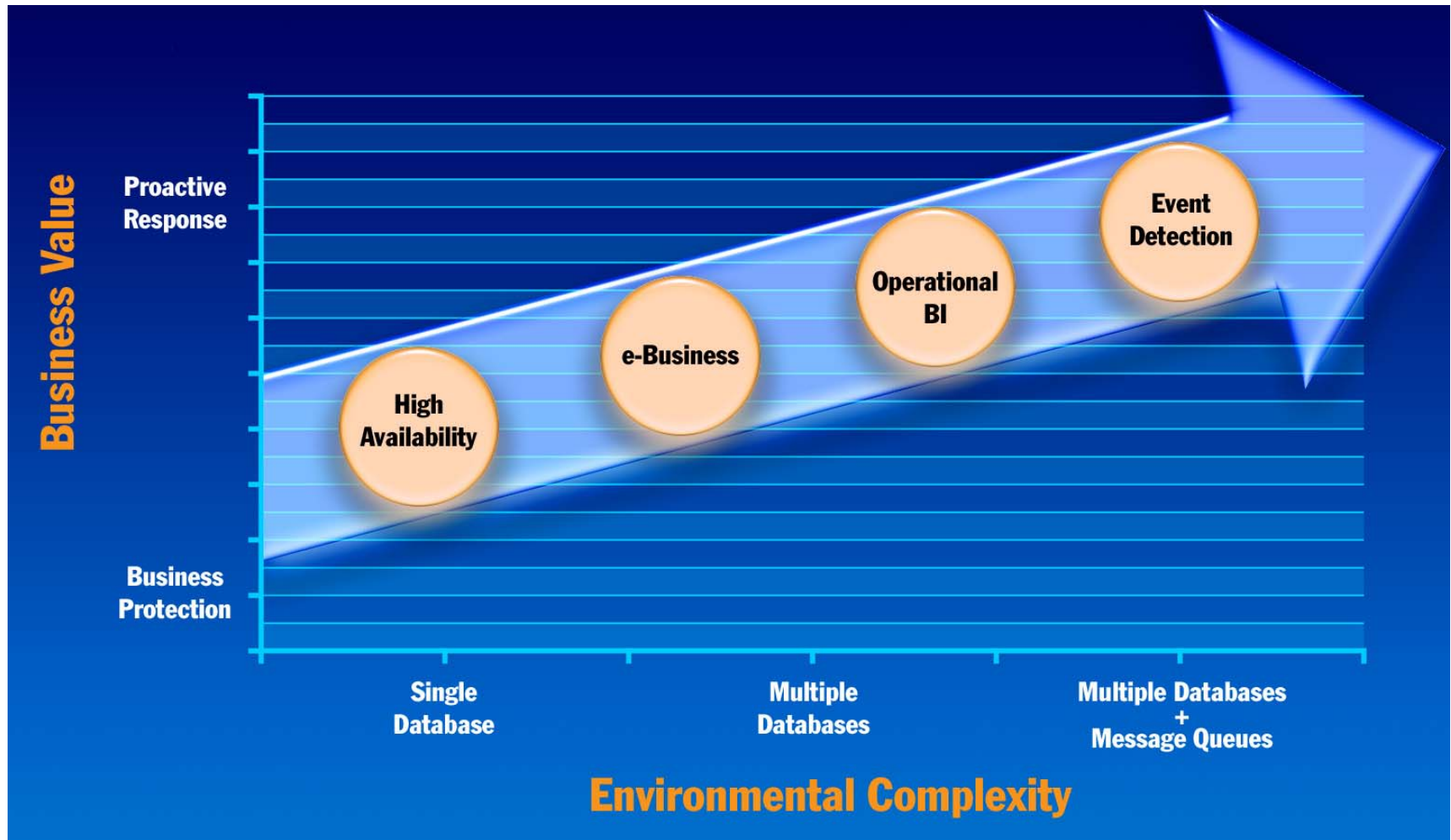


## IBM InfoSphere™ CDC

원천 시스템에서 발생하는 트랜잭션 데이터를 다양한 형태의 타겟 시스템으로 실시간으로 변환 및 전달  
실시간 액세스를 통해 애플리케이션의 가치 확대



# CDC : 비즈니스 가치를 높이는 실시간 데이터 통합



# CDC : 실시간 데이터 통합 솔루션 활용

*Dynamic Warehousing &  
Business Intelligence  
and Reporting*



실시간 성격의 데이터를 통한 분석 및 리포팅

*생산 시스템과 e-Business  
어플리케이션간의 데이터 통합*



e-Business를 위한 실시간 데이터 제공

*Real-time  
Event Detection*



특정한 이벤트에 대한 실시간 모니터링

*HA 구성 및 장애 복구*



데이터 복제를 통한 효율적인 HA 구현



# 메타데이터 서비스 : IBM InfoSphere Metadata Workbench

- Information Server 기반의 메타데이터를 조회하고 활용할 수 있는 Web-기반의 솔루션
- 영향도 분석 및 상호 의존성을 쉽게 분석할 수 있는 기반 솔루션
- BI 보고서에 대한 데이터의 생성 경로를 추적할 수 있는 솔루션



Data Integration Managers



Developers

Understand



IBM InfoSphere™ Metadata Workbench®

IT 담당자에게 Information Server 플랫폼에서 생성되고 활용되는 정보 자산에 대한 탐색과 이해를 제공

Help | About | Lo

▼ Report Finder - Data Movement

- Choose the information to view: View where data comes from
- Choose a type of information asset to report on: BI Report Attribute
- Choose a specific information asset to report on: Customer Name Attribute
- Choose the type of report output: Graph
- Create Report

Data Movement for: Customer Name Attribute

```

    graph LR
      JFV[Job Flow Variable  
CUST_NAME] -- "Is Source Of Flow Variable -  
Created From Data Field -  
Created From Data Field" --> DF[Data Field  
DWCUSTNAME]
      DF -- "Report Attribute" --> RA[BI Report Attribute  
Customer Name Attribute]
    
```



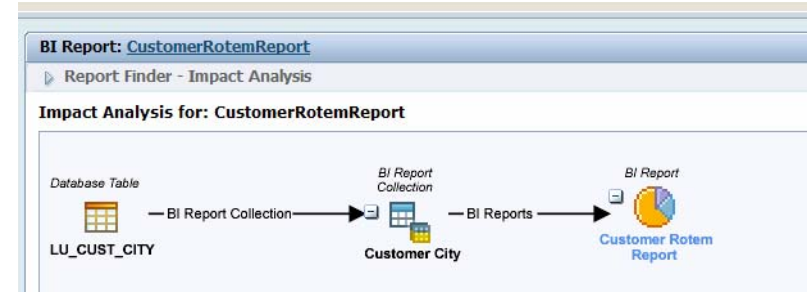
IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**



# 메타데이터 서비스 – Impact Analysis & Data Lineage

영향도 분석(Impact Analysis) 및 데이터 계보 분석(Data Lineage)을 통해서 보고서의 근간이 되는 데이터의 변화에 따른 보고서 결과의 변화 및 영향도를 사전에 파악함으로써 의사 결정 과정의 리스크를 줄일 수 있는 솔루션



BI 정형 보고서  
OLAP 물리 스키마 및 논리 스키마 개체



소스 데이터베이스

DataStage Job

데이터 웨어하우스 / 데이터 마트

OLAP 개체

OLAP 보고서



IBM InfoSphere Information Server



**COGNOS**  
AN IBM COMPANY



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**

# 정보 통합 최적화 : InfoSphere DataStage Balanced Optimization

- 데이터 흐름 상의 변환 로직을 최적화된 로직으로 자동 변환
- DBMS 내부에서 데이터 통합을 수행하도록 설계 최적화를 통해서 DBMS 하드웨어 투자를 적극 활용
- 개발자가 Job 자체 또는 Job의 일부분이 실행될 위치를 통제함으로써 실행 시간을 최적화



Developers

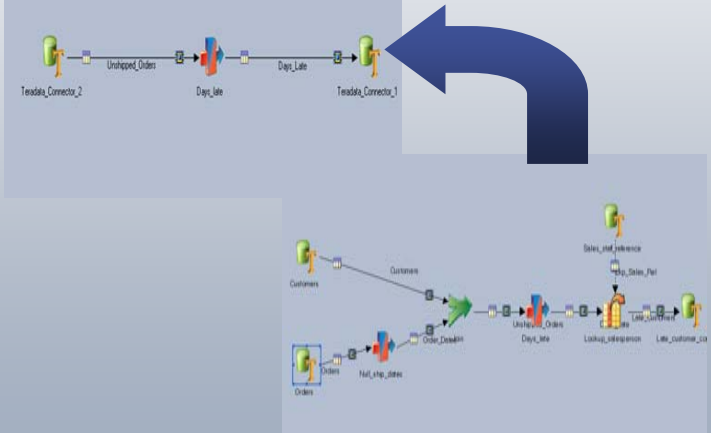


Architects

Transform



GUI기반 환경에서 정보 통합 과정을 설계하여 대량의 데이터를 배치 및 실시간으로 변환하고 통합, 요약

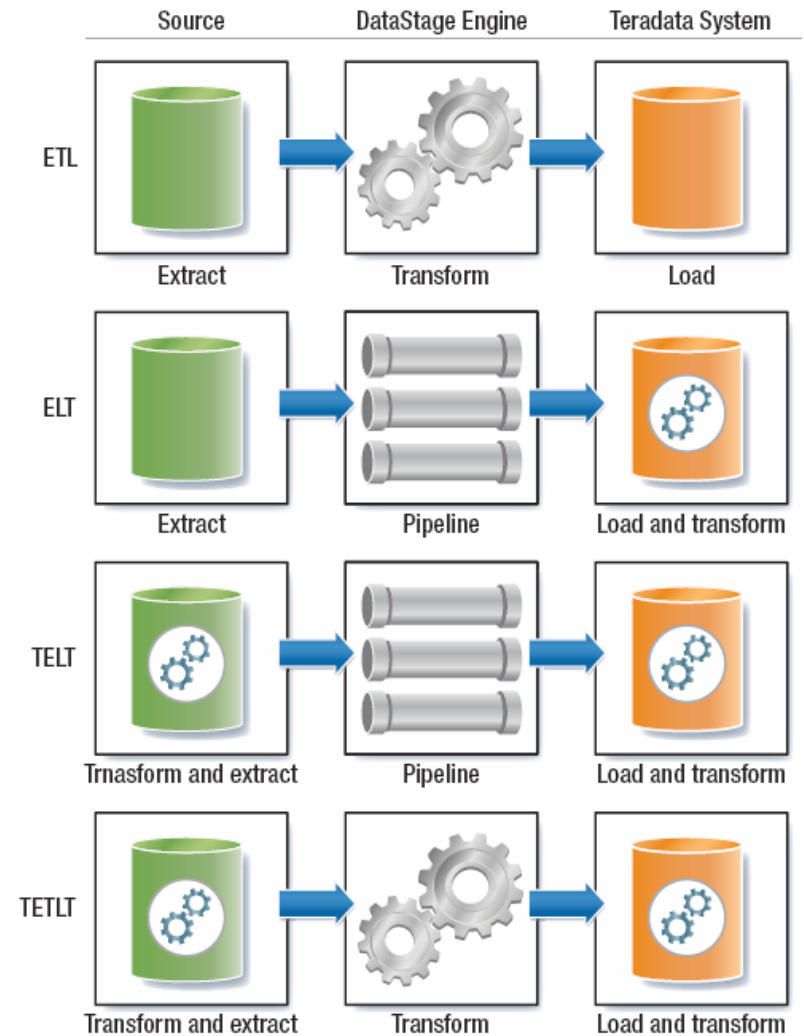


DBMS 하드웨어를 지능적으로 활용함으로써 실행 시간을 최적화



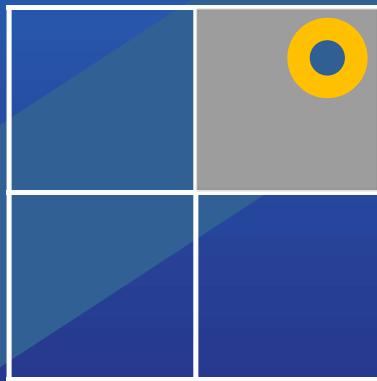
# 정보 통합 최적화 : 통합 프로세스 최적화 및 투자 효과 극대화

- 정보 통합 프로세스를 단순화 및 최적화하여 성능 개선
- 기술 재활용성을 확대하여 신뢰도 높은 정보를 일관성 있게 제공
- 기존 방식과 동일하게 설계하고 **DBMS** 자원을 활용하여 실행하는 유연성 제공
- 병렬 처리 **DBMS**에 대한 투자 효과 극대화



# Leading the way

## Industry Analysts Rank IBM as Leaders....



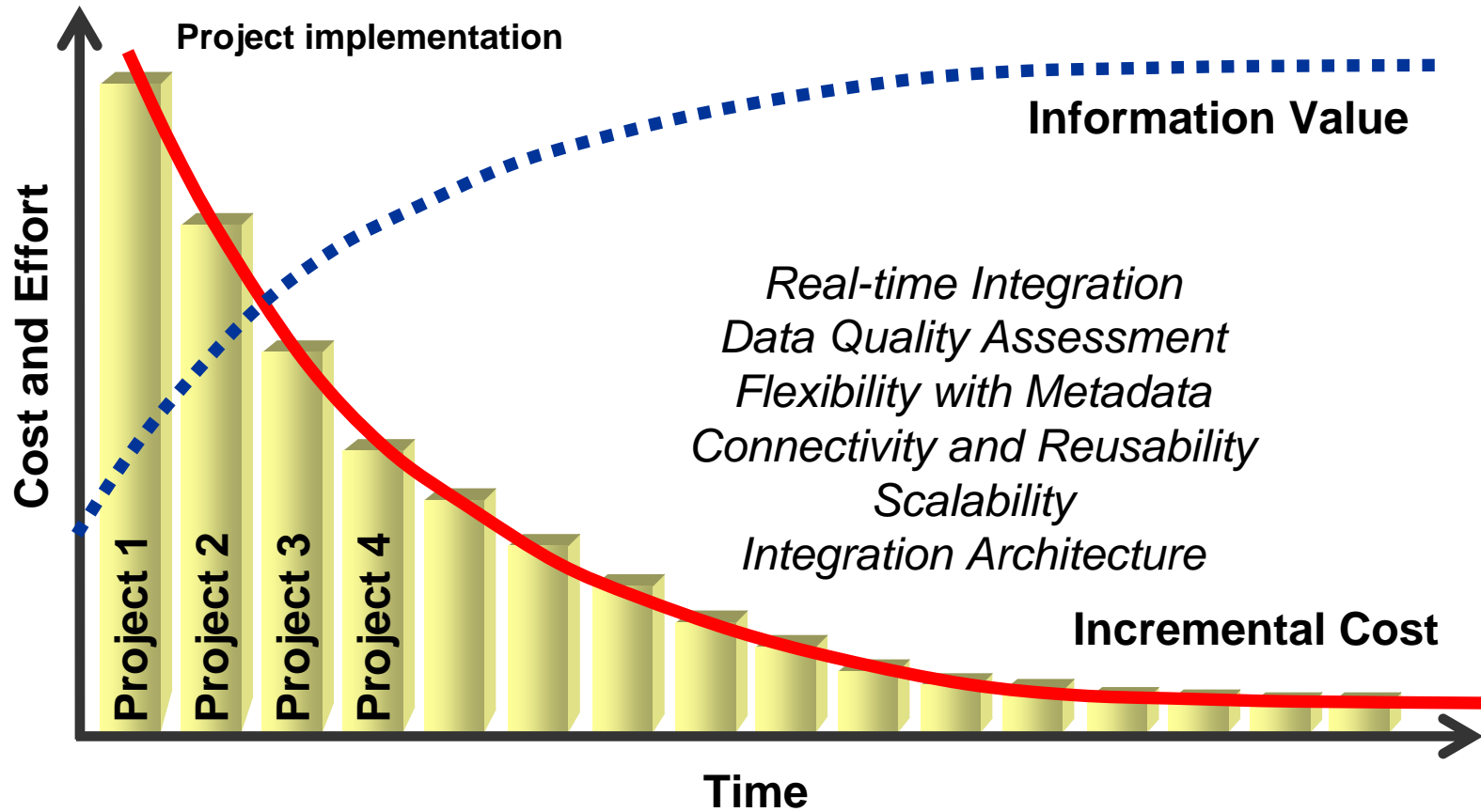
*and many other firms...*

- *Data Integration Tools*
- *ETL Tools*
- *Data Quality Tools*
- *Business Intelligence Platforms*
- *Data Warehousing*
- *Customer Data Integration*

**InfoSphere™**  
software



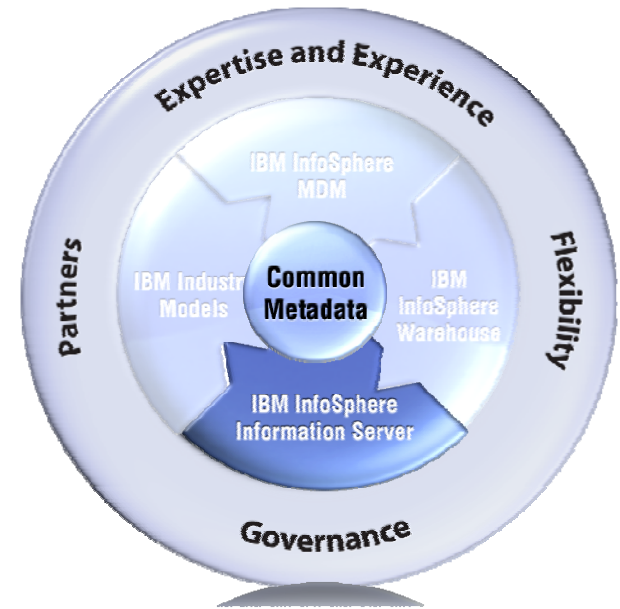
# Break-through Productivity with Information Server



# IBM InfoSphere™ Information Server

**Leading The Market In Innovation**

- CDC를 통한 실시간 정보 통합
- 자동화되고 반복적인 데이터 품질 관리 인프라
- 통합된 메타데이터에 기반한 유연성
- 일관된 연결성 및 재활용성
- 병렬처리를 통한 성능 및 확장성 보장
- 포괄적이고 통합적인 전사 정보 아키텍처 확보



IBM INFORMATION ON DEMAND COMES TO YOU

INNOVATE. OPTIMIZE.  
**PERFORM**