

MDM으로 기업의 데이터 한 눈에 보기

IBM SWG, 송정숙, 차장 (jussong@kr.ibm.com)



InfoSphere Addresses Big Data Governance Concerns

Integrate & Link Big Data

- Big Data as a Source
- Big Data as a Target
- Data Transformations
- Data Movement
- Integrate w/existing Enterprise
- Data Lineage & Impact Analysis
- Metadata Integration w/Analytics
- Realtime & Data Federation



Cleanse and Validate Big Data

- Accuracy and Entity Matching with Social Data
- De-duplication and Standardization of Machine Data
- In-line Cleansing with Integration
- Trusted Data Dashboard and Reporting on Data Quality



Protect Big Data

- Activity Monitoring
- Data Masking
- Data Encryption
- On-Demand / In-Place Protection
- In-Line Protection (w/ETL etc.)
- Active Detection & Alerting



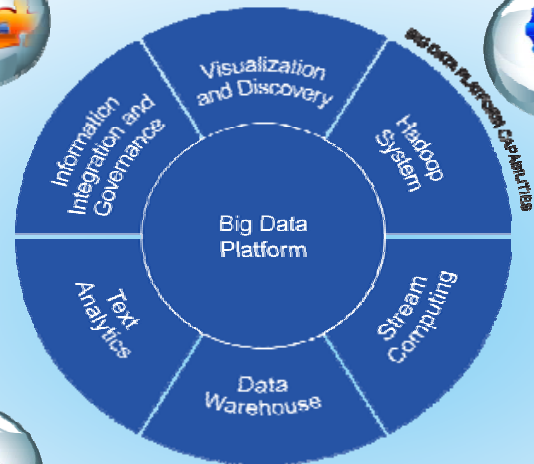
Master Big Data

- Big Data as a Supplier
- Big Data as a Consumer
- Links between Big Data and Trusted Golden Records
- Leverage Master Data in Big Data Analytics
- Entity Resolution at Extreme Scale Out Levels
- Probabilistic Entity Matching



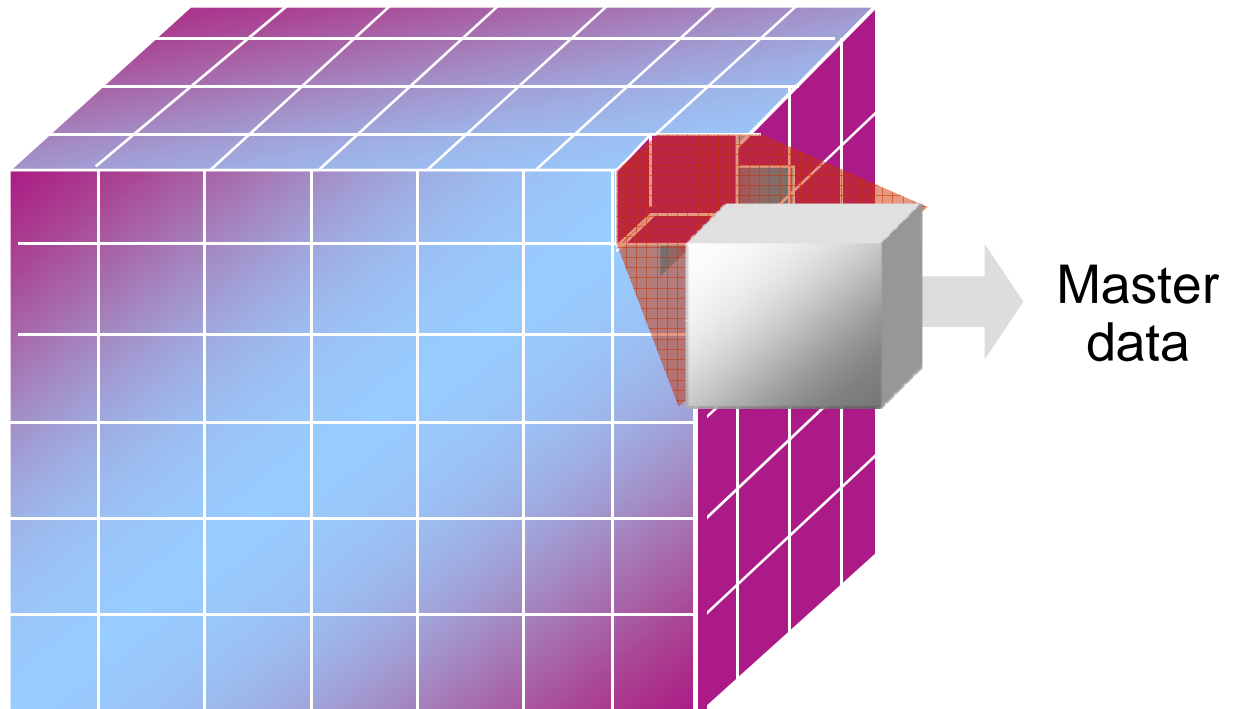
Audit & Archive Big Data

- Queryable Archive
- Structured and Semi-Structured
- Optimized Connectors to existing Apps
- Hot-Restorable On-the-Fly
- Immutable and Secure Access
- Automated Legal Hold Capability for Data Freeze



마스터데이터란?

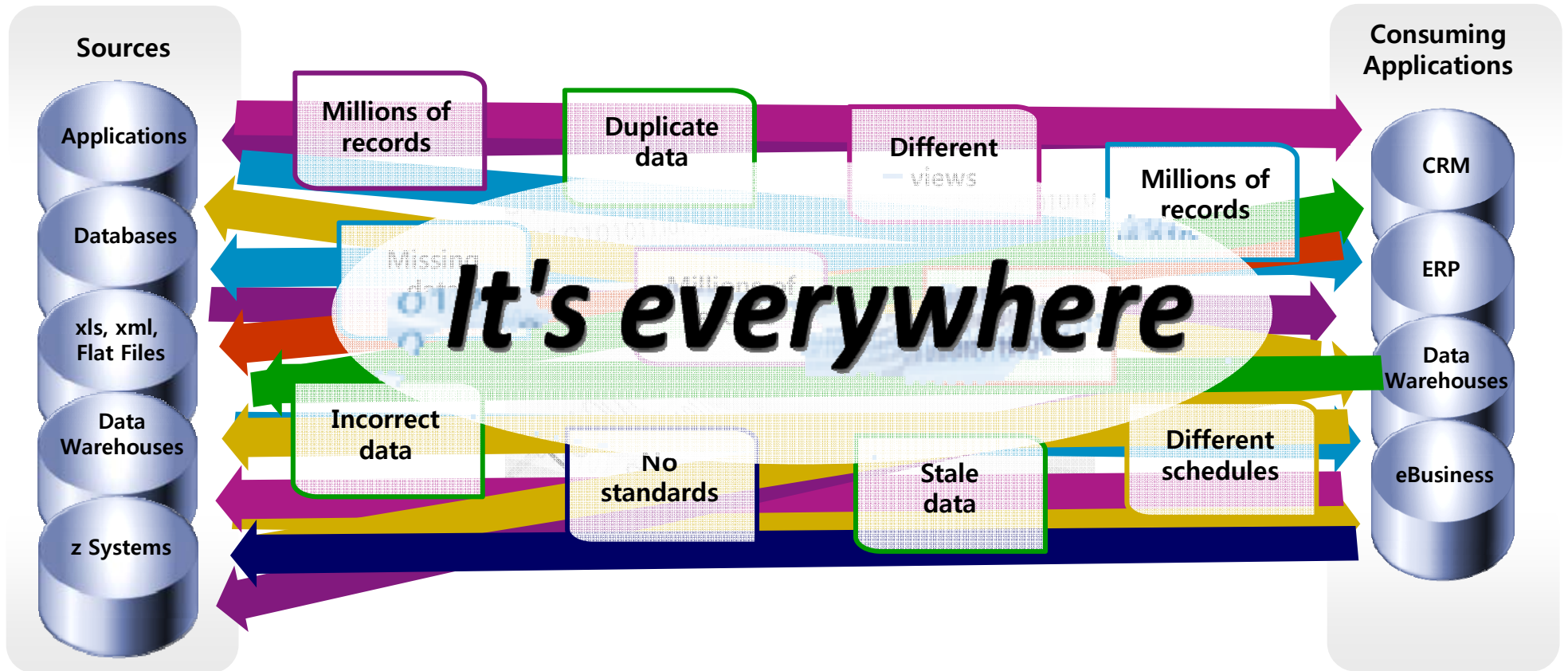
- ◆ **Master data** is the high-value, core information used to support critical business processes across the enterprise
- ◆ **Master data** is a subset of all enterprise data
- ◆ **Master Data** is information about customers, suppliers, partners, products, materials, employees, accounts and more
- ◆ **Master Data** is at the heart of every business transaction, application and decision



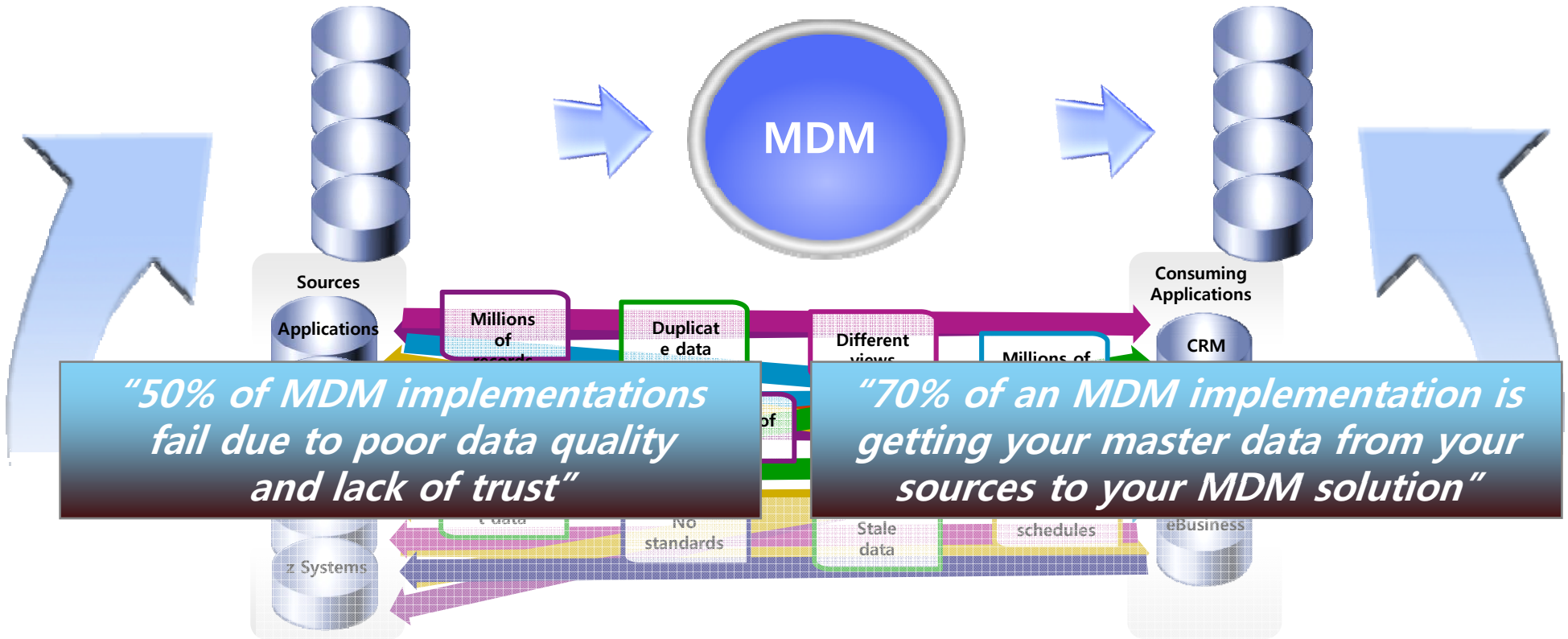
마스터데이터관리의 현실 #1

What systems contain master data?

Who Consumes Master Data?

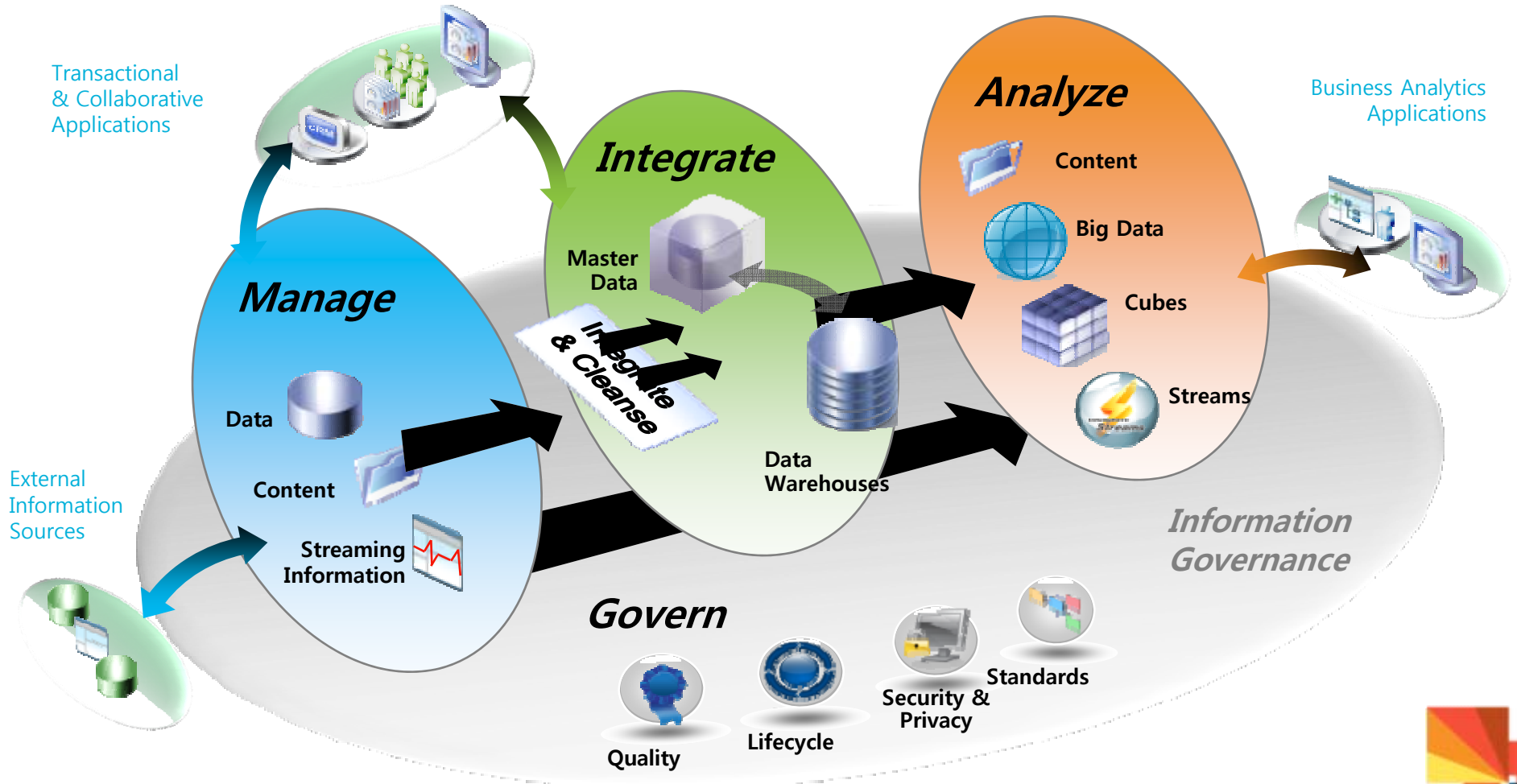


마스터데이터관리의 현실 #2



IBM이 말하는 마스터데이터관리란?

Trusted ♦ Relevant ♦ Governed



IBM InfoSphere Master Data Management 솔루션

InfoSphere Master Data Management

Creates trusted views of your master data for improving your applications and business processes.



The Most Comprehensive

- All domains
- All styles
- All use cases
- All industries
- Fully Pre-Built or Build it Yourself

Best Solution for Real Time Processes

- Optimized for real time
- Business Services
- BPM Express
- MDM Powered Applications

The Quickest Time to Value

- For your chosen implementation
 - Transactional
 - Registry
 - Collaborative Authoring

Proven – The Deepest Experience

- 600 + customers
- Lowest risk
- Global reach
- Information Governance Workshops

마스터데이터관리를 통해 얻는 비즈니스 효과는?



Improve business agility
Integrate new systems (M&A) in weeks



Lower operational costs
Eliminate 2 of 3 CRM systems

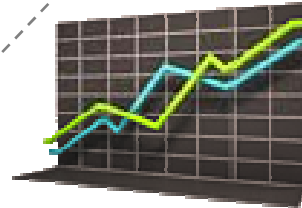


Improve customer service
Reduce time to search for customers by 58%

Align MDM goals with business priorities



New market and product strategy
Increase campaign yields by 5%



Cross-sell & Up-sell
Increase average deal size by 2%

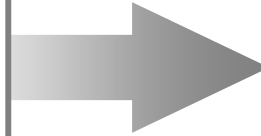


Comply with regulations
Improve privacy preference compliance by 10%



MDM 주요 해외 적용 사례

Make real-time product offers based on accurate & complete customer data



Access clinical patient data on demand



Know-your-customer compliance & auditing



Accurate client identification for software purchasing processes



Reduce time to market with new product introductions



MDM's Big Data Challenge!

12+ TBs
of tweet data
every day



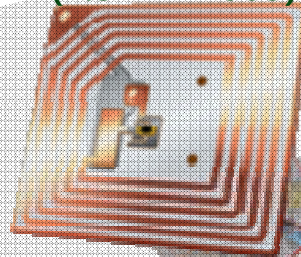
? TBs of
data every
day



25+ TBs of
log data
every day



30 billion RFID
tags today
(1.3B in 2005)



4.6 billion
camera
phones
world
wide



100s of millions
of GPS
enabled
devices
sold
annually



76 million smart
meters in 2009...
200M by 2014

2+ billion
people on
the Web
by end
2011



MDM & Big Data 통합방안

Entity Matching on Big Data Sets

- Probabilistic matching on large volumes of data

Mining Big Data to Populate MDM

- Option 1: start from MDM profile, then search through unstructured big data to augment MDM
- Option 2: mine big data to populate MDM

Social Media Analytics & MDM

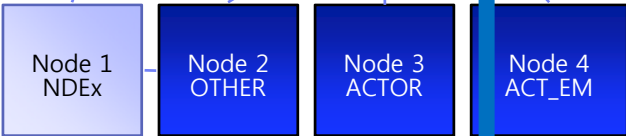
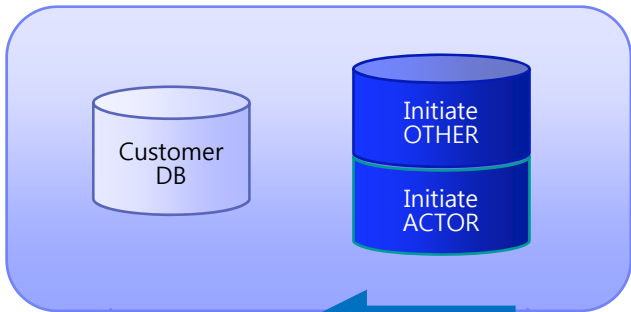
- Understanding customer's social media profiles
- Analyzing big data for targeted customers



Big Data & MDM 통합방안 #1

Entity Matching on Big Data Integrated with MDM

InfoSphere MDM
Entity Management turned off



Copy raw data files to the distributed filesystem (GPFS-SNC), which causes them to be split into large blocks and spread out over the cluster.



Key	Family	Qualifier	Value
IBM:1	Attribute	FNAME	Thomas
IBM:1	Attribute	LNAME	Watson
IBM:1	Attribute	ADDR1	11501 Burnet Rd
IBM:1	Attribute	CITY	Austin
IBM:1	Attribute	STATE	TX
IBM:1	Attribute	ZIP	78758
IBM:1	Comparison		TOM^WATSON^11501^...
IBM:1	Bucket	12345	
IBM:1	Bucket	54321	

Load BXM Entities into mpi_entlink_tables

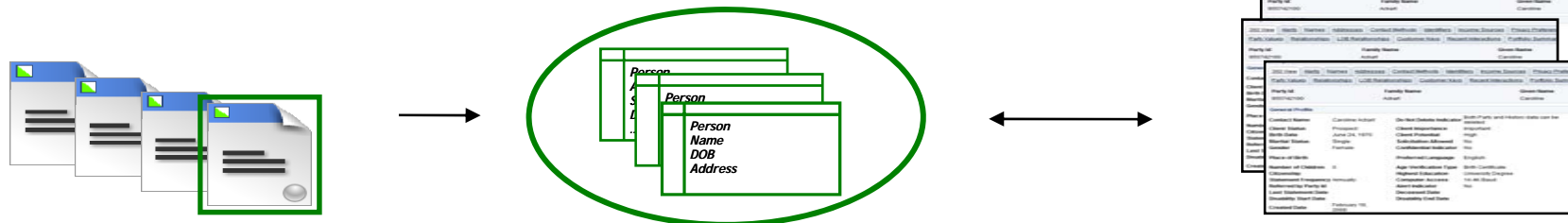
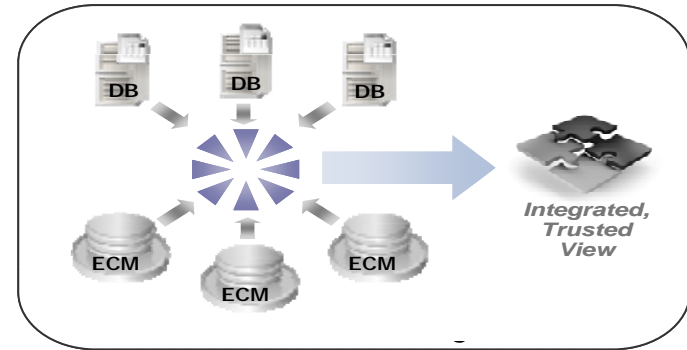
Load data into Apache HBase using a MapReduce job that standardizes and derives the data using the IBM Initiate Probabilistic Matching Engine.



Big Data & MDM 통합방안 #2

Mining Big Data to Populate MDM

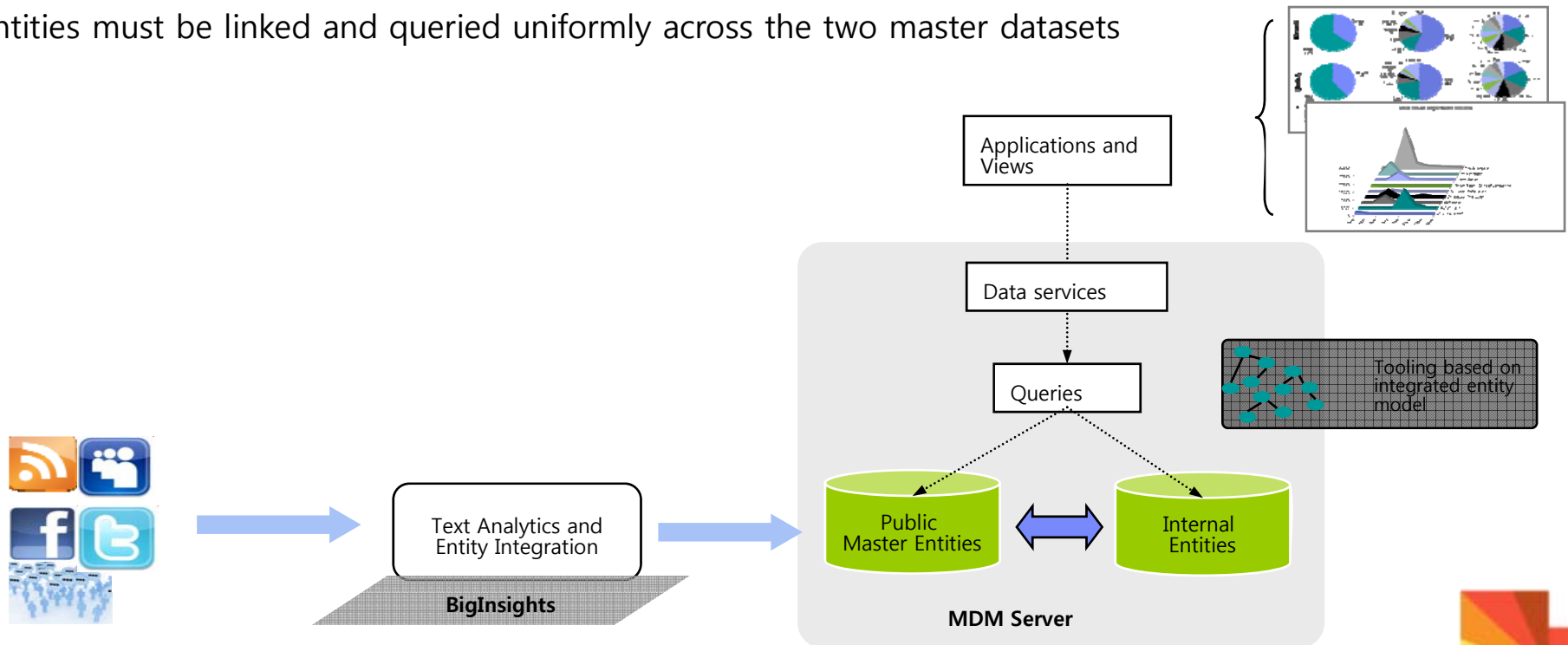
- Bridges the gap between unstructured content and master data by automatically extracting entity information from documents.
- Makes use of existing structured data from MDM to extract information about one or more entities from each document
- The extracted information is grouped into entity records that are provided to MDM to enhance existing master data and to improve entity resolution and relationship discovery based on evidence found in unstructured data



Big Data & MDM 통합방안 #3

Social Media Analytics & MDM

- Extraction, integration and management of entities (profiles) from Social Media
 - Tens of Millions of profiles (e.g., from Twitter, Blogs and Forums)
 - Use Cases: **Customer Retention/Lead Generation, Measuring Campaign Management Effectiveness**
- Integrating Enterprise-specific data with the Public master entities
 - Entities must be linked and queried uniformly across the two master datasets



Social Media Data



Big Data & MDM 통합에 따른 기대효과

- ◆ Augment your traditional product information data with dynamically derived product traits based on web & social media feedback
- ◆ Improve the '360-degree customer view' that you provide to customer service by adding in unstructured content alongside structured master data
- ◆ Discover additional relationship links between your master entities based on insights from unstructured documents
- ◆ Enhance the '360-degree customer view' that you provide to marketing & sales by deriving additional insight from social media

