




**ORACLE®**

## **Data Management, Governance and Data Quality**

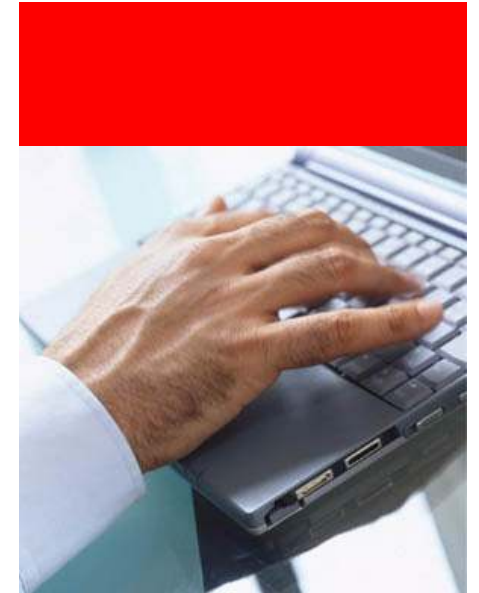
**Haidong Song**  
**Product Strategy Director**  
**Master Data Management**  
**Oracle Corp**



The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

# Program **Agenda**

- Governance and Stewardship Overview
- Customer Success Stories
- Governance, Stewardship with Oracle  
MDM Solutions



# Governance and Stewardship Overview



# Poor Data Quality is the #1 enemy of MDM-CDI Solutions

## Garbage

Typos, misspellings, transposed numbers, etc.



## Out of Date

Rapid changes in a dynamic society: marriages, divorces, births, deaths, moves



## Fraud

Purposeful misrepresentation of data: identity theft, wrong information (bankruptcies, occupation, education, etc)

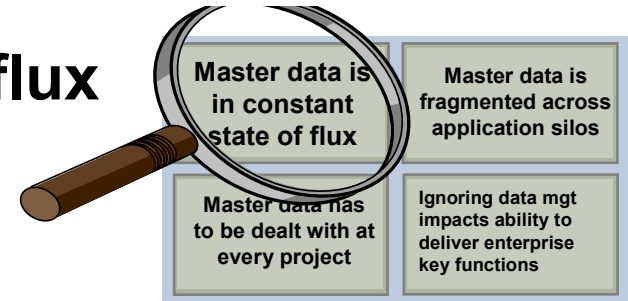


## Missed Opportunities

Information that we do not know about (customer relationships, upsells, cross-sells)



# Master Data is in constant state of flux



## Individuals

### *In one hour...*

- **5,769 individuals in the US will change jobs**
- **2,748 individuals will change address**
- **515 individuals will get married**
- **263 individuals will get divorced**
- **186 individuals will declare a personal bankruptcy**

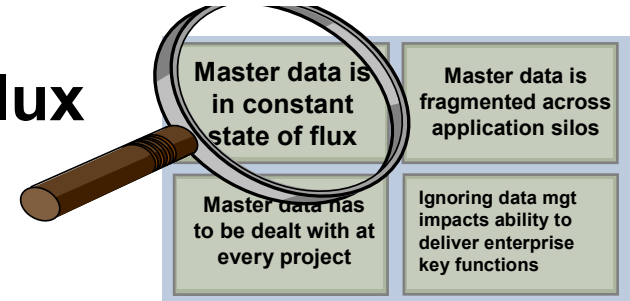
### *In one year...*

- **4.7 Million Marriages**
- **1.53 Million First Births**
- **2.04 Million First-time Home Buyers**
- **1.9 Million Divorces**
- **43 Million Residential Moves**
- **1.4 Million Work Retirements**

Source: D&B, US Census Bureau, US Department of Health and Human Services, Administrative Office of the US Courts, Bureau of Labor Statistics, Gartner, A.T Kearney, GMA Invoice Accuracy Study

**Master data changes at rate of 2% per month.**

# Master Data is in constant state of flux



## US Companies

### *In one hour...*

- 183 business telephone numbers will change or be disconnected
- 43 business addresses will change
- 36 directorship (CEO, CFO, etc.) changes will occur
- 33 new businesses will open their doors
- 8 companies will change their names

### *In one year...*

| % Change<br>in 12<br>Months | New<br>Businesses<br>(<2 yrs old) |       |
|-----------------------------|-----------------------------------|-------|
|                             | All                               |       |
| Company<br>Names            | 12.4%                             | 36.4% |
| Physical<br>Address         | 20.7%                             | 27.3% |

**Master data changes at rate of 2% per month.**

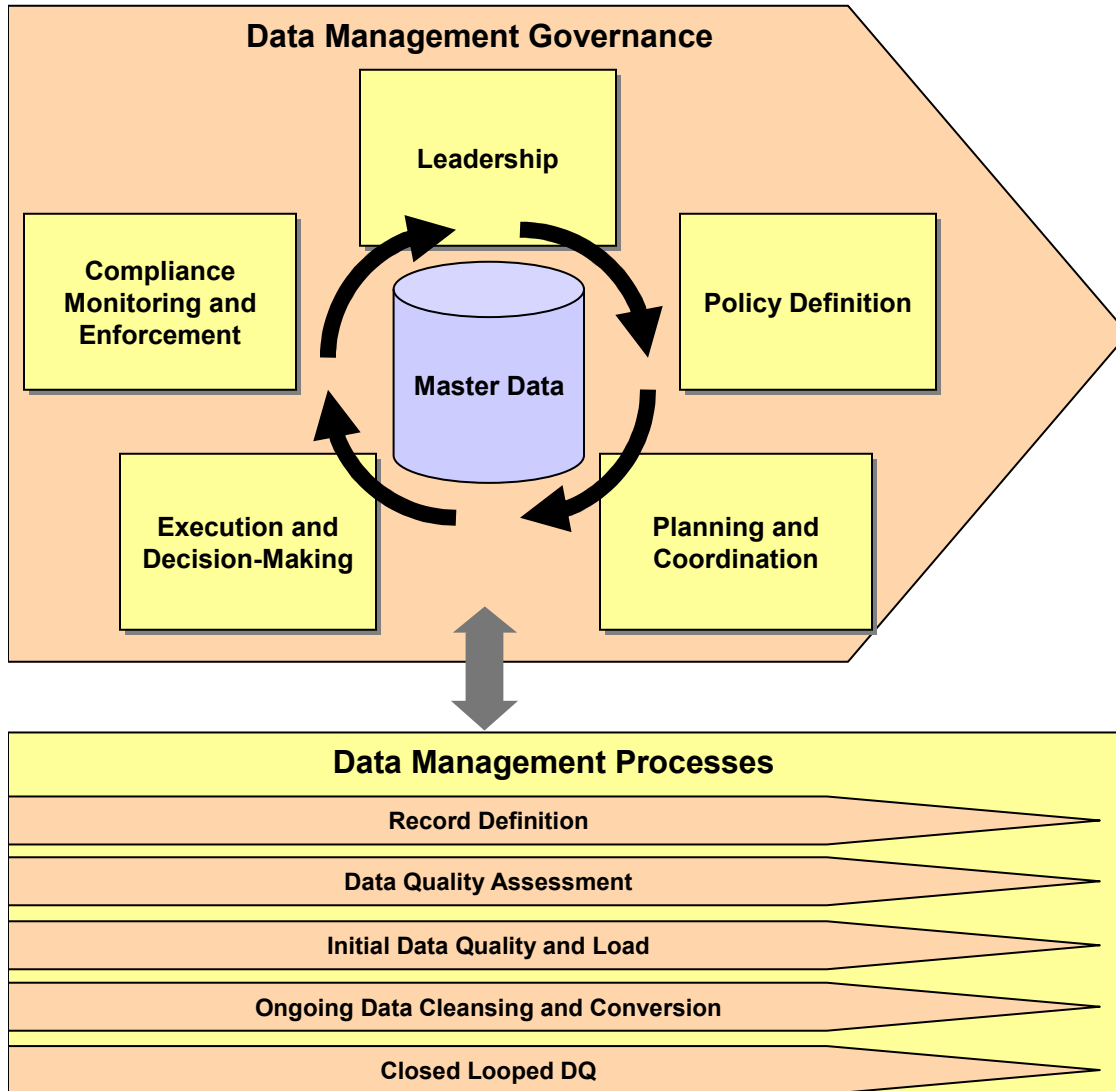


# Why Data Governance is Needed

- Compliance
  - SOX, Basel II
  - Consumer Privacy
  - Information Security
- Legacy Data Migration and M&A Activity
  - Fixing Problems instead of Moving them from one system to another
- Operational Business Intelligence
  - Getting it right 'now' matters
- Data Quality
  - Inhibitor to Application Acceptance



# Formalize a Governance Framework



- Central executive leadership
- Enterprise steering committee to arbitrate issues and enforce the rules
- Coordination and compliance
- Define & communicate data quality expectations
- Establish policies, procedures, success metrics and processes to maintain quality data
- Identify all business and application stakeholders across the enterprise – data owners
- Conduct audit and control
- Communication and change management

# Roles & responsibilities of the Data Steward

- Responsible to perform and support **all data lifecycle activities**
- Has **in-depth expertise** of architecture and tools
  - Masters all data quality tools involved in process
  - Has all creation, deletion and alteration privileges
  - Always performs in coordination with data owners
- Performs **everyday upkeep** of data quality
  - Monitoring, profiling, cleansing, merging, unmerging
- Defines overall data governance by creating **clear and unambiguous** definitions of all data elements



**Monitor  
& Profile**



**Consolidate**



**Cleanse**



# Customer Success Stories



# Loyalty & Compliance



*“We selected Siebel CDI because of its out-of-the-box, rich customer master functionality, its industry-specific best practices, and its ability to integrate many different applications”*

*Shaun Coyne, VP & CIO*

## COMPANY OVERVIEW

- 7<sup>th</sup> largest finance company in US
- Revenue: \$1B
- Customers: 10M

## CHALLENGES/OPPORTUNITIES

- Need to increase **customer service & loyalty** for dealers & consumers
- Need to **comply to complex privacy management rules**
- Need for **new generation IT architecture** (Flexible, scalable, less costly, integrated)
- Need for **phased path** to new architecture

## SOLUTIONS

- Oracle Customer Hub
- TIBCO & UAN, Informatica
- 12 integrated Applications

## RESULTS

- **Support for corporate priorities**
  - Growth, Info control & access, new services introduction, respect of privacy rules
- **Deepen dealer & customer loyalty**
  - Improved services, customer insight & customer satisfaction & retention
- **Create operational gains: IT costs & process efficiency**

# Holistic Customer View



Count on more.™

*“The project yielded CRM success, some of which UMB links directly to Trillium Software and Oracle UCM. UMB’s information aggregation model is highly streamlined and has gained accolades from peer group customers and fellow commercial entities.”*

*The Data Warehouse Institute Best Practice Award for Data Governance*

## COMPANY OVERVIEW

- Assets: Over USD 8 Billion
- 3500+ Associates
- 140 Branches in 8 States

## CHALLENGES/OPPORTUNITIES

- Create a **360 Degree, holistic view** of UMB’s Customer Relationships
- **Increase efficiency** of front office by consolidation of customer data
- Improve **cross selling** opportunities
- **Empower employees** to anticipate, advocate and advise; innovate and surprise

## SOLUTIONS

- Oracle Customer Hub
- Trillium Data Quality
- 7 Primary Sources of Customer Data
- Data Stewards report to the Business, not IT

## RESULTS

- **Excellent Performance**
  - No SLA’s missed since go-live
  - Usually ‘waiting’ for Mainframe batch processes to end
- **Deeper Understanding** of the definition of a customer (ongoing)

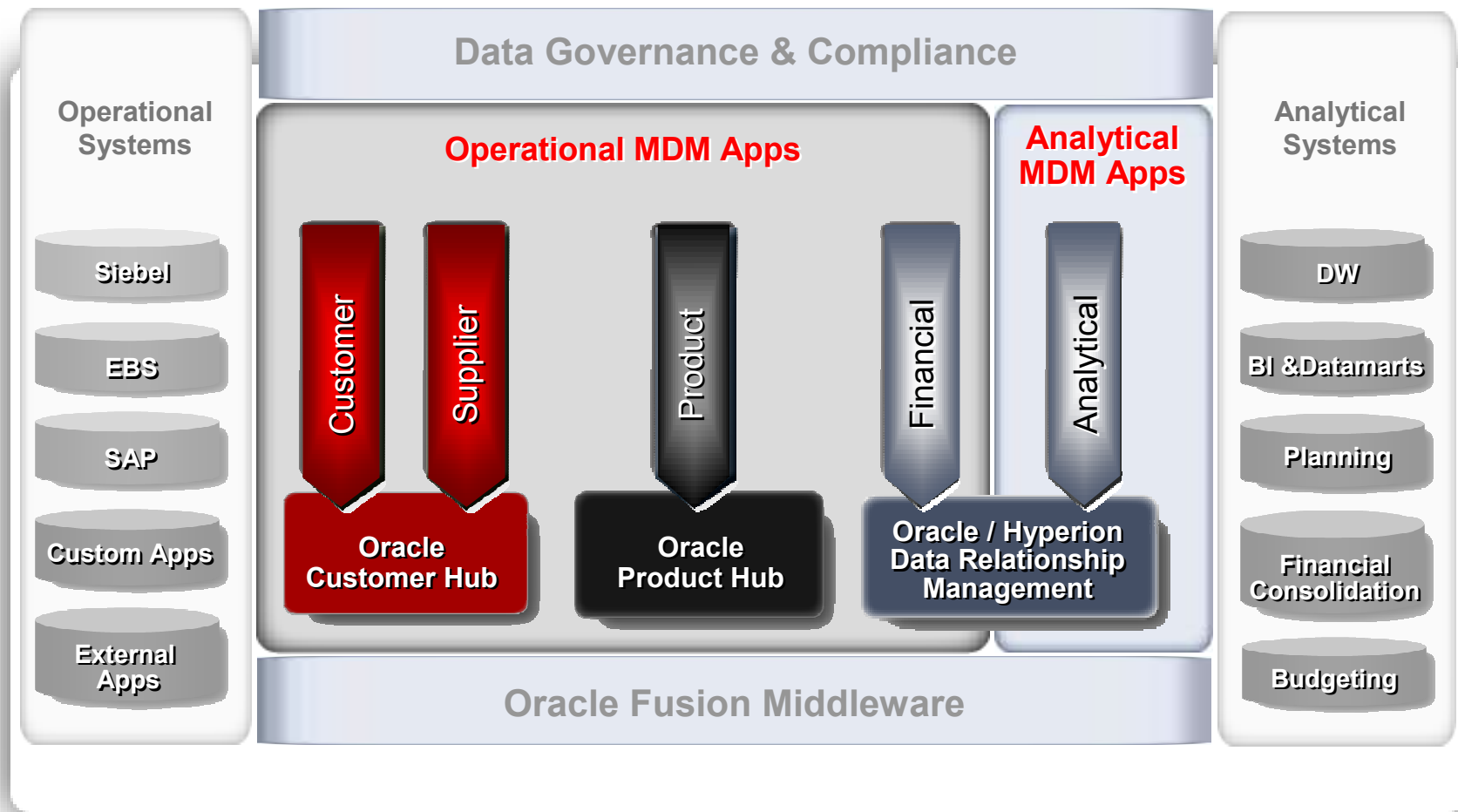


# Governance and Stewardship with Oracle MDM Solution



# Oracle Enterprise Master Data Management

The most complete MDM offering in the market today





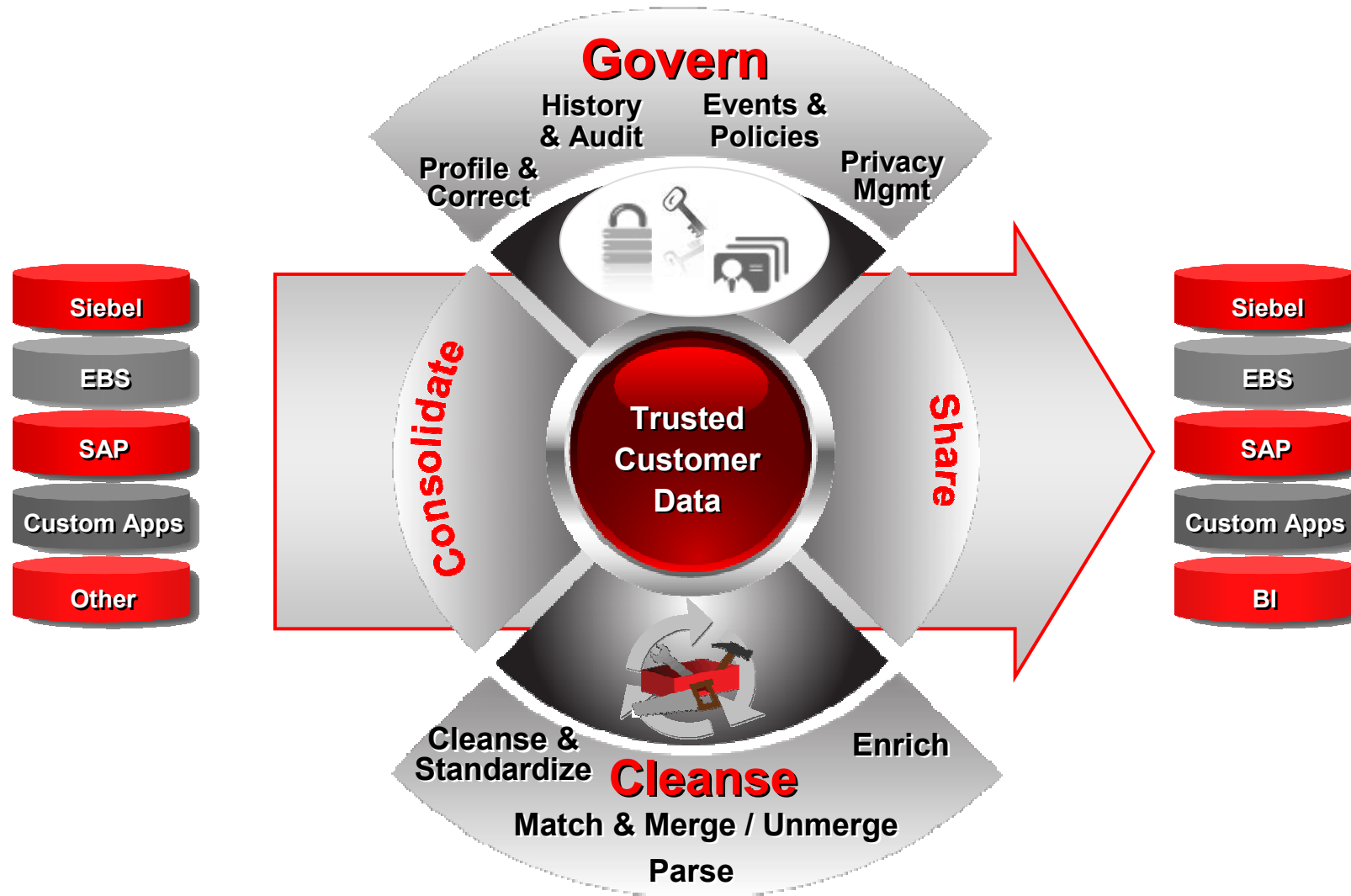
# A Day in the Life of a **Data Steward**

## *Ensuring Data Quality across MDM hub*

1. Runs **profiling routines** to monitor overall DQ within application
  - Inspects most crucial or known problem areas
  - Gains deep-level understanding of data (*e.g. min, max, # nulls..*)
2. Creates and applies new **data rule** based on profiling results
3. Resolves **duplicates** and creates **links**
4. Reviews **history** and **audit trail**
5. Defines **compliance rules** and **policies**
6. Defines event and policies for ongoing monitoring and management
7. Executes **corrective action**: recover, unmerge, etc.
8. Performs ongoing **monitoring** of data quality



# Key Components of Oracle Customer Hub



# Data Watch and Repair

## *A Closed Loop Process*

### **1. Connect:**

**OWB  
and MDM**



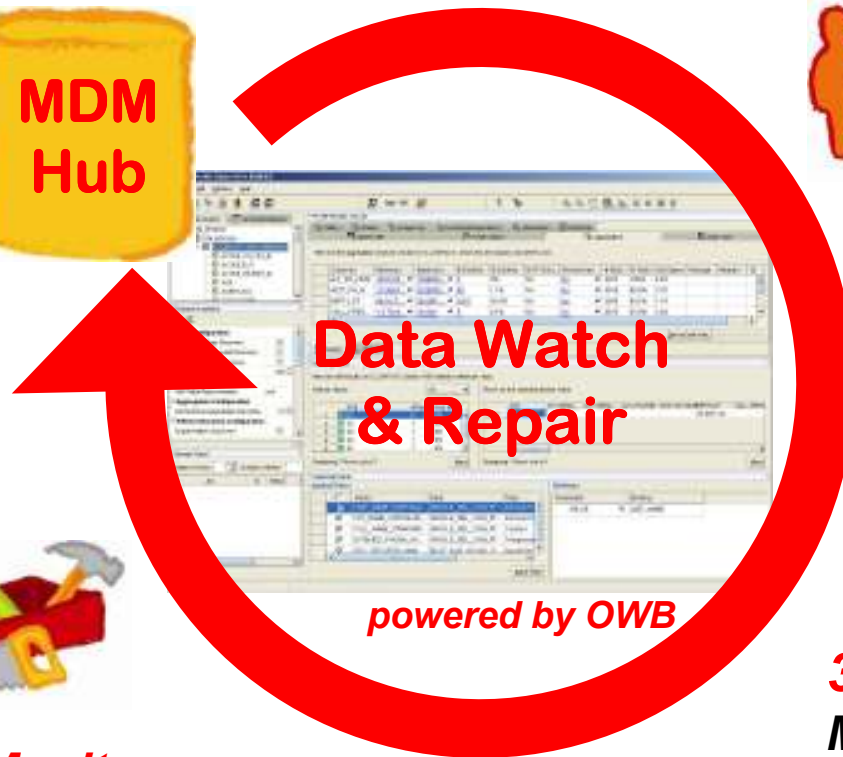
**2. Profile:** Discover structure, meaning, and quality of data



**4. Fix and Monitor:**  
Schedule corrective actions for future



**3. Assess:**  
Measure data compliance using data rules



# Connect: OWB and MDM Hub

The screenshot displays the Oracle Design Center interface with the following components and annotations:

- Project Explorer:** Shows a tree view of a project named 'MY\_PROJECT'. Under 'Applications', 'SIEBEL\_UCM' is expanded to show 'UCM\_MODULE', which is circled in red. Below it, a list of metadata tables is visible, including S\_ACCNT\_ATT, S\_ADDR\_PER, S\_APPLD\_DISCNT, S\_AREA\_LS, S\_ASGN\_GRP, S\_ASSET, S\_ASSETCON\_ADDR, S\_ASSET\_CON, S\_ASSET\_VWRNTY, S\_BU, S\_CASE, S\_CIF\_AST\_MAP, S\_CIF\_CON\_MAP, S\_CIF\_EXT\_SYST, S\_CIF\_ORGRP\_MAP, S\_CIF\_ORG\_MAP, S\_CIF\_SYS\_DTL, S\_CM\_TELESET, S\_CONTACT, and S\_CONTACT\_FNXM.
- Connection Explorer:** Shows a tree view of 'Locations' and 'Databases'. Under 'Locations', 'CORRECTED\_PROFILE\_LOCATION' is circled in red.
- Global Explorer:** Shows a tree view of 'Public Transformations', 'Public Experts', 'Public User Defined Modules', 'Public Data Rules', 'Icon Sets', and 'Security'.
- Annotations:**
  - Red text: "Connector imports metadata for MDM schema into OWB" with a red arrow pointing to the 'UCM\_MODULE' folder.
  - Red text: "MDM instance location and connection details" with a red circle around 'CORRECTED\_PROFILE\_LOCATION'.
  - Red text: "Connectors developed for" followed by a list:
    - CDH
    - UCM / Siebel CRM
    - PIM

# Profile:

## Discover data structure, meaning, and quality

### Profiling discovers in data:

- min/max values, ranges
- domains and patterns
- % of distinct, NOT NULL values

Click Next

Property Inspector

Load Configuration

Aggregation Configuration

Pattern Discovery Configuration

Monitor Panel

Number of Days: 7

Number of Rows:

Job: | | Status: |

Profile Results Console

Here are the aggregation analysis results for S\_CONTACT:

| Columns    | Minimum  | Maximum | #Distinct | % Distinct | NOT NULL | Recommen. | #Nulls | % Nulls | St. Devia | Average | Median |
|------------|----------|---------|-----------|------------|----------|-----------|--------|---------|-----------|---------|--------|
| ALT_PH_NUM | 2487244  | 7888083 | 2         | 0%         | No       | No        | 4973   | 100%    | 4.85      |         |        |
| ASST_PH_N  | 2124567  | 9259260 | 56        | 1.1%       | No       | No        | 4918   | 98.6%   | 3.76      |         |        |
| BIRTH_DT   | 09-OCT-  | 05-SEP- | 1917      | 36.5%      | No       | No        | 2959   | 59.5%   | 1.74      |         |        |
| CALL_FREQ  | 1-3 Time | No See  | 4         | 0.1%       | No       | No        | 4670   | 93.9%   | 3.04      |         |        |

Derive Data Rule

Data Drill Panel

Here are drill results on S\_CONTACT column AGE related to Minimum value:

Distinct values:

| AGE | # Rows | % of 4975 |
|-----|--------|-----------|
| 20  | 1      | 0%        |
| 22  | 1      | 0%        |
| 29  | 1      | 0%        |
| 40  | 1      | 0%        |
| 42  | 1      | 0%        |

Rows for the selected distinct value:

| AGE | ALT_EMAIL | ALT_EMAIL | ALT_PH_NUM | ASST_PH_NUM | BIRTH_DT | CALL      |
|-----|-----------|-----------|------------|-------------|----------|-----------|
| 20  |           |           |            |             |          | 05-MAY-92 |

Displaying 7 Rows out of 7

Displaying 1 Rows out of 1

Data Rule Panel

| Name               | Rule              | Type     |
|--------------------|-------------------|----------|
| LAST_NAME_CAPITALD | ORACLE_SBL_UCM_PP | Domain P |

Bindings:

| Parameter | Binding   |
|-----------|-----------|
| VALUE     | LAST_NAME |

# Data Profiling Results

Tabs show different types of profiling results

Here are the aggregation analysis results for S\_CONTACT, which has 49 columns and 4975 rows.

| Columns      | Minimum     | Maximum    | # Distinct | % Distinct | NOT NULL | Recommen... | # Nulls | % Nulls | Six-Sigma | Average | Median | St |
|--------------|-------------|------------|------------|------------|----------|-------------|---------|---------|-----------|---------|--------|----|
| ALT_PH_NUM   | 2487244...  | 7899083... | 2          | 0%         | No       | No          | 4973    | 100%    | 4.85      |         |        |    |
| ASST_PH_N... | 2124567...  | 9259980... | 56         | 1.1%       | No       | No          | 4916    | 98.8%   | 3.76      |         |        |    |
| BIRTH_DT     | 09-OCT-...  | 06-SEP-... | 1817       | 36.5%      | No       | No          | 2959    | 59.5%   | 1.74      |         |        |    |
| CALL_FREQ... | 1-2 Time... | No See     | 4          | 0.1%       | No       | No          | 4670    | 93.9%   | 3.04      |         |        |    |

e.g. Min/Max values, # and % null records, ...

Here are drill results on S\_CONTACT column AGE related to Minimum value.

Distinct values:

| AGE | # Rows | % of 4975 |
|-----|--------|-----------|
| 20  | 1      | 0%        |
| 22  | 1      | 0%        |
| 29  | 1      | 0%        |
| 40  | 1      | 0%        |
| 42  | 1      | 0%        |

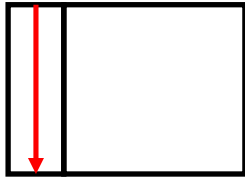
Rows for the selected distinct value:

| AGE | ALT_EMAIL | ALT_EMAIL | ALT_PH_NUM | ASST_PH_NUM | BIRTH_DT  | CALL_FREQ |
|-----|-----------|-----------|------------|-------------|-----------|-----------|
| 20  |           |           |            |             | 05-MAY-82 |           |

Drill on hyperlinks in results to show corresponding values and rows for selected value



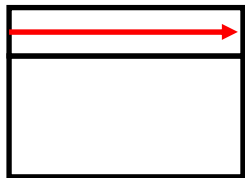
# Data Profiling Types



## Attribute Analysis

Column Level

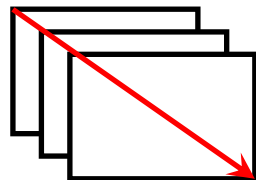
- Min/max values, value domain (list, range)
- Value frequency, special values
- Data type, length, precision
- Dominant patterns and common formats (phone, email SSN, ...)



## Dependency Analysis

Cross-column

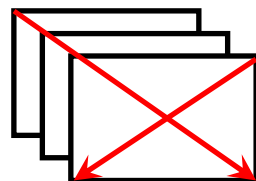
- Functional Dependency
- Candidate Key Discovery



## Referential Analysis

Cross-table

- Unique Key Discovery
- Row Relationship



## Business Rule Analysis

Custom

- Content-specific custom rules
  - Suppression List
  - Name Standardization (N&A)
  - Record Completeness
  - Domain values



## **Assess:**

### ***Measure data compliance using data rules***

- Data rules can be easily created by
  - Deriving from profile results of data
  - Manually defining them, using Oracle general expression  
e.g.: specify numerous forms of accepted capitalizations, such as: “Aaaa..”, “Aaaa Bbbb”, “Aa-Bbbb”, etc...
- Quick customizations can be implemented to adjust to new DQ needs
  - Rules can be modified or extended to meet any implementation-specific business requirements
  - Custom data rules that address MDM common DQ needs are included

# Defining and Applying a Data Rule

**Patterns for valid values are defined using Oracle Database regular expression syntax**

| Value   |
|---|
| <code>^[[:space:]]*[A-Z][a-z]*[[:space:]]*\$</code>                       |
| <code>^[[:space:]]*[A-Z][a-z]*-[A-Z][a-z]*[[:space:]]*\$</code>           |
| <code>^[[:space:]]*[A-Z][A-Z][a-z]*[[:space:]]*\$</code>                  |
| <code>^[[:space:]]*[A-Z][a-z]*[[:space:]][A-Z][a-z]*[[:space:]]*\$</code> |

**Applying a data rule**

- % of compliance is shown
- values can be drilled into

| Name   | Rule                    | Type       |
|--|-------------------------|------------|
| <input checked="" type="checkbox"/> LAST_NAME_CAPITALIZ... | NAME_CAPITALIZATION ... | Domain Pa  |
| <input checked="" type="checkbox"/> FST_NAME_CAPITALIZA... | ORACLE_SBL_UCM_PF...    | Domain Pa  |
| <input checked="" type="checkbox"/> FULL_NAME_STANDARI...  | ORACLE_SBL_UCM_PF...    | Custom     |
| <input checked="" type="checkbox"/> EXTENDED_PHONE_NUM...  | ORACLE_SBL_UCM_PF...    | Telephone  |
| <input checked="" type="checkbox"/> SOC_SECURITY_NUM       | BUILT_IN_IS_SOCIAL_S    | Social Sec |

| Parameter | Binding     |
|-----------|-------------|
| VALUE     | % LAST_NAME |





## **Repair and Monitor:**

### ***Schedule corrective actions for future***

- Correction schemas create a corrected set of data
- Non-compliant entries are ignored, reported or corrected
- Correction strategies
  - Remove
  - Similarity Match
  - Soundex Match
  - Custom functions (written in PL/SQL)



# User Scenario: Data Steward

## *Ensuring DQ maintenance across MDM hub*

1. Runs **profiling routines** to assess overall DQ within application
  - Inspects most crucial or known problem areas
  - Gains deep-level understanding of data (*e.g. min, max, # nulls..*)
2. DS finds many incomplete records with name match
  - Often several incomplete records are entered by mistake by a sales representative before a new one is successfully created
3. DS creates and applies new **data rule** that checks for record completeness
  - A contact must now include: full name, common-pattern SSN & complete address within US
4. Compliant and non-compliant values are examined
  - DS defines cleansing strategy for non-compliant values: **Delete** all incomplete records with a matching name to a complete record.
5. DS executes **corrective action** and writes back to MDM hub



## **Solution: Constant and easy-to-perform** *Data Watch and Repair*

Data Quality solution should be complemented with a module that is

- ***Non-intrusive:*** has to be able to run in parallel to normal data traffic without slowing it down or interrupting
- ***Quick and easy:*** data watch needs to quickly and reliably identify any data inconsistencies
- ***Flexible and customizable:*** fixes data to changes that arise in a dynamic business environment

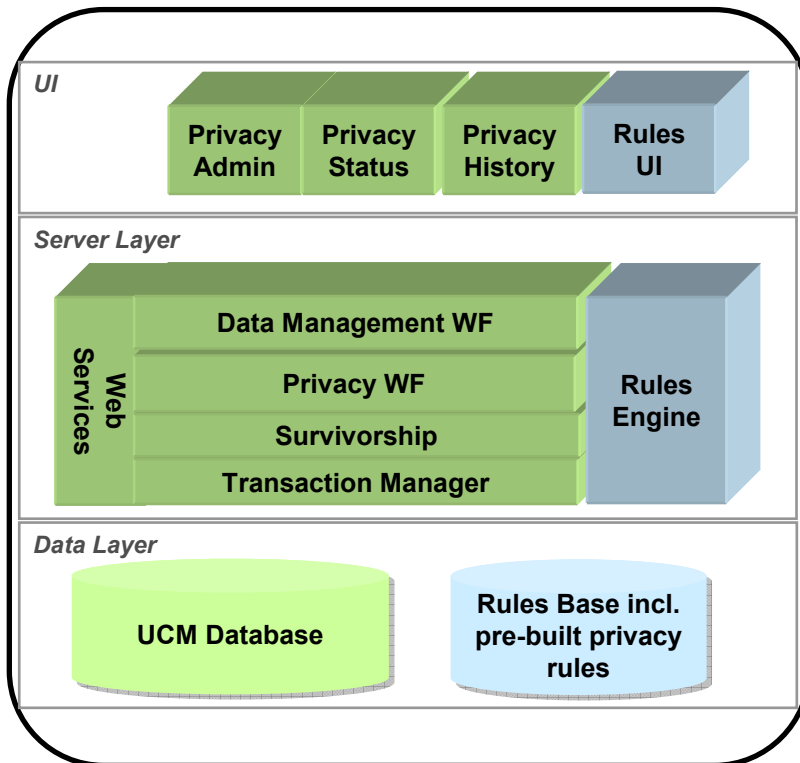


# Resolution, History and Audit

## *Lineage of Data Life Cycle*

- **Resolve** pending update or merge actions
  - Data steward needs to resolve pending actions when records are below automatic update or merge thresholds
- During any **resolution** activity,
  - A copy of current best version is recorded in the record's **History**
  - Current best version and incoming record are **linked**
  - **Survivorship rules** are invoked to produce new best record
- An **audit trail** is a history of the changes made to a record throughout these processes
  - Includes source data and of previous best versions

# Privacy Management

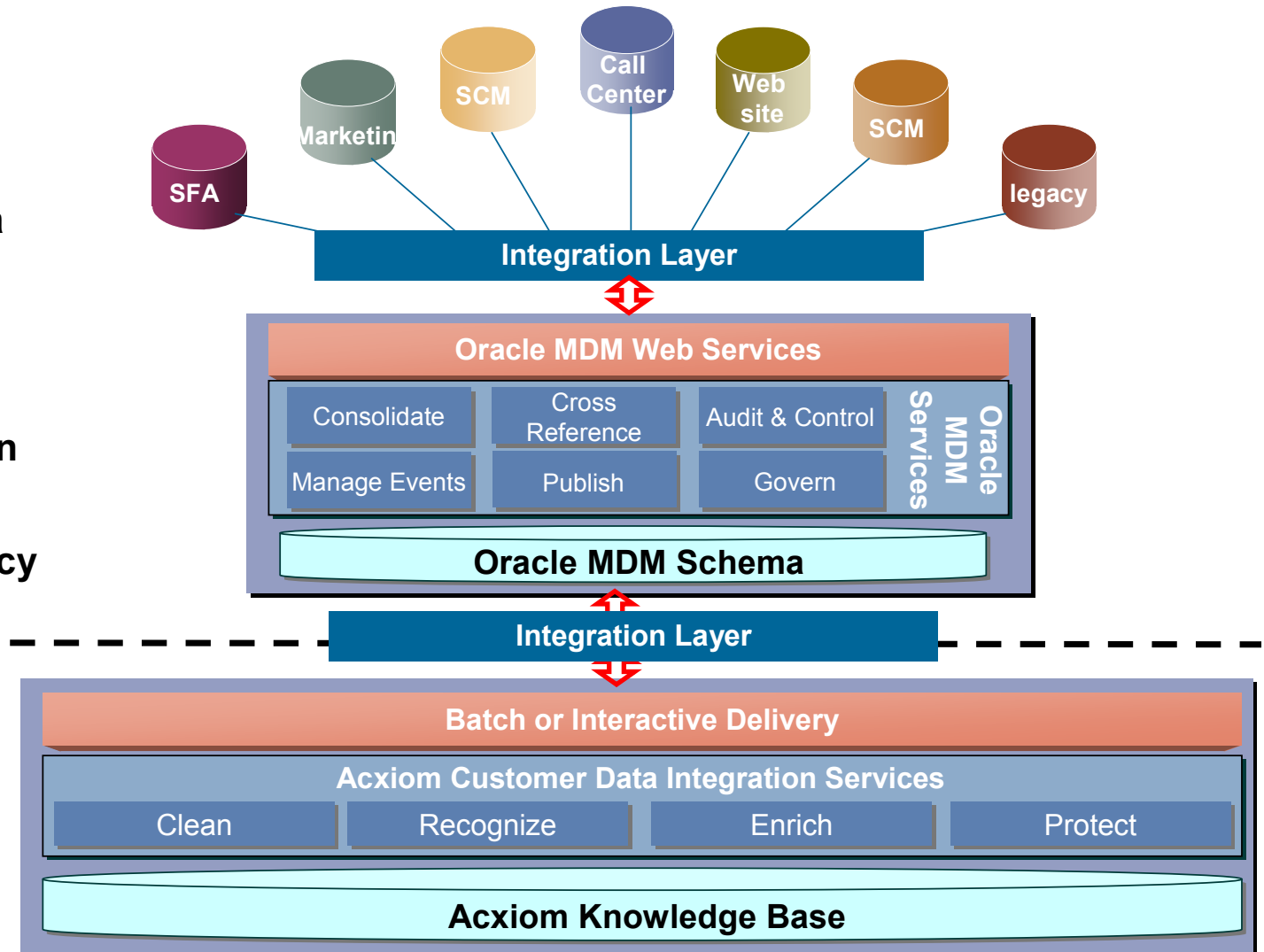


- Privacy specific data model and metadata
- Data Steward / Privacy analyst views
- Modular Framework
- Sample Rules
- Pre-Built Workflows
- Support Fair Information Practice Principles
- Real Time Services, Batch Integration



# Knowledge Based MDM – World class CDI with unprecedented DQ

- Enterprise MDM with effective data
- Rapid implementation
- Reduced cost of content integration
- Improved Risk Mitigation & Privacy



# Knowledge-Based MDM: Key benefits

## *Highest Data Quality*

Ensure accurate consolidated & complete single view of customer and prospects data

- Better sales and customer retention abilities
- Reduced costs of data management (e.g. IT, marketing mailing)
- Easier risk & fraud management

## *Lower Cost of Ownership*

Reduce implementation time and integration costs

- Real time & batch integration
- Easier consolidation of existing customer & prospect databases
- Leverage of true Enterprise Operational MDM-CDI solution
- Reduced implementation risk & enhanced support

## *Security & Privacy*

Take into account all regulatory and compliance mandates

- Privacy data is populated in the master and disseminated across the enterprise:
  - Do not call and do not mail lists
  - Fraud detection

# Prospect Mastering

## with Knowledge-Based MDM

### Campaign Planning

- Perform segmentation within Siebel Marketing application
- Generate prospect selection criteria

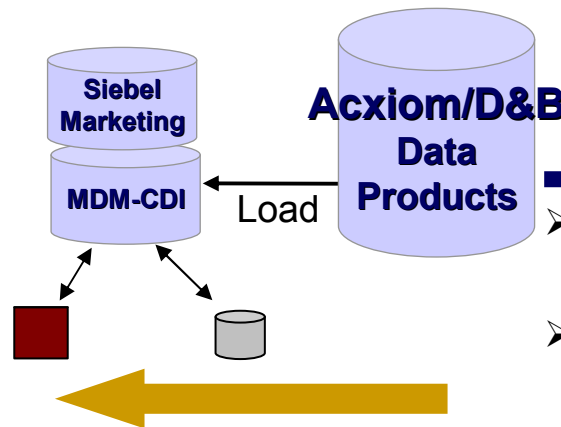


### Prospect Acquisition

- Send criteria and list of existing cust/prospect to Acxiom/D&B etc
- Acxiom/D&B produces the net new prospect list and send to customer
  - Contact information
  - Demographic data
  - Wealth/income classifications
  - Segmentation groupings
  - Lifestyle indicators

### Campaign Execution

- Plug & Play Market Campaign Execution



### Loading & Matching

- Load selected prospect records into Oracle MDM-CDI solution
- Consolidate existing customer info with prospects from other sources



# Prospect Mastering: Key benefits

## *Plug & Play Marketing*

Pre-population of Siebel Marketing with Acxiom content

- Reduced time to market for marketing campaign
- Marketing information including: Contact, Demographic Data, Wealth & Lifestyle Indicators & Segmentation, ...

## *Better consumer intimacy*

Easier ability to consolidate with existing customer & prospect databases

- Prevention of duplicate/inappropriate offers sent out to same prospect / customer
- Better customer targeting

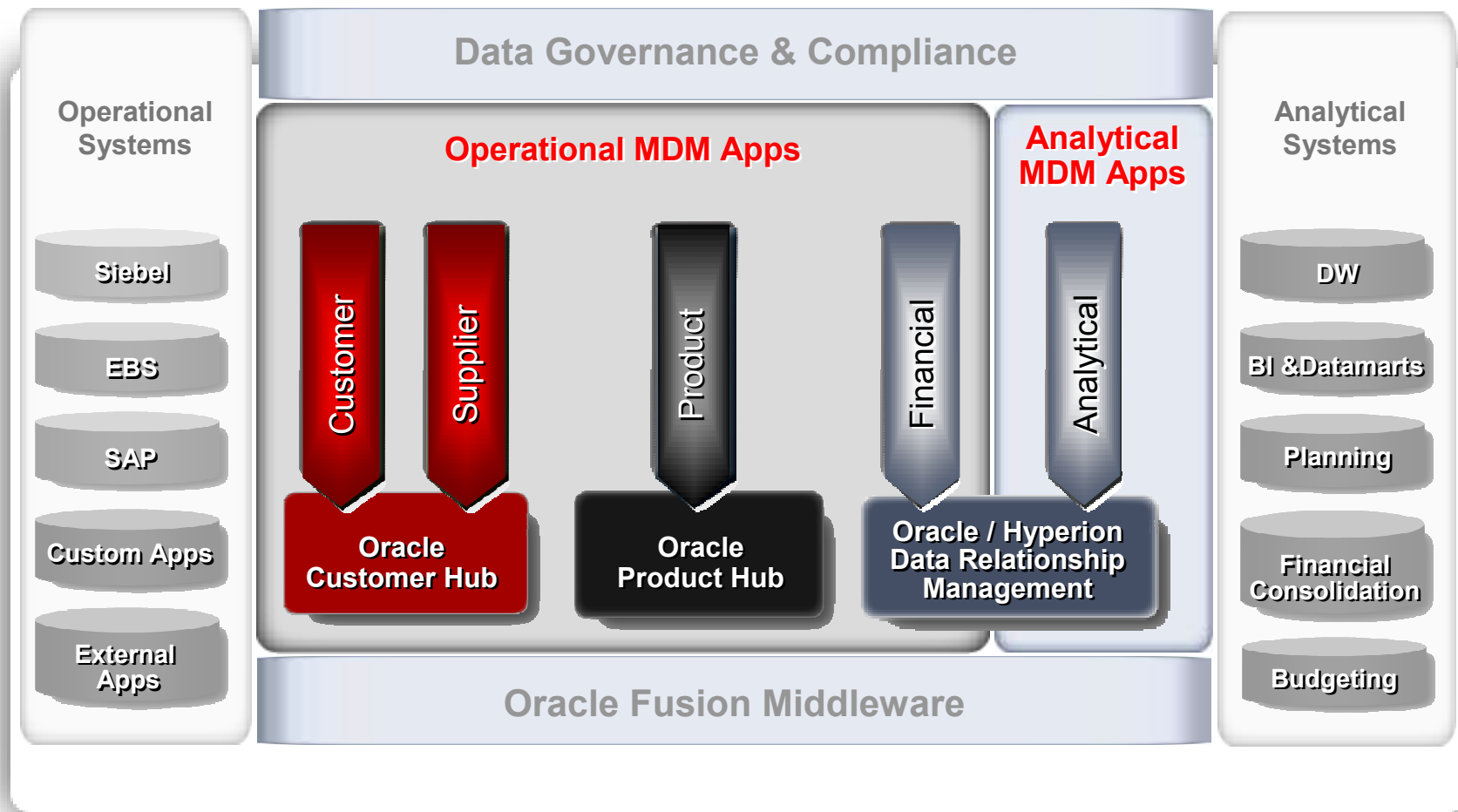
## *Lower Cost of Ownership*

Reduce implementation time and integration costs of Acxiom content

- Reduced marketing campaign costs and increased ROI
- Improved IT agility, reduced implementation risk & enhanced support

# Oracle Enterprise Master Data Management

The most complete MDM offering in the market today





## For More Information

[search.oracle.com](https://search.oracle.com)

Master data management



or

[oracle.com](https://oracle.com)

**ORACLE®**