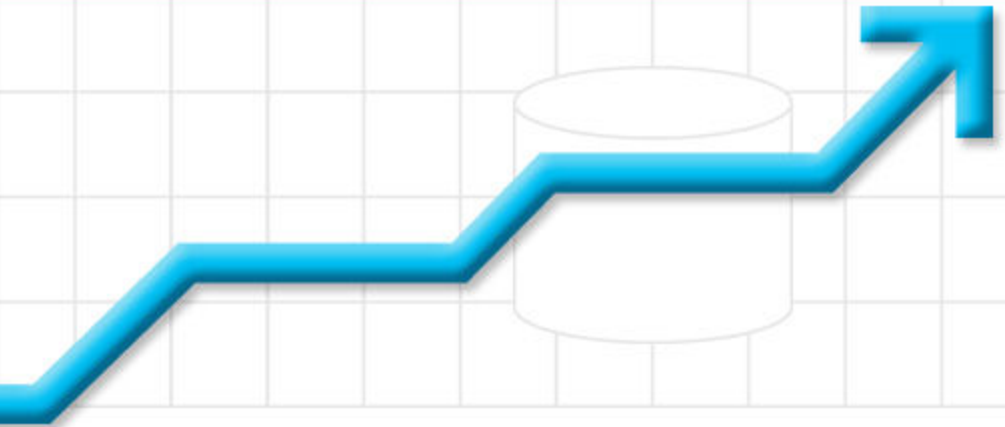
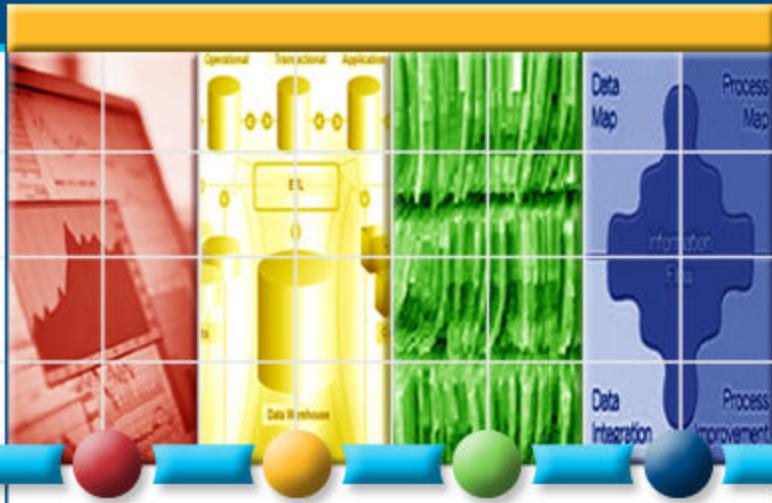


MDM for the Enterprise

Thank you for attending the MDM for the Enterprise Seminar Series!

- Please do not distribute these presentations without permission from the speaker (see contact information within.) This is just intended as an additional reference for seminar participants.
- Keep in touch! What other MDM events, topics, articles or resources will help you? Email editor@searchdatamanagement.com with your thoughts.



baseline
C O N S U L T I N G
Data Mastered. Value Unleashed.™

www.baseline-consulting.com

Bullet-Proof MDM: Designing a World-Class Development Environment

Evan Levy

Partner

Baseline Consulting

TechTarget MDM for the Enterprise Series





Views of MDM from the Industry

"MDM is much more than a single technology solution; it requires an ecosystem of technologies to allow the creation, management, and distribution of high-quality master data throughout the organization"



MDM is a set of disciplines and strategies that can be combined with information management products and services in order to provide a single view of customers, products, or other business entities.



"MDM is a workflow-driven process in which business units and IT collaborate to harmonize, cleanse, publish and protect common information assets that must be shared across the enterprise.



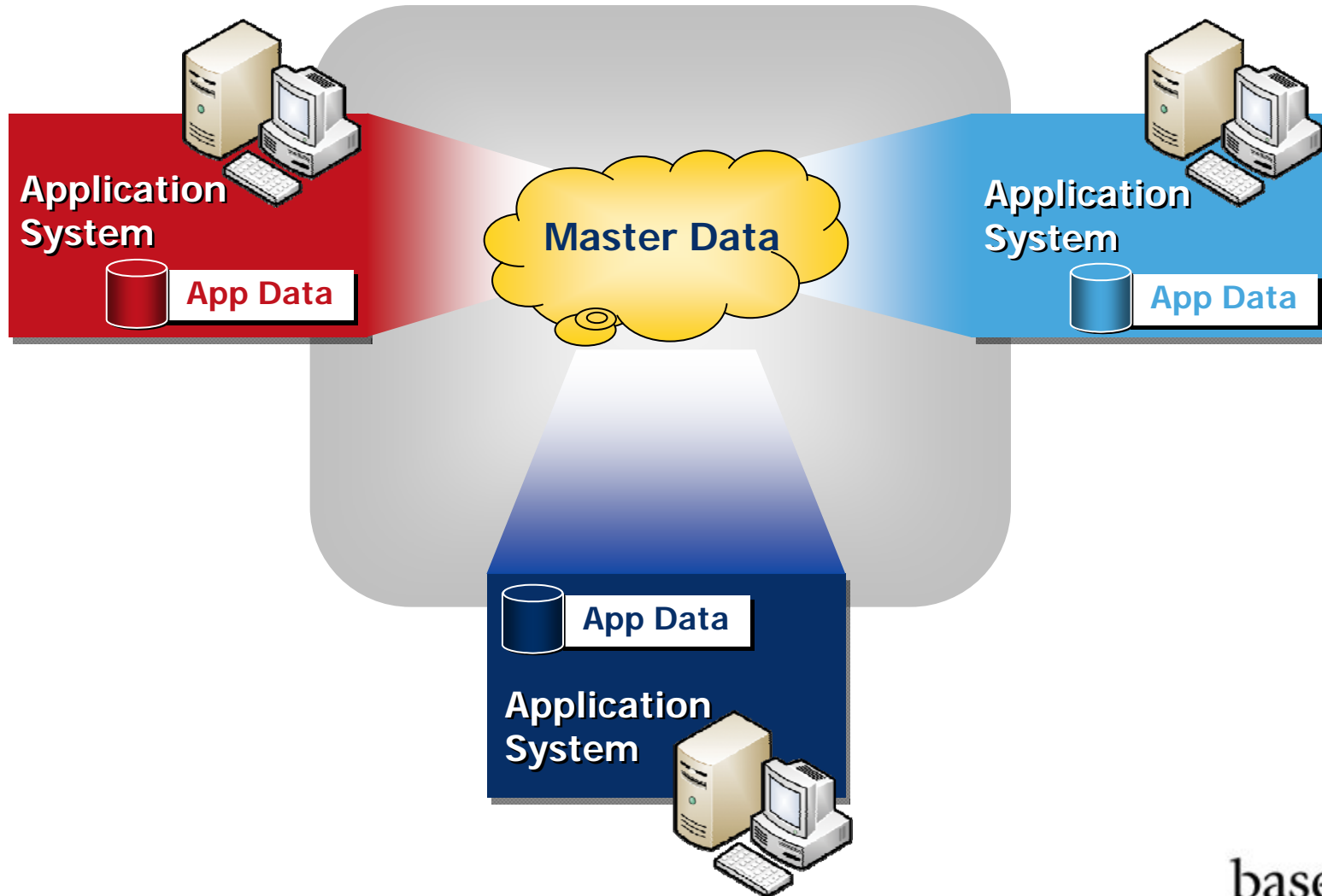
"The discipline in IT that focuses on the management of reference or master data that is shared by several disparate IT systems and groups"



WIKIPEDIA



MDM: The Convergence of Data

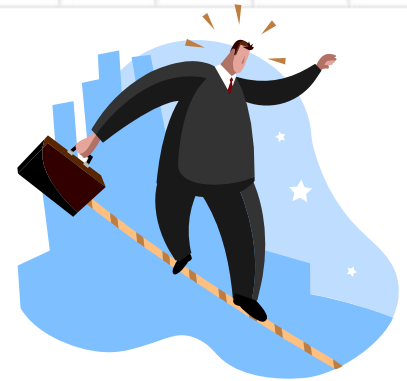




The Development Environment

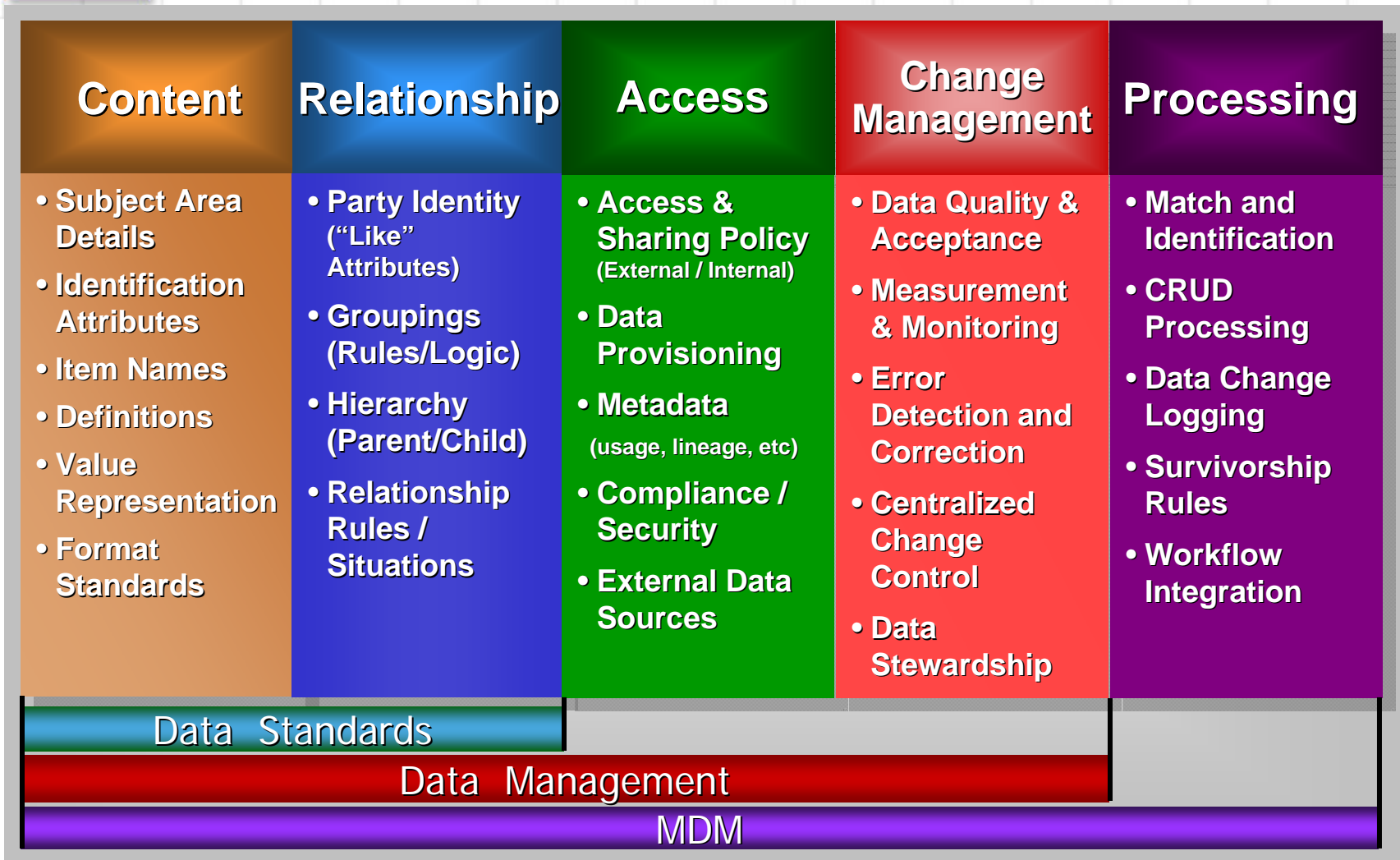
Common MDM Development Challenges

- ❏ Determining the MDM Capability Requirements
- ❏ Positioning MDM as an Infrastructure Component
- ❏ Understanding the Differences in MDM Development
- ❏ Staffing the Development Team



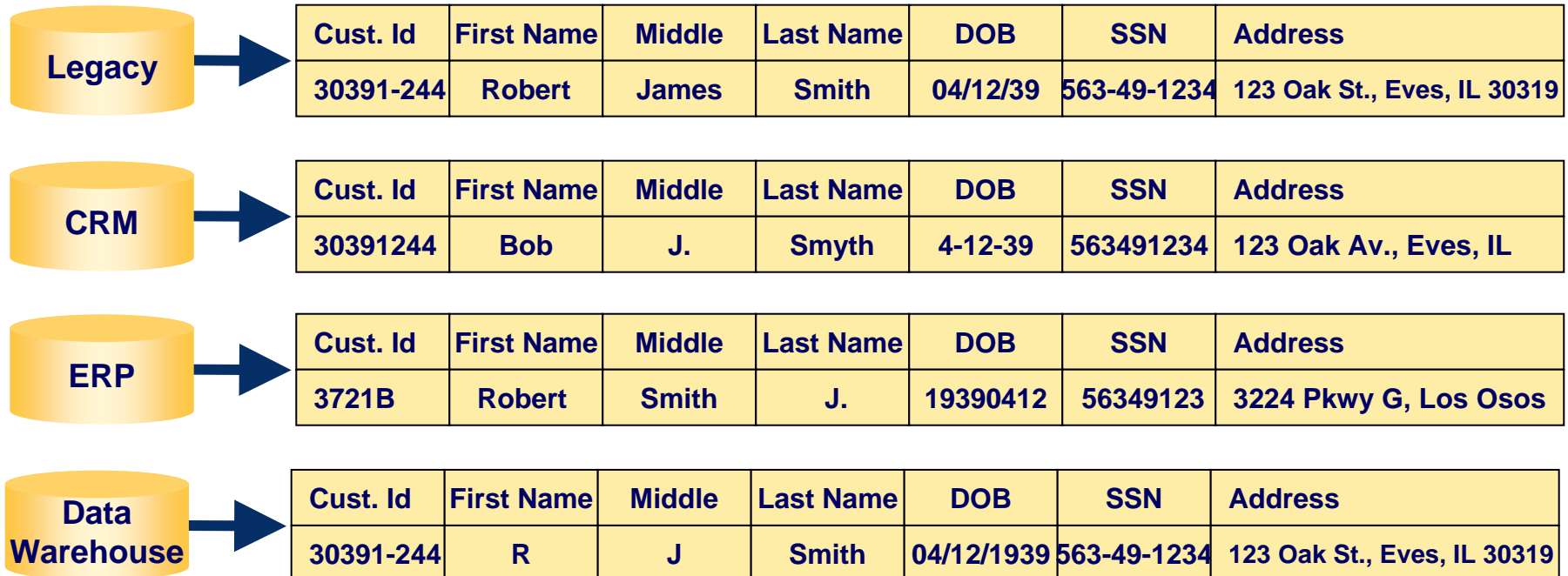


The Elements of MDM





Data Identification



- Attributes used to recognize unique (and non-unique) values
 - Used for storage and organization purposes
 - Surrogate keys are often used to simplify id retention
- Multiple descriptive attributes are frequently used for identification



Data Standardization

How is the customer ID represented?

Is there a standard name format? (Bob, Robert, Rob, etc.)

Do we store this with or without dashes

Cust. Id	First Name	Middle	Last Name	DOB	SSN	Address
30391244	Bob	J.	Smyth	4-12-39	563491234	123 Oak St., Eves, IL

What happens if the name changes (acquisition, legal activity, etc.)?

MM/DD/YY or DD/MM/YY?

Which address is this (home, business, etc.)



Cleanse

Cust. Id	First Name	Middle	Last Name	DOB	SSN	Address
30391244	Bob	J.	Smyth	4-12-39	563491234	123 Oak St., Eves, IL

Callouts for the table above:

- Smith Smyth (points to Last Name)
- 123 Oak St., Eves, IL 30324 (points to Address)
- Bob Robert Rob (points to First Name)
- April 12, 1939
19390412
04/12/39 (points to DOB)

- “Cleansing” attributes focuses on adjusting data to conform to accepted, accurate values.
 - Typically based on business-oriented logic
 - Adjustment reflects formatting and corrections



Match



Cust. Id	First Name	Middle	Last Name	DOB	SSN	Address
30391-244	Robert	James	Smith	04/12/39	563-49-1234	123 Oak St., Eves, IL 30319



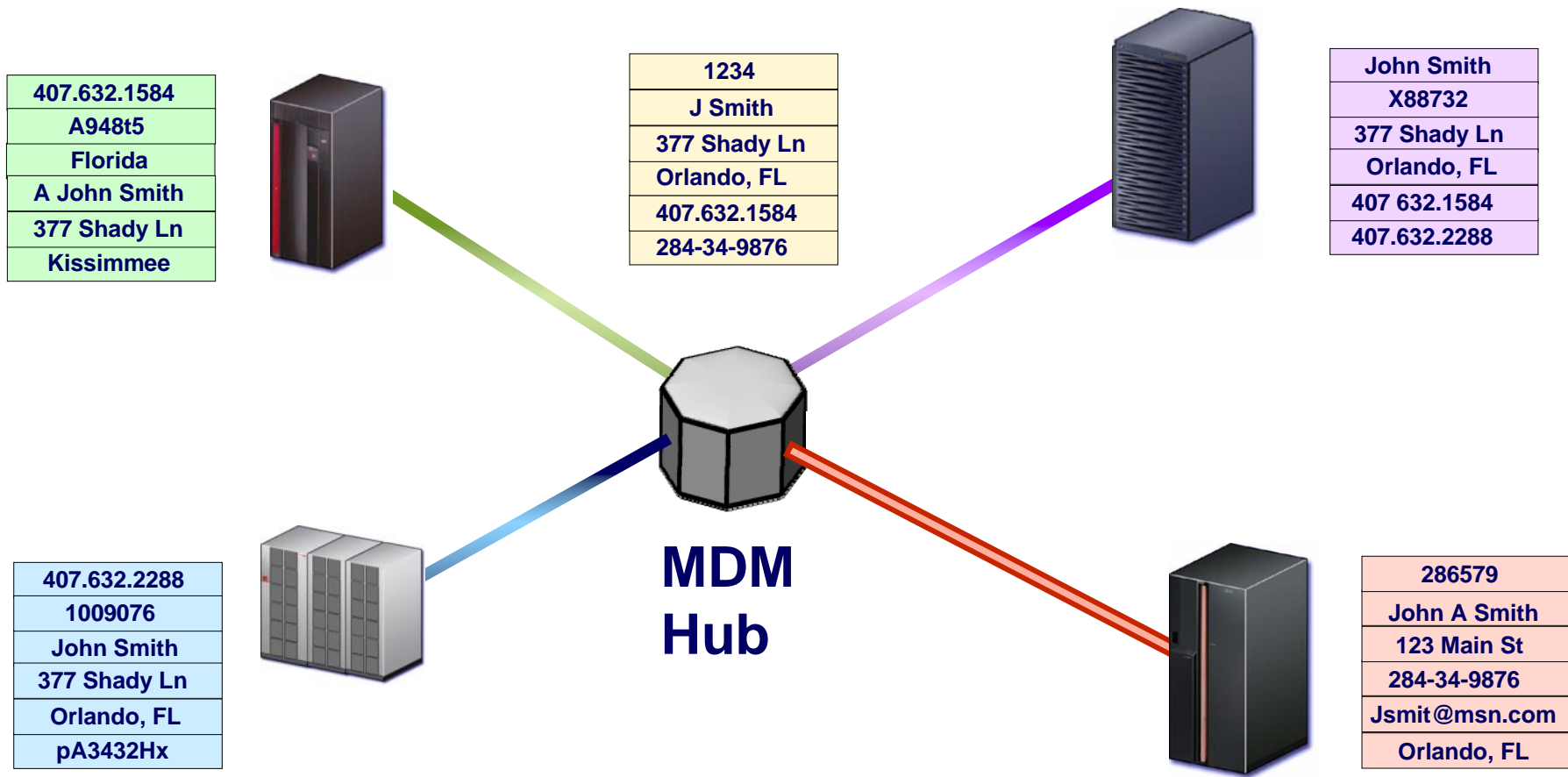
Cust. Id	First Name	Middle	Last Name	DOB	SSN	Address
30391244	Bob	J.	Smyth	4-12-39	563491234	123 Oak Av., Eves, IL

Match Processing The checking of two or more units of data for common characteristics.

- Match determination is typically based on business-oriented rules to determine if different data elements reflect the same meaning
 - Matching is not the same as equivalence (e.g. Bob = Robert)
 - Different data elements (and their associated types) will have different match rules
 - Different subject areas will require different match rules



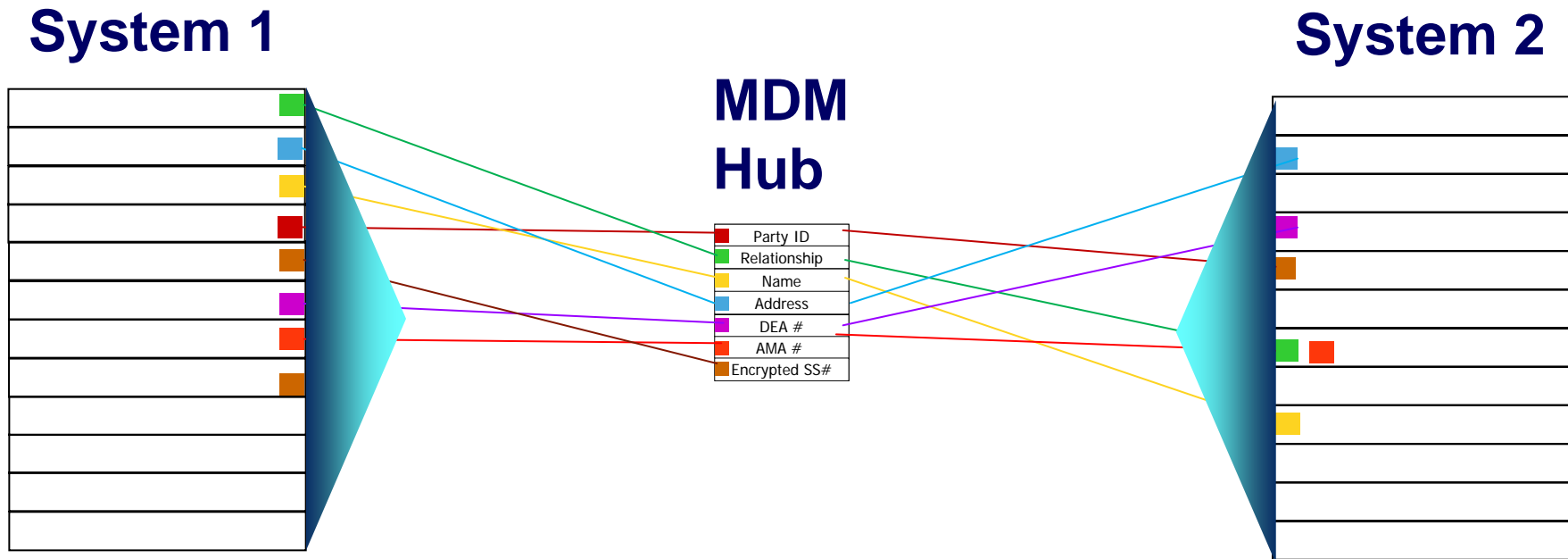
Integration in a Different Way



Integration with MDM doesn't reformat the source



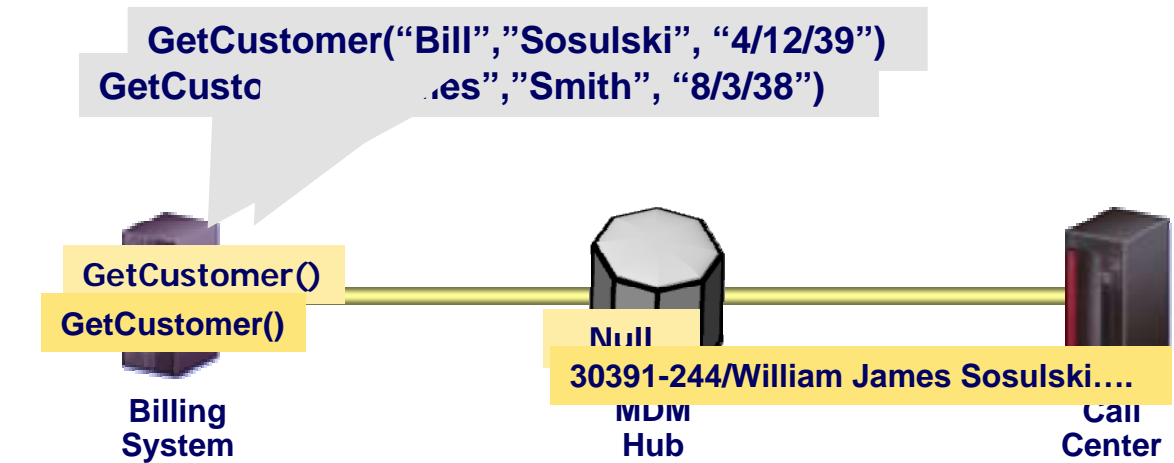
Integration: MDM Hub Integration



Systems can be linked without requiring source changes



Identification



Cust. Id	First Name	Middle	Last Name	DOB	SSN	Address
30391-244	William	James	Sosulski	04/12/39	563-49-1234	123 Oak St., Eves, IL 30319



Data Correction

GetCustomer("James","Smith")

h", ...,42 Windy Dr, Eves, IL, 30345)

CreateCustomer(JamesSmith 30345)

10593 James Smith 30319...

Billing System

MDM Hub

Call Center

Cust. Id	First Name	Middle	Last Name	DOB	SSN	Address
10593	James	Aaron	Smith	06/24/28	527-32-6741	42 Windy Dr., Eves, IL 30319

Create a customer "James Smith, 42 Windy Dr, Eves, IL 30345"

- Check if James Smith exists.
- If not, review data elements
- If acceptable and correctible, add the elements



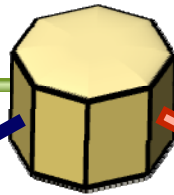
The Operational Functions of MDM

- ☐ Master Index Storage
 - Master record (and identification elements) are stored and linked back to system of origin
- ☐ Centralized Lookup and Matching
 - The hub contains all logical/processing for cleansing, searching, and match
 - Application's don't need to know standard values or match criteria
- ☐ Centralized Integration/Survivorship Process
 - Rules defined for source/attribute precedence
- ☐ Data Standardization and Correction
 - Subject area attributes are standardized and cleansed by hub
 - All groupings and hierarchies will be managed and defined centrally
- ☐ Change Management and Logging
 - All subject identification element changes (deduplication, inadvertent merges, etc.) is centrally processed (either via rules or manual administration)
- ☐ Security and Access Control
 - Logic is managed and executed centrally. CRUD (create,read,update,delete) access varies based upon record and element details



Integration in a Different Way

407.632.1584
A948t5
Florida
A John Smith
377 Shady Ln
Kissimmee



John Smith
X88732
377 Shady Ln
Orlando, FL
407 632.1584
4/12/1939

407.632.2288
1009076
John Smith
377 Shady Ln
Orlando, FL
pA3432Hx



**MDM
Hub**



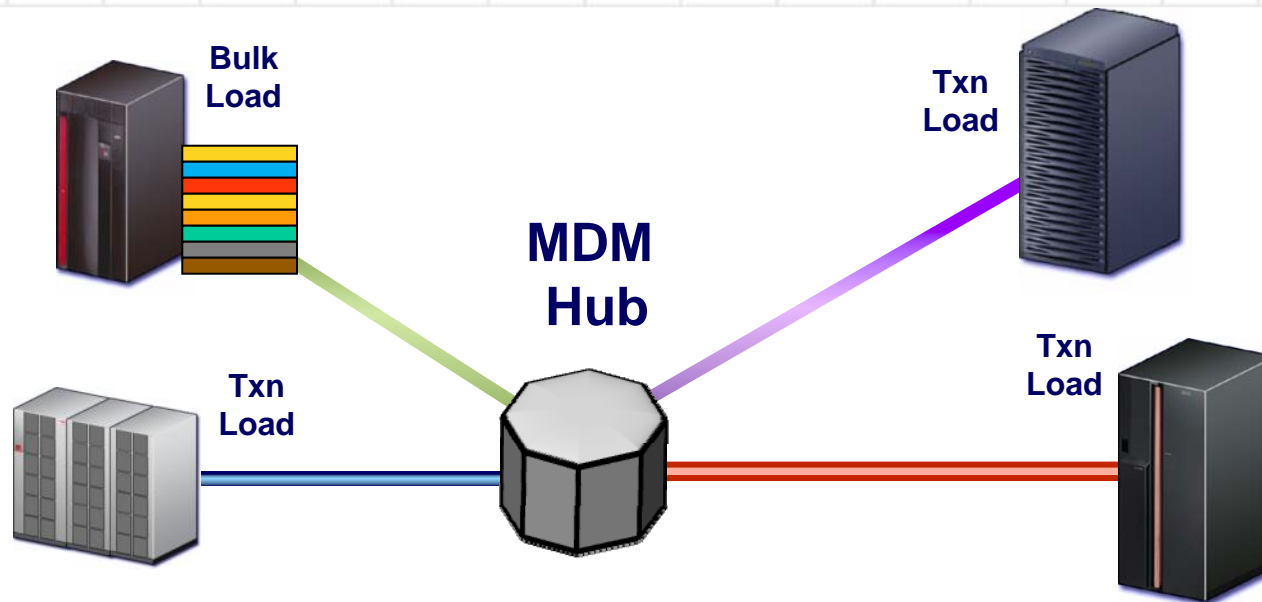
286579
John A Smith
123 Main St
284-34-9876
Jsmith@msn.com
Orlando, FL

1001	a94815	1009076	x88732	286f79	John	Alan	Smith	04/12/1939	4076321584	123 Main St	Orlando	FL	32805
------	--------	---------	--------	--------	------	------	-------	------------	------------	-------------	---------	----	-------

- Stores and maintains links to each system for each unique value contained within the hub.
- The Hub contains the "master id" for John Smith along with the individual system ids from each onboarded system
- The master index provides linkage support to other application systems



Loading the Hub

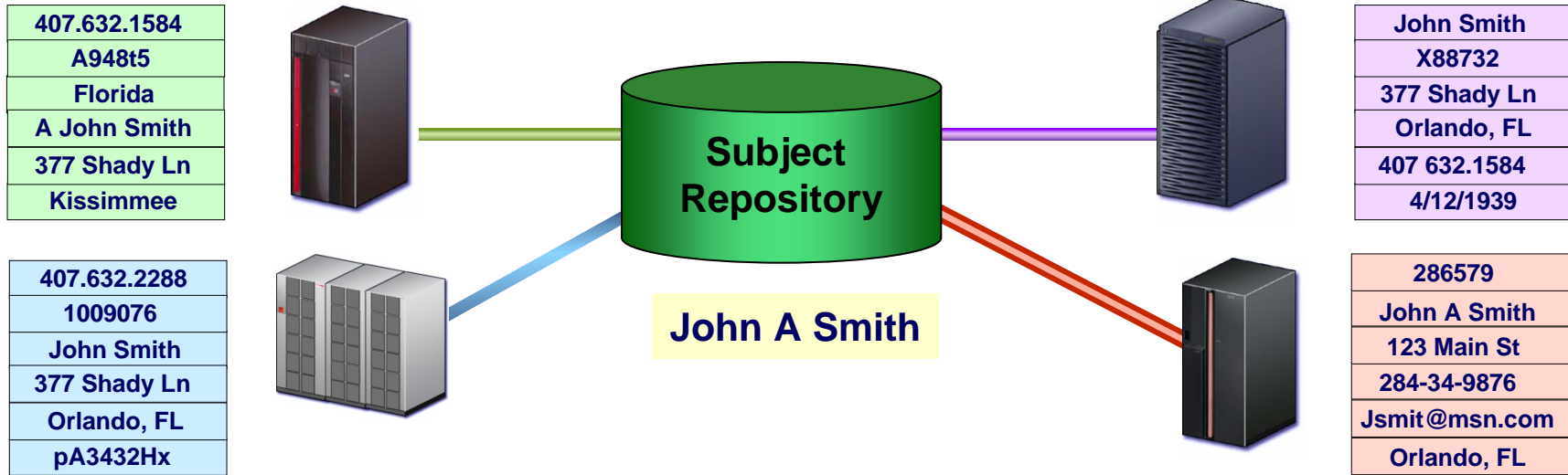


1	Yellow	Green	Red	White	Purple
2	Blue	Green	White	Blue	White
3	Red	Green	Red	Blue	Purple
4	Orange	Green	White	White	White
5	Green	Green	White	Blue	Purple
6	Grey	Green	White	White	White
7	Brown	Green	Red	White	White
8	Purple	White	White	Blue	Purple



Differentiating Master Index and Subject Repository

The Subject Repository

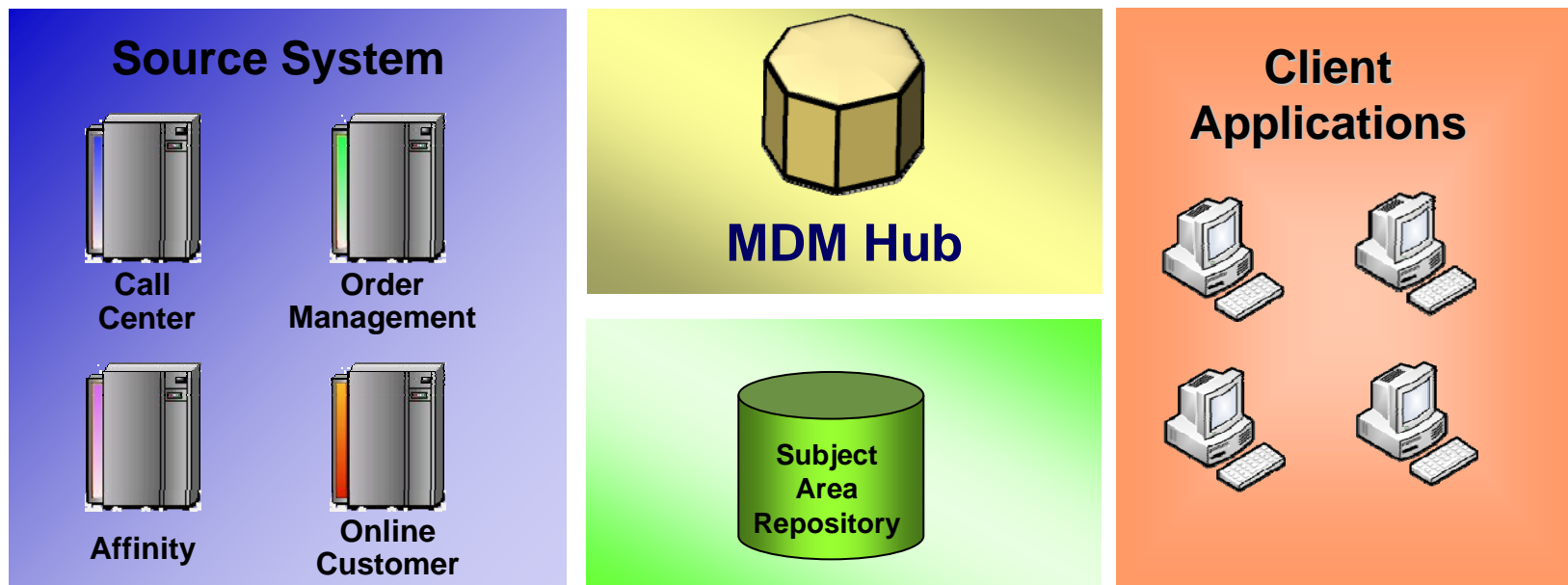


- The Subject Repository acts as a repository for all of the descriptive details
 - Current identification and descriptive information (name, address, spouse's name, home address, business address, cell phone, email, etc.)
 - Historical content (current and prior addresses, past employers, credit history, purchase history, customer contact records, etc.)
- The Subject Repository provides subject content to applications (and users) to support operational and analytical needs



Differentiating Master Index and Subject Repository

MDM Hub: Logical Architecture



- ❑ The source systems are the “system of creation” for subject data
- ❑ The MDM Hub provides an operational integration point linking subject data across multiple sources
- ❑ Client Applications request subject identification services and request data location information
- ❑ The Subject Repository contains descriptive detail and history

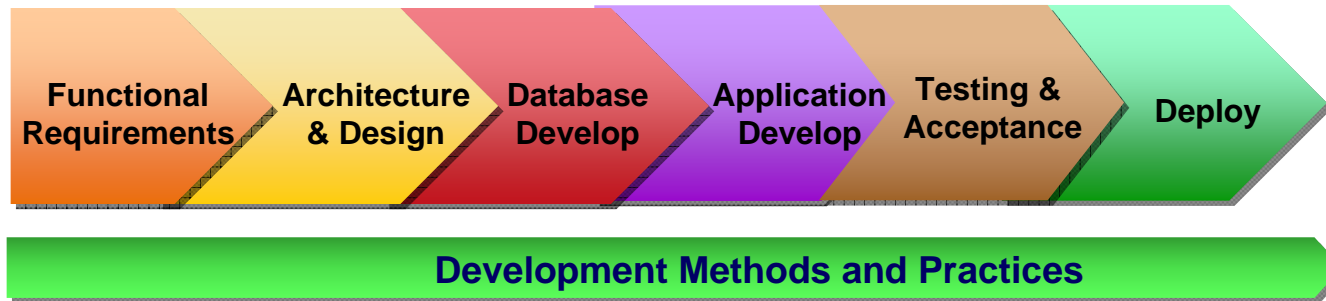


The Challenge with MDM Development

- Most hub projects require both master index and subject repository capabilities
 - Operational (response time) issues dictate that the hub contains descriptive detail (for application support)
 - The “out of the box” hub products expect a blend of identification and descriptive content
- MDM doesn't support query federation
 - It doesn't track and manage every attribute associated with the subject area
 - It can't replace an EII system or a data warehouse
- The challenge is to balance content vs. performance
 - The more non-identification data the hub contains, the slower it will perform
 - The more descriptive data the hub contains, the greater business value it can deliver to the applications



A Simple Implementation Framework



System Development Lifecycle

- Exists to support a repeatable, consistent development process
- Includes participation from all stakeholders (management, development, and end users)
- Most SDLC's are IT-built and focused on application processing
 - Typically lacks data rigor details (data acceptance, data certification)
 - Can easily master data management activities with minor adjustment



An MDM Implementation Framework

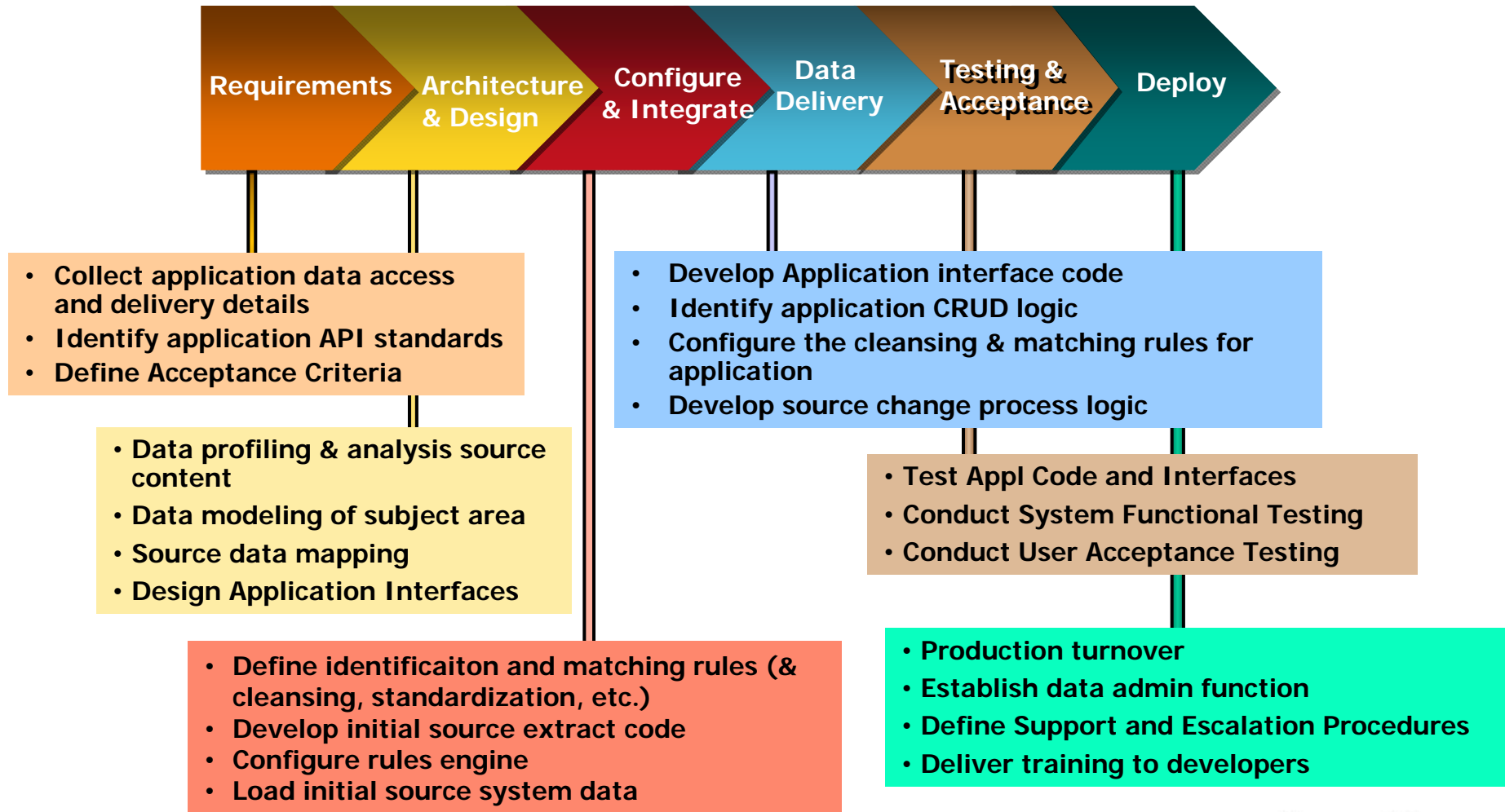


6 Steps for Delivery:

- Requirements
- Architecture & Design
- Configure & Integrate
- Data Delivery (Application Onboarding)
- Testing and Acceptance
- Deployment



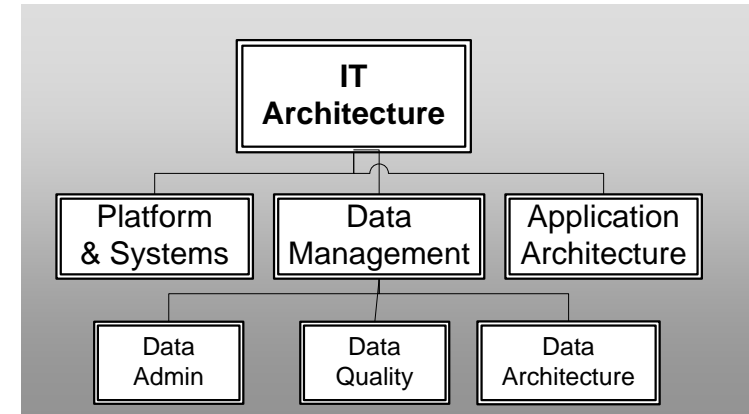
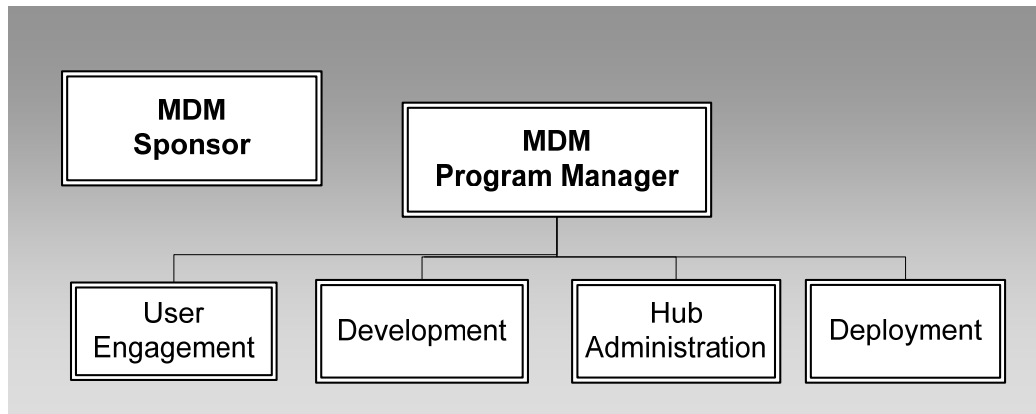
MDM Development Tasks





The Development Team

The MDM Development Team



The MDM Development Team

- Works with stakeholders to identify requirements
- Builds/configures the MDM hub
- Works with application teams to onboard applications
- Owns ongoing hub administration
- Leverages standards from IT architecture

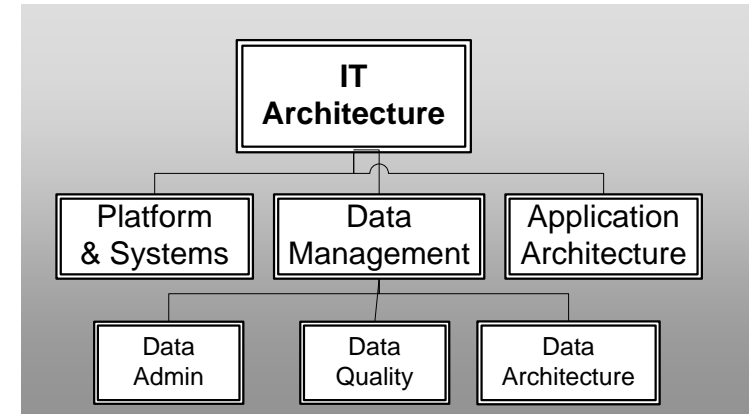
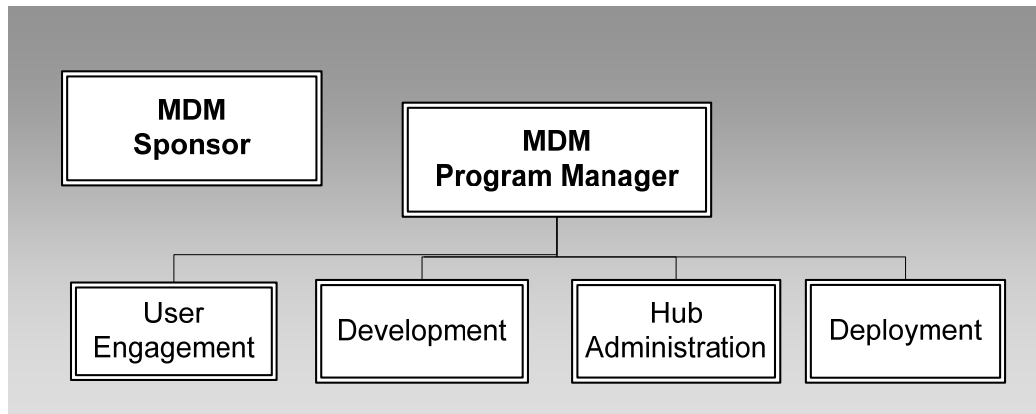
IT Architecture

- Owns application interface standards (SOA)
- Includes Data Management function
- Establishes data standards (value, hygiene, etc.)
- Owns development standards



The Development Team

The MDM Development Team



Program Manager Manages team to deliver technically effective MDM solution

User Engagement Works with business users and application developers to establish business, data, and functional requirements

Development Typically contains 2 core functions: MDM Development (matching, ID, interfaces, etc.) and Data Development (correction, hierarchy, etc.)

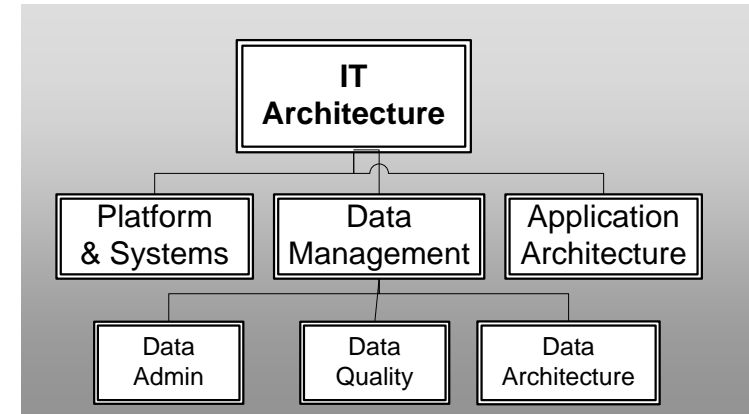
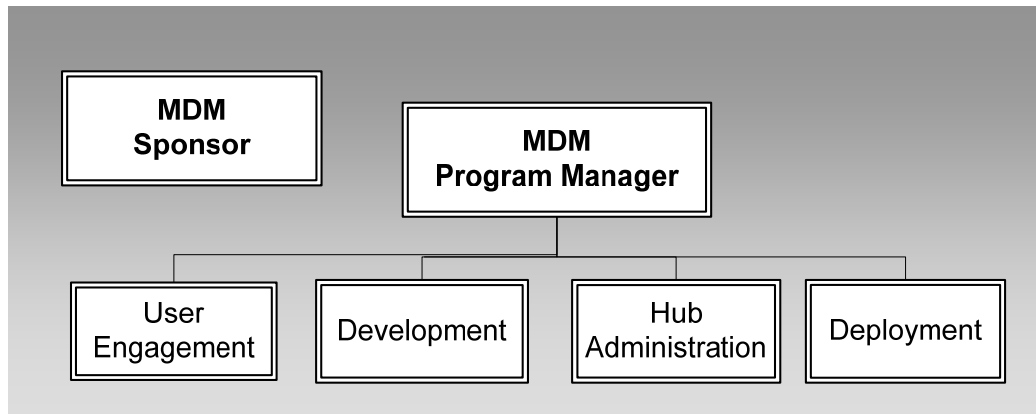
Hub Admin Addresses policy/rule conflicts and violations as well as and hub operational issues.

Deployment Works with application developers to onboard applications. Supports initial release and production support issues



The Development Team

The MDM Development Team



Data Management Team typically within architecture organization. Focused on data standards (sometimes called data architecture)

Data Admin Manages standard business terms and definitions established by business

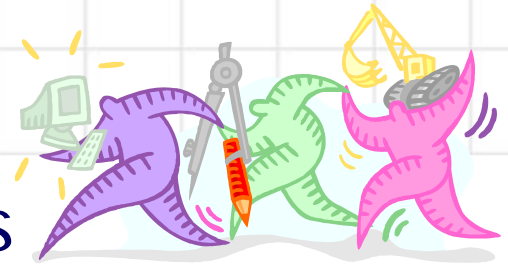
Data Quality Measures conformance of data values to accepted metrics. Publishes production metrics and recommend corrective actions

Application Arch Team within IT architecture organization focused on development standards. Owns SOA standards



The Development Team

Shifting the Paradigm



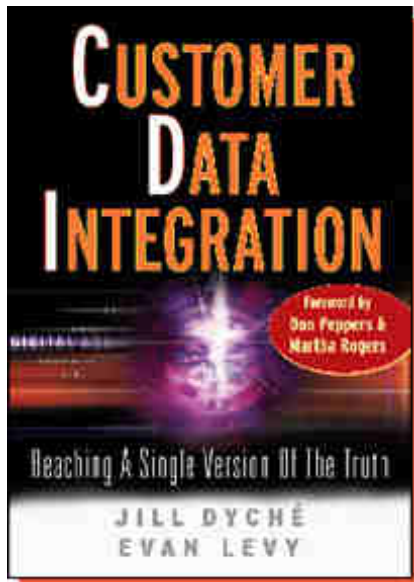
- MDM development success requires traditional development rigor
 - Requirements include functional and data details
 - Success requires a different stakeholder engagement model
- Developer skills will extend beyond traditional development experience
 - Identification and match processing requires business-level data knowledge
 - Data profiling and production data quality measurement will become a constant
- MDM isn't feasible without data management
 - Data standards and data administration is a dependency (not an option)



Thank You!

*For more on MDM, see
Baseline's website!*

See our latest MDM white papers
on SearchDataManagement.com



EvanLevy@baseline-consulting.com

