Enterprise Data and Information Architecture: Why is it important? What could be learned from IBM internal information transformation journey?

Jamshid A. Vayghan, Ph.D.

Senior Technical Staff Member, Chief Enterprise Information Architect
Master Certified IT Architect: The Open Group

April 3, 2008
Before we get started

● Relevance to “Software Process Improvement”
● Doing the right things VS doing the things right
● IBM as a customer not a vendor
Topics

● Why is Enterprise Information Architecture important?
  – Business drivers
  – Enterprise Data and Information Challenges

● IBM’s internal information transformation journey
  – Business drivers
  – Enterprise Data Strategy
  – Enterprise Data and Information Architecture Program
  – Accomplishments & future steps

● Relevance IBM’s experience and lessons learned to Your Company
  – Assist you to define its target environment
  – Define practical steps that you can take now

● Questions

● Next step?
Defining the Challenge
Information Quality is Critical to Business Integrity

Inconsistent customer information can cause reduced customer satisfaction, decreased revenue, and hinder relationships.

Jane Smith...

- Is a high value web customer
- Yet... to the call center she is completely unknown
  - Poor customer service
  - High cost of service due to “multi call resolution”
- Inability to act on customer insight leads to missed sales opportunities

“77% of 144 CIOs surveyed identified single view of customer as the single most important benefit of MDM”
Information Quality is Critical to Business Integrity

Inconsistent information across channels can cause inconsistent shopping experience and reduce customer satisfaction

- **Product SKU 11111**
- **Product short description:** Outdoor gas grill
- **Features:** auto-shut off, rubber wheels, rotisserie, sound system
- **Price:** Regular $700
- **Price:** Sale $550 Expiration Sep. 30
- **Warranty:** 1 year
- **Return Policy:** 30 days

- **Web Site**
  - Product: Outdoor Grill
  - Price: $550 *Special*
  - X-Sell / Up-Sell Items: 5432, 4355
  - Warranty: 1-year
  - Features: auto shut-off, rubber wheels, rotisserie
  - Return Policy: 30 days

- **Contact Center**
  - Product: Gas Grill
  - Price: $550 *Special*
  - Warranty: 1-year
  - Features: sound system, rotisserie
  - Return Policy: 30 days

- **Store**
  - Product: Gas Grill
  - Price: $700
  - X-Sell / Up-Sell Items: 6400, 9345
  - Warranty: 1-year
  - Features: sound system, rotisserie
  - Return Policy: 30 days

79% of Retailers and 61% of CPG manufacturers rank “item management” as their top priority
A Complementary View of the Enterprise

An view of the enterprise in terms of core data entities and associated interdependencies with processes and policies is critical to managing integrity.

Understanding the business services of the enterprise

Understanding the core data entities of the enterprise

**Imperative Model**

**Declarative Model**

**Processes**

**Core Data Entities**

**Business Policies**
- GAAP
- Privacy
- SOX
- HIPAA

**Governing Policies**

**Customer**
- attribute

**Employee**
- attribute

**Product**
- attribute

**Service**
- attribute
Managing Business Integrity

Business integrity links fundamental issues of high-level business policy to practical, operational matters of processes, data and IT

Business integrity is:

- Having the right information and basis for doing things – **Information Integrity**
  - Information integrity ensures correct and consistent information with appropriate access control
  - Needs information management that properly and completely represents core business entities

- Doing things in the right way – **Process Integrity**
  - Process integrity ensures actions follow intended paths and applicable policy
  - Needs regular and complete monitoring and management of appropriate metrics

- A basis for defining “doing the right thing” – **Policy Integrity**
  - Policies include business practices, government regulations, and ethical policies
  - Needs well-defined policies that can be translated to be measurable and actionable
Enterprise Information Architecture: Business Drivers

**Business Drivers**
- Globally Distributed Operations, Shortage of Skilled Resources, Market Conditions, Competitive Threats, Increasing and diverse Government Regulations, Merger and Acquisitions, Selling business

**Business Integrity**
1. **Policy Integrity**: Doing the right thing
2. **Process Integrity**: Doing things in the right way
3. **Information Integrity**: Having the right basis for doing things

- Increase agility, responsiveness and flexibility
- Reduce and contain cost to rapidly create new business solutions
- Improve partner interoperability
- Simplify enterprise end-to-end integration to adopt new business models
- Legacy: Modernize and extend ROI
- Efficiently integrate 3rd party products

Disconnect between Structured and Unstructured data
Business Integrity Management: Evolution

- Today information, processes and policies are typically inconsistent and incomplete
- Emerging solutions address consistencies and management of policies, processes and core data entities independently
- An integrated and automated approach to policy, process and core data entity management will emerge to ensure business integrity is maintained at all levels of the enterprise
Defining the Challenge:

1. Lack of a holistic approach to enterprise information management which is essential to establishment of business integrity in an enterprise.

2. Much of research on software architecture excludes data as an important architectural element.
The Elephant Metaphor of Reality

Observations of 6 blind men who touched an elephant:
1. "Hey, the elephant is a pillar," said the first man who touched his leg.
2. "Oh, no! it is like a rope," said the second man who touched the tail.
3. "Oh, no! it is like a thick branch of a tree," said the third man who touched the trunk of the elephant.
4. "It is like a big hand fan" said the fourth man who touched the ear of the elephant.
5. "It is like a huge wall," said the fifth man who touched the belly of the elephant.
6. "It is like a solid pipe," Said the sixth man who touched the tusk of the elephant.

Their observations were different because each one of them touched the different part of the elephant. So, actually the elephant has all those features but none of those features describe the elephant.

What does this story have to do with describing enterprise data challenges?
Enterprise Data Challenges: Observation 1

Information connects people to their work but there are challenges

- **Information overload** – World has just hit the 161 billion gigabytes of data!
  - IDC found that the largest chunk of the 161 billion gigabytes of data the world created in 2006 is e-mail
  - IDC found that person-to-person communications alone accounted for six exabytes -- or six billion gigabytes -- of data in 2006, and growth continues to be rapid.

- **Fragmented user experiences**
  - Understanding what the company and its employees are paying attention to and responding to are critical
  - Different brands, different portals, different businesses (looking for our attention)
  - Web is an everyday tool – not a marvel of innovation so we differentiate ourselves by providing both better content and better solutions to users' problems

- **Grass roots collaboration actually increase the amount of information (feeds, blogs, wikis)**
  - Finding information is critical to understanding the current state and making decisions about the future
Enterprise Data Challenges: Observation 2

- There are multiple web content management tools:
  - Each group has its own content management ideas and their own web sites:
    - Tools from a variety of vendors
    - Home grown CM Tools
    - Hand Coded HTML Pages
    - Mergers and acquisitions increasingly fragmented the content and introduced new CM systems

- Issues with multiple CM systems:
  - No content sharing
  - Difficulty with cross-branding
  - Inconsistent navigation and content
  - Many broken links
  - Overlap in spending; Associated cost of duplicate systems, including vendor products
  - Reduced operational efficiency
  - Redundant effort and content
  - Low customer satisfaction
Enterprise Data Challenges: Observation 3

- Tight coupling of applications to data sources (due to embedded data access code in application logic) results in inflexibility to evolve data model and to use new data sources with changing business needs

- Lack of data quality
  - Tied to a fixed data source that may not provide quality data anymore
  - Inconsistent use of data validation and cleansing rules across applications

- Lack of reuse of information processing code

- Explicit audit and control of access to data (Data Governance)

- Lack of ease-of-development of information processing tasks
Enterprise Data Challenges: Observation 4

● Lack of information standards
  – Different formats & structures across different systems

● Data surprises in individual fields
  – Data misplaced in the database

● Information buried in free-form fields

● Data myopia
  – Lack of consistent identifiers inhibit a single view

● The redundancy nightmare
  – Duplicate records with a lack of standards
Enterprise Data Challenges: Observation 5

- Need to provide more accurate inventory availability information to their dealers in order to better optimize inventory.

- Required the integration of external market and parts information with internal order forecasts and shipments in their Supply Chain Management system.
Enterprise Data Challenges: Observation 6

Order Processing Application
- Customer Lookup - 1

Invoicing Application
- Customer Lookup - 2

CRM Application
- Customer Lookup - 3

How to replace redundant services with a reusable and equivalent functionality?
Enterprise Data Challenges: Observation 7

Without managed data standards and a corporate data dictionary..

- Each application defines their own data names, structures, relationships, …
- The meaning of the data will vary.. based on the developer’s definition..
- From a messaging perspective, each interfacing application will need to go through a process of understanding the data definitions for the interfacing program, and then develop a “mapping program” to translate to and from the interfacing application.
- There can be up to n*(n-1) message interfaces between n applications.
- Making changes are difficult, expensive, and might have unpredictable side effects.
IBM’s Internal Information Transformation Journey
IBM Internal transformation simplified infrastructure and governance

<table>
<thead>
<tr>
<th></th>
<th>1992</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIOs</td>
<td>128</td>
<td>1</td>
</tr>
<tr>
<td>Host Data Centers</td>
<td>155</td>
<td>6</td>
</tr>
<tr>
<td>Web Hosting Centers</td>
<td>80</td>
<td>6</td>
</tr>
<tr>
<td>Network</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>Applications</td>
<td>16,000</td>
<td>4,190</td>
</tr>
</tbody>
</table>
Business and Technical Drivers for IBM Internal Information transformation Journey

- For some environments, data and information architecture were not explicitly defined. When existed, it was a reflection of organizational structure.
- Lack of well defined trusted data sources.
- Missing clarity on ownership of information.
- Poor quality of information.
- Inconsistent information across applications.
- Existing tower solutions that result in end-to-end solutions which are complex, costly and redundant. Lack of end-to-end visibility of information.
- Excessive reporting solutions.
- Poor governance of information.
- Difficult to reuse and share information assets.
- Too many data warehouses and data marts.
- Most of IT budget spent on operations not transformation.
- Shortage of critical data and information skills.
- Difficulty to make changes in business operations to support emerging business models.
Enterprise Architecture: Alignment of business strategy with IT strategy to realize business integrity
Enterprise Business Information Center of Excellence…
EVOLUTION

2002: Individual IT Domains Everywhere

2003: Consolidated Information Domains Starts

2004: Integrated Information Management

2005 – Present COE
Matured in Enterprise Integration, Enterprise Data Architecture, Enterprise Information Management and Governance

2005: Enterprise Data Strategy Published & Approved by all LOBS
EBI COE evolution tied to IBM’s Information On Demand strategy

Innovative Use of Information Drives New Business Value

- Information On Demand
  - Information as a Competitive Differentiator
  - Information to Enable Innovation
  - Information as a Strategic Asset
  - Information In Business Context
  - Basic Information Interaction
  - Focus on Data And Reporting

Operations & Reporting
- Data to Run the Business
- Information to Manage the Business

Business Value

Governance
Technology
Metrics
Maturity of Information Use
We learned that …

If your enterprise looks like this …

You need an enterprise information strategy to drive definition and realization of a solution that breaks information silos and opens up enterprise data for reuse while maintaining business integrity. Elements of that solution are:

- **Business: Governance and Process**
  - Integrated governance program
  - Integrated business integrity

- **Organization: People and Culture**
  - Data and information professionals
  - Eliminate silo mentality

- **Information Technology**
  - Enterprise Information Management & Architecture
  - Unification of structured & unstructured data
  - Information integration
  - Information discovery
  - Integrated Enterprise Metadata Management
We also learned that enterprise information management is a competency!

- **Governance is key**
  - Top down and bottom up
  - Enterprise integrated governance
  - Open standards

- **End-to-end data/information management is important in large enterprise projects**
  - Lacking in complex enterprise projects
  - Tight Coupling between data, application, process
  - Importance of Explicit Enterprise Metadata

- **Invest in critical data skills**
  - Data Architects, Data Analysts
  - Operational Competencies
  - Process Expertise

- **Integrate structured and unstructured data**
  - XML and Hybrid (XML & Relational) databases are key enabling technologies.

- **Information as a service**

- **Invest in critical data skills**

- **Three myths of Information quality**
  1. Data quality is an IT issue
  2. We cleaned the data last year, so it is good this year
  3. Data quality means the same thing to everyone

- **Try Enterprise Information Mash-ups ➔ “Data is Inside”**
  - Use data services to aggregate and integrate data to create data consumables for Web 2.0 applications.

- **We are not alone!**
  - Customers have the same challenges!
We have accomplished a lot, but journey continues! Here are examples of accomplishments:

- **Enterprise Data Strategy approved by all LOBs.**
- **Enterprise Data Architecture Program in operation.**
  - Project reviews have prevented many redundant initiatives and helped compliance with business and technical data standards.
  - Current landscape is captured. Information assets are profiled.
  - Data Warehousing target architecture defined.
  - Trusted data services and associated processes are defined and being implemented.
  - Roadmap for critical areas have been defined and implemented.
  - Enterprise integration messaging specification based on Open Application group being implemented for B2B and A2A within IBM.
  - Enterprise vocabulary defined.
  - SAP systems are being consolidated.
- **Significant improvements in information quality.**
- **Information discovery capabilities being deployed.**
  - Much easier to find what you need.
  - Subscribe to information that you need.
- **Significant accomplishments in consolidation content repositories.**
- **Data governance**
  - Data stewardship program is implemented.
  - Data stewards for critical subject areas identified.
  - Enterprise owners for product, customer and worker data named.
  - Business data standards are defined by data stewards. They are stored in an enterprise repository.
- **Consolidation of content management tools in progress.**
- **Enterprise information mash-ups has started to open up the enterprise data for reuse.**
- **Many legacy systems were sunset.**
- **Enterprise wide reporting system for sales opportunities deployed to all level of management.**
- **Enterprise analytic applications being developed and deployed.**
- **Cost reductions in operation, maintenance, and development of information assets.**
Enterprise Information Management in IBM: Enterprise Business Information Center of Excellence (EBI CoE)
Our lessons learned, Your action plan!
Some Strategic Recommendations

1. Define an enterprise data strategy to enable business integrity, unification of structured and unstructured data, plus data archiving.

2. Define an enterprise information maturity model to measure the progress.

3. Complete your enterprise information architecture program. TOGAF and Zachman frameworks do not explicitly define all necessary steps.

4. Define and deploy a comprehensive information discovery program.

5. Define and implement your integrated data governance program through a data stewardship program.

6. Create a common framework for information delivery for analytic applications.

7. Develop a strategy to extend your existing data warehouses instead of creating new ones.

8. In addition to the creation of the enterprise information architecture, the enterprise information architecture team could also be responsible for implementation of the architecture.

9. Use SOA and Event driven architecture principles to extend your enterprise information architecture. Deploy trusted data services that could aggregate and integrate information without creating new copies of the data.
Data Governance in Practice: IM Management System

**IM Review Purpose**
1. Direction and standards for the IM Disciplines
2. Ensure consistent application of IM disciplines in projects
3. Enable integration & reuse of Architecture, Metadata and the data itself

**Result:**
- Go, Contingent
- Go, Redirect, No Go

**Architecture Council Review process**

**Business Planning**
- Initiative Definition

**BTOP**
- Concept
- Plan
- Qualify

**Operations Mgmt**

**Project Lifecycle**

**Approval Process**

- Information Architecture
- Data Req.
- Migration
- Information Quality
- Metadata
- Business Data Stnd
- Data Stewardship
- Security & Privacy

**Project**

**Executive Steering Committee (IPMT)**
Hybrid Architecture Style: Loose Coupling, Agility, Reuse

● **SOA**
  - Provide an effective foundation to integrate application, process, and data by creating loosely coupled components.
  - Enable delivery of information as a service by providing the following:
    - One to one communication.
    - Consumer-based triggers.
    - Synchronous operations.
    - Standard specification for information exchange.

● **Enterprise Information Architecture (EIA)**
  - Knowing where to find trusted sources of information so trusted data services can be established.
  - Metadata to consolidate explicit data about location, structure, context and usage of data. This data would be essential to creation of metadata services that are needed to establish discoverable trusted data services.
  - Data standards to establish a forum for resolving differences in the meaning of data.
  - Defining and managing a proactive data quality program.
  - Integrate and federate data from separate data sources to create a common view of data when needed.
  - Establish a mechanism for transforming data into meaningful business information.
  - Establish an integrated governance program for data and SOA.

● **Event Driven Architecture**
  - Provide event-based triggers.
  - Decouple interactions between components.
  - Many to Many communication (e.g., Publication/Subscription).
  - Asynchronous operations.
  - Sense and response capability.
Some Tactical Recommendations that Could be Taken Now:

- **Improve Agility:**
  - Try enterprise information mash-ups for some key business challenges
  - Develop REST Web services to open up data. Eliminate unnecessary access control.
  - Use Web 2.0 technologies including Google Map to rapidly develop solutions.

- **Accelerate Information Sharing**
  - Implement social networking (e.g. Wiki, Blogging).
  - Enable information discovery
  - Social tagging
  - Information aggregation

- **Prevent creation of new redundant information assets when possible**
  - Define a process to review the rational for creation of new copies of existing data and creation of new data warehouses
  - Use the technology to avoid creation of redundant assets
  - Remember that duplicate copies of the same asset can cause security and quality challenges while increasing operational costs.

- **Enable Data Sharing:**
  - Open up critical enterprise data for reuse by creating REST web services to access the data
  - Identify a few critical business problems that could benefit from using REST web services to create data services.
Top 10 lessons learned

10. Create a “sense of urgency” that the company can rally around

9. Create a revisionist history – you’ll be surprised at how far you’ve come

8. While you’re creating, define short-term projects with near-term results

7. Review business processes to see if changes are needed before you deploy technology

6. Technology enables and hastens transformation

5. Set your milestones and metrics with an end-to-end lifecycle view

4. Sunset legacy systems/applications/tools as new ones are deployed

3. Can NOT over-emphasize the importance of culture and fostering innovation

2. Transform constantly or risk extinction – there is no other option

1. Always, always, always listen to your customers and make sure you have the right perspective
Questions?

Comments?

Next Step?

Thank you!

Jamshid A. Vayghan, Ph.D.

vayghan@us.ibm.com