Building Actionable Data Governance through Data Models & Metadata

Donna Burbank
Global Data Strategy Ltd.

Data Modeling Zone
October 16th, 2017
Donna Burbank

Donna is a recognised industry expert in information management with over 20 years of experience in data strategy, information management, data modeling, metadata management, and enterprise architecture. Her background is multi-faceted across consulting, product development, product management, brand strategy, marketing, and business leadership.

She is currently the Managing Director at Global Data Strategy, Ltd., an international information management consulting company that specializes in the alignment of business drivers with data-centric technology. In past roles, she has served in key brand strategy and product management roles at CA Technologies and Embarcadero Technologies for several of the leading data management products in the market.

As an active contributor to the data management community, she is a long time DAMA International member, Past President and Advisor to the DAMA Rocky Mountain chapter, and was recently awarded the Excellence in Data Management Award from DAMA International in 2016. She was on the review committee for the Object Management Group’s Information Management Metamodel (IMM) and the Business Process Modeling Notation (BPMN). Donna is also an analyst at the Boulder BI Train Trust (BBBT) where she provides advices and gains insight on the latest BI and Analytics software in the market.

She has worked with dozens of Fortune 500 companies worldwide in the Americas, Europe, Asia, and Africa and speaks regularly at industry conferences. She has co-authored two books: *Data Modeling for the Business* and *Data Modeling Made Simple with ERwin Data Modeler* and is a regular contributor to industry publications. She can be reached at donna.burbank@globaldatastrategy.com

Donna is based in Boulder, Colorado, USA.

Follow on Twitter @donnaburbank
• Data Governance is often referred to as the people, processes, and policies around data and information, and these aspects are critical to the success of any Data Governance implementation.

• But just as critical is the technical infrastructure that supports the diverse data environments that run the business.

• Data Models can be the critical link between business definitions and rules and the technical data systems that support them. Without the valuable Metadata these models provide, Data Governance often lacks the “teeth” to be applied in operational and reporting systems.

• This session how Data Models and Metadata-driven Data Governance can be applied in your organization.
Agenda
What we’ll cover today

• Definitions & Overview
• How data modeling & metadata support data governance
• Making data governance “actionable” through data model metadata
• Summary & questions
Survey: Data Modeling

DMZ Attendees

Who is currently using a data model in their organization?

Industry Survey Results*

Over 96% of DATAVERSITY respondents are using a data model*.

* Sneak preview of DATAVERSITY “Trends in Data Architecture” research paper to be published October/November 2017.
Data Modeling is part of a Wider Data Strategy

A Successful Data Strategy links Business Goals with Technology Solutions

“Top-Down” alignment with business priorities

Managing the people, process, policies & culture around data

Leveraging & managing data for strategic advantage

Coordinating & integrating disparate data sources

“Bottom-Up” management & inventory of data sources

Global Data Strategy, Ltd. 2017
Survey: Data Governance

Who is currently has a data governance effort underway?

DMZ Attendees

Industry Survey Results*

Approx 60% of DATAVERSITY cited Data Governance as a key driver for their data management initiatives*.

* Sneak preview of DATAVERSITY “Trends in Data Architecture” research paper to be published October/November 2017.
Data Governance – A Basic Framework

Vision & Strategy

- Business Goals & Objectives
- Data Issues & Challenges

Tools & Technology

- Organization & People
- Process & Workflows
- Data Management & Measures
- Culture & Communication
# Building the Data Governance Framework

<table>
<thead>
<tr>
<th>Vision &amp; Strategy</th>
<th>Organization &amp; People</th>
<th>Processes &amp; Workflows</th>
<th>Data Management &amp; Measures</th>
<th>Culture &amp; Communications</th>
<th>Tools &amp; Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a clear understanding of the strategic goals of your organization &amp; the need for enterprise data governance?</td>
<td>Who are the key data stakeholders within and outside your organization?</td>
<td>Do business process design and operations management take data needs into account?</td>
<td>Has key data been identified, defined and analyzed?</td>
<td>Has the importance of data been communicated across the organization? Is there a data communications plan?</td>
<td>Is there a coherent data architecture in place to define and guide how data is captured, processed, stored and used?</td>
</tr>
<tr>
<td>How does your organization rely on data – now and in the future?</td>
<td>Who are the primary data producers, consumers &amp; modifiers?</td>
<td>Are there any specific data management / improvement processes in place?</td>
<td>Have data models been built – conceptual / logical / physical?</td>
<td>Is the value of good data management understood and championed by senior managers?</td>
<td>What primary IT systems and platforms are used to store and process key data?</td>
</tr>
<tr>
<td>What impact are data problems currently having on your organization?</td>
<td>Are individuals formally accountable for data ownership?</td>
<td>Are there issue and workflow management processes to address data problems?</td>
<td>Has the relationship between business processes and data been mapped?</td>
<td>Do all employees and third parties receive data awareness and improvement education and training?</td>
<td>Do design gateways exist to ensure data needs are taken into account in new &amp; modified platforms?</td>
</tr>
<tr>
<td>Do you have a data governance policy?</td>
<td>Are employees trained in good data management practices?</td>
<td>Has there been any analysis of the efficiency and effectiveness of how data is managed within operational business processes?</td>
<td>Are data shortcomings known, measured &amp; recorded?</td>
<td>Are there communication channels for communicating best practice in data management?</td>
<td>What specialist data management tools are currently in use?</td>
</tr>
<tr>
<td>What are the overall expected benefits of better data governance?</td>
<td>Are there any channels through which data shortcomings can be highlighted and investigated?</td>
<td>How does the business and IT interact to manage data improvement?</td>
<td>Are there formal standards &amp; rules specifying how data should be managed and improved?</td>
<td>Are there internal success stories that could be used to promote better data management across the organization?</td>
<td>What metadata is captured and stored?</td>
</tr>
</tbody>
</table>

Data Models support many areas of the Data Governance Framework.
What is a Data Model?

Translates Business Rules & Definitions... ...to the Technical Data Systems & Structures that Support Them
What is a Data Model?

Translates Regulations, Policies & Procedures...

...to the Technical Data Systems & Structures that Support Them

Policy

“All Personally Identifiable Information (PII) must be anonymized for the purpose of information sharing between departments."
## Metadata is the “Who, What, Where, Why, When & How” of Data

<table>
<thead>
<tr>
<th>Who created this data?</th>
<th>What is the business definition of this data element?</th>
<th>Where is this data stored?</th>
<th>Why are we storing this data?</th>
<th>When was this data created?</th>
<th>How is this data formatted? (character, numeric, etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who is the Steward of this data?</td>
<td>What are the business rules for this data?</td>
<td>Where did this data come from?</td>
<td>What is its usage &amp; purpose?</td>
<td>When was this data last updated?</td>
<td>How many databases or data sources store this data?</td>
</tr>
<tr>
<td>Who is using this data?</td>
<td>What is the security level or privacy level of this data?</td>
<td>Where is this data used &amp; shared?</td>
<td>What are the business drivers for using this data?</td>
<td>How long should it be stored?</td>
<td></td>
</tr>
<tr>
<td>Who “owns” this data?</td>
<td>What is the abbreviation or acronym for this data element?</td>
<td>Where is the backup for this data?</td>
<td></td>
<td>When does it need to be purged/deleted?</td>
<td></td>
</tr>
<tr>
<td>Who is regulating or auditing this data?</td>
<td>What are the technical naming standards for database implementation?</td>
<td>Are there regional privacy or security policies that regulate this data?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Global Data Strategy, Ltd. 2017
Metadata is Hotter than ever
A Growing Trend

In a recent DATAVERSITY survey, over 80% of respondents stated that: Metadata is as important, if not more important, than in the past.
Data Governance is a Key Driver for Metadata Usage

A Key Use Case for Metadata Management

In a recent DATAVERSITY survey, over 60% of respondents stated that: Data Governance is a key driver for their use of Metadata.
Technical & Business Metadata

- **Technical Metadata** describes the structure, format, and rules for storing data.
- **Business Metadata** describes the business definitions, rules, and context for data.
- **Data** represents actual instances (e.g. John Smith)

### Technical Metadata

```sql
CREATE TABLE EMPLOYEE (  
  employee_id       INTEGER NOT NULL,  
  department_id     INTEGER NOT NULL,  
  employee_fname    VARCHAR(50) NULL,  
  employee_lname    VARCHAR(50) NULL,  
  employee_ssn      CHAR(9) NULL);
```

```sql
CREATE TABLE CUSTOMER (  
  customer_id       INTEGER NOT NULL,  
  customer_name     VARCHAR(50) NULL,  
  customer_address  VARCHAR(150) NULL,  
  customer_city     VARCHAR(50) NULL,  
  customer_state    CHAR(2) NULL,  
  customer_zip      CHAR(9) NULL);
```

### Business Metadata

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employee</td>
<td>An employee is an individual who currently works for the organization or who has been recently employed within the past 6 months.</td>
</tr>
<tr>
<td>Customer</td>
<td>A customer is a person or organization who has purchased from the organization within the past 2 years and has an active loyalty card or maintenance contract.</td>
</tr>
</tbody>
</table>
Business vs. Technical Metadata

- The following are examples of types of business & technical metadata.

<table>
<thead>
<tr>
<th>Business Metadata</th>
<th>Technical Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitions &amp; Glossary</td>
<td>Column structure of a database table</td>
</tr>
<tr>
<td>Data Steward</td>
<td>Data Type &amp; Length (e.g. VARCHAR(20))</td>
</tr>
<tr>
<td>Organization</td>
<td>Domains</td>
</tr>
<tr>
<td>Privacy Level</td>
<td>Standard abbreviations (e.g. CUSTOMER -&gt; CUST)</td>
</tr>
<tr>
<td>Security Level</td>
<td>Nullability</td>
</tr>
<tr>
<td>Acronyms &amp; Abbreviations</td>
<td>Keys (primary, foreign, alternate, etc.)</td>
</tr>
<tr>
<td>Business Rules</td>
<td>Validation Rules</td>
</tr>
<tr>
<td>Etc.</td>
<td>Data Movement Rules</td>
</tr>
<tr>
<td></td>
<td>Permissions</td>
</tr>
<tr>
<td></td>
<td>Etc.</td>
</tr>
</tbody>
</table>
Human Metadata
Avoid the dreaded “I just know”

• Much business metadata and the history of the business exists in employee’s heads.
• It is important to capture this metadata in an electronic format for sharing with others.
• Avoid the dreaded “I just know”
Better Definitions Drive Better Communication

• Wouldn’t it be helpful if we did this in daily life, too?
• i.e. “Let’s go on a family vacation!”

<table>
<thead>
<tr>
<th>Person</th>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father</td>
<td>Vacation</td>
<td>An opportunity to take the time to achieve new goals</td>
</tr>
<tr>
<td>Mother</td>
<td>Vacation</td>
<td>Time to relax and read a book</td>
</tr>
<tr>
<td>Jane</td>
<td>Vacation</td>
<td>A chance to get outside and exercise</td>
</tr>
<tr>
<td>Bobby</td>
<td>Vacation</td>
<td>Time to be with friends</td>
</tr>
<tr>
<td>Donna</td>
<td>Vacation</td>
<td>More time to build data models</td>
</tr>
<tr>
<td>Ian</td>
<td>Holiday</td>
<td>You Americans use crazy words for things</td>
</tr>
</tbody>
</table>
Discussion: What is in a Name?

The importance of good definitions

• How many different terms & definitions can we come up with for the general term “address”?

• For example, here are some obvious ones:
The Importance of Business Definitions

OK, we're almost done with user acceptance testing and everything looks great with this new marketing application. Just one small question - what is a Customer?

From Data Modeling for the Business by Hoberman, Burbank, Bradley, Technics Publications, 2009
Publishing Business Definitions in a Data Model

• Data Models are a great place to store business definitions
  • Display them on the model for a business audience
  • Store them in the model repository for reuse across the organization (various users, tools, etc.)
Creating a Technical Data Inventory
Linking business definitions to technical implementations

- Data models & the associated metadata can create a real-world inventory of the data storage associated with key business data domains in the control of a data governance program.
Data Modeling Creates an “Active Inventory” of Data Assets

- **Data Models can manage both the “As-Is” and “To Be” environments**
  - **Know what data you have:** Create a visual inventory of database systems
  - **Know what your data means:** Communicate key business requirements between business and IT stakeholders
  - **Support data consistency:** Build consistent database structures & support data governance initiatives

![Data Models Diagram]
Supporting Consistency & Metadata Sharing
Publication & Interchange

- Once business & technical metadata is stored in data models and the associated metadata repositories, it can be shared with other key tools & technologies across the enterprise.
Domains
Supporting Consistency & Reuse

• Creating Domains within data models are a helpful way to
  • Enforce Standards
  • Support Impact Analysis
**Naming Standards**
Supporting Consistency & Reuse

“I need a report showing total revenue from all regions.”

- Creating naming standards help create consistent, agreed upon naming conventions for data objects
  - Increasing Reuse
  - Improving Data Quality
  - Assisting with Impact Analysis
Governance Properties in a Data Model
Stewardship, Security, Privacy, etc.

• Key aspects of data governance can be stored in data model metadata:
  • Data Stewardship
  • Privacy
  • Security
  • Etc.
Standardization, Efficiency & Reuse

- Data Models & Metadata Management can help rationalize data storage throughout the organization, leading to significant efficiencies & cost reduction.

1. Identify the issue via Metadata Discovery & Inventory

   Customer information is stored 175 different ways across the organization.

   Customer, CUST, CDB1, Client, Party, etc, etc.

2. Resolve & Prevent issues via published metadata standards.

   Standard Reference Metadata

   Customer
Data Lineage

• In the data warehouse example below, metadata for CUSTOMER exists in a number tools & data stores.

• This lineage can be tracked in most data modeling tools.
Impact Analysis & Where Used

- Impact Analysis shows the relationship between a piece of metadata and other sources that rely on that metadata to assess the impact of a potential change.
- For example, if I change the length & name of a field, what other systems that are referencing that field will be affected?

What happens if I change the name & length of the “Brand” field?

- Brand CHAR(10)
- MyBrand VARCHAR(30)
Technical Metadata Makes Data Governance Actionable

Data models are a good vehicle for this

- Data models can help take the business rules & definitions defined in policies and make them actionable in physical systems, maintaining a lineage & audit trail.
Summary

• Data Governance is a large driver in the growth in demand for data modeling & metadata management

• Data models are a rich source of metadata

• Data Models supports the policies & procedures defined by data governance
  • Business definitions
  • Technical data structures
  • Data lineage & impact analysis

• Data model metadata supports actionable data governance through
  • Linking business & technical definitions & business rules
  • Providing standardization & consistency
  • Supporting data lineage & audit trails

• Technical data governance provides support for the policies, people, & procedures of data governance
About Global Data Strategy, Ltd
Data-Driven Business Transformation

• Global Data Strategy is an international information management consulting company that specializes in the alignment of business drivers with data-centric technology.

• Our passion is data, and helping organizations enrich their business opportunities through data and information.

• Our core values center around providing solutions that are:
  • Business-Driven: We put the needs of your business first, before we look at any technology solution.
  • Clear & Relevant: We provide clear explanations using real-world examples.
  • Customized & Right-Sized: Our implementations are based on the unique needs of your organization’s size, corporate culture, and geography.
  • High Quality & Technically Precise: We pride ourselves in excellence of execution, with years of technical expertise in the industry.

Visit www.globaldatastrategy.com for more information

Global Data Strategy, Ltd 2017
Contact Info

• Email: donna.burbank@globaldatastrategy.com
• Twitter: @donnaburbank
  @GlobalDataStrat
• Website: www.globaldatastrategy.com
• Company Linkedin: https://www.linkedin.com/company/global-data-strategy-ltd
• Personal Linkedin: https://www.linkedin.com/in/donnaburbank
White Paper: Emerging Trends in Metadata Management

Free Download

• Download from www.globaldatastrategy.com
  • Under ‘Whitepapers’
DATAVERSITY Lessons in Data Modeling Series

This Year’s Line Up - www.dataversity.net

• January - on demand  How Data Modeling Fits Into an Overall Enterprise Architecture
• February - on demand  Data Modeling and Business Intelligence
• March - on demand  Conceptual Data Modeling – How to Get the Attention of Business Users
• April - on demand  The Evolving Role of the Data Architect – What does it mean for your Career?
• May - on demand  Data Modeling & Metadata Management
• June - on demand  Self-Service Data Analysis, Data Wrangling, Data Munging, and Data Modeling
• July - on demand  Data Modeling & Metadata for Graph Databases
• August - on demand  Data Modeling & Data Integration
• Sept - on demand  Data Modeling & Master Data Management (MDM)

• October 26  Agile & Data Modeling – How Can They Work Together?
• December 5  Data Modeling, Data Quality & Data Governance
DATAVERSITY Training Center
Online Training Courses

Metadata Management Course

• Learn the basics of Metadata Management and practical tips on how to apply metadata management in the real world. This online course hosted by DATAVERSITY provides a series of six courses including:
  • What is Metadata
  • The Business Value of Metadata
  • Sources of Metadata
  • Metamodels and Metadata Standards
  • Metadata Architecture, Integration, and Storage
  • Metadata Strategy and Implementation

• Purchase all six courses for $399 or individually at $79 each. Register [here](http://training.dataversity.net/lms/)
  • Use code “GDS” for 20% off
  • Other courses available on Data Governance, Data Quality, etc.

Visit: [http://training.dataversity.net/lms/](http://training.dataversity.net/lms/)
Questions?
Thoughts? Ideas?