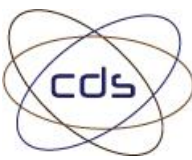


DATA STEWARDSHIP BODY OF KNOWLEDGE (DSBOK)

Release 2.0. May 2012.

This document was created in collaboration of the leading experts and educators in the field and members of the Certified Data Steward (CDS) Advisory Council.



1. Data Stewardship Fundamentals

1.1. Definitions

- 1.1.1. Stewardship
- 1.1.2. Data Stewardship

1.2. Data Stewardship Organizations

- 1.2.1. Kinds of Data Stewards
 - 1.2.1.1. Kinds of Business Stewards
 - 1.2.1.2. Kinds of IT Stewards
 - 1.2.1.3. Differences of Business and IT Stewards
- 1.2.2. Roles and Responsibilities
- 1.2.3. Reporting and Relationships

1.3. Data Steward Characteristics

- 1.3.1. Knowledge and Experience
- 1.3.2. Skills and Competencies
 - 1.3.2.1. Data Skills
 - 1.3.2.2. Technical Skills
 - 1.3.2.3. Human Skills
 - 1.3.2.4. Facilitation Skills
 - 1.3.2.5. Communication Skills

2. Data Quality (DQ)

2.1. Quality Management Basics

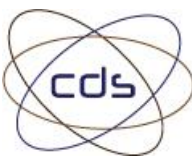
- 2.1.1. Quality Perspectives
 - 2.1.1.1. Expectations
 - 2.1.1.2. Purpose
 - 2.1.1.3. Specifications
 - 2.1.1.4. Defects
- 2.1.2. Quality Management Terminology
 - 2.1.2.1. Quality Control (QC)
 - 2.1.2.2. Quality Assurance (QA)
 - 2.1.2.3. Waste and rework
 - 2.1.2.4. Inspection, Correction, Prevention
- 2.1.3. Quality Management Methodologies
 - 2.1.3.1. Total Quality Management (TQM)
 - 2.1.3.2. Statistical Process Control (SPC)
 - 2.1.3.3. Six Sigma

2.2. Data Quality Concepts and Principles

- 2.2.1. DQ Definitions
 - 2.2.1.1. Defect Free
 - 2.2.1.2. Conform to Specifications
 - 2.2.1.3. Fit to Purpose
 - 2.2.1.4. Meet Customer Expectations
- 2.2.2. Common Causes of DQ Problems
- 2.2.3. Costs and Benefits of DQ
- 2.2.4. Bad Habits and Good Practices of DQ Management
- 2.2.5. DQ Culture



- 2.3. *Data Quality Dimensions*
 - 2.3.1. Content Quality (Correctness)
 - 2.3.1.1. Accuracy
 - 2.3.1.2. Completeness
 - 2.3.1.3. Precision
 - 2.3.1.4. Granularity
 - 2.3.1.5. Consistency
 - 2.3.2. Structural Quality (Integrity)
 - 2.3.2.1. Identity
 - 2.3.2.2. Reference
 - 2.3.2.3. Cardinality
 - 2.3.2.4. Dependency
 - 2.3.2.5. Inheritance
 - 2.3.2.6. Domain of Values
 - 2.3.3. Temporal Quality (Timeliness)
 - 2.3.3.1. Currency
 - 2.3.3.2. Retention
 - 2.3.3.3. Continuity
 - 2.3.3.4. Precedence
 - 2.3.4. Business Quality (Value)
 - 2.3.4.1. Aligned
 - 2.3.4.2. Trusted
 - 2.3.4.3. Understandable
 - 2.3.4.4. Reliable
 - 2.3.4.5. Compliant
 - 2.3.4.6. Useful
 - 2.3.5. Usage Quality (Usability)
 - 2.3.5.1. Available
 - 2.3.5.2. Accessible
 - 2.3.5.3. Navigable
 - 2.3.5.4. Recoverable
 - 2.3.5.5. Secure
 - 2.3.5.6. Private
 - 2.3.6. Presentation Quality (Communication)
 - 2.3.6.1. Clear
 - 2.3.6.2. Organized
 - 2.3.6.3. Non-Ambiguous
 - 2.3.6.4. Tool-Integrated
 - 2.3.6.5. Media-Fit
- 2.4. *Data Quality Processes and Projects*
 - 2.4.1. Data Profiling
 - 2.4.2. DQ Assessment
 - 2.4.3. Root Cause Analysis
 - 2.4.4. Data Cleansing
 - 2.4.5. Data Enrichment
- 2.5. *Data Quality in IT Processes and Projects*
 - 2.5.1. DQ in Application and Database Projects
 - 2.5.2. DQ in Data Conversion and Consolidation Projects
 - 2.5.3. DQ in Data Sharing and Data Interfaces



3. Data Governance

3.1. Data Governance Basics

- 3.1.1. Definitions
- 3.1.2. Purpose
- 3.1.3. What to Govern and What not to Govern?

3.2. Components of Data Governance

- 3.2.1. Governance Goals
 - 3.2.1.1. Quality
 - 3.2.1.2. Security
 - 3.2.1.3. Compliance
 - 3.2.1.4. Standardization
 - 3.2.1.5. Usage
 - 3.2.1.6. Value
 - 3.2.1.7. Business Impact
- 3.2.2. Governance Elements
 - 3.2.2.1. Roles
 - 3.2.2.2. Responsibilities
 - 3.2.2.3. Decision Authority
 - 3.2.2.4. Accountability
- 3.2.3. Policies
 - 3.2.3.1. Quality
 - 3.2.3.2. Sensitivity
 - 3.2.3.3. Privacy
 - 3.2.3.4. Security
 - 3.2.3.5. Access
 - 3.2.3.6. Retention
 - 3.2.3.7. Disposal
- 3.2.4. Standards
 - 3.2.4.1. Naming
 - 3.2.4.2. Definition
 - 3.2.4.3. Architecture
 - 3.2.4.4. Transfer
 - 3.2.4.5. Integration
- 3.2.5. Technology
 - 3.2.5.1. DG Enabling
 - 3.2.5.2. Data Quality
 - 3.2.5.3. Workflow
 - 3.2.5.4. Collaboration

3.3. Data Governance Programs

- 3.3.1. Business Drivers for Governance
 - 3.3.1.1. Legal
 - 3.3.1.2. Regulatory
 - 3.3.1.3. Financial
 - 3.3.1.4. Operational
 - 3.3.1.5. Competitive



- 3.3.2. Data Governance Roles
 - 3.3.2.1. Sponsors
 - 3.3.2.2. Owners
 - 3.3.2.3. Data Stewards
 - 3.3.2.4. Custodians
 - 3.3.2.5. Stakeholders
 - 3.3.2.6. Data Officers
 - 3.3.2.7. Councils, Committees, Competency Centers

3.4. *Executing Data Governance*

- 3.4.1. Getting Started
- 3.4.2. Funding and Sponsorship
- 3.4.3. Day-to-Day Governance
- 3.4.4. Sustaining Governance
- 3.4.5. Monitoring Governance: Measures, Metrics, Maturity
- 3.4.6. Evolving Governance
- 3.4.7. Stakeholder Communications

4. **Metadata Management**

4.1. *Metadata Concepts*

- 4.1.1. Metadata Definition
- 4.1.2. Roles of Metadata
 - 4.1.2.1. Classify
 - 4.1.2.2. Describe
 - 4.1.2.3. Guide
 - 4.1.2.4. Control

4.2. *Business Metadata*

- 4.2.1. Types of Business Metadata
- 4.2.2. Business Metadata Standards and Practices
- 4.2.3. Data Steward Roles in Business Metadata

4.3. *Technical Metadata*

- 4.3.1. Types of Technical Metadata
- 4.3.2. Technical Metadata Standards and Practices
- 4.3.3. Data Steward Roles in Technical Metadata

4.4. *Data Modeling*

- 4.4.1. Data Model Principles
- 4.4.2. Types of Data Models
 - 4.4.2.1. ER vs. Dimensional
 - 4.4.2.2. Logical vs. Physical
 - 4.4.2.3. Who and Why for each Type of Model
- 4.4.3. Complementary Models
 - 4.4.3.1. Subject Area Models
 - 4.4.3.2. State Transition Models
- 4.4.4. Reading Data Models
- 4.4.5. Data Stewards and Data Modeling



5. Data Management Processes

5.1. Architectural Processes

- 5.1.1. Enterprise and Subject Area Modeling
- 5.1.2. Data Mapping and Consolidation
- 5.1.3. Data Flow
- 5.1.4. Role of the Data Steward

5.2. Utilization Processes (CRUD)

- 5.2.1. Create Processes and Practices
- 5.2.2. Retrieval and Reporting Processes and Practices
- 5.2.3. Update Processes and Practices
- 5.2.4. Delete and Archive Processes and Practices
- 5.2.5. Business Process Management
 - 5.2.5.1. End-to-end data flow
 - 5.2.5.2. Data interfaces
 - 5.2.5.3. Data transfer
 - 5.2.5.4. Data exchange
 - 5.2.5.5. Process gaps
- 5.2.6. Role of the Data Steward

5.3. Custodial Processes

- 5.3.1. Database Administration
- 5.3.2. Security Administration and Access Authorization
- 5.3.3. Backup, Recovery, and Business Continuity
- 5.3.4. Role of the Data Steward

6. Information Management Concepts

6.1. Types of Data and Information

- 6.1.1. Operational and Analytical
- 6.1.2. Event and Reference
- 6.1.3. Structured and Unstructured
- 6.1.4. Transactional Data
- 6.1.5. Master Data

6.2. Types of Data Stores

- 6.2.1. Application Databases
- 6.2.2. Departmental and End-User Databases
- 6.2.3. Data Warehouses and Data Marts
- 6.2.4. Operational Data Stores
- 6.2.5. Master Data Hubs

6.3. Common Uses of Data

- 6.3.1. Record Keeping and Audit Trail
- 6.3.2. Reporting and Information
- 6.3.3. Measurement and Monitoring
- 6.3.4. Analysis and Discovery



- 6.4. *Business Data Flow*
 - 6.4.1. The Business Value of Data
 - 6.4.2. Data Sources and Data Acquisition
 - 6.4.3. Data Providers and Consumers
 - 6.4.4. Data Flow through Organizations and Systems
 - 6.4.5. Data Conversion and Consolidation
 - 6.4.6. Data Replication and Redundancy
 - 6.4.7. Data Sharing and Interfaces
 - 6.4.7.1. Batch and Real Time
 - 6.4.7.2. Feeds, Hubs, Messages, Virtualization, etc.
 - 6.4.8. Data Disposal and Destruction

- 6.5. *Information Management Disciplines*
 - 6.5.1. Understanding the Data
 - 6.5.1.1. Data Modeling
 - 6.5.1.2. Metadata Management
 - 6.5.2. Information Supply and Demand
 - 6.5.2.1. Content Management
 - 6.5.2.2. Enterprise Information Management
 - 6.5.3. Data Utility
 - 6.5.3.1. Data Quality
 - 6.5.3.2. Data Governance
 - 6.5.4. Data Resource Consolidation
 - 6.5.4.1. Data Integration
 - 6.5.4.2. Data Warehousing
 - 6.5.4.3. Master Data Management
 - 6.5.5. Applied Information
 - 6.5.5.1. Business Intelligence
 - 6.5.5.2. Business Analytics
 - 6.5.5.3. Performance Management
 - 6.5.6. Discovery and Inference
 - 6.5.6.1. Data Mining
 - 6.5.6.2. Predictive Analytics

Note: While this document goes into four levels of detail, the fourth level is not exhaustive and in many cases only includes some examples for topic clarification.