DATA MANAGEMENT
GLOSSARY

The Standard Glossary of Data Management Concepts

Developed by professional data practitioners to establish standard terminology and meaning for the practice of data management, with definitions, related terms and commentary

Version 0.2.1
August 2017

ENTERPRISE DATA MANAGEMENT
EDM
C O U N C I L

INTRODUCTION

About the Data Management Glossary

On the request of the membership, the EDM Council has initiated the formal creation of the Data Management Glossary. The goal is to formalize the business terminology used to describe the practice of data management. To achieve this goal, we organized a group of industry practitioners and asked them to impose rigor on the definition of terms and to provide clarity needed to ensure consistent collaboration among organizational stakeholders.

All EDM Council members are invited to offer critiques and alternative suggestions to the Data Management Glossary. For each term in the glossary, there is a hyperlink (*Send feedback* on this term) which will assist you in providing feedback.

We are starting with 120 business concepts in six core data management categories (data architecture, roles & responsibilities, provisioning, metadata, governance and data quality). We want to make sure the terms, definitions, synonyms and descriptions align with your data management programs. We encourage you to contribute additional terms and internal glossaries in support of this initiative. You can also direct your critiques and alternative suggestions to John Bottega, Head of Best Practice for the EDM Council (908-501-3826, jbottega@edmcouncil.org).

About the EDM Council

The EDM Council is a non-profit trade association established by participants in the financial industry to define content standards and document best practices for implementing data management. The 200+ global member firms and more than 7,000 professionals of the EDM Council represent banks, asset managers, data vendors, technology companies, consultants, academia and regulators across the global financial system.

Data Management Best Practice

The EDM Council defines the capabilities and processes that are required to implement sustainable data management, based on what does and does not work in the real world. Our best practice agenda is created and verified based on the collective experience of individuals that have earned their expertise by addressing the realities of financial service operations. The results of our research and documentation are available to our constituents as a benefit of membership.

A key component of the data management best practices is DCAM – the Data Management Capability Assessment Model. DCAM is the industry standard guideline on the practice of data management – used as a framework for establishing data management programs, obtaining alignment from stakeholders and benchmarking progress. DCAM defines the core criteria for data management strategy, business case, operating model, content engineering, data quality and governance.

For more information:

www.EDMCouncil.org

ABOUT THIS EDITION

Published August 2017

Categories of Terms

*This edition includes definitions for the first 2 categories of terms – **Function/Role** and **Sourcing/Provisioning,** totaling 25 terms. Additional categories and terms will be added in subsequent releases.*

DATA MANAGEMENT GLOSSARY

Terms are organized alphabetically by category

FUNCTION / ROLE

Function: Business Data Stewardship	
Definition	*The function that implements and/or executes the data management program for its respective business* data domain*(s).*
Description	The Business Data Stewardship function must understand the business and data manufacturing processes to ensure that the data is fit for its intended purposes. The function includes (but not limited to): • Defining and establishing the domain boundaries in collaboration with peer Business Data Stewards • Ensuring that all data management processes comply with organizational policy and standards • Capturing and verify data requirements in collaboration with data consumers • Source and provision data for use by business or other consuming applications • Ensuring that data is consistently defined, aligned to business concepts and captured as metadata • Designing, executing and monitoring data controls and transformation processes • Managing data quality including profiling, remediation, business rules and validation processes • Coordinating across functions to define and implement data sharing agreements, security levels, privacy restrictions and data retention policies
Possible Role Titles	• Business Data Steward (Executive) • Front Line Data Officer • Functional Data Officer • Data Domain Steward • Data Custodian

Send feedback *on this term*

Function: Data Architect

Definition	*The function that defines and implements the data content strategy for a given subset of data.*
Description	The Data Architecture function provides the "content engineering" bridge between business applications and technology implementation. The focus is on content management including how the data will be identified/defined as well as how to access it across the organizational ecosystem. The function of Data Architecture includes understanding the scope of data needed to satisfy business requirements as well as ensuring that the data is aligned to its precise meaning. The function includes (but is not limited to): • Designing the data architecture processes • Establishing the operating model required to execute the defined data engineering processes • Establishing and implementing the framework for conceptual and logical data modeling • Defining logical domains of data in collaboration with business and IT • Implementing a unified view of data meaning across the enterprise
Possible Role Titles	• Data Architecture Executive • Head of Data Architecture • Data Architect Manager • Data Architect

Send feedback on this term

Function: Data Control

Definition	*The function that sets the enterprise data risk controls framework and strategy.*
Description	The Data Control function is the first line of defense against risk from ineffective data management practices. This includes the identification of key risk indicators as well as reporting of key risk and performance indicators to executive management. The function includes (but not limited to): • Designing and implementing an enterprise-wide data risk governance framework • Establishing data policy and standards for mitigating risk from data • Defining and implementing risk profiling, assessment and oversight processes • Developing risk training and tools to ensure compliance with risk mitigation objectives • Monitoring and enforcing compliance with data policies and standards
Possible Role Titles	• Data Control Executive • Head of Data Control • Data Control Manager

Send feedback on this term

Function: Data Governance

Definition	*The function that defines and implements the standards, controls and best practices of the data management program in alignment with strategy.*
Description	Data Governance is responsible for creating and implementing a "data control" environment. According to the Basel Committee on Banking Supervision – a data control environment consists of a set of policies governing all aspects of data acquisition, distribution, integration and usage that are sanctioned by executive management, based on standards, implemented across the data lifecycle, with clear accountability and monitored by audit. The function includes (but is not limited to): • Designing and implementing the framework (including associated processes) necessary to sustain a data control environment • Establishing the operating model required to achieve governance objectives • Defining and implementing policy, standards and operating procedures • Establishing and implementing the data accountability mechanisms • Developing and implementing metrics needed to monitor/report on data management progress • Designing and implementing data governance training programs
Possible Role Titles	• Data Governance Executive • Head of Data Governance • Data Governance Manager

Send feedback on this term

Function: Data Officer

Definition	*The function performed by an executive responsible for the strategy and implementation of the data management program at a geographic, line-of-business or functional level.*
Description	Depending on the structure of the organization, the Data Officer function may fulfill the same responsibilities as the CDO. In addition, the function includes (but not limited to): • Ensuring the data management program is on track to deliver against objectives, goals and expectations • Ensuring business data ownership/stewardship • Securing the resources required for execution • Ensuring that the data management program is implemented in accordance with standards, policies and procedures • Managing collaboration with technology, operations and other cross-organizational control functions
Possible Role Titles	• Regional Data Officer • Group Data Officer • [*Business name*] Data Officer

Send feedback on this term

Function: Data Quality

Definition	*The function that ensures the data being used for operating processes is fit for its intended purpose.*
Description	The function includes (but not limited to): • Defining data quality requirements in collaboration with Stakeholders • Designing the data quality processes (i.e. validation and remediation) to ensure that data is fit-for-purpose • Establishing and implementing the data quality operating model • Facilitating the prioritization and escalation of cross-domain data remediation • Defining the data quality measurement criteria and reporting framework • Delivering training to achieve adoption of the data quality operating model and associated processes
Possible Role Titles	• Data Quality Executive • Head of Data Quality • Data Quality Manager

Send feedback on this term

Function: Metadata Management

Definition	*The function that ensures the documentation, definition, and implementation of metadata.*
Description	The Metadata Management function is responsible for the quality, implementation, recording and use of metadata. The metadata management function supports the business and technical data stewards to ensure compliance with policy and the adoption of metadata standards. The function includes (but not limited to): • Establishing the metadata management framework and associated processes • Defining and implementing the operating model for metadata management • Selecting and implementing the metadata management tool sets in accordance with internal guidelines • Documenting content in the metadata repository • Monitoring the completeness and accuracy of metadata • Developing and delivering training to achieve adoption of the operating model associated processes
Possible Role Titles	• Metadata Management Executive • Head of Metadata Management • Metadata Manager

Send feedback on this term

Function: Technical Data Stewardship

Definition	*The function that manages the technical implementation of the data management program.*
Description	The Technical Data Stewardship function manages the technology (i.e. databases; data marts; data warehouses) and executes the physical implementation of the data elements associated with selected data domains. The function includes (but not limited to): • Designing, building and managing the technical infrastructure associated with a selected data domain • Aligning business elements with their associated data components • Translating business and data elements into technical specifications • Defining and managing technical service level agreements • Defining the technical aspects of data quality, transformation and movement controls • Monitoring and remediating data defects against established quality thresholds • Establishing and implementing root cause analysis of data defects • Managing technical metadata
Possible Role Titles	• Technical Data Steward (Executive) • Data Custodian

Send feedback on this term

Function: Technology Architecture

Definition	*The function that supports the technical design, approach and toolset selection in alignment with the data strategy.*
Description	The Technology Architecture function ensures the objectives of the data management program can be made operational. The function includes (but not limited to): • Translating business requirements into the technology architecture and systems design (blueprint) • Aligning the business technology blueprint to enterprise architectural policy and guidelines • Manage infrastructure capacity, systems design, transmission capability and analytical platforms • Evaluate and recommend vendor solutions
Possible Role Titles	• Technology Architecture Executive • Technology Architect Manager • Technology Architect • Solution Architect

Send feedback on this term

Role: Chief Data Officer	
Definition	*The most senior executive in the enterprise responsible for the strategy and implementation of the data management program.*
Description	The function includes (but not limited to): • Establishing and maintaining the data management program • Understanding business requirements and aligning them to the objectives of the data management program • Defining the strategy and operational framework needed to implement the data management program • Documenting the business case and aligning it to the budget/funding mechanisms needed for program implementation and sustainability • Defining the data quality framework necessary to ensure trust and confidence in data assets • Defining the data architecture framework for ensuring a unified view of meaning across the organization • Establishing the standards, policies and procedures for the data management program • Defining and implementing the data governance and organizational structures needed to support the objectives of the data management program • Ensuring stakeholder understanding, alignment and commitment to the objectives of the data management program

Send feedback on this term

Role: Data Consumer	
Definition	*A process, application or stakeholder that receives or uses data from a data producer.*
Description	The Data Consumer establishes requirements and quality expectations for the data. Consumers need assurance that the data is fit for its intended use and that the appropriate use of the data is aligned to data management, governance and risk management policy.

Send feedback on this term

Role: Data Owner	
Definition	*The person with overall accountability for the meaning, content, quality and distribution of a given set of data.*
Description	The Data Owner is accountable for the quality of a given data domain or set of data. This includes the quality of the data as well as how the data is defined, manufactured, identified, maintained, delivered and consumed. The Data Owner may not be directly involved in the curation and maintenance of the data, but they are accountable for ensuring that it meets quality criteria and is in alignment with organizational standards.

Send feedback on this term

Role: Data Sponsor

Definition	*An executive responsible for ensuring adequate funding of the data management program.*
Description	The Data Sponsor owns the P&L for the data management program and the allocation of resources needed to mitigate the risk of service interruption. In addition to budgets and staff, this includes understanding data dependencies as well as the ensuring that the contractual obligations associated with third party data procurement are fulfilled.
Possible Role Titles	• Data Domain Sponsor

Send feedback on this term

Role: Stakeholder

Definition	*An interested participant (producer, consumer, supporting process) in the data ecosystem.*
Possible Role Titles	• Involved Party

Send feedback on this term

Role: Subject Matter Expert (SME)

Definition	*An individual who is an authority in a business process, data manufacturing process or application.*
Possible Role Titles	• SME

Send feedback on this term

SOURCING/PROVISIONING

Authoritative Data Domain	
Definition	A Data Domain that has been designated, verified, approved and enforced by the data management governing body.
Synonyms	Authorized Data Domain

Send feedback on this term

Authoritative Data Source	
Definition	An officially designated system or repository for data that has been determined to be reliable and accurate.
Description	An Authoritative Data Source has been designated by the data management governing body as the official source of a specific Data Domain. Required use of the authorized source is driven by established policy and standards.
Synonyms	• Golden Source • Authorized Data Source • Single Authoritative Source • Designated Repository • Publication Point

Send feedback on this term

Authoritative Provisioning Point (APP)	
Definition	A Provisioning Point that has been designated by the relevant data management governing body as providing data from an Authoritative Data Domain.
Description	Authoritative Provisioning Points are designated by the organization's governing body after the content has been rationalized against internal data engineering standards for meaning, structure and format. As part of that designation the provisioning point includes adequate data controls to ensure data remains fit for purpose. Best practice is that an Authoritative Provisioning Point would be registered in the Enterprise Provisioning Registry.
Synonyms	• Provisioning Point (Authorized, Approved) • Data Service Point (Authorized, Approved) • Data Delivery Service (Authorized, Approved) • Data Access Point (Authorized, Approved)

Send feedback on this term

Data Domain

Definition	*A logical representation of a category of data that has been designated and named.*
Description	Data Domains are not physical repositories or databases. Instead, they are "logical" categories or groupings of data that are deemed important and necessary to a firm's normal business operation. Data Domains include both internally generated data as well as externally acquired data. Examples of Data Domains might include "product data;" "customer data;" "trade data;" "pricing data;" "index data;" "risk data," etc. It is imperative that these strategic categories of data are identified, defined and inventoried to ensure their proper maintenance and use throughout the organization.
Synonyms	• Data Subject Area

Send feedback on this term

Data Lineage

Definition	*Documentation of the sequence of movement and/or transformation of data as it flows between the consumer and the source(s).*
Description	Data Elements may have multiple sources and end consumers. Data Lineage describes the chronology of ownership, custody and location of data. Data Lineage provides a visual mapping of the movement and changes in data from system to system. The goal is to ensure that the data consumed is equivalent to the data delivered. Data Lineage provides a mapping of data for use in impact analysis and operational risk integrity. The complete lineage will document the full data flow and capture metadata about the movement and transformation of the data element. Lineage may include a mapping of the data controls. Data Lineage is commonly confused with Data Traceability and Data Provenance and should be understood in relationship to one another.

Send feedback on this term

Golden Record

Definition	*A single, precisely defined, verified and officially designated version of data.*
Description	The Golden Record is designated by a business or operational process to indicate that the data has been validated as fit for its intended purpose. A Golden Record should be used for all applications and enforceable by policy.
Synonyms	• Master Record • Golden Copy • Single Version of the Truth

Send feedback on this term

Provisioning Point

Definition	A service from which data associated with a Data Domain can be acquired after it has been sourced.
Description	The purpose of a Provisioning Point is the distribution of data from a given data domain to ensure the appropriate source of data throughout the organization. Provisioning Points can be executed in either physical repositories or via virtual access and help establish a 'control environment' for data throughout the organization.
Synonyms	- Data Service Point - Data Delivery Service - Data Access Point

Send feedback on this term

Source of Origin

Definition	The genesis of data content prior to being captured electronically.
Description	The Source of Origin refers to the source of a Data Element. The data value may have been created by an individual in the business process, manually captured in a document, sourced manually or captured electronically from an external provider. For example, the prospectus would be the Source of Origin for the instrument record in the Security Master Database.

Send feedback on this term

System of Origin

Definition	Any application or repository where data is initially captured.
Description	A System of Origin is the point at which information has been introduced (without validation or remediation) into the organization. If the System of Origin is considered valid without reconciliation, then the System of Origin could also be classified as the System of Record.

Send feedback on this term

System of Record

Definition	The Authoritative Data Source for the specified Data Element after it has been remediated and validated.
Description	SORs are repositories of data that have been screened, validated and exceptions remediated. To ensure data integrity there must be only one System of Record for a logical category of data.

Send feedback on this term

DATA QUALITY

Accuracy	
Definition	*A measurement of the veracity of data to its authoritative source.*
Description	Accuracy is a measurement of the precision of data. It can be measured against either original documents or authoritative sources and validated against defined business rules. Accuracy is one of the seven Data Quality Dimensions. Examples: • Records that are wrong at a specified time (i.e. a record with an incorrect maturity date) • Records that haven't been refreshed or updated • Records at the wrong level of precision (i.e. prices that were originally quoted at three decimal places, but cut-off and stored at two decimal places)

Send feedback on this term

Completeness	
Definition	*A measurement of the availability of required data attributes.*
Description	Completeness measures the existence of required data attributes in the population of data records. Completeness is one of the seven Data Quality Dimensions. Examples: • A missing ticker symbol, CUSIP, or other identifier • A fixed income instrument record with a null coupon value • A benchmark or index that is missing a dividend notice or stock split • A record with missing attributes

Send feedback on this term

Conformity	
Definition	*A measurement of the alignment of content with the required standards.*
Description	Conformity measures how well the data aligns to internal, external or industry-wide standards. Conformity is one of the seven Data Quality Dimensions. Examples: • Invalid ISO currency codes • Violation of allowable values (i.e. a state code for a country that does not have states) • Inconsistent date formats

Send feedback on this term

Consistency	
Definition	*A measurement of compliance with required formats, values or definitions.*
Description	Consistency provides assurance that data values, formats and definitions in one population agree with those in another data population Consistency is one of the seven Data Quality Dimensions.

Send feedback on this term

Coverage	
Definition	*A measurement of the availability of required data records.*
Description	Coverage refers to the breadth, depth and availability of data that exists but is missing from a data provider. Coverage is one of the seven Data Quality Dimensions. Examples: • A group of securities (i.e.: corporate bonds) not included in a vendor feed • Quoted prices from an emerging market that are missing • Legal entity and hierarchy data missing from a country or region

Send feedback on this term

Data Harmonization	
Definition	*The process of aligning all representations of data to precise and consistent meaning.*

Send feedback on this term

Data Normalization	
Definition	*The process of aligning data to its defined parameters.*
Description	This is a business centric perspective and not to be confused with the process of normalization in a data modeling context.

Send feedback on this term

Data Profiling

Definition	*The process of evaluating and grading a given source of data to determine whether it is fit for its intended purpose.*
Description	Profiling is a methodology for determining the current state of data quality in a repository. Data profiling would include a review against all data quality dimensions to identify data anomalies and evaluate data variance.
Synonyms	• Data Quality analysis • Data quality assessment

Send feedback on this term

Data Quality

Definition	*A measurement of qualitative and quantitative conditions that determine whether the data is fit for its intended use in a business process or operation.*
Description	The quality of data can be evaluated against defined dimensions (i.e. accuracy, completeness, conformity to standards, consistency, coverage, timeliness and uniqueness) as well as against the business processes associated with its production.

Send feedback on this term

Data Quality Decay Rate

Definition	*The rate that a data attribute loses its level of quality over time if not maintained. Decay can be measured against all dimension of data quality.*

Send feedback on this term

Data Quality Dimensions

Definition	*Categories of measurement used to evaluate the degree to which data is fit for its intended purpose.*
Description	The EDM Council recognizes the meaning of seven core dimensions used to define and measure the quality of data. The seven dimensions include: • Accuracy • Completeness • Conformity • Consistency • Coverage • Timeliness • Uniqueness
Synonyms	Dimension of Quality

Send feedback on this term

Data Quality Rule

Definition	*Defined and documented criteria for measuring whether data is fit for its intended purpose.*

Data Transformation

Definition	*The process of converting the meaning and format of data from one system to another.*

Exception Handling

Definition	*A process by which records are catalogued, queued and remediated after failing a data quality rule.*

Timeliness

Definition	*A measurement of the degree to which data is both representative of current conditions and available for use.*
Description	Timeliness is a measurement of how well content represents current market/business conditions as well as whether the data is functionally available when needed. Timeliness is one of the seven Data Quality Dimensions.

Examples:

- A file delivered too late for a business process or operation
- An issuance or corporate action not delivered when it was announced
- A credit rating change not updated on the day if was issued
- A new prospectus not given an official number from the national numbering agency

Uniqueness

Definition	*A measurement of the degree that no record or attribute is recorded more than once.*
Description	Uniqueness refers to the singularity of records and or attributes. The objective is a single (unique) recording of data. Uniqueness is one of the seven Data Quality Dimensions.

Examples:

- Two instances of the same security with different identifiers or spellings
- A preferred share represented as both an equity and debt object in the same database

For More Information

For questions about the Data Management Glossary,
to provide comments, or to inquire about membership programs,
please contact the EDM Council.

UK Main Office: + 44 (0) 208 399 7189

US Main Office: +1 (301) 933 2945

info@edmcouncil.org

ENTERPRISE DATA MANAGEMENT
EDM
C O U N C I L