

Data Steward Course - Glossary

Term	Definition
Accessibility	A dimension of data quality pertaining to the degree to which required data be accessed.
Accuracy	A dimension of data quality pertaining to the degree to which required data is consistent with original intent
Acquired Data	Data brought into an existing system from one or more other systems.
Agile Method	An iterative approach that breaks down the full set of business needs for a system into small segments that can be developed quickly to provide portions of a working system.
Allowed Values	Predefined values that should be used to populate a cell in a database. For example, codes, flags, statuses, min-max values.
Attribute	A fact about an entity type. For example, Client Last Name, Client Phone Number, etc. (See Column)
Business Data Rules	Requirements for processes employed to control data content to ensure it will meet quality and functional requirements.
Business Glossary	A repository or collection of business terms that are included in the language associated with the business. A glossary typically contains business terms, the definition of the terms, and the person or agency that owns the responsibility for managing business associated with the terms.
Business Metadata	Represents data entered and managed for users to perform business processes, such as ordering and billing.
Business Process	A set of activities and tasks that, once completed, will accomplish an organizational goal.
Business Term	A word conveying shared meaning, that is commonly used while conducting ordinary business activities, e.g., client, account, purchase, etc.
Cardinality	Fundamental principle of one data aspect with respect to another, i.e., "one-to-many", "many-to-many" "one-to-zero-or-one", etc.
Column	A data value of a particular type that populates one value for each row of a table. (See Attribute) C50
Common Data Matrix	A template that assists in identification of the appropriate stakeholders and group members for shared data.
Completeness	A dimension of data quality pertaining to the availability of required data elements captured in a record.
Conformity	A dimension of data quality pertaining to the alignment of data with required standards.
Consistency	A dimension of data quality pertaining to the degree to which data follows specified patterns and uniformity rules across data stores.

Critical Data Elements	Data elements that highly impact the performance of important business processes or are needed for a specific business purpose; for example data elements contained in an audited financial statement.
Data	A set of facts the evidence and serve to describe observed phenomena, such as business events.
Data asset	Anything that exists in or assists in the governance of data and that must be managed to improve the value to the business.
Data Catalog	A metadata repository or inventory of information about data. Typical data catalogs contain inventories of data stores, reports associated with the data stores, business rules and information about the ownership of data.
Data Classification	The process of categorizing data to enable the organization to effectively handle different types of data, locate data, grant access to the data and communicate the rules associated with compliance and regulatory concerns. Typical classifications include highly confidential, confidential, sensitive and public and require different handling rules depending on the classification.
Data Cleansing	A single field of data associated with a data store, information system or application through a data file or database structure.
Data Collection	A set of data that has been gathered and stored through systems and processes in order to satisfy requirements.
Data Definitions	The specification of data, including definition, domain values, and business rules.
Data Element	A single field of data associated with a data store, information system or application through a data file or database structure.
Data Governance	The execution and enforcement of authority over the definition, production and use of data and data-related assets. These assets typically focus on metadata about how the data is defined, produced and used.
Data Handling Rules	Rules associated with data availability including how data can be talked about, printed, shared (physically or electronically), stored and eliminated. Data handling rules are often directly associated with how the data is classified.
Data Initiative	A project or a work effort focused on improving the value received from the data. Typical data initiatives include new system or data store development or maintenance, software package implementation, data integration and transformation efforts, data quality improvement efforts, development of analytical environments and efforts to improve understanding and shareability of data.
Data Interoperability	The ability of systems and services that create, exchange and consume data to have clear, shared expectations for the contents, context and meaning of that data, such that data can be easily exchanged.

Data Inventory	The formal collection and recording of all data assets and data stores available for business purposes. A formal data inventory typically records the name, location, ownership, stewardship and use of each data asset and data store.
Data Management	The development and execution of architectures, policies, practices, and procedures to manage information lifecycle needs in an effective manner.
Data Model	An organization of data concepts (metadata) that is intended to clarify and specify the fundamental data needed to satisfy business objectives.
Data Modifications	Changes made to the metadata and data content, usually to improve usability of the data as an input to another system.
Data Owners	The person responsible for a business data asset , typically the line of business individual who is in charge of a business process or application data store.
Data Production	Processes deployed to create data that is collected, stored, transmitted between systems, and used by data consumers.
Data Profiling	The analysis of information for use in a specific data store to clarify the structure, content, relationships, and derivation rules of the data. Profiling helps to understand anomalies and assess data quality, but also to discover, register, and assess enterprise metadata.
Data Protection Rules	Requirements for processes employed to control data content to ensure it will meet privacy and other security requirements.
Data Provisioning	Ensuring that data provided meets requirements.
Data Quality	The practice of improving the data to ensure business requirements are met through assessment and application of the dimensions of quality including data accessibility, accuracy, completeness, conformity, integrity, reasonableness, timeliness, uniqueness and validity.
Data Quality Assessment	Evaluation of data that support the development and attainment of predefined quality expectations and measure data quality for the most important business data attributes, organized by subject area.
Data Quality Dimensions	The language of data quality facilitates a shared understanding that is essential for ensuring business requirements are met by established quality criteria, measuring defects, and communicating within and across organizations priorities, measurement scores, thresholds, and targets. Dimensions typically include the following quantifiable concepts: accuracy, completeness, validity, timeliness, integrity, consistency, conformity, and uniqueness.
Data Quality Requirements	Definition of the quality of the data needed to achieve business objectives. Data requirements should be stated in approved data quality dimensions that are stored and available from an authoritative business glossary, and consistently expressed in requirements documentation.
Data Relationships	Codified database rules that are derived from business processes.

Data Requirements	Definition of the data needed to achieve business objectives. Data requirements should be stated in business language and should reuse any approved, available standard business terms from an authoritative business glossary.
Data Risk	Increased complexity leading to rework, higher maintenance costs, and a data landscape that is harder to navigate.
Data Sharing	Data sourced from one or more systems that is utilized for multiple business purposes. If the organization also includes a technical specification for formats and methods, this can be referred to as data interoperability.
Data Standards	The accepted definition, formatting, values and usage rules associated with a specific data element. Data standards can be defined and enforced internally or externally to the agency.
Data Steward	A person with formal accountability for the definition, production and use of data. Depending on the stewardship approach, these individuals may be assigned, identified or recognized for their relationship for data. Different levels of stewards exist including operational stewards who define, produce and use data as part of their everyday job, and tactical stewards who are recognized as subject matter experts associated with a domain or subject matter of data.
Data Store	A data store is a repository for storing and managing collections of data. Data stores are most often associated with an information system, packaged software application, or database capability including master data solutions, data warehouses and open data sets.
Data Working Group	A group of selected individuals who are stakeholders in a specific set of data or application of data. They meet regularly to discuss, document and resolve data issues and improve the overall quality and value of that set of data. Working groups traditionally focus on data quality, data understanding, value or protection of this set of data.
Database	A structured set of data stored in a computer.
Database Administrators	Ensures that a database is available as needed and maintains security, monitoring, troubleshooting and planning for growth.
Entity Type	A data object that refers to a person, place, thing, concept, or event of interest to the organization, about which information is kept. Examples: Employee, Customer, Quarterly Report, Annual Conference.
Information System	An organized system for the collection, organization, processing, storage, and communication of information.
Integrity	"accuracy of the data's relationships to other data, for instance, if a stock purchase order is canceled, there should be an original buy order – this dimension is also known as Parent / Child, or Referential Integrity"

Master Data Management	The method used to consistently define and manage critical data of an organization to provide, with data integration, a single point of truth or reference. Data that is mastered may include reference data - the set of permissible values, and the analytical data that supports business intelligence and decision making.
Metadata	Data about data. Metadata is information stored in information technology tools that improved the business and technical understanding of data and data-related resources such as databases, table, reports, queries, ...
Operational Metadata	Represents information about the origins, movement, and history of the data, for example, capturing a modify date for a nightly data refresh, and database performance statistics. Process metadata is information about the outcome of automated loading and maintenance tasks, such as a system log.
Outcome-Driven Processes	Processes that are measured on the results or outcomes of a specific decision or set of actions. In terms of data management, outcome-driven data working group efforts are measured on the impact of their results on business operations.
Physical Data	Content stored in databases needed to perform business processes
Policy	A guiding principle typically established by senior management that is adopted by an organization to influence and determine activities and decisions.
Properties	Descriptions of information about data (metadata). For example, who is the data owner, what is the maximum range, etc.
Quality Data	Data is generally considered high quality if it is fit for its intended uses, in operations, decision making and planning.
Records Management	The practice of managing official records of an organization throughout their lifecycle, from creation through to archiving or deletion. It includes identifying, classifying, storing, securing, retrieving, tracking, and destroying or preserving records.
Security Controls	Safeguards or countermeasures to avoid, detect, counteract, or minimize security risk to physical data.
Shared Data	Data that can be used for more than one business function whether that function resides within one agency or across agencies. Shared data can be passed from one function to another function without the need for data manipulation. It is important that the rules associated with the data must be shared as well as part of the process of sharing data for business use.
Sprints	A series of mostly self-contained functions with the data they use, and the corresponding system behavior and screens required.
Stakeholder	An individual or organization that is affected by a decision about data, an application, a data store or a set of data.

Subtyping	Used to distinguish different levels of entity type. The 'supertype' contains attributes common to its subtypes, and each subtype may have distinguishing attributes applicable only to it. For example, for the supertype 'Person' there may be a subtype 'Employee' and another subtype 'Client.'
Table	A set of related data stored in cells that are organized by a column and row structure.
Technical Metadata	Represents information needed to access the data, for example, where the data resides, the structure of data in databases, interface logic, and servers.
Timeliness	A dimension of data quality pertaining to the degree to which the update schedule for data meets requirements.
Uniqueness	A dimension of data quality pertaining to the ability to assure that data is non-redundant.
User Stories	Describes how a user interacts with the system and how it behaves.
Validity	A dimension of data quality pertaining to the data's overall fitness for purpose has been verified by the business unit that creates and manages the data.
Waterfall Method	A software development approach that requires a linear, sequential flow of phases, whereby each phase must be completed before the next phase begins.