11ED
Glossary

Chapter 2

3 Vs	Three basic characteristics of Big Data databases: volume, velocity, and variety.
American National Standards Institute (ANSI)	The group that accepted the DBTG recommendations and augmented database standards in 1975 through its SPARC committee.
attribute	A characteristic of an entity or object. An attribute has a name and a data type.
Big Data	A movement to find new and better ways to manage large amounts of Web-generated data and derive business insight from it, while simultaneously providing high performance and scalability at a reasonable cost.
business rule	A description of a policy, procedure, or principle within an organization. For example, a pilot cannot be on duty for more than 10 hours during a 24-hour period, or a professor may teach up to four classes during a semester.
class	A collection of similar objects with shared structure and behavior. A class encapsulates an object's data representation and a method's implementation. Classes are organized in a class hierarchy.
class diagram	A diagram used to represent data and their relationships in UML object notation.
class diagram notation	The set of symbols used in the creation of class diagrams.
class hierarchy	The organization of classes in a hierarchical tree in which each parent class is a superclass and each child class is a subclass.
client node	One of three types of nodes used in the Hadoop Distributed File System. The client node acts as the interface between the user application and the HDFS.
conceptual model	The output of the conceptual design process. The conceptual model provides a global view of an entire database and describes the main data objects, avoiding details.
conceptual schema	A representation of the conceptual model, usually expressed graphically.
connectivity	The classification of the relationship between entities. Classifications include 1:1, 1:M, and M:N.
constraint	A restriction placed on data, usually expressed in the form of rules.
Crow's Foot notation	A representation of the entity relationship diagram that uses a three-pronged symbol to represent the many sides of the relationship.
data definition language (DDL)	The language that allows a database administrator to define the database structure, schema, and subschema.
data manipulation language (DML)	The set of commands that allows an end user to manipulate the data in the database. The commands include SELECT, INSERT, UPDATE, DELETE, COMMIT, and ROLLBACK.
data model	A representation, usually graphic, of a complex ?real-world? data structure. Data models are used in the database design phase of the Database Life Cycle.
data modeling	The process of creating a specific data model for a determined problem domain.
entity	A person, place, thing, concept, or event for which data can be stored.
entity instance	In ER modeling, a specific table row.
entity relationship (ER) model (ERM)	A data model that describes relationships, 1:1, 1:M, and M:N, among entities at the conceptual level with the help of ER diagrams. The model was developed by P. Chen in 1975.

entity relationship diagram (ERD)	A diagram that depicts an entity relationship model's entities, attributes, and relations.
entity set	In a relational model, a grouping of related entities.
eventual consistency	A model for database consistency in which updates to the database will propagate through the system so that all data copies will be consistent eventually.
extended relational data model (ERDM)	A model that includes the object-oriented model's best features in an inherently simpler relational database structural environment.
Extensible Markup Language (XML)	A metalanguage used to represent and manipulate data elements. Unlike other markup languages, XML permits the manipulation of a document's data elements. XML facilitates the exchange of structured documents such as orders and invoices over the Internet.
external model	The application programmer's view of the data environment. Given its business focus, an external model works with a data subset of the global database schema.
external schema	The specific representation of an external view; the end user's view of the data environment.
Hadoop	A Java based, open source, high speed, fault-tolerant distributed storage and computational framework. Hadoop uses low-cost hardware to create clusters of thousands of computer nodes to store and process data.
Hadoop Distributed File System (HDFS)	A highly distributed, fault-tolerant file storage system designed to manage large amounts of data at high speeds.
hardware independence	A condition in which a model does not depend on the hardware used in the model's implementation. Therefore, changes in the hardware will have no effect on the database design at the conceptual level.
hierarchical model	An early database model whose basic concepts and characteristics formed the basis for subsequent database development. This model is based on an upside-down tree structure in which each record is called a segment. The top record is the root segment. Each segment has a 1:M relationship to the segment directly below it.
inheritance	In the object-oriented data model, the ability of an object to inherit the data structure and methods of the classes above it in the class hierarchy.
internal model	In database modeling, a level of data abstraction that adapts the conceptual model to a specific DBMS model for implementation. The internal model is the representation of a database as seen by the DBMS. In other words, the internal model requires a designer to match the conceptual model's characteristics and constraints to those of the selected implementation model.
internal schema	A representation of an internal model using the database constructs supported by the chosen database.
key-value	A data model based on a structure composed of two data elements: a key and a value, in which every key has a corresponding value or set of values. The key-value data model is also called the associative or attribute-value data model.
logical design	A stage in the design phase that matches the conceptual design to the requirements of the selected DBMS and is therefore software-dependent. Logical design is used to translate the conceptual design into the internal model for a selected database management system, such as DB2, SQL Server, Oracle, IMS, Informix, Access, or Ingress.
logical independence	A condition in which the internal model can be changed without affecting the conceptual model.
many-to-many (M:N or *..*) relationship	Associations among two or more entities in which one occurrence of an entity is associated with many occurrences of a related entity and one occurrence of the related entity is associated with many occurrences of the first entity.

MapReduce	An open-source application programming interface that provides fast data analytics services; one of the main Big Data technologies that allows organizations to process massive data stores.
method	In the object-oriented data model, a named set of instructions to perform an action. Methods represent real-world actions, and are invoked through messages.
name node	One of three types of nodes used in the Hadoop Distributed File System. The name node stores all the metadata about the file system.
network model	A data model standard created in the late 1960s that represented data as a collection of record types and relationships as predefined sets with an owner record type and a member record type in a 1:M relationship.
NoSQL	A new generation of database management systems that is not based on the traditional relational database model.
object	An abstract representation of a real-world entity that has a unique identity, embedded properties, and the ability to interact with other objects and itself.
object-oriented data model (OODM)	A data model whose basic modeling structure is an object.
object-oriented database management system (OODBMS)	Data management software used to manage data in an object-oriented database model.
object/relational database management system (O/R DBMS)	A DBMS based on the extended relational model. The ERDM, championed by many relational database researchers, constitutes the relational model's response to the OODM. This model includes many of the object-oriented model's best features within an inherently simpler relational database structure.
one-to-many (1:M or 1..*) relationship	Associations among two or more entities that are used by data models. In a 1:M relationship, one entity instance is associated with many instances of the related entity.
one-to-one (1:1 or 1..1) relationship	Associations among two or more entities that are used by data models. In a 1:1 relationship, one entity instance is associated with only one instance of the related entity.
physical independence	A condition in which the physical model can be changed without affecting the internal model.
physical model	A model in which physical characteristics such as location, path, and format are described for the data. The physical model is both hardware- and software-dependent.
relation	In a relational database model, an entity set. Relations are implemented as tables. Relations are related to each other through the sharing of a common entity characteristic.
relational database management system (RDBMS)	A collection of programs that manages a relational database. The RDBMS software translates a user's logical requests into commands that physically locate and retrieve the requested data. A good RDBMS also creates and maintains a data dictionary to help provide data security, data integrity, concurrent and easy access, and system administration to the data through a query language and application programs.
relational diagram	A graphical representation of a relational database's entities, the attributes within those entities, and the relationships among the entities.
relational model	Developed by E. F. Codd of IBM in 1970, it represented a major breakthrough for users and designers because of its conceptual simplicity. The relational model is based on mathematical set theory and represents data as independent relations. Each relation is conceptually represented as a matrix of intersecting rows and columns. The relations are related to each other through the sharing of common entity characteristics.
relationship	An association between entities.

schema	A logical grouping of database objects, such as tables, indexes, views, and queries, that are related to each other. Usually, a schema belongs to a single user or application.
segment	In the hierarchical data model, the equivalent of a file system's record type.
semantic data model	The first of a series of data models that more closely represented the real world, modeling both data and their relationships in a single structure known as an object. The SDM, published in 1981, was developed by M. Hammer and D. McLeod.
software independence	A property of any model or application that does not depend on the software used to implement it.
sparse data	A case in which the number of table attributes is very large but the number of actual data instances is low.
subschema	In the network model, the portion of the database seen by the application programs that produce the desired information from the data in the database.
table	A matrix composed of intersecting rows and columns that represents an entity set in the relational model.
tuple	In the relational model, a table row.
Unified Modeling Language (UML)	A language based on object-oriented concepts that provides tools such as diagrams and symbols to graphically model a system.