Chapter 1

Here are brief definitions of the database system terms you provided:

**Integrated database** A database that combines data from multiple sources into a single, unified view. This can make it easier to access and analyze data, and to identify relationships between different data sets.

**Enterprise** A large organization with multiple users and complex data needs. Enterprise databases are typically designed to be scalable, reliable, and secure.

**Metadata** Data about data. Metadata can describe the structure of a database, the types of data it contains, and the relationships between different data sets.

**Concurrent use** The ability of multiple users to access and modify a database at the same time. This is important for enterprise databases, which need to be able to support a large number of users.

**Query** A request for data from a database. Queries can be used to retrieve all of the data in a database, or to filter the data based on specific criteria.

**End user** A person who uses a database, but is not involved in its design or administration. End users typically use database applications to access and manage data.

**Data redundancy** The storage of the same data in multiple locations. Data redundancy can be useful for performance and reliability, but it can also lead to data inconsistency.

**Data consistency** The state of ensuring that all copies of the same data are identical. Data consistency is important for ensuring the accuracy and reliability of data.

**Integrity constraint** A rule that defines how data in a database can be structured and modified. Integrity constraints help to maintain data accuracy and consistency.

**Data encryption** The process of converting data into a format that is unreadable to unauthorized users. Data encryption is important for protecting sensitive data from unauthorized access.

**Economy of scale** The cost advantage that businesses achieve when they produce or purchase goods or services in large quantities. Economy of scale can be applied to database systems by using specialized hardware and software that is designed to handle large volumes of data.

**Backup** A copy of data that is used to restore the data in the event of a data loss or corruption. Backups should be stored regularly and in a safe location.

**Recovery log** A record of all changes made to a database. Recovery logs can be used to restore the database to a previous state in the event of a data loss or corruption.

**User view** A subset of a database that is designed for a specific user or group of users. User views can be used to restrict access to certain data or to present data in a specific format.

**Semantic model** A model that describes the meaning of the data in a database. Semantic models can be used to improve the accuracy and reliability of data queries

.

**SQL** Structured Query Language. A programming language that is used to interact with relational databases. SQL is the most widely used database language.

**XML** Extensible Markup Language. A standard format for representing data in a structured form. XML can be used to exchange data between different database systems and applications.

**Data mining** the process of extracting knowledge from large data sets. Data mining techniques can be used to identify patterns and trends in data that would be difficult or impossible to see manually.

**Data warehouse** A database that is designed to store and analyse large volumes of data. Data warehouses are typically used for business intelligence and reporting purposes.

**Big data** Data sets that are too large or complex to be processed by traditional data processing applications. Big data sets can be characterized by the 5 V's: volume, velocity, variety, veracity, and value.

**SQL database** A relational database that uses SQL to interact with data. SQL databases are the most widely used type of database.

**The 5 V’s of big data** The five V's of big data are volume, velocity, variety, veracity, and value.

* Volume: Big data sets are typically very large, containing petabytes or even exabytes of data.
* Velocity: Big data sets are often generated at a very high rate, such as in real time.
* Variety: Big data sets can come from a wide variety of sources, such as sensors, social media, and financial transactions.

Veracity: Big data sets can often be noisy or incomplete, and it can be challenging to ensure their accuracy and reliability.  
Veracity in the 5 V's of big data refers to the quality and trustworthiness of the data. It is important to ensure that big data is accurate and reliable before using it to make decisions.

One way to improve the veracity of big data is to use data quality checks. This involves identifying and correcting errors in the data.

Another way to improve veracity is to use data provenance. This involves tracking the source of the data and how it has been transformed over time. This can help to ensure that the data is still accurate and reliable when it is used.

Here are some real-world examples of the importance of veracity in big data:

* **Healthcare:** Big data is being used to develop new medical treatments and improve the quality of healthcare. However, it is important to ensure that the data is accurate and reliable before using it to make decisions about patient care.
* Value: Big data sets can contain valuable insights that can be used to improve business operations, customer service, and product development.

Chapter 2

**Operational data** Data that is used to support the day-to-day operations of a business. Operational data is typically stored in transactional systems, such as ERP and CRM systems.

**Corporate resource** An asset that a company uses to generate revenue or produce goods or services. Corporate resources can include tangible assets, such as land, buildings, and equipment, as well as intangible assets, such as intellectual property and trademarks.

**Metadata** Data about data. Metadata can describe the structure of a database, the types of data it contains, and the relationships between different data sets.

**Entity** A person, place, thing, or event that is of interest to a business. Entities are typically represented in a database as tables.

**Attribute** A characteristic of an entity. Attributes are typically represented in a database as columns in a table.

**Data item** A single piece of data. Data items are the smallest unit of data that can be stored in a database.

**Data aggregate** A summary of data that has been calculated from a larger set of data. Data aggregates are often used to improve the performance of database queries and to generate reports.

**Data record** A collection of data items that describe a single entity. Data records are typically stored in a database as rows in a table.

**Data file** A collection of data records that are stored together. Data files are typically stored on disk or tape.

**Data sublanguage** A subset of a database language that is used to perform specific operations on a database. For example, the SQL data sublanguage is used to query and manipulate relational databases.

**Prototype** A working model of a system that is used to test the system's design and functionality. Prototypes are often used to gather feedback from users and to make changes to the system before it is deployed.

**System tuning** the process of optimizing a system to improve its performance. System tuning can involve adjusting system parameters, making changes to the system's code, or upgrading the system's hardware.

**CASE** Computer-aided software engineering (CASE) is a set of tools and methodologies that are used to develop and maintain software systems. CASE tools can be used to generate code, automate tasks, and create documentation.

**Integrated data dictionary** A data dictionary that stores all of the metadata for a database in a single location. Integrated data dictionaries can help to improve the accuracy and consistency of metadata.

**Data dictionary synonym** an alternative name for a data dictionary object, such as a table, column, or view. Data dictionary synonyms can be used to make it easier for users to access and understand the database.

**Data dictionary homonym** A data dictionary object that has the same name as another data dictionary object, but is different in terms of its type or structure. Data dictionary homonyms can be confusing for users and can lead to errors.

**Data standards** A set of rules and guidelines that define how data should be represented and used. Data standards can help to improve the consistency and quality of data.

**DBA** Database administrator (DBA). A DBA is responsible for the administration of a database system. This includes tasks such as installing and configuring the database, managing database users and permissions, and backing up and restoring the database.

I hope this helps!