

Data Storage

The i2b2 data is stored in a relational database, usually either Oracle or SQL Server and always in a **star schema** format. A star schema contains one fact and many dimension tables. The fact table contains the quantitative or factual data, while the dimension tables contain descriptors that further characterize the facts. Facts are defined by concept codes and the hierarchical structure of these codes together with their descriptive terms and some other information forms the i2b2 ontology (also called metadata).

i2b2 ontology data may consist of one or many tables. If there is one table, it will contain all the possible data types or categories. The other option is to have one table for each data type. Examples of data types are: diagnoses, procedures, demographics, lab tests, encounter (visits or observations), providers, health history, transfusion data, microbiology data and various types of genetics data. All metadata tables must have the same basic structure. This document will discuss the case of using one ontology table that holds all data types.

The structure of the metadata is integral to the visualization of concepts in the i2b2 workbench as well as for querying the data. The next two sections are a representation of the i2b2 ontology table and a discussion of the fields therein.