i2b2 Data Mart & Ontology

This page is a brief overview of the i2b2 Data Mart and Ontology in relation to the Multi-fact table project. It only touches the surfaces of the i2b2 database schema and is not meant to be a full tutorial.

Data Storage

The i2b2 data is stored in one of three supported relational databases, and always in a star schema format. A star schema contains one fact and many dimension tables. This statement although true is no longer completely accurate for the i2b2. With the release of version 1.7.09, the i2b2 server can now support multiple fact tables. The standard i2b2 demo data is released following the star schema format. However, sites have the option to use more than one fact table in their own environments.

Ontology Data

i2b2 ontology data consists of one or more metadata 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 diverse types of genetics data.

All metadata tables must have the same basic structure. The structure of the metadata is integral to the visualization of concepts in the i2b2 workbench as well as for querying the data.

Data Mart

The i2b2 data mart is a data warehouse consisting of one fact table and several dimension tables that provide additional information about fields in the fact table.

What is a fact?

In healthcare, a logical fact is an observation on a patient. It is important to note that an observation may not represent the onset or date of the condition or event being described, but instead is simply a recording or a notation of something. For example, the observation of ‘diabetes’ recorded in the database as a ‘fact’ at a particular time does not mean that the condition of diabetes began exactly at that time, only that a diagnosis was recorded at that time (there may be many diagnoses of diabetes for this patient over time).

Fact Table

The fact table contains the basic attributes about the observation, such as the patient and provider numbers, a concept code for the concept observed, a start and end date, and other parameters described in this document. In the i2b2, the fact table is called **observation_fact**.

Important

In the i2b2, you are required to have one fact table called observation_fact.
When the multi-fact table feature is turned off or other fact tables are not defined in the metadata tables, the i2b2 server will continue to search the observation_fact table.

Dimension Table

Dimension tables contain further descriptive and analytical information about attributes in the fact table. A dimension table may contain information about how certain data is organized, such as a hierarchy that can be used to categorize or summarize the data. In the i2b2 data mart, there are several dimension tables that provide additional information about fields in the fact table.