



Data Warehouse

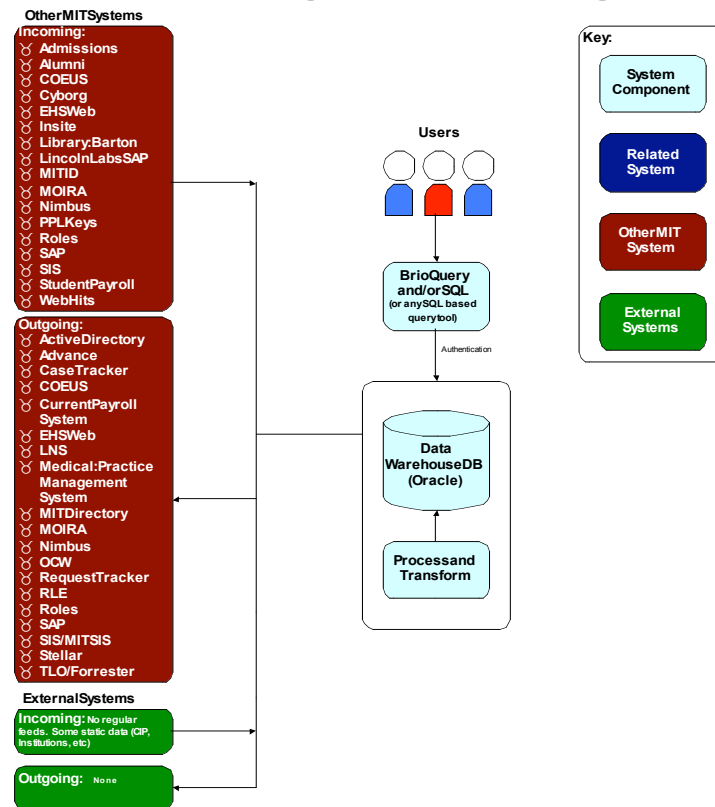
The Data Warehouse provides the MIT community with integrated data from various administrative systems stored in one location. It is a read only database and therefore, information represented in it is maintained by other systems, called "systems of record". The main purpose of the Data Warehouse is to serve as a central reporting and data distribution environment for Departments, Labs and Centers at MIT. The Data Warehouse acts as a hub, to facilitate the exchange of information between systems and therefore serves as the enterprise information infrastructure at MIT.

Reports can be created from the Data Warehouse using any SQL query tool. The tool most often used at MIT is BrioQuery. The Data Warehouse website contains standard reporting templates which can be downloaded and run by DLC users. Users can also create their own adhoc reports as needed. Reports and/or query results can be exported into other software programs, such as Microsoft Excel. DLC users can also join local data with data from the Data Warehouse to produce custom report of interest to their particular DLC.

There are various advantages of using the warehouse as a reporting tool. The Data Warehouse tables are specifically designed for end user adhoc reporting. Since the Warehouse is a central repository for MIT Data, a user can obtain all the information he/she needs from various systems at one place instead of having to obtain it from each system separately.

The operational hours of the Data Warehouse are as follows. The Data Warehouse is available seven days a week, except on Saturday afternoon from 6:00 -10:00 P.M. This time is used for weekly backups. The Normal User hours of the Warehouse are between 8:00 A.M. and 8:00 P.M daily. During these hours all data is stable and is not altered by any incoming feeds. Data is loaded daily into the Warehouse between 8:00 P.M. and 6:00 A.M. Although the Warehouse can be accessed during these hours, the data maybe in an inconsistent state since it is being refreshed by the system of records. Data is extracted from the Warehouse between 6:00 A.M. and 8:00 A.M. This time is set aside to allow systems refresh their data extracts from the Data Warehouse without affecting end user response time. Users may also access the Warehouse during this time, but the response time may be slow.

Data Warehouse Logical Architecture Diagram



Data Warehouse Physical Architecture Diagram

