Jenna Ficula

Professor Labouseur

Database Systems CMPT 308

<p align="center">Lab 1</p>

2. Short essay: Data vs. Information - Select a database in use today (real or imagined) and identify the elements of "data" stored therein and describe how the database organizes the "data" into "information". Give contrasting examples of "data" and "information" that illustrate the meaninglessness of "data" without context and organization. Talk about the value the "information" provides once the component data is given context.

Netflix is one of the most popular databases in use today. The "data" stored in the Netflix database are files for movies and television shows. Data becomes information when the files are given attributes such as genres, cast members, directors, length, descriptions, reviews, ratings, and even season numbers. All of this additional information groups the content into categories and provides context for the different files. If the data is not given context, Netflix would be meaningless list of endless movie and television show files with no organization. The website also needs the database to provide organization and meaning to the data and to make the website navigable for users. People would not want to search through endless files without titles, descriptions or genres.

It is because of the different context and organization that Netflix is able to provide users with catered recommendations and suggestions. If data was in the form of movie files without being organized into information, it would be nearly impossible to perform analytics. Context is necessary to connect the different files and analyze the user's personal preferences to be able to present recommendations bases off of prior user information. This information has tremendous value for a company like Netflix because they can cater content directly to their customers, present a more advanced user experience, and generate more revenue.

3. Short Essay: Data Models - Briefly describe the hierarchical and network pre relational data models. Explain their shortcomings in relation to the relational model.
Considering this, what do you think of XML as a model for data storage?

The hierarchical model is a database model that uses a tree shaped structure to depict a database. In this model every node relates to a single root. This model was the first to achieve physical data independence. To navigate the hierarchical model it is necessary to first start from the root and navigate downward to each of the nodes. This design is flawed because it is possible to have duplicate sets of data which is redundant. Due to its tree-like structure it is also slow to search for information that may reside lower on the model because it is necessary to search from top to bottom.

The network model is similar to a hierarchical model in the sense that they are graphed networks with roots which branch off into nodes. The difference with this model is that one node can be connected to more than one root. This model is not hierarchical in the sense that shared nodes exist and it is more web-like with more interconnectivity between nodes. Thus, the network model resolves the problem of duplicated and redundant data.

Both the network model and hierarchical model have shortcomings when compared to the relational data model because a relational model is based on the math of set theory. Through the use of relational calculus, relational algebra, and set theory, it is possible to prove the relational model. Relational models ensure more accurate data because of the mathematical support.

XML is a good model for storing smaller sets of data. However, relational databases can store more data and they are more reliable because they are less prone to redundancy. Relational databases are also easier to manipulate and can be more complex. Despite this fact, it is true that XML may be more useful when data is inherently hierarchical or when it is necessary to have data and its context all on a more centralized document.

pgAdmin III

File   Edit   Plugins   View   Tools   Help

Object browser
Server Groups
Servers (1)
PostgreSQL 9.5 (localhost:5432)

Properties   Statistics   Dependencies   Dependents

Properties
No properties are available for the current selection

SQL pane

Retrieving details on servers... Done.

2 msec

5:31 PM
9/1/2016