Megan Makini
Alan Labouseur
Database Management
7 September 2016

Lab 1

1. Data vs. Information

A database that I can think of is the one that the Registrar uses to store all of the data and information about courses, which students and faculty are able to access. It offers you the subjects, the instructors, the time slots, the names of the courses, and the descriptions of the courses. Everything a student needs to know about the course during registration is available because of Registrar's database. But when you look at the bits of data separately, it means nothing because there is no context. Take this class for example. The CRN number is 12238. What does that mean? It's an identification number clearly, but for what class, when is the class, who teaches the class, where is the class, does this ID number pertain to a class or is it just a random number? Presented independently, the number has no meaning or real value and vaguely represents anything. But as it is stored in the database it relates to a specific course, a specific time, a specific location, and a specific instructor. Other pieces of data coupled with the CRN number provide information about a course. The other bits of data are organized in classifications, as mentioned earlier, such as subjects, instructors, time, and the name of the courses. This organization gives the data meaning and transforms it into information. The data is structured to relate certain pieces of data in order to illustrate an individual course. For example, this class, CRN number 12238 relates to the instructor Alan Labouseur, location Hancock 2005, time Thursday 11:00 - 1:45. With the data organized and given context, a simple number now has meaning behind it.

2. Data Models

The hierarchical model is exactly how it sounds, a hierarchy. The data is organized in a "tree-based" model stemming from a "parent" and spreading out to connect to the "children" data. Each "parent" has its own "children" and do not share. Which causes the issue of duplication. As for the network model, it is generally the same structure but the "parent" data are able to share the "children" data, which eliminates the duplication issue. Although the issue of duplication was solved by the network model, the models possessed shortcomings such as the inability to support high-level query languages resulting in a considerable effort in writing programs for simple queries. While the relational model could express queries in a very high-level language, which increases the efficiency and decreases the amount of effort needed by a programmer. The relational model has the database system setup so that the data is

organized as tables called relations. Within the tables (‘relations’) there are unique columns and rows. The relational model achieves logical data independence.

As for XML (extensible modeling language), it is designed to carry data and focus on what the data is. It is up to the programmer to create the XML tags and the structure, allowing the programmer to have full liberty in writing but also simplifies everything to the point where the XML data can be exchange between systems of different sorts. The XML does advance data storage because it no longer limits the data to one system.